Dynamical systems

Oliver Knill

Harvard University, Spring semester, 2005

Abstract

This course Math 118r was taught in the spring 2005 at Harvard university. The first lecture took place Febrary 2 2005, the last lecture on May 6. There were 13 weeks. Except of the first week with an introduction and the last week with a final quiz and project presentation, the course covered each week a different and independent topic.

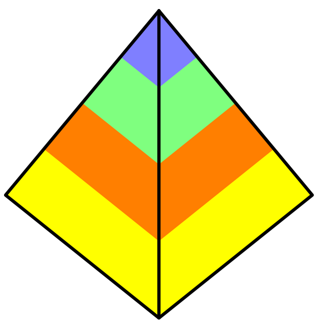
INTRODUCTION — Math118, O. Knill

ABSTRACT. We discuss the methodology and organization of the course.

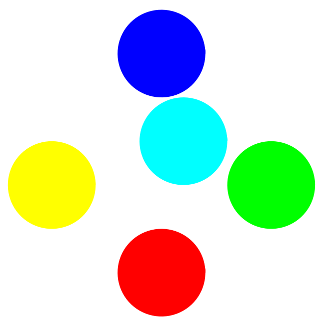
The subject. Dynamical system theory has matured into an independent mathematical subject. It is linked to many other areas of mathematics and has its own AMS classification which is 37-xx. The subject has grown so fast, that already specific subareas of dynamical systems like one-dimensional dynamics or ergodic theory have become independent research areas. As in other mathematical subjects, like topology, geometry or analysis which have "settled down", the teaching of the subject from the bottom up needs a lot of time. It makes more sense to study the subject by picking a few interesting subtopics.

The case method. This course is taught with an adaptation of the 'case method'. Each week, we pick a topic and use it to discuss some aspect of dynamical systems theory. The advantage of the 'case method' approach is that one can start early with mentioning open research topics. Furthermore, there are frequent fresh starts. We used this style for a course called "Mathematical Chaos Theory" in 1994 at Caltech, where an integral part were computer demonstrations using Mathematica and special software. It was also the first course, where I had used a course web-site.

The "case method" style has been used in Mathematics for a long time. Examples are the booklet **pearls of number theory** by Khinchin or Bowen's lectures in dynamical systems theory. The case method is a traditional Russian presentation style which can be found in many books. It is also used in research summer schools, where the breakup into different subjects and lectures comes naturally.



Systematic approach



Case approach

The history of a mathematical subject. Each part of the course has its own theme and flavor and is labeled by the name of a "protector", which is either a mathematician or physicist. We try to keep each subject independent of the others but of course, we will cross reference and relate to older topics. We also aim to give a glimpse into the history and gossip of the given subject. Because many different topics are covered, you will be able to get an idea, what dynamical systems is about and pick your favorite theme for a final project. which can either be of experimental or theoretical nature.

The level of difficulty. The course should be attractive for people who are interested in the applications of dynamical systems theory as well as for students, who want to see more mathematics beyond calculus. Some of the mathematical facts mentioned in class will be proven in full mathematical rigor and illustrated with live experiments in class. Participants of the course will be provided tools to experiment using online applications, computer algebra systems or using their own favorite programming language. No programming knowledge is required. More theoretically inclined or application oriented students will be given the opportunity to read some hand-picked survey articles if they wish.

Other fields Many introductory books on dynamical systems theory give the impression that the subject is about iterating maps on the interval, watching pictures of the Mandelbrot set or looking at phase portraits of some nonlinear differential equations in the plane. This is far from the reality. The topic can be seen as an interdisciplinary approach to many mathematical and nonmathematical areas. The field has matured and is successfully used in other fields like game theory, it is used to approach difficult unsolved problems in topology, and helps to see number theoretical problems with different eyes. There is hardly any mathematical field, which is not involved. For example: iterating smooth map or evolving smooth flows on manifolds is rooted in geometry, a sequence of independent random variables in probability theory can be modeled as a Bernoulli shift, the law of large numbers a special case of the ergodic theorem, the learning process in artificial intelligence can be seen as a discretized gradient flow. Dynamical systems are used heavily in number theory. For example, to understand the frequency of decimal digits occurring in the real number $\pi = 3.14159...$, where a dynamical systems approach looks the most promising one. The practical applications of the theory of dynamical systems are enormous: it ranges from medical applications like bifurcations of heartbeat patterns to explain the synchronous rhythmic flashing of fireflies. And then there are the obvious applications in population dynamics, fluid dynamics, quantum dynamics or statistical mechanics.

Prerequisites. To follow this course, a one semester multi-variable calculus like math21a, applied math21a, math23b, as well as a one semester of linear algebra course like math21b, applied math23b, math23b is required.

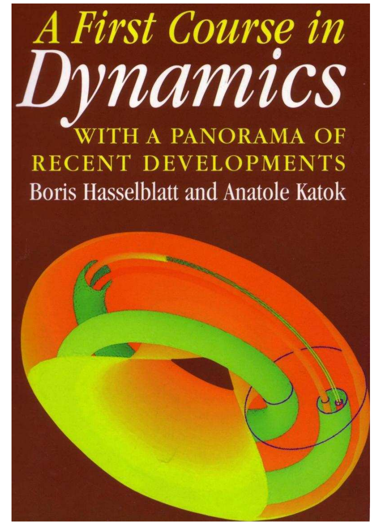
Exams. We plan to do several small quizzes. This, the homework, a final project and participation will make up the grade.

Syllabus. The difficulty and pace of the course will somehow be adjusted according to the audience. For a modular course like that, the theme structure allows an easy adaption of the pace.

- **1. Week: Introduction**.
  - What are dynamical systems?
  - Organization of the course
  - Examples of dynamical systems
- **2. Week: Feigenbaum: maps in one dimensions**.
  - Maps on the interval
  - Periodic points and their stability.
  - Bifurcation of periodic points
  - The dynamical zeta function
  - Invariant measures
  - The Lyapunov exponent
- **3. Week: Henon: maps in two dimensions**.
  - Area preservation
  - Periodic points and their nature
  - Stable manifold theorem and homoclinic points
  - Construction of stable manifolds
  - Lyapunov exponents and random matrices
  - Definitions of chaos
- **4. Week: Hilbert: Differential equations in two dimensions**
  - Differential equations in the plane and torus
  - Poincare-Bendixon theorem
  - Limit cycles
  - Hopf bifurcations
  - The Hilbert problem on limit cycles
- **5. Week: Lorentz: ODEs in higher dimensions**
  - Differential equations in space
  - The attractor in the Lorentz system
  - Forced oscillators
  - Lyapunov functions for ODE's
  - Strange attractors

- **6. Week: Birkhoff: billiards**
  - Billiards
  - The variational setup
  - Existence of periodic points
  - Polygonal billiards
  - Chaos for the stadium billiard
- **7. Week: Hedlund: cellular automata**
  - Curtis-Hedlund-Lyndon theorem
  - Topological entropy for CA
  - Attractors
  - Higher dimensional automata
  - Special solutions
- **8. Week: Mandelbrot: maps in the complex plane**
  - Mandelbrot and Julia sets
  - Basics of complex dynamics
  - Some topological notions
  - Connectivity of Mandelbrot set
- **9. Week: Bernoulli: subshifts of finite type**
  - Bernoulli shift
  - Subshifts of finite type
  - Sophic subshifts
  - Normal numbers and randomness
  - Normal numbers and randomness
- **10. Week: Weyl: dynamical systems in number theory**
  - Irrational rotation on the torus
  - Dirichlet theorem
  - Continued fractions
  - Diophantine lattice point problems
  - Unique and strict ergodicity
- **11. Week Poincare: many body problems**
  - The equations of the n-body problem
  - Integrals and the solution of the 2 body problem
  - The Sitnikov problem
  - Changing into rotating coordinate system
  - The planar restricted three body problem
  - Non-collision singularities and special solutions
- **12. Week: Einstein: geodesic flows**
  - Geodesic flows examples
  - Surfaces of revolutionn
  - surface billiards
  - Wave fronts and caustics
- **13. Week: Review**
  - Review
  - Open problems in dynamical systems
  - Projects

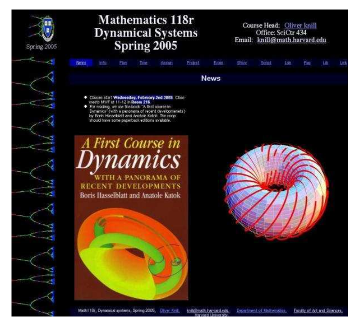
The book. It is important to have a 'second opinion' on things. We will not follow a book but the "First course in Dynamics, with a panorama of recent developments" by Boris Hasselblatt and Anatole Katok comes closest. It is written by leading experts in the area of dynamical systems.



More literature suggestions can be found on the course web-site.

The website. All of the material will be available on the course website:

http://www.courses.fas.harard.edu/ math118r.



2/2/04 WHAT ARE DYNAMICAL SYSTEMS? Math118, O.Knill

ABSTRACT. We discuss in this lecture, what dynamical systems are and where the subject is located within mathematics.

A FIRST DEFINITION.
The theory of dynamical systems deals with the **evolution of systems**. It describes **processes in motion**, tries to **predict the future** of these systems or processes and understand the **limitations of these predictions**.

RELEVANCE OF DYNAMICAL SYSTEMS.
To see that dynamical systems are relevant, one has just to look at a few news stories which broke during the last few weeks:

- Tsunami damage prediction
- Metor path computation
- Currents in the sea
- Landing of the Cassini probe on Titan
- Roulette ball prediction
- Statistics of digits in π
- Global earth temperature prediction

A FANCY DEFINITION.
Mathematically, any semigroup G acting on a set is a dynamical system. A **semigroup** $(G, \star)$ is a set G on which we can add two elements together and where the **associativity law** $(x \star y) \star z = x \star (y \star z)$ holds. The action is defined by a collection of maps T_t on X. It is assumed that $T_{t \star s} = T_t \circ T_s$, where $\star$ is the operation on G (usually addition) and $\circ$ is the composition of maps.

CLASSES OF DYNAMICAL SYSTEMS:

Time G (semigroup)	**Action**
Natural numbers $(N, +)$	Maps
Integers $(Z, +)$	Invertible maps
Positive real numbers $(R^+, +)$	Semiflows (some PDE's)
Real numbers $(R, +)$	Flows (Differential equations)
Any group $(G, \star)$	Representations
Lattice $(Z^n, +)$	Lattice gases, Spin systems
Euclidean space $(R^n, +)$	Tiling dynamical systems
Free group $(F_n, \circ)$	Iterated function systems

TWO IMPORTANT CASES OF ONE DIMENSIONAL TIME. We mention the general definition to stress that the ideas developed for one dimensional time generalize to other situations. Because physical time is one dimensional, the important cases for us are definitely **discrete and continuous dynamical systems**:

dynamics of **maps** defined by transformations	dynamics of **flows** defined by differential equations

DYNAMICAL SYSTEMS AND THE REST OF MATH. All areas of mathematics are linked together in some way or an other. Intersections of fields like algebraic topology, geometric measure theory, geometry of numbers or algebraic number theory can be considered full blown independent subjects. The theory of dynamical systems has relations with all other main fields and intersections typically form subfields of both.

Algebra	Measure theory	Analysis
Topology	Probability theory	Geometry
Logic	Dynamics	Number Theory

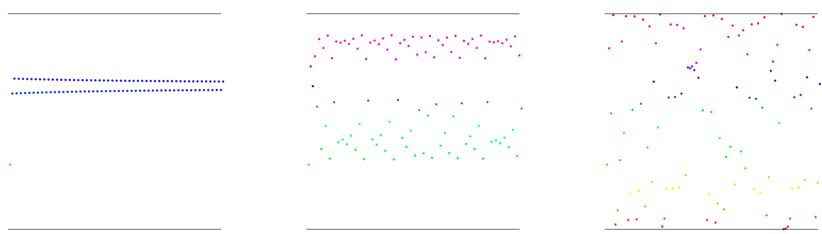
EXAMPLES OF INTERSECTIONS OF DYNAMICS WITH OTHER FIELDS:

- Link with **algebra**: group theorists often look at the action of the group on itself. The action of the group on vector spaces defines a field called **representation theory**.
- Link with **measure theory**: in **ergodic theory** one studies a map T on a measure space (X, μ). Measure theory is one foundation of ergodic theory.
- Link with **analysis**: the study of **partial differential equations** or **functional analysis** as well as **complex analysis** or **potential theory**.
- Link with **topology**: the **Poincare conjecture** states that every compact three dimensional simply connected manifold is a sphere. The problem is currently attacked using a dynamical system on the space of all surfaces which is called the **Ricci flow**.
- Link with **geometry**: **Kleins Erlanger program** attempted to classify geometries by its symmetry groups. For example, the group of projective transformations on a projective space. A concrete dynamical system in geometry is the geodesic flow. An other connection is the relations of partial differential equations with intrinsic geometric properties of the space.
- Link with **probability theory**: sequences of **independent random variables** can be obtained using dynamical systems. For example, with $T(x) = 2x \bmod 1$ and with the function f which is equal to 1 on $[0, 1/2]$ and equal to 0 on $[1/2, 1]$, $f(T^n(x))$ are independent random variables for most x.
- Link with **logic**: **logical deductions** in a proof or doing computations can be modeled as dynamical systems. Because every **computation** by a **Turing machine** can be realized as a dynamical system, there are fundamental limitations, what a dynamical system can compute and what not.
- Link with **number theory**: some problems in the theory of **Diophantine approximations** can be seen as problems in dynamics. For example, if you take a curve in the plane and look at the sequence of distances to nearest lattice points, this defines a dynamical system.
- A final link: a **category** X of mathematical objects has a semigroup G of **homomorphisms** acting on it (topological spaces have continuous maps, sets have arbitrary maps, groups, rings fields or algebras have homomorphisms, measure spaces have measurable maps). We can view each of these categories as a dynamical system. One can even include the category of dynamical systems with suitable homomorphisms. But this viewpoint is not a very useful in itself.

2/4/04 EXAMPLES OF DYNAMICAL SYSTEMS Math118, O.Knill

ABSTRACT. In this lecture, we look at examples of dynamical systems. Most examples in this zoo of systems belong to the "hall of fame". They are "stars" in the world of all dynamical systems and will appear later in this course.

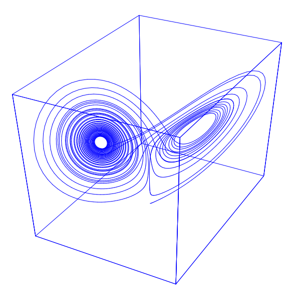
THE LOGISTIC MAP. $T(x) = cx(1-x)$. This is an example of an **interval map** . The **parameter** c is fixed in the interval $[0, 4]$. Lets look at some **orbits**. To compute an orbit say for $c = 3.0$, start with some **initial condition** like $x_0 = 0.3$, and iterate the map $x_1 = T(x_0) = 3x_0(1-x_0) = 0.63$, $x_2 = T(x_1) = 2x_1(1-x_1) = 0.6993$ etc. Lets do this with the computer. We show a few orbits for different parameters c. We always start with the initial condition $x_0 = 0.3$. Time is the horizontal axes and the interval $[0, 1]$ is on the vertical axes.



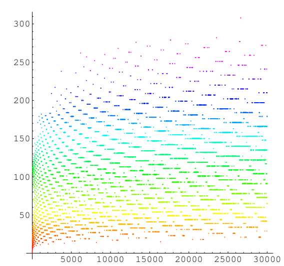
THE LORENTZ SYSTEM. The system of differential equations

$$
\begin{aligned}
\dot{x} &= 10(y-x) \\
\dot{y} &= -xz + 28x - y \\
\dot{z} &= xy - \frac{8z}{3}
\end{aligned}
$$

is called the **Lorentz system**. We see a numerically integrated orbit $(x(t), y(t), z(t))$ which is attracted by a set called the **Lorentz attractor**. It is an example of what one calls a **strange attractor**. Orbits behave chaotically on that set in the sense that one observes sensitive dependence on initial conditions The set is also measured to be a fractal, of dimension strictly between 1 and 2.

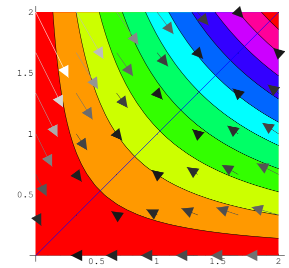


THE COLLATZ PROBLEM. Define a map T on the positive integers as follows. If n is even, then define $T(n) = n/2$, if n is odd, then define $T(n) = 3n + 1$. One believes that every orbit $n, T(n), T(T(n))$ will end up at 1 but one does not have a proof and there are people who think that mathematics is not ready for this problem. Theoretically, it would be possible that an orbit escapes to infinity, or that there exists a **periodic orbit** $n, T(n), T^2(n), ..., T^k(n) = n$. The problem is also called **Ulam problem** or $3n + 1$ **problem**. It is a notorious open problem. The picture to the right shows how long it takes to get from n to 1.



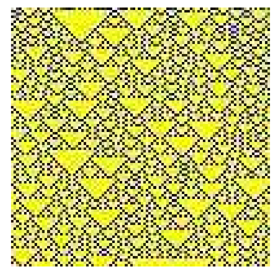
COMPUTING SQUARE ROOTS. Look at the map

$$T(x, y) = (\frac{2xy}{x+y}, \frac{x+y}{2})$$



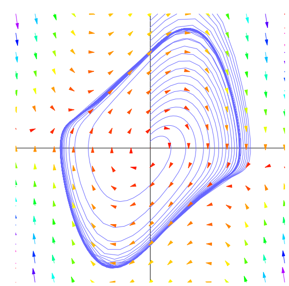
which assigns to two numbers a new pair, the **harmonic means** as well as the **algebraic mean**. You can easily check that the quantity $F(x, y) = xy$ is preserved: $F(T(x, y)) = F(x, y)$. It is called an **integral**. A map in the plane with such an integral is called an **integrable system**. All orbits converge to the line $x = y$ which consists of **fixed points**. Why is this useful? Start with $(1, 5)$ for example. The sequence (x_n, y_n) will converge to the diagonal and so to $(\sqrt{5}, \sqrt{5})$. Lets do it: we have $(1, 5), (\frac{5}{3}, 3), (\frac{15}{7}, \frac{7}{3}), (\frac{105}{47}, \frac{47}{21})$. We know that $\sqrt{5}$ is in the interval $[x_n, y_n]$ for all n. For example, $47/21 = 2.238...$ is already a good approximation to $\sqrt{5} = 2.236....$

CELLULAR AUTOMATA. Given a infinite sequence x of 0's and 1's, define a new sequence $y = T(x)$, where each entry y_n depends only on x_{n-1}, x_n, x_{n+1}. There are 256 different automata of this type. The picture below shows an orbit of "Rule 18". One of the interesting features of this automaton is that its evolution is linear on parts of the phase space. The nonlinear and interesting behavior is the motion of the **kinks**, the boundaries between regions with linear motion. A sequence x has a kink at n, if for some $k \geq 0$, $[x_{n-k}, \dots, x_{n+k+1}] = [1, 0, \dots 0, 1]$, like the pattern 10001.



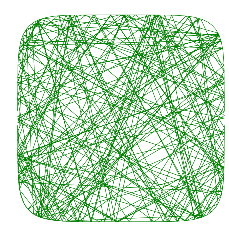
DIFFERENTIAL EQUATIONS IN THE PLANE. Second order differential equations can be written as differential equations in the plane. An example is the **van der Pool oscillator**

$$
\begin{aligned}
\frac{d}{dt}x &= y \\
\frac{d}{dt}y &= -x - (x^2 - 1)y;
\end{aligned}
$$



which shows a **limit cycle**. All orbits (except with the initial condition $(0, 0)$ converge to that limit cycle.

BILLIARDS. Let us take a table like the region $x^6 + y^6 \leq 1$. A ball reflects at the boundary. What is the long time behavior of this system? Is it possible that the angles a light ray makes with the boundary of the table become arbitrarily close to 0 and arbitrarily close to 180 degrees? Are there paths which come arbitrarily close to any point? The billiard flow defines a smooth map on the annulus. The study of this system has relations with elementary differential geometry. For example, the curvature of the boundary plays a role. The study of billiards is also part of a mathematical field called **calculus of variations** which deals with finding extrema of functions.

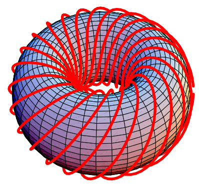


STANDARD MAP. The map

$$T(x,y) = (2x + \gamma \sin(x) - y, x)$$

on the plane is called the **Standard map**. Because $T(x+2\pi, y) = T(x, y+2\pi) = T(x,y)$, one can take both variables x, y modulo 2π and obtain a map on the **torus**. The real number γ is a parameter. The map appeared around 1960 in relation with the dynamics of electrons in microtrons and was first studied numerically by Taylor in 1968 and by Chirikov in 1969. The map can be completely analyzed for $\gamma = 0$. The map exhibits more and more "chaos" as γ increases. The picture to the right shows a few orbits in the case $\gamma = 1.3$.

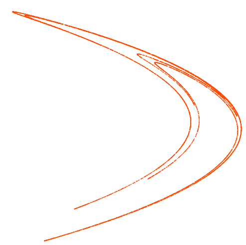
GEODESIC FLOW. Light on a surface takes the shortest possible path. These paths are called **geodesics**. On the plane, the geodesics are lines, on the round sphere, the geodesics are **great circles**, on a flat torus (see picture), the geodesics are lines too, but they wind around the surface. On some surfaces like surfaces of revolution or the ellipsoid, the geodesic flow can be analyzed completely on. On other surfaces, the flow can become very complicated. There are bumpy spheres on which each geodesic path is dense in the sense that the curve comes close to every point and also every direction at that point.



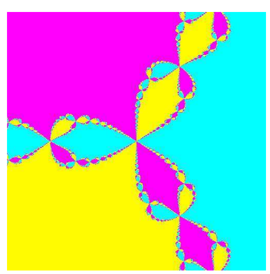
THE HENON MAP. One of the simplest nonlinear nonlinear maps on the plane is the **Henon map**

$$T(x,y) = (ax^2 + 1 - by, x) \ .$$

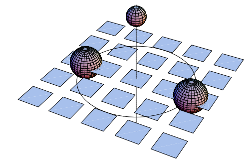
For $|b| = 1$ the map is area-preserving. For $|b| < 1$, it contracts area and produces **attractors**. The **Henon attractor** is obtained for $a = -1.4, b = -0.3$. The Henon map is equivalent to the **nonlinear recursion** $x_{n+1} = ax_{n-1}^2 + 1 - bx_{n-1}$. While **linear recursions** like the Fibonacci recursion $x_{n+1} = x_n + x_{n-1}$ can be solved explicitly using linear algebra, nonlinear recursions do no more lead to explicit formulas for x_n.



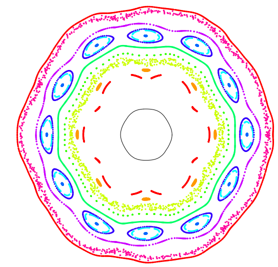
SOLVING EQUATIONS. To solve the equation $f(x) = 0$ numerically, one can start with an approximation x_0, then apply the map the **Newton iteration map** $T(x) = x - f(x)/f'(x)$. If $T(x) = x$, then $f(x) = 0$. As long as the root y satisfies $f'(x) \neq 0$, this algorithm works for x_0 near y. The method also works in the complex. In the case of several roots, is an interesting question to explore the **basin of attraction** of a root. The picture to the right shows this in the case of $f(z) = z^3 - 2$, where one has 3 roots. Depending on the initial point z_0, one ends up on one of the three roots. The Newton map for polynomials f defines a **rational map**. Its study is part of **complex dynamics**.



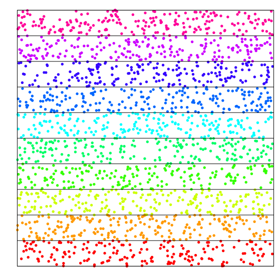
THREE BODY PROBLEM. Celestial mechanics determines very much the timing of our lives. Our calendar is based on it. While the motion of 2 bodies is understood well since Kepler, the **three body problem** is very complicated. Part of modern Mathematics, like topology have been developed in order to understand it. The **Sitnikov problem** is a **restricted three body problem** where the motion of a planet moves with negligible mass in a binary star system. The two suns circle each other on ellipses. The planet moves on the line through the center of mass, perpendicular to the plane in which the stars are located. For this system, there is a mathematical proof of some chaotic motion.



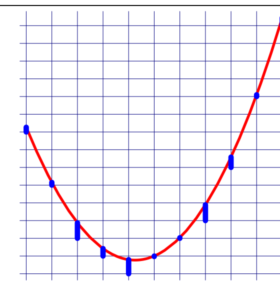
EXTERIOR BILLIARS. A geometrically defined dynamical system has been used to capture the main difficulties of the three body problem. The system is defined by a convex table as in billiards but this time, the a point outside the table is reflected at the table boundary: take the tangent to the table in the anti-clockwise direction and take the other point which has the same distance to the touching point. The map defined on the exterior of the table is area preserving and in general very complicated. It is not known, whether there exists a table for which there are unbounded orbits.



THE DIGITS OF PI. The digits of the number $\pi =$ 3.14159265358979323846264338327950288419716939937510... appear random. With $T(x) = 10x$ mod1 and $f(x) = [10x]$, where $[x]$ is the integer part of a number, the number $f(T^n(x))$ is the n'th digit of π. It appears that every digit appears with the same frequency and also all combinations of digit sequences. It is an open problem, whether this is true. One would call π normal. The picture to the right shows the sequence $x_n = f(T^n(x))$.



LATTICE POINTS NEAR GRAPHS. Given the graph of a function f on the real line, one can look at the distances to the nearest lattice points. This defines a sequence of numbers which can be generated by a dynamical system. For polynomials of degree n, the system is a map on the n dimensional torus. For the parabola $f(x) = ax^2 + bx + c$ we obtain which leads to a map of the type $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} x+\alpha \\ x+y \end{bmatrix}$ on the two dimensional torus.



2/11/2005: ENTROPY AND CHAOS   Math118, O. Knill

ABSTRACT. We look today at some notions of "chaos". One definition is the positivity of a number called the **positive entropy**, an other is the positivity of the Lyapunov exponent for every orbit which is not eventually periodic. An other definition is "chaos in the sense of Devaney". The Ulam map or the tent maps are examples for which we know that this type of chaos happens.

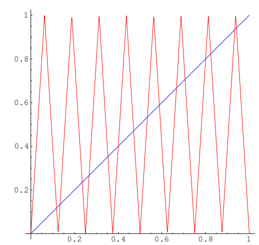
A DEFINITION OF CHAOS.
A purely topological notion of chaos which applies also to map with no differentiability is the **definition of Devaney**:

A map $T : [0,1] \to [0,1]$ is called **chaotic**, if there is a dense set of periodic orbits and if there exists an orbit which is dense.

A set Y is **dense** in $[0,1]$ if there is no interval which has empty intersection with Y.
EXAMPLES. a) the set of rational numbers is dense in $[0,1]$.
b) the set of irrational numbers is dense in $[0,1]$.
c) The set of numbers $\{1/n \mid n = 1,2,...\}$ is not dense in the interval.
d) Consider **Champernown's number** $x = 0.12345678910111213141516171819202122324...$ (do you see the pattern?), and the map $T(x) = 10x$ mod1. Then $T(x) = 0.2345678910111213141516171819202122324...$ and $T^2(x) = 0.345678910111213141516171819202122324...$ etc. Can you see why the orbit of x under the map T is dense in the interval $[0,1]$?

THE ULAM MAP IS CHAOTIC. We only state this theorem now. We will put later in this course put together some tools to prove it.

THEOREM. The map $f_4(x) = 4x(1-x)$ is chaotic in the sense of Devaney.



To have Devaney chaos, one needs to have an initial point, which visits each interval as well as to find for each interval a periodic orbit which visits that interval.
Because the Ulam map is conjugated to the tent map, we need only to prove the claim for the tent map.
In the homework, you see the density of the periodic points by understanding the graphs of the iterates of the map.

The problem to construct a dense periodic point will be solved later.

TOPOLOGICAL ENTROPY. Let $P_n(f)$ be the number of periodic points of true period n. Define the topological entropy of the map as

$$p(f) = \limsup_{n\to\infty} \frac{1}{n} \log |P_n(f)| ,$$

where the limits $p(f) = +\infty$ and $p(f) = -\infty$ are also allowed.
The topological entropy measures the growth of the number of periodic points. Similarly as the Lyapunov exponent, it measures how "complex" the map is.

EXAMPLES. The map defined by $\tilde{f}(x) = 2x$ mod 1 has the topological entropy $p(f) = \log(2)$ because $P_n(f) = 2^n$.

DYNAMICAL ZETA FUNCTION Related to topological entropy is the **dynamical ζ function** which is defined as

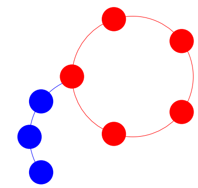
$$\log(\zeta_f(z)) = \sum_{n=1}^{\infty} \frac{P_n(f)}{n} z^n ,$$

where z is a real (or complex) variable. The series converges if $P_n(f)$ is finite for all n and for all complex numbers $|z| < e^{-p(f)}$. If $p(f) = -\infty$ and $P_n(f)$ is finite for all n, then $\zeta_f(z)$ is defined for all z.

EXAMPLE. The dynamical zeta function satisfies $\log(\zeta_f(z)) = \sum_{n=1}^{\infty} \frac{2^n}{n} z^n$. Because $\sum_{n=1}^{\infty} x^n = 1/(1-x) - 1$ and integration gives $\sum_{n=2}^{\infty} x^n/n = -\log(1-x) - x$ we have $\sum_{n=1}^{\infty} x^n/n = -\log(1-x)$. We see that $\log(\zeta_f(z)) = -\log(1-2z)$ and

$$\zeta_f(z) = 1/(1-2z) .$$

EVENTUALLY PERIODIC ORBITS. If an orbit has the structure $x_0, x_1, x_2, ..., x_m, x_{m+1}, x_{m+2}, ..., x_{m+n} = x_m$, it is called **eventually periodic**. Eventually periodic orbits appear often in dynamical systems which are not invertible.



EXAMPLES:
1) The point $x_0 = 1$ of the logistic map $f_c(x) = cx(1-x)$ is eventually periodic. It is actually eventually fixed. We have

$$x_0 = 1, x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0$$

2) The point $x_0 = 7/10$ is eventually periodic for $T(x) = 1 - 2|x - 1/2|$.

EVENTUALLY PERIODIC POINTS FOR THE TENT MAP.

THEOREM. x is eventually periodic if and only if $x = p/q$ is rational.

PROOF.
Since $T(x) = 2x$ or $T(x) = 2 - 2x$ we have

$$T(x) = integer + 2x$$

$$T^2(x) = integer + 2^2 x$$

$$T^n(x) = integer + 2^n x$$

If $T^n(x) = T^m(x)$ then $k + 2^n x = l + 2^m x$ so that $x = (k-l)/(2^n - 2^m)$ and x is a rational number.

(ii) To see the other direction, lets assume now that $x = p/q$ is rational. Then, $T(x) = 2p/q$ or $T(x) = 2 - 2p/q = 2(q-p)/q$. In any case, $T(x)$ is again of the form k/q. Repeating this argument shows that $T^n(x)$ is of the form k/q. There are only finitely many fractions of the form k/q and x therefore has to be eventually periodic.

REMARK. It needs a bit of combinatorial thought to figure out, when an orbit is eventually periodic and when it is actually periodic. Here is the answer (without a proof):

THEOREM. $x = p/q$ is periodic for the tent map if and only p is an even integer and q is an odd integer.

EXAMPLES:
1) $x = 4/5$ is a periodic point of period 2.
$x_0 = 4/5, x_1 = 2/5, x_2 = 4/5$ etc

2) $x = 5/7$ is an eventually periodic point.
$x_0 = 5/7, x_1 = 4/7, x_2 = 6/7, x_3 = 2/7, x_4 = 4/7$.

EVENTUALLY PERIODIC POINTS FOR THE ULAM MAP. The conjugation between the two maps T and S matches periodic points of T to periodic points of S and "eventually periodic points" of T with eventually periodic points of S.

EXAMPLE: Because $x_0 = 5/7$ is an eventually periodic point for the tent map, the point $y_0 = U^{-1}(5/7) = (1 - \cos(5\pi/7))/2$ is the initial condition for an eventually periodic point for the Ulam map.

$$\begin{array}{ccccc} x_0 = 5/7 & x_1 = 4/7 & x_2 = 6/7 & x_3 = 2/7 & x_4 = 4/7 \\ \uparrow \mathrm{U} & \uparrow \mathrm{U} & \uparrow \mathrm{U} & \uparrow \mathrm{U} & \uparrow \mathrm{U} \\ y_0 = 0.811745 & y_1 = 0.61126 & y_2 = 0.950 & y_3 = 0.1882 & y_4 = 0.61126 \end{array}$$

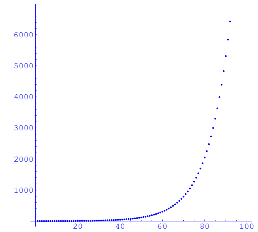
ABSTRACT. Our first dynamical system is the **logistic map** $f(x) = cx(1-x)$, where $0 \leq c \leq 4$ is a **parameter**. It is an example of an **interval map** because it can be restricted to the interval $[0,1]$.

You can read about this dynamical system on pages 14-16, pages 57-60, pages 198-199 as well as from page 299 on in the book. On this lecture, we have a first look at interval maps. We will focus on the logistic map, study periodic orbits, their stability as well as stability changes which are called bifurcations.

A FIRST POPULATION MODEL. In a simplest possible population model, one assumes that the population growth is proportional to the population itself. If x_n is the population size at time n, then $x_{n+1} = T(x_n) = x_n + ax_n = cx_n$ with some constant $a > 0$. We can immediately give a closed formula for the population x_n at time n:

$$x_n = T^n(x) = c^n x_0 .$$

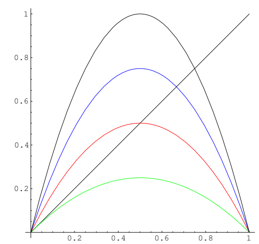
We see that for $c > 1$, the population grows **exponentially** for $c < 1$, the population shrinks exponentially.



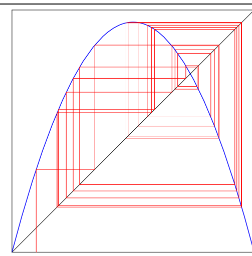
DERIVATION OF THE LOGISTIC POPULATION MODEL. If the population gets large, food becomes sparse (or the members become too shy to reproduce ...). In any case, the growth rate decreases. This can be modeled with $y_{n+1} = cy_n - dy_n^2$. Using the new variable $x_n = (c/d)y_n$, this recursion becomes

$$x_{n+1} = cx_n(1-x_n) .$$

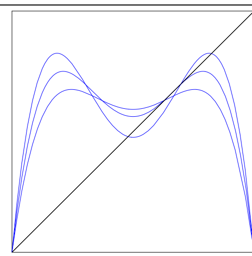
To the right, we see a few graphs of $f_c(x) = cx(1-x)$ for different c's. The intersection of the graph with the diagonal reveals **fixed points** of f_c. You see that 0 is always a fixed point. The graph has there the slope $f'(0) = c$. For $c > 1$, there exists a second fixed point $x = 1 - \frac{1}{c}$.



INTERVAL MAPS. If $f : [0,1] \to [0,1]$ is a **map** like $T(x) = 3x(1-x)$ and $x \in [0,1]$ is a point, one can look at the successive **iterates** $x_0 = x, x_1 = T(x), x_2 = f^2(x) = T(T(x)), ...$ The sequence x_n is called an **orbit**. If $x_n = x_0$, then x is called a **periodic orbit** of period n or n-cycle. If there exists no smaller $n > 0$ with $x_n = x$, the integer n is called the **true period**. A **fixed point** of f is a point x such that $f(x) = x$. Fixed points of f^n are periodic points of period n. The fixed points of f are obtained by intersecting the graph $y = f^n(x)$ with the graph $y = x$. The iterates of an **interval map** can visualized with a **cobweb construction**: connect (x,x) with $(x, f(x))$, then go back to the diagonal $(f(x), f(x))$ and iterate the procedure.

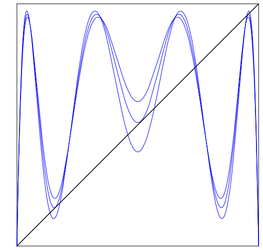


STABILITY OF PERIODIC POINTS. If x_0 is a fixed point of a differentiable interval map f and $|f'(x_0)| > 1$, then x_0 is **unstable** in the following sense: a point close to x_0 will move away from x_0 at first because linear approximation $T(x) \sim x_0 + f'(x_0)(x - x_0)$ shows that $|f(x) - x_0| \sim |f'(x_0)||x - x_0| > |x - x_0|$ near x_0. On the other hand, if $|f'(x)| = 1$, then x_0 is stable. For periodic points of period n, the stability is defined as the stability of the fixed point of f^n. The picture to the right shows situations, where $f'(x_0) < 1, f'(x_0) = 1, f'(x_0 > 1$ at a fixed point. The parameter at which the stability changes will be denoted a bifurcation.

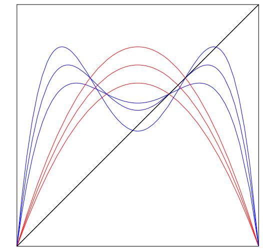


REMEMBER THE IMPORTANT FACT: if $f(x) = x$ is a fixed point of f and $|f'(x)| < 1$, then the fixed point is **stable**. It attracts an entire neighborhood. If $|f'(x)| > 1$, then the fixed point is **unstable**. It repells points in a neighborhood.

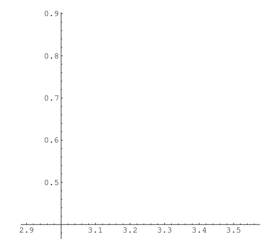
BIFURCATIONS. Let f_c be a **family of interval maps**. Assume that x_c is a fixed point of f_c. If $|f_c'(x_c)| = 1$, then c_0 is called a **bifurcation parameter**. At such a parameter, the point x_c can change from stable to unstable or from unstable to stable if c changes. At such parameters, it is also possible that new fixed points can appear. Different type of bifurcations are known: **saddle-node bifurcation** (also called **blue-sky** or **tangent** bifurcation). They can be seen in the picture to the right). **Flip bifurcations** (also related to **pitch-fork bifurcation**) lead to the **period doubling bifurcation** event seen below.



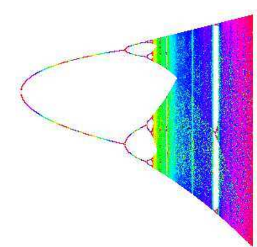
PERIOD DOUBLING BIFURCATION. Period doubling bifurcations happen for parameters c for which $(f_c^n)'(x_c) = -1$. The graph of f_c intersects the diagonal in one point, but the graph of f_c^2 which has slope 1 at x_0 starts to have three intersections. Two of the intersections belong to newly formed periodic points, which have twice the period. To the right, we see a simultaneous view of the graphs of f_c and f_c^2 for $c = 2.7, c = 3.0, c = 3.3$. You see that f_c keeps having one fixed point throughout the bifurcation. But f_c^2, which has initially one fixed point starts to have 3 fixed points! The middle one has minimal period 1, the other two are periodic points with minimal period 2.



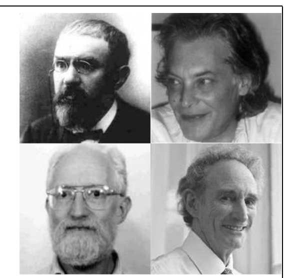
BIFURCATION DIAGRAM. The **logistic map** $f_c(x) = cx(1-x)$ always has the fixed point 0. For $c > 1$, there is an additional fixed point $x = 1 - 1/c$. Because $f'(0) = c$, the origin is stable for $c < 1$ and unstable for $c > 1$. At the other fixed point, $f'(x) = c - 2cx = c - 2c(1-1/c) = 2 - c$. It is stable for $1 < c < 3$ and unstable for $c > 3$. The point $c = 3$ is a bifurcation. It is called a **flip bifurcation**. Because a periodic point of period 2 is created, it is called a **period doubling bifurcation**. To see what happens with the periodic point of period 2, we look at $f^2(x) - x = c^2x(1-x)(1-cx(1-x)) - x$ which has the roots $(c + 1 \pm \sqrt{(c-3)(c+1)})/(2c)$ which are real for $c > 3$. Its stability can be determined with $(f^2)'(x) = f'(x)f'(f(x)) = 4 + 2c - c^2$. This shows that the 2-cycle is stable for $3 < c < 1 + \sqrt{6}$. At the parameter $1 + \sqrt{6}$ it bifurcates and gives rise to a periodic orbit of period 4.



FEIGENBAUM UNIVERSALITY. We have computed the first bifurcation points $c_1 = 3, c_2 = 1 + \sqrt{6}$. The successive period doubling bifurcation parameters c_k have the property that $\frac{c_{k+1} - c_k}{c_{k+2} - c_{k+1}}$ converges to a number $\delta = 4.69920166$. It was **Mitchell Feigenbaum**, who realized that this number is **universal** and conjectured how it could happen using a **renormalization picture**. In 1982, **Oscar Lanford** proved these Feigenbaum conjectures: for a class of smooth interval maps with a single quadratic maximum, the limit δ exists and is universal: that number does not depend on the chosen family of maps. It works for example also for the family $f_c(x) = c\sin(\pi x)$. The proof demonstrates that there is a fixed point g of a **renormalization map** $\mathcal{R}f(x) = \alpha f^2(\alpha x)$ in a class of interval maps. The object which is mapped by $\mathcal{R}$ is a map!



HISTORY. Babylonians considered already the rotation $f(x) = x + \alpha$ mod 1 on the circle. Since the 18th century, one knows the Newton-Rapson method for solving equations. Already in the 19th century **Poincaré** studied circle maps. Since the beginning of the 20th century, there exists a systematic theory about the iteration of maps in the complex plane (Julia and Fatou), a theory which applies also to maps in the real. In population dynamics and finance growth models $x_{n+1} = f(x_n)$ appereard since a long time. It had been popularized by theoretical biologists like Robert May in 1976. Periodic orbits of the logistic map were studied for example by N. Metropolis, M.L. Stein and P.R. Stein in 1973. Universaility was discovered numerically by Feigenbaum (1979) and Coullet-Tresser (1978) and proven by Lanford in 1982.



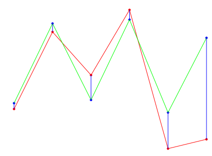
2/9/2005: LYAPUNOV EXPONENT — Math118, O. Knill

ABSTRACT. We demonstrate that the **logistic map** $f(x) = 4x(1-x)$ is chaotic in the sense that the Lyapunov exponent, a measure for sensitive dependence on initial conditions is positive.

LYAPUNOV EXPONENT. For an orbit of f with starting point x, we define the **Lyapunov exponent** as

$$\boxed{\lambda(f,x) = \lim_{n\to\infty} \frac{1}{n} \log |(f^n)'(x)|}$$

where $(f^n)'$ is the derivative of the n'th iterate (f^n). Remark. It turns out that usually, the limit exists. If not, one should replace lim with lim inf, the smallest accumulation point of the sequence. Choosing lim inf instead of lim sup has nicer analytic properties.



A BETTER FORMULA. The function $f^n(x)$ becomes complicated already for small n. The following formula is more convenient to compute the Lyapunov exponent of an orbit through x_0:

$$\boxed{(f^n)'(x) = f'(x_{n-1})...f'(x_1)f'(x_0)}$$

PROOF. Use induction: for $n = 1$, the claim is obvious. If we differentiate $f^n(x_0) = f^{n-1}(f(x_0))$ we get $(f^{n-1})'(x_1)f'(x_0)$, then use the induction assumption $(f^{n-1})'(x_1) = f'(x_{n-1})...f'(x_1)$. Therefore:

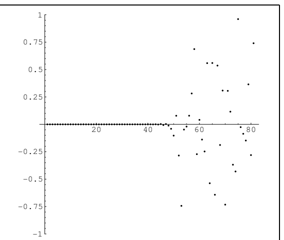
$$\boxed{\lambda(f,x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \log |f'(x_k)|}$$

EXAMPLE. For the logistic map, we compute the Lyapunov exponent by taking a large n and form

$$\frac{1}{n} [\log|c(1-2x_{n-1})| + \log|c(1-2x_{n-2})| + \cdots + \log|c(1-2x_0)|] \; .$$

WHAT DOES THE LYAPUNOV EXPONENT MEASURE? If x and y are close, then $|f(y)-f(y)| \sim |f'(x)||x-y|$, if x and y are close. because Taylors formula assures $f(y) = f(x) + f'(x)(y-x)$ plus something of the order $(y-x)^2$. If x_n is the orbit of x and y_n is the orbit of y, then for a fixed n, we have $|x_n - y_n| \sim |(f^n)'(x)||x_0 - y_0|$ if x_0 and y_0 are close together.
The Lyapunov exponent is a quantitative number which indicates the **sensitive dependence on initial conditions**. It measures the exponential rate at which errors grow. If the Lyapunov exponent is $\log|c|$ then you can expect an error $c^n\epsilon$ after n iterations, if ϵ was the initial error.

EXAMPLE. We will see below that the Lyapunov exponent of $f(x) = 4x(1-x)$ is $\log|2|$. If your initial error is $\epsilon = 10^{-16}$, then we have after n iterations an error $2^n\epsilon$ which is of the order 1 for $n = 53$. To the right we see the difference $x_n - y_n$ between two orbits of the map $f = f_4$ which have an initial condition $|x_0 - y_0| = 10^{-16}$. You see that after about 50 iterations, the error has grown so much that it becomes visible.



LYAPUNOV EXPONENT OF PERIODIC ORBIT. If $x_0, x_1, ..., x_n = x_0$ is a periodic orbit of period n, then

$$\lambda(f,x) = \frac{1}{n}(f^n)'(x) = \frac{1}{n}(\log|f'(x_{n-1})| + \log|f'(x_{n-2})| + ... + \log|f'(x_0)|) \; .$$

PROOF. We we have to show that the sequence $s_k = \frac{1}{k}(\log|f'(x_0)| + \log|f'(x_1)| + ... + \log|f'(x_k)|)$ converges to the right hand side which is s_n. If k is a multiple of n, then $s_k = s_n$. If M is the maximum of all the numbers $\log|f'(x_i)|$, then $|s_j| \leq jM$ for $k = 1, ..., n$ and $s_k - \lambda(f,x) \leq (nM)/k$.

EXAMPLES.

- The Lyapunov exponent of the fixed point 0 is $\log(c)$. It is negative for $c < 1$ and positive for $c > 1$.
- The Lyapunov exponent of the fixed point $1 - 1/c$ of the logistic map f_c is $\log|f'(1-1/c))| = \log|2-c|$.

LYAPUNOV EXPONENT OF AN ATTRACTIVE PERIODIC ORBIT.
The Lyapunov exponent of an attractive periodic orbit is negative.

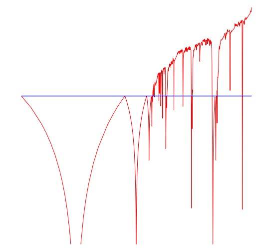
PROOF. We have $\lambda(f,x) = \frac{1}{n}\log|(f^n)'(x)|$. We have seen that for an attractive periodic orbit, $|(f^n)'| < 1$.

It follows that the Lyapunov exponent of an orbit which is attracted to a periodic orbit is negative too.

LYAPUNOV EXPONENT AND BIFURCATION. A periodic point can only bifurcate if its Lyapunov exponent is zero.

LYAPUNOV EXPONENT OF THE LOGISTIC MAP.
The picture to the right shows the Lyapunov exponent of an orbit starting at x_0 in dependence of c. You see that this graph looks very complicated. If the Lyapunov exponent is negative, we typically have an attractive periodic orbit.

It is difficult to say something about the Lyapunov exponent of a specific parameter. We know what happens for $c = 4$ and we know what happens in case of an attractive periodic orbit. If an attractive periodic orbit exists, there is an entire interval, where the Lyapunov exponent is negative. It has only recently been shown that there is a dense set of parameters for which the Lyapunov exponent is negative. This means, we don't find a single interval in $[0,4]$ on which the Lyapunov exponent is positive.



CONJUGATION OF MAPS. Two interval maps T and S are **conjugated**, if there exists an invertible map U from the interval onto itself such that $T(U(x)) = U(S(x))$. If both maps are differentiable maps, one usually requires the map U to be smooth too.

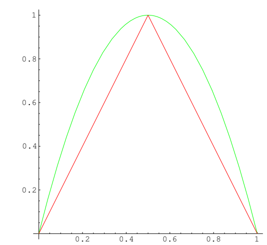
> COROLLARY: The Lyapunov exponents of corresponding orbits of two conjugated interval maps are the same.

More precisely, if $\lambda(f,x)$ is the Lyapunov exponent of the orbit of f through x, and $\lambda(g,y)$ is the Lyapunov exponent of the orbit of g through $y = h(x)$, and $gh(x) = hf(x))$ is the conjugation, then $\lambda(f,x) = \lambda(g,y)$.

Proof: This is an application the chain rule.

ULAM AND TENT ARE CONJUGATED MAPS: the Ulam map $T(x) = 4x(1-x)$ is conjugated to the tent map $S(x) = 1 - 2|x - 1/2|$ with the conjugation $U(x) = \frac{1}{2} - \frac{1}{\pi}\arcsin(1-2x)$ and $U^{-1}(x) = \frac{1}{2} - \frac{1}{2}\cos(\pi x)$.

PROOF. To check that $UTU^{-1}(x) = S(x)$, we show $UT(x) = S(U(x))$. One can get rid of the absolute value by distinguishing the cases $x > 1/2$ and $x < 1/2$. We have $U(T(x)) = \frac{1}{2} - (\arcsin(1 - 8(1-x)x)))/\pi$ and $SU(x)) = 1 + 2(\arcsin(1-2x))/\pi$ for $x \in [1/2, 1]$. To verify the identity, we check that both sides are 1 for $x = 1/2$ and that $\frac{d}{dx}\arcsin(1 - 8x + 8x^2) = -\frac{d}{dx} 2\arcsin(1-2x)$. The last identity is best checked by squaring both sides and using $\arcsin'(x) = 1/\sqrt{1-x^2}$. The identity on $[0, 1/2]$ is solved in the same way.



LYAPUNOV EXPONENT OF THE ULAM MAP.

> THEOREM. For all but a countable set of initial conditions x_0, the Lyapunov exponent of $f(x) = 4x(1-x)$ with initial condition x_0 is equal to $\log(2)$.

The tent map $S(x) = 1 - 2|x - 1/2|$ is piecewise linear. The derivative $S'(x)$ is either 2 or -2. Since $\log|S'(x_k)| = \log(2)$, the map has the Lyapunov exponent $\log(2c)$ for orbits, which do not hit one of the discontinuities. Most initial points do not hit the discontinuity because there is only a countable set of initial conditions for which this can happen.

Because the map is conjugated to T_4, the Lyapunov exponent of f_4 is $\log(2)$ too by the corollary.

PERIODIC POINTS AND LYAPUNOV EXPONENT — Math118, O. Knill

ABSTRACT. After distinguishing different types of periodic orbits of two dimensional maps, we look at the possible nature of periodic points and distinguish between elliptic, parabolic and hyperbolic cases, sources and sinks. We further introduce the Lyapunov exponent.

PERIODIC POINTS AND LINEARIZATION. Fixed points of the map T in the plane are called periodic points of period 1. Fixed points of T^n are periodic points of period n.

The Jacobean $DT(x,y) = T'(x,y)$ of a fixed point (x_0,y_0) plays an important role. It defines a linear map A which is called the **linearization** of T at the fixed point.

THEOREM. Near a fixed point (x_0,y_0), the map $T(x,y)-(x_0,y_0)$ is close to $DT(x_0,y_0)(x-x_0,y-y_0)$.

PROOF. The functions $f(x,y)$ and $g(x,y)$ have a Taylor expansion like $f(x,y) = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x,y)(y-y_0) + f_{xx}(x_0,y_0)(x-x_0)^2/2 + f_{xy}(x_0,y_0)(x-x_0)(y-y_0) + f_{yy}(x_0,y_0)(y-y_0)^2/2 + \dots$. Terms like $(x-x_0)^2$ are small near the fixed point (x_0,y_0).
It follows that that if we iterate only for a fixed number of points, we can approximate the real map with the linearized map. However, because orbits will in general move away from the fixed point, where the linearization will no more be a good approximation, we can not expect a global correspondence. We will see that under some conditions, we can deduce something from the knowledge of the linearization. It also follows from this result that T is invertible if $\det(DT(x,y)) \neq 0$ for all point (x,y).

EXAMPLE: THE STANDARD MAP. The map $T(x,y) = (2x + c\sin(x) - y, x)$ is a map on the plane. It can also be considered a map on the torus because $T(x+2\pi, y) = T(x,y) + (4\pi, 2\pi)$, $T(x, y+2\pi) = T(x,y) + (-2\pi, 0)$. The map is called the Standard map. The Jacobean matrix at a point (x,y) is

$$DT(x,y) = T'(x,y) = \begin{bmatrix} 2 + c\cos(x) & -1 \\ 1 & 0 \end{bmatrix} .$$

Because the determinant of the Jacobean is 1 at all points, the map is area-preserving for all parameters c.

At the fixed point $(0,0)$, the Jacobean matrix is

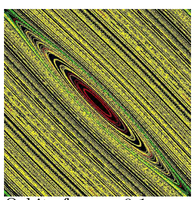
$$T'(0,0) = \begin{bmatrix} 2+c & -1 \\ 1 & 0 \end{bmatrix}$$

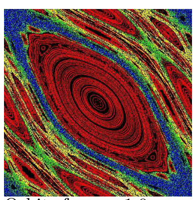
The eigenvalues are real and different for $c > 0$.

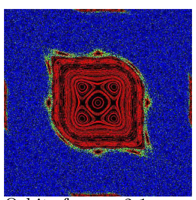
At the fixed point (π,π), the Jacobean matrix is

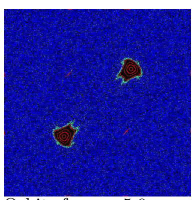
$$T'(\pi,\pi) = \begin{bmatrix} 2-c & -1 \\ 1 & 0 \end{bmatrix}$$

λ_i imaginary for $0 < c < 4$ and real for $c > 4$.


Orbits for $c = 0.1$.


Orbits for $c = 1.0$.


Orbits for $c = 2.1$.


Orbits for $c = 5.0$.

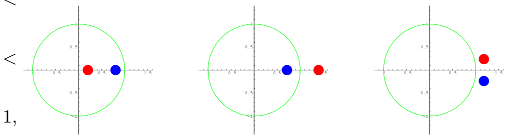
THE STABILITY QUESTION.

For nonlinear dynamical systems, the question of stability of fixed points can be very difficult. A pioneer in stability theory was Aleksandr Lyapunov (1857-1918). It turns out that already for simple cases like the Henon map or the Standard map, the stability of points, where the linearization is a rotation is difficult to establish. In the case, when the eigenvalues are real and both have not absolute value 1, then one can conjugate the map near the fixed point to its linearization. In those cases, the linearized picture essentially gives the real picture near the fixed point.

EIGENVALUES OF LINEAR MAPS. The characteristic polynomial of a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $f_A(\lambda) = \lambda^2 - \mathrm{tr}(A)\lambda + \det(A)$. If $c \neq 0$, the eigenvalues are $\lambda_\pm = \mathrm{tr}(A)/2 \pm \sqrt{(\mathrm{tr}(A)/2)^2 - \det(A)}$.

TYPICAL FIXED POINTS. If $T(x,y)$ is a differentiable map and $T(x_0,y_0) = (x_0,y_0)$ is fixed point with Jacobean $A = DT(x_0,y_0)$. Using the eigenvalues λ_1, λ_2 of A, we define the following **typical cases**, typical in the sense that the property is stable under small changes of parameters of the map:

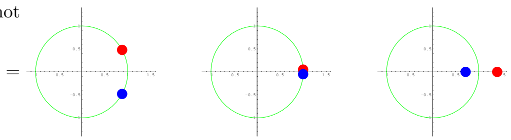
- **hyperbolic sink** $|\lambda_1| < 1, |\lambda_2| < 1$.
- **hyperbolic saddle** $|\lambda_1| < 1 |\lambda_2| > 1$
- **hyperbolic source** $|\lambda_1| > 1$, $|\lambda_2| > 1$.



EXAMPLE. Fixed points of the quadratic Henon map $T(x,y) = (1 - ax^2 - y, bx)$ are of the form (x, bx). Lets look at the case $a = 1.4, b = 0.3$. Solving $1 - ax^2 - bx = x$ gives the fixed points $(-10/7, -3/7)$ and $(1/2, 3/20)$. At the fixed point $(-10/7, -3/7)$ the eigenvalues are $\lambda_1 = (20 + \sqrt{370})/10 = 3.92..$ and $\lambda_2 = (20 - \sqrt{370})/10 = 0.07646...$ At the fixed point $(1/2, 3/20)$ the eigenvalues are $\lambda_1 = (-7 - \sqrt{19})/10 = -1.13589, \lambda_2 = (-7 + \sqrt{19})/10 = -0.26411$. We see that both fixed points are hyperbolic.

TYPICAL FIXED POINTS OF AREA-PRESERVING MAPS. If $\det(DT(x,y)) = 1$ for all (x,y) then T is area-preserving by the change of variable formula. In that case $\lambda_1\lambda_2 = 1$ and sinks or sources are no more possible. Cases with $|\lambda_i| = 1$ can now persist under parameter changes, if the deformation happens in the class of area-preserving maps. We distinguish now between the following cases:

- **elliptic** $|\lambda_1| = |\lambda_2| = 1$, λ_i not real.
- **parabolic** $\lambda_1 = \lambda_2 = -1$ or $\lambda_1 = \lambda_2 = 1$.
- **hyperbolic** $|\lambda_1| < 1 |\lambda_2| > 1$



Parameter values, for which a periodic orbit changes from hyperbolic to elliptic or in the other direction are called **bifurcation parameters**.

THEOREM. A fixed point of an area preserving map is elliptic if $|\mathrm{tr}(DT)| < 2$. It is parabolic if $|\mathrm{tr}(DT)| = 2$ and hyperbolic, if $|\mathrm{tr}(DT)| > 2$.

PROOF. Distinguish between the cases $\det(DT) = 1$ and $\det(DT) = -1$.

THE NORM OF A MATRIX. For a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ define the **norm** $||A|| = \sqrt{tr(AA^T)} = \sqrt{a^2 + b^2 + c^2 + d^2}$. Remember that A^T is the **transpose** of the matrix and $tr(A)$ denotes the **trace** of a matrix A.

Side remark. There are different ways to define the norm. The usual norm $||A|| = \max_{|v|=1} |Av|$ is known to be the square root of the largest eigenvalue of $A^T A$ but is less convenient to compute.

LYAPUNOV EXPONENT. The exponential growth rate of $||DT^n(x,y)||$ is

$$\boxed{\lambda(T,(x,y)) = \liminf_{n\to\infty} \frac{1}{n} \log ||DT^n(x,y)||}$$

is called the **Lyapunov exponent** of T at the point (x,y). For area preserving maps T on the torus, define $\lambda(T) = \lim_{n\to\infty} \frac{1}{n} \int\int \log ||DT^n(x)||\, dxdy$ is known to be to a quantity called the **entropy** of the map.

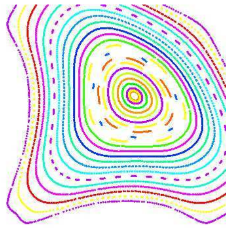
Examples:
a) If $T(x,y) = (x+\alpha, y+\beta)$, then the Lyapunov exponent is zero for every orbit.
b) If $T(x,y) = (2x+y, x+y)$ is the cat map on the torus, then the Lyapunov exponent is $\log(|3+\sqrt{5}|/2)$ for all orbits
c) In the case of the Standard map, one does not know the Lyapunov exponent for most orbits. One numerically measures an entropy $\geq \log(c/2)$.

2/14/2005: MAPS IN TWO DIMENSIONS — Math118, O. Knill

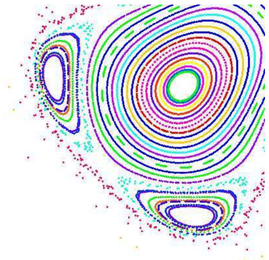
ABSTRACT. In this first lecture, we look at the dynamics of maps in the plane and introduce some terminology related to the Jacobian matrix $DT(x, y)$.

THE HENON MAP. A map of the form $T(x, y) = (-ax^2 + 1 - y, bx)$ with parameters b, c is called a **Henon map**. For $b = 0$, the map restricted to the first coordinate is the one-dimensional quadratic map $f(x) = -ax^2+1$ and not invertible. For $b \neq 0$, the map is invertible.

Orbits of $T(x, y) = (-1.5x^2 - 0.3y, x)$ accumulate on an attractor.



Orbits of the map $T(x, y) = (-0.5x^2 + 1 - y, x)$.



Orbits of the map $T(x, y) = (0.4x^2 + 1 - y, x)$.

THE JACOBEAN. Two smooth functions $f(x, y)$ and $g(x, y)$ of two variables, define a map

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f(x, y) \\ g(x, y) \end{bmatrix}$$

in the plane. We say, a map is **area preserving** or **conservative** if $T(A)$ has the same area then A for any rectangle A. We write partial derivatives as $f_x(x, y), f_y(x, y)$.

THEOREM. T is area-preserving if and only if the determinant of the Jacobian matrix $DT(x, y) = \begin{bmatrix} f_x(x, y) & f_y(x, y) \\ g_x(x, y) & g_y(x, y) \end{bmatrix}$ is equal to 1 or -1 at all points (x, y).

Proof. This is the change of variable formula in multi-variable calculus. An elegant way to verify the formula is to interpret the map to define a parameterization of a surface $(u, v) \to \vec{r}(u, v) = (f(u, v), g(u, v), 0)$ which is also called the uv-map. You know the surface area element as $|\vec{r}_u \times \vec{r}_v|$ which is $|f_x g_y - g_x f_y| = |\det(DT(x, y))|$.

If $|\det(DT(x, y))| < 1$ everywhere, then the map is called **dissipative**. It shrinks volume. If $\det(DT(x, y))| > 0$ for all (x, y), the map is called **orientation preserving**.

EXAMPLE. The Henon map $T(x, y) = (1 - ax^2 - y, bx)$ is area preserving if and only if $|b| = 1$. and orientation preserving for positive b because the Jacobian matrix is

$$DT(x, y) = \begin{bmatrix} -2a & -1 \\ b & 0 \end{bmatrix}$$

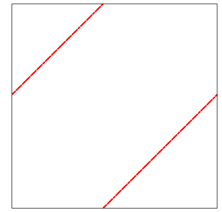
has the determinant $\det(DT) = b$.

SECOND ORDER DIFFERENCE EQUATIONS. A recursion like $x_{n+1} = x_{n-1} + \sin(x_n)$ defines a map if we introduce $y_n = x_{n-1}$:

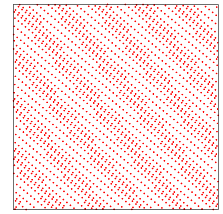
$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} y_n + \sin(x_n) \\ x_n \end{bmatrix} .$$

You might have seen the **Fibonacci recursion** $x_{n+1} = x_n + x_{n-1}$. The Henon map can be written as a recursion.

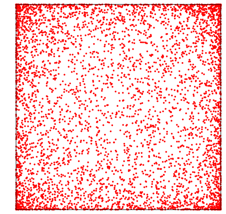
PRODUCT OF 1D MAPS. If f, g are one dimensional maps, then $T(x, y) = (f(x), g(y))$ is the product of these maps. The orbits of T is determined by the orbits of f and g. We just run the two dynamical systems in parallel. Examples:



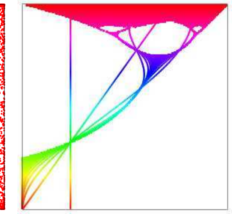
$T(x, y) = (x + \alpha, y + \alpha)$. The map is area preserving. $DT(x, y)$ is the identity matrix. Orbits are on line of slope 1.



$T(x, y) = (x + \alpha, y + \beta)$. Again $j(x, y)$ is the identity matrix. Orbits can be dense. Also this map is area preserving.



$T(x, y) = (4x(1 - x), 4y(1 - y)$ is not conservative. $\det(DT(x, y)) = 16(1 - 2x)(1 - 2y)$.

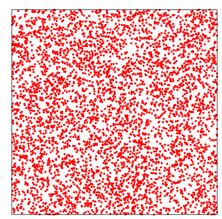


$T(x, y) = (4yx(1 - x), y)$. This map is not area-preserving.

THE CAT MAP. If A is a matrix with integer entries, then $T\vec{x} = A\vec{x}$ defines a map on the **torus** R^2/Z^2, which means we take x mod 1 and y mod 1. The example

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + y \\ x + y \end{bmatrix}$$

is called the "cat map". Arnold had illustrated the map using a cat. It belongs to a class of dynamical systems which can be understood completely. They are extremely "chaotic".



An orbit of the "cat map".

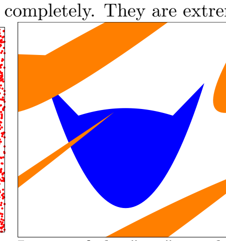


Image of the "cat" on the torus.

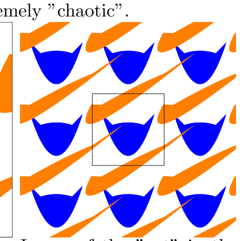
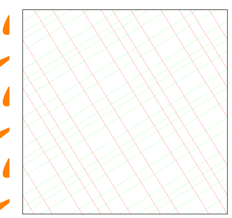
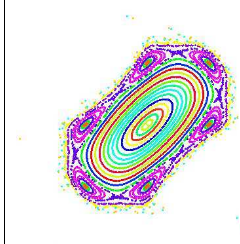


Image of the "cat" in the plane.

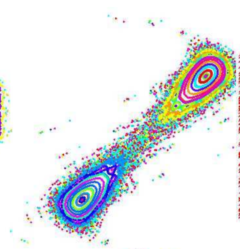


Invariant directions.

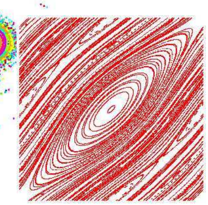
CUBIC HENON MAP AND THE STANDARD MAP. The map $T(x, y) = (cx - x^3 - y, x)$ with parameter c is called **cubic Henon map**. It is area preserving. The map $T(x, y) = (2x + c\sin(x) - y, x)$ is a map on the torus called the **Chirikov Standard map**. It is area preserving for all parameters c.



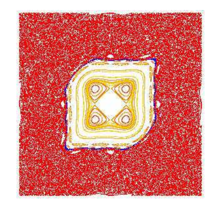
Orbits of $T(x, y) = (cx - x^3 - y, x)$ for $c = 1.5$.



Orbits of $T(x, y) = (cx - x^3 - y, x)$ for $c = 2.5$.



Orbits of $T(x, y) = (2x + c\sin(x) - y, x)$ for $c = 0.5$.

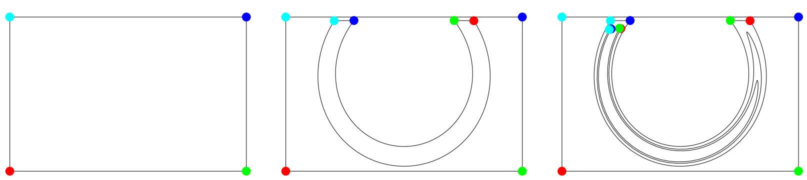


Orbits of $T(x, y) = (2x + c\sin(x) - y, x)$ for $c = 2.1$.

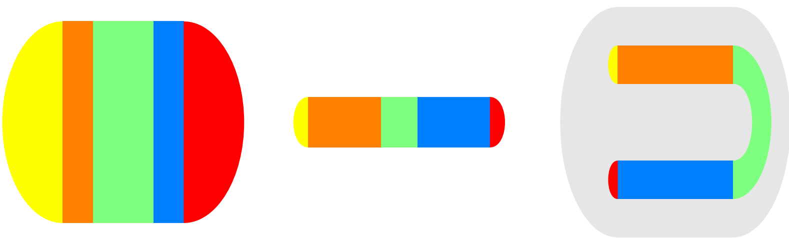
HORSE SHOES AND HOMOCLINIC TANGLE — Math118, O. Knill

ABSTRACT. The horse shoe map is a map in the plane with complicated dynamics. The horse shoe construction applies also to conservative maps and occurs near homoclinic points.

THE HORSE SHOE MAP. We construct a map T on the plane which maps a rectangle into a horse-shoe-shaped set within the old rectangle. The following pictures show an actual implementation with an explicit map



A better picture, which can be found in the book of Gleick shows a now rounded region which is first stretched out, then bent back into the same region.



THE HORSE SHOE ATTRACTOR. The map T maps the region G into $T(G)$ which is a subset of G. The image $T(T(G))$ is then a subset of $T(G)$ etc. The intersection K^+ of all the sets $T^n(G)$ is called the **horse shoe attractor**. It is invariant under the map T but T is not invertible on K^+. But now look at the set of points K which do not leave the original rectangle when applying T^{-1}. This set is now T invariant.

THEOREM (Smale) The map T restricted to K is conjugated to the shift map $S(x)_n = x_{n+1}$ on the space X of all $0-1$ sequences. With a suitable distance function defined on X, the conjugating map is continuous, invertible and its inverse is also continuous.

Similar as with the tent rsp. Ulam map, we will prove this only later when we cover the shift map. The conjugation is a coding map: for a point $z \in K$, define the sequence x_n as follows. Call C_0 the left of the three rectangles and C_1 the right one.

$x_n = \left[\begin{array}{c} 1,\ if\ T^{-n}(z) \in C_1 \\ 0\ if\ T^{-n}(z) \in C_0 \end{array} \right.$. It has to be shown that every point $z \in K$ is associated to exactly one $0-1$ sequence.

HORSE SHOES IN REAL MAPS. The horse shoe map often occur in an iterate of maps. One can see this sometimes directly. In the picture to the right, we iterated the points in a disc using the Standard map

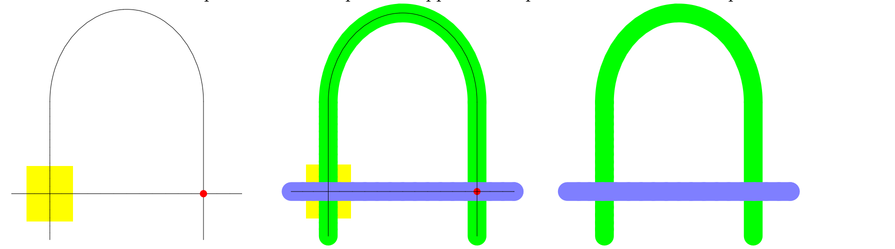
$$T(x,y) = (2x + y + c\sin(x), x)$$

on the torus. The picture has been made with the parameter value $c = 2.4$. We took a disc and applied the map T 5 times.

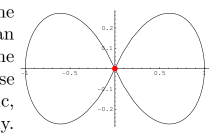
HORSESHOE FROM HOMOCLINIC POINTS.

THEOREM. A transverse homoclinic point leads to a horseshoe. There exists then an invariant set, on which the map is conjugated to a shift on two symbols.

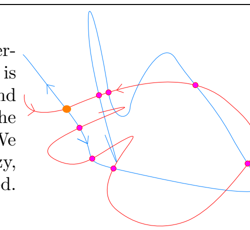
Take a small rectangle A centered at the hyperbolic fixed point. Some iterate of T will have the property that $T^n(A)$ contains the homoclinic point. Some iterate of the inverse of T will have the property that $B = T^m(A)$ contains the homoclinic point to. The map T^{n+m} applied to B produces a horseshoe map.



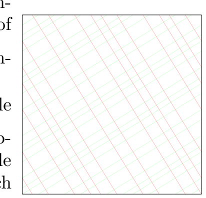
MANY HYPERBOLIC PERIODIC POINTS. The conjugation shows that periodic points are dense in K and that it contains dense periodic orbits. The map T restricted to K is chaotic in the sense of Devaney. Actually, one can show that each of the periodic points in K form hyperbolic points again. The stable and unstable manifolds of these hyperbolic points form again transverse homoclinic points and the story repeats again. This story is pretty generic, but there are cases, where stable and unstable manifolds come together nicely. This must be the case in integrable systems.



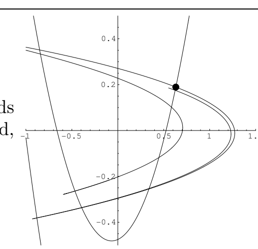
THE HOMOCLINIC TANGLE. If stable and unstable manifolds of a hyperbolic fixed point intersect, then they must intersect a lot more. The reason is that the image of the intersection produces an other intersection of stable and unstable manifolds because both curves are invariant. Note however that the stable manifold can not intersect with itself, except at hyperbolic points. We have in general two curves in the plane, both of which wind around like crazy, produce a lot of hyperbolic points due to all the horse-shoes which are created.



CAT MAP For the Cat map, one can compute the stable and unstable manifolds explicitly. The point $(0,0)$ is a fixed point. The Jacobean matrix of $T(x,y) = (2x+y, x+y)$ is $T'(x,y) = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. The eigenvector to the eigenvalue $(3+\sqrt{5})/2$ is $\begin{bmatrix} (1+\sqrt{5})/2 \\ 1 \end{bmatrix}$. The eigenspace is a line. It is the unstable manifold. When wrapped around the torus and plotted on the square, it appears as an infinite sequence of parallel line segments. Similarly, the stable manifold is a curve which winds around indefinitely around the torus. Each intersection of these lines is a homoclinic point.



HENON MAP In the case of the Henon map, stable and unstable manifolds intersect transversely in general. The map has all the complexity described, and especially, infinitely many hyperbolic points.



The notes below were added after the notes were distributed in class.

NICE INTEGRALS. Let us call smooth function $F(x,y)$ **nice** if any intersection of the set $F(x,y) = c$ with a bounded rectangle consists of a finite union of curves and points only. Let us call a map **nicely integrable** if it has a nice integral.

THEOREM (Poincare). A map with a transverse homoclinic point can not be nicely integrable.

PROOF. If T is nicely integrable, then also each iterate T^n is nicely integrable. The invariant horse shoe of some iterate of T is a set in which each point is accumulated by other points of the set. The horse show set can not be contained in a finite union of curves. Since the invariant function F must be constant on the horse shoe, the function can not be nice. The level set either contains infinitely many points or infinitely many curves in a rectangle which contains the horse shoe.

(Poincare knew this result only for analytic integrals).

THE GINGERBREADMEN MAP. The map

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1-y+|x| \\ x \end{bmatrix}$$

is an area-preserving map in the plane. For this map, one knows that it is stochastic on part of the phase space. The map had been studied by Devaney in 1992 and is often called the **Gingerbreadman map**. Around the fixed point $(1,1)$ and the periodic orbit $(-1,3),(-1,1),(3,-1),(5,3),(3,5)$, the motion is integrable. The iterates of the y axes produce a finite set of lines which bound an invariant region on which the map is hyperbolic.

(Related is the dissipative version, called the **Lozi map**

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1-y+a|x| \\ bx \end{bmatrix}$$

which shows for $a = 1.4, b = 0.3$ a "flat version" of the Henon attractor.)

INTEGRABLE MAPS IN TWO DIMENSIONS — Math118, O. Knill

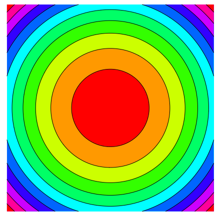
ABSTRACT. A map T in the plane is called integrable, if there is a non-constant continuous function $F(x,y)$ which is invariant under T. We give examples of integrable maps.

INTEGRABILITY. A map T is called **integrable**, if there exists a real valued continuous function $F(x,y)$ called **integral** for which the level sets $F=c$ are curves, or points and for which the identity

$$F(T(x,y)) = F(x,y)$$

holds for all (x,y).

EXAMPLES.



1. Let $T(x,y) = (\cos(\alpha)x - \sin(\alpha)y, \sin(\alpha)x + \cos(\alpha)y)$ be a rotation in the plane. It is integrable: the function $F(x,y) = x^2 + y^2$ is an integral.

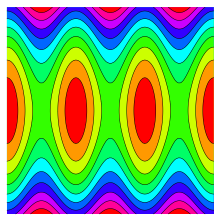
2. The map $T(x,y) = (3x, y/3)$ is integrable with integral $F(x,y) = xy$.

3. The map $T(x,y) = (x + \sin(y), y)$ is integrable with integral $F(x,y) = y$.

4. The cat map $T(x,y) = (2x+y, x+y)$ on the two dimensional torus is not integrable as you will demonstrate in a homework.

AN EXAMPLE FROM PHYSICS.

THEOREM. For every smooth function F, we can find a map, which has this function as an integral.



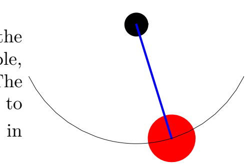
Consider the system of differential equations $\frac{d}{dt}x = F_y(x,y)$, $\frac{d}{dt}y = -F_x(x,y)$. By the chain rule, we have

$$\begin{aligned}\frac{d}{dt}F(x(t),y(t)) &= F_x(x(t),y(t))\frac{d}{dt}x(t) + F_y(x(t),y(t))\frac{d}{dt}y(t) \\ &= -\frac{d}{dt}y(t)\frac{d}{dt}x(t) + \frac{d}{dt}y(t)\frac{d}{dt}x(t)\end{aligned}$$

so that F does not change along a solution of the system. Define the map

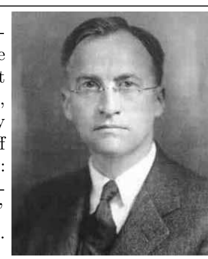
$$T(x,y) = (x(1), y(1))$$

if $x = x(0), y = y(0)$. This map has F as an integral.



In physics, the function F is often called the **energy** or **Hamiltonian** of the system. The fact that F is an integral is then energy conservation. For example, for $F(x,y) = \cos(x) + y^2/2$, one obtains the energy of the **pendulum**. The differential equations are then $\frac{d}{dt}x = y$, $\frac{d}{dt}y = \sin(x)$. They are equivalent to the Newton equations $\frac{d^2}{dt^2}x = \sin(x)$. We will look at differential equations in the plane in the next week.

BIRKHOFF ON INTEGRABILITY. Like "Chaos", "Integrability" is a mathematical term, which has many different definitions. One has to specify what one means with "integrable". The fact that one has to deal with several different definitions for integrability" expressed Birkhoff in the following way "When, however, one attempts to formulate a precise definition of integrability, many possibilities appear, each with a certain intrinsic theoretical interest". Birkhoff suggested his own (as he admits not very precise) definition of integrability: there exists a finite set of periodic orbits, around which the formal series development converge and which allow to represent any solution of the system." This Birkhoff integrability is probably hard to check in a specific applications.



THE SBKP MAP. For $|k| < 1$, lets call the map

$$T(x,y) = (2x + 4\cdot\arg(1 + k\cdot e^{-ix}) - y, x)$$

on the torus the **Suris-Bobenko-Kutz Pinkall map**. It had been found by Bobenko,Kutz, Pinkall and independently by Suris. Even so the map uses complex numbers for its definition, it is real. The argument $arg(z)$ of a complex number $z = x + iy = r\cos(\alpha) + ir\sin(\alpha) = re^{i\alpha}$ is defined as the angle α.

THEOREM. The SBKP map is integrable.

PROOF. The function

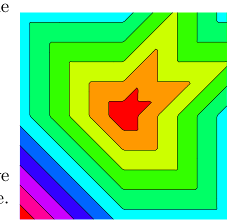
$$F(x,y) = 2(\cos(x) + \cos(y)) + k\cdot\cos(x+y) + k^{-1}\cdot\cos(x-y)$$

is an **integral**. It is not easy to verify that. Don't ask how F was found!

THE COHEN-COLINE-DE VERDIERE MAP. The map

$$T(x,y) = (\sqrt{x^2+\epsilon^2} - y, x)$$

in the plane is called the **Cohen-Coline-de Verdiere map**. By rescaling coordinates in R^2, we can assume $\epsilon = 0$ or $\epsilon = 1$. For $\epsilon = 0$, the map has the form



$$T(x,y) = (|x| - y, x) .$$

We call it the **Knuth map**.

THEOREM (KNUTH) The Knuth map is integrable.

PROOF. We check that $T^9 = Id$. Note that the map is piecewise linear, we only have to look at the orbits of the x axes to understand the entire picture. Actually, every orbit is periodic with period $1, 3$ or 9.

LEMMA. A map in the plane for which there exists n such that $T^n(x,y) = (x,y)$, must be integrable.

PROOF. Take $f(x,y) = y$ for example. Then $F(x,y) = \sum_{k=0}^{n-1} f(T^k(x,y))$ is an integral.

If we apply this lemma to the Knuth map, we get an explicit integral

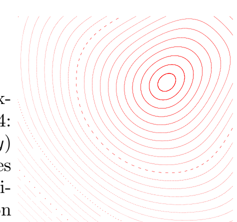
$$F(x,y) = y + |y - |x|| + |x - |y - |x||| + |y - |x - |y||| + |x - |y| + |y - |x - |y|||| .$$

The level curves of this function are shown in the graphics above. For every value $c > 0$ the level set $F(x,y) = c$ is a closed gingerman shaped curve on which T is conjugated to a rotation by an angle $1/9$.

Remark: The problem of proving periodicity of the map has been posed by Morton Brown in the American Mathematical Monthly 90, 1983, p. 569. The Monthly published the elegant solution of Donald Knuth in the volume 92, 1985 p. 218.

INTEGRAGBLE OR NOT? Lets look at the case $\epsilon = 1$, where

$$T(x,y) = (\sqrt{x^2+1} - y, x)$$



All orbits seem all to lie on invariant curves. The map looks integrable.
It had been communicated to me by M. Rychlik in 1998, that numerical experiments by John Hubbard revealed a hyperbolic periodic orbit of period 14: $(x,y) = (u,u)$ with $u = 1.54871181145059$. The largest eigenvalue of $dT^{14}(x,y)$ is $\lambda = 1.012275907$. The existence of a hyperbolic point of such a period makes integrability unlikely since homoclinic points might exist, but it is not impossible. It is difficult to find other hyperbolic periodic points. An other indication for non-integrability is a result of Rychlik and Torgenson who have shown that this map has no integral given by algebraic functions.

What follows was added after the handout was distributed in class:

HOW TO FIND AN INTEGRAL?

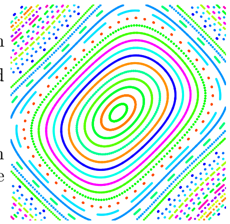
If we know a map is integrable, we could recover the invariant function F by taking $f(x,y) = y$ and defining $F(x,y) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k(x,y))$.

This invariant function is called the **time average** along the orbit. In the case of nonintegrability, this function is constant on complicated sets or even be infinite on some part of the plane. If the map is integrable with a nice analytic function, one could expect the integral to found using time averages.

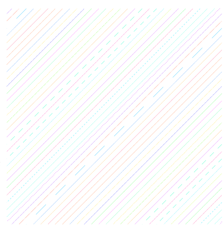
THE McMILLAN MAP $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{2kx}{(1+x^2)-y} \\ x \end{bmatrix}$ is an other example of an integrable map, where k is a parameter. It is called the McMillam map and has the integral

$$F(x,y) = x^2 + y^2 + x^2y^2 - 2kxy \ .$$

It is especially interesting to study because T is a rational function, a fraction of two polynomials. I don't think, one has a complete list of all integrable rational maps in the plane.



WHAT HAPPENS CLOSE TO THE INTEGRABLE CASE? In general, integrability gets lost when making small changes to an integrable map. For example, the Standard map $T(x,y) = (2x - y + \epsilon\sin(x), x)$ can for small ϵ be considered as a **perturbation** of the integrable map $T(x,y) = (2x-y,x)$ which has the integral $F(x,y) = x - y$. A study of the stable and unstable manifolds of the hyperbolic fixed point $(0,0)$ shows that they intersect transversely for small ϵ. One usually studies the map in an other form. Because $H(x,y) = (-x, y-x) = H^{-1}(x,y)$ satisfies $H(S(H(x,y)))$, where $T(x,y) = (2x - y + \epsilon\sin(x), x)$ and $S(x,y) = (x + y + \epsilon\sin(x), y + \epsilon\sin(x))$, we can look also at the map S instead. This map has the integral $F(x,y) = y$ for $\epsilon = 0$ and the invariant curves are horizontal.



KAM. Near integrable maps, remnants of integrability still exist. These traces of integrability persist in the form of smooth **invariant curves** which are now called KAM curves. The acronym KAM stands for Kolmogorov-Arnold-Moser. The proof that invariant curves persist after the perturbation is not easy. To find an invariant curve on which the map is conjugated to an irrational rotation with angle α, we need to find a periodic function $q(x)$ such that $q_n = q(n\alpha)$ satisfies the nonlinear recursion $q_{n+1} - 2q_n + q_{n-1} = \epsilon\sin(q_n)$. This means

$$\boxed{q(x+\alpha) - 2q(x) + q(x) = \epsilon\sin(q) \ .}$$

Naively, one could try to find q using the **implicit function theorem**: if one could invert the linear map $L(q) = q(x+\alpha) - 2q(x) + q(x)$.

SMALL DIVISORS. Lets look at this inversion problem If $q(x) = \sum_n c_n e^{inx}$ is the Fourier series of q, then $Lq(x) = \sum_n c_n(e^{i\alpha} - 2 + e^{-i\alpha})e^{inx}$. If $L(q) = p = \sum_n d_n e^{inx}$. then

$$q = L^{-1}p = \sum_n d_n \frac{1}{e^{in\alpha} - 2 + e^{-in\alpha}} e^{inx} = \sum_n d_n \frac{2}{\cos(n\alpha) - 1} e^{inx} \ .$$

You see the appearance of **small divisors** $\frac{2}{\cos(n\alpha)-1}$. In order that the Fourier series of the inverse converges, one needs α to be far away from rational numbers. Such numbers are called **Diophantine numbers**. Evenso, one is able to invert L in certain cases, the map L is not invertible as required for the implicit function theorem. One needs a so called **hard implicit function theorem**.

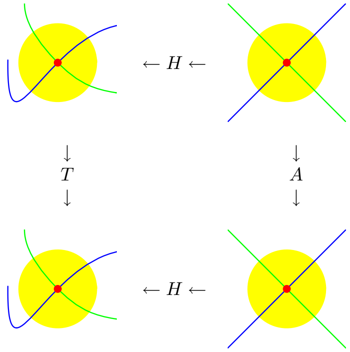
STABLE AND UNSTABLE MANIFOLDS — Math118, O. Knill

ABSTRACT. Near a hyperbolic point, one can conjugate the map by its linearization. This conjugation defines local curves through the origin which are invariant. These stable and unstable manifolds intersect in general to form homoclinic points. We will not prove the linearization theorem in class.

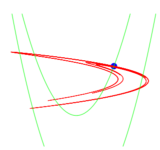
STERNBERG-GROBMAN-HARTMAN LINEARIZATION THEOREM. If $T(x)$ is smooth map with a hyperbolic fixed point x_0, then T is conjugated to its linearization DT near x_0.

Near the fixed point x_0, the dynamics can be computed by first going into a new coordinate system $H^{-1}(x_0)$, applying the linear map A, and undoing the coordinate change by applying H.

More precisely, there exists a small disc D around x_0 and a map H in the plane such that in D the identity $H \circ A(x) = T \circ H(x)$ holds.



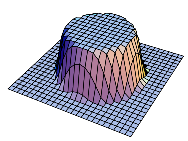
INVARIANT MANIFOLDS. The linear equation $x \mapsto Ax$ has two invariant curves, the lines spanned by the eigenvectors v_i of A. The conjugation defines two invariant curves $r_i(t) = H(tv_i)$ through a hyperbolic fixed point. These curves are called **stable and unstable manifolds** of the hyperbolic fixed point. The picture shows the stable and invariant manifolds for one of the fixed points of the Henon map. The unstable manifold lies in the attractor. Note that the unstable manifold of $T(x,y) = (1-ax^2+y, bx)$ is the stable manifold for $T^{-1}(x,y) = (y/b, (x-1+ay^2/b^2)$.



Here is the proof of the linearization theorem in its simplest case. The conjugation can actually be proven to be smooth too. The theorem had first be proven by S. Sternberg in 1958 (smooth conjugation for smooth T) and P. Hartman in 1960 (C^1 conjugation for C^2 maps T). The proof (not done in class) is not so easy and requires the language of linear operators.

PROOF PART 0: Some notations and preparations.
The proof works in any dimension. So x is now a vector in n-dimensional space $X = R^n$. Write $C(X,X)$ for the linear space of all continuous maps from X to X. The norm on this space is defined as $||f|| = \sup_{x \in X} |f(x)|$. For example: $||\sin|| = 1$ The norm of a linear operator U from $C(X,X)$ to $C(X,X)$ is defined as $||U|| = \sup_{||f||=1} ||U(f)||$. A linear map is called a contraction if $||U|| < 1$. If U is a contraction, then $(1-U)$ is invertible: the inverse is given by a geometric series $(1-U)^{-1} = \sum_{n=0}^\infty U^n$. For a hyperbolic matrix A, we write $X = E_+ \oplus E_-$, where E^+ is the linear space spanned by the eigenvectors of A belonging to the eigenvalues $|\lambda_i| < 1$ and E^- is the space spanned by the eigenvectors of A belonging to the eigenvalues $|\lambda_i| > 1$.

PROOF PART I: Reduction to a global conjugation problem.
Take first a smooth scalar function $\phi_\epsilon(x)$, which satisfies $\phi_\epsilon(x) = 1$ for $|x - x_0| > 2\epsilon$ and $\phi_\epsilon(x) = 0$ for $|x - x_0| < \epsilon$ (see picture to the right). The map $S = T + \phi_\epsilon \cdot (A - T)$ is equal to T for $|x - x_0| < \epsilon$ and equal to A for $|x - x_0| > 2\epsilon$. If can write $S(x) = Ax + f(x)$, where f is a smooth map satisfying $||f'||_\infty \to 0$ for $\epsilon \to 0$. Using this surgery, we can solve a global problem.



PROOF PART II: The conjugating equation and its linearization.
The aim is to show that S is conjugated by a map $H(x) = x + h(x)$ to the linear map A if $S = A + f$ if $||f'||_\infty$ is small enough. Remember that $f' = Df$ is the Jacobean matrix of f. The condition $H \circ A(x) = S \circ H(x)$ can be rewritten with $S(x) = Ax + f(x), H(x) = x + h(x)$ as

$$h(A(x)) - Ah(x) = f(x + h(x)) .$$

It is an equation for the unknown map $h \in C(X,X)$. We first consider the **linearized problem**

$$(Lh)(x) := h(A(x)) - Ah(x) = f(x) .$$

PROOF PART III: Solving the linearized problem.
We can decompose the problem into two parts

$$h_\pm(A(x)) - Ah_\pm(x) = f_\pm(x) ,$$

where $h = h_+ + h_-, f = f_+ + f_-$ is the decomposition satisfying $f_\pm, h_\pm \in E^\pm$. The linear map on continuous functions on the plane $U : C(X) \mapsto C(X), h \mapsto h(A)$ as well as its inverse U^{-1} have norm $||U|| = ||U^{-1}|| = 1$. We write $Af = A_+ f_+ + A_- f_-$. Because

$$\begin{aligned} ||(U - A_+)^{-1}|| &= ||U^{-1} \sum_{n=0}^\infty A_+^n U^{-n}|| \leq \frac{1}{1-\lambda} \\ ||(U - A_-)^{-1}|| &= ||A_-^{-1} \sum_{n=0}^\infty A_-^{-n} U^n|| \leq \frac{\lambda}{1-\lambda} < \frac{1}{1-\lambda} \end{aligned}$$

with $\lambda = \max\{||A_+||, ||A_-^{-1}||\} < 1$, we can find h using the formula

$$h = h_+ + h_- = (U - A_+)^{-1} f_+ + (U - A_-)^{-1} f_- .$$

PROOF PART IV: Solving the nonlinear problem.
Define $\Phi(h)(x) = f(x + h(x)) - f(x)$. We need to solve the equation

$$Lh = \Phi h + f$$

in for the unknown h in $C(X)$. The solution to this equation $(L^{-1}\Phi - 1)h = L^{-1} f$ is

$$h = (1 - L^{-1}\Phi)^{-1} L^{-1} f$$

if $1 - L^{-1}\Phi$ is invertible. Sufficent to invertibility is that $L^{-1}\Phi$ is a contraction. This is indeed the case if ϵ is small that is if $||f'||_\infty$ is small:

$$||(L^{-1}\Phi)h_1 - (L^{-1}\Phi)h_2|| \leq \frac{1}{1-\lambda} \cdot ||\Phi h_1 - \Phi h_2||_\infty \leq \frac{1}{1-\lambda} \cdot ||f'||_\infty \cdot ||h_1 - h_2|| .$$

COMPUTATION OF MANIFOLDS. The stable and unstable manifolds of a hyperbolic fixed point can be computed using power series. This calculation is due to Francescini and Russo. To get one of the manifolds, construct a curve $r(t) = (x(t), y(t))$ satisfying $r(0) = (x_0, y_0)$ and

$$T(x(t), y(t)) = (1 - ax(t)^2 + y(t), bx(t)) = (x(\lambda t), y(\lambda t))$$

for all $t \in R$. Here λ is an eigenvalue of the Jacobean matrix at the fixed point. Because $y(\lambda t) = bx(t)$, it is enough to calculate $x(t)$. With a Taylor series $x(t) = \sum_{n=0}^\infty a_n t^n$, the invariance condition $1 - ax(t)^2 + y(t) = x(\lambda t)$ or equivalently $x(\lambda t) + ax(t)^2 - bx(\lambda^{-1} t) = 1$ becomes

$$\sum_{n=0}^\infty [a_n \lambda^n - b a_n \lambda^{-n} + a \sum_{j=0}^n a_j a_{n-j}] t^n = 1 .$$

This equation allows to calculate the Taylor coefficients a_n recursively. Comparing coefficients of t^n gives $a(a_0 a_n + a_1 a_{n-1} + ... + a_{n-1} a_1 + a_n a_0) - b\lambda^{-n} a_n = -\lambda^n a_n$ and so

$$\boxed{a_n = \frac{a(a_1 a_{n-1} + ... + a_{n-1} a_1)}{-\lambda^n - 2aa_0 + b\lambda^{-n}}}$$

once $a_0, ..., a_{n-1}$ are given. The first coefficient a_0 is just x_0. Because a_1 satisfies $2aa_0 a_1 - b\lambda^{-1} a_1 = a_1 \lambda$, it can be chosen arbitrary like $a_1 = 1$. For the parameters $a = 1.4, b = 0.3$ the unstable manifold is $r(t) = (0.631354 + t - 0.25986t^2 + ..., 0.189406 - 0.155946t - 0.0210654t^2 + ...)$, the stable manifold is $r(t) = (0.631354 + t + 0.13278t^2 + ..., 0.189406 + 1.92374t + 1.63796t^2 + ...)$.

HOMOCLINIC POINTS. The intersection points of stable and unstable manifolds different from the fixed point itself are called **homoclinic points**. It has been realized already by Poincaré that the existence of homoclinic points produces a horrible mess. We will see why soon.

EXISTENCE OF SOLUTIONS TO ODE's — Math118, O. Knill

ABSTRACT. This is a proof of local existence of solutions of ordinary differential equations.

METRIC SPACES. Let X be a set on which a distance $d(x,y)$ between any two points x,y is defined. The function d must have the properties $d(y,x) = d(x,y) \geq 0$, $d(x,x) = 0$ and that $d(x,y) > 0$ for two different points x,y. Furthermore, one requires the triangle inequality $d(x,z) \leq d(x,y) + d(y,z)$ to hold for all x,y,z. A pair (X,d) with these properties is called a **metric space**.

EXAMPLES. 1) The plane R^2 with the usual distance $d(x,y) = |x-y|$. An other metric is the Manhattan or taxi metric $d(x,y) = |x_1 - y_1| + |x_2 - y_2|$.

2) The set $C([0,1])$ of all continuous functions $x(t)$ on the interval $[0,1]$ with the distance $d(x,y) = \max_t |x(t) - y(t)|$ is a metric space.

CONTRACTION. A map $\phi : X \to X$ is called a **contraction**, if there exists $\lambda < 1$ such that $d(\phi(x), \phi(y)) \leq \lambda \cdot d(x,y)$ for all $x,y \in X$. The map ϕ shrinks the distance of any two points by the contraction factor λ.

EXAMPLES. 1) The map $\phi(x) = \frac{1}{2}x + (1,0)$ is a contraction on R^2.
2) The map $\phi(x)(t) = \sin(t)x(t) + t$ is a contraction on $C([0,1])$ because $|\phi(x)(t) - \phi(y)(t)| = |\sin(t)| \cdot |x(t) - y(t)| \leq \sin(1) \cdot |x(t) - y(t)|$.

CAUCHY SEQUENCE. A **Cauchy sequence** in a metric space (X,d) is defined to be a sequence which has the property that for any $\epsilon > 0$, there exists n_0 such that $|x_n - x_m| \leq \epsilon$ for $n \geq n_0, m \geq n_0$.

COMPLETENESS. A metric space in which every Cauchy sequence converges to a limit is called **complete**.

EXAMPLES 1) $(R^n, d(x,y) = |x-y|)$ is complete. The rational numbers $(Q, d(x,y) = |x-y|)$ are not.

2) $C[0,1]$ is complete: given a Cauchy sequence x_n, then $x_n(t)$ is a Cauchy sequence in R for all t. Therefore $x_n(t)$ converges point-wise to a function $x(t)$. This function is continuous: take $\epsilon > 0$, then $|x(t) - x(s)| \leq |x(t) - x_n(t)| + |x_n(t) - y_n(s)| + |y_n(s) - y(s)|$ by the triangle inequality. If s is close to t, the second term is smaller than $\epsilon/3$. For large n, $|x(t) - x_n(t)| \leq \epsilon/3$ and $|y_n(s) - y(s)| \leq \epsilon/3$. So, $|x(t) - x(s)| \leq \epsilon$ if $|t-s|$ is small.

BANACH's FIXED POINT THEOREM. A contraction ϕ in a complete metric space (X,d) has exactly one fixed point in X.

PROOF.

(i) We first show by induction that

$$d(\phi^n(x), \phi^n(y)) \leq \lambda^n \cdot d(x,y)$$

for all n.

(ii) Using the triangle inequality and $\sum_k \lambda^k = (1-\lambda)^{-1}$, we get for all $x \in X$,

$$d(x, \phi^n x) \leq \sum_{k=0}^{n-1} d(\phi^k x, \phi^{k+1} x) \leq \sum_{k=0}^{n-1} \lambda^k d(x, \phi(x)) \leq \frac{1}{1-\lambda} \cdot d(x, \phi(x)) \; .$$

(iii) For all $x \in X$ the sequence $x_n = \phi^n(x)$ is Cauchy because by (i),(ii),

$$d(x_n, x_{n+k}) \leq \lambda^n \cdot d(x, x_k) \leq \lambda^n \cdot \frac{1}{1-\lambda} \cdot d(x, x_1) \; .$$

By completeness of X it has a limit $\tilde{x}$ which is a fixed point of ϕ.

(iv) There is only one fixed point. Assume, there were two fixed points $\tilde{x}, \tilde{y}$ of ϕ. Then

$$d(\tilde{x}, \tilde{y}) = d(\phi(\tilde{x}), \phi(\tilde{y})) \leq \lambda \cdot d(\tilde{x}, \tilde{y}) \; .$$

This is impossible unless $\tilde{x} = \tilde{y}$.

THE CAUCHY-PICARD EXISTENCE THEOREM. Assume $f : \mathbf{R}^n \to \mathbf{R}^n$ has a continuous derivative. For every initial condition x_0 there exists $\tau > 0$ such that on the time interval $[0, \tau)$ there exists exactly one solution of the initial value problem

$$\dot{x}(t) = f(x(t)), x(0) = x_0 \; .$$

PROOF.

(i)
Consider for every $\tau > 0$ and $r > 0$ the complete metric space

$$X = X_\tau(r) = \{x \in C[0,\tau] \mid \max_{0 \leq t \leq \tau} ||x(t) - x_0|| \leq r\}$$

with metric $d(x,y) = \max_{0 \leq t \leq \tau} ||x(t) - y(t)||$. With $c(t) = x_0$, we can write also $X = \{x \mid d(x,c) \leq r\}$.

Define a map ϕ on $C[0,\tau]$ by

$$\boxed{\phi(y) : t \mapsto x_0 + \int_0^t f(y(s)) \, ds \; .}$$

(ii) Define the constant

$$\lambda = \max\{\frac{||f(u) - f(v)||}{||u - v||} \mid ||u - x_0|| \leq 1, ||v - x_0|| \leq 1, u \neq v\} \; .$$

For every $x, y \in X_\tau(r)$ and $\tau \leq 1$, one has then

$$||f(x(s)) - f(y(s))|| \leq \lambda \cdot ||x(s) - y(s)|| \leq \lambda \cdot d(x,y)$$

for every $0 \leq s < \tau$. Therefore

$$d(\phi(x), \phi(y) = \max_{0<t<\tau} || \int_0^t f(x(s)) - f(y(s)) \, ds|| \leq \int_0^t ||f(x(s)) - f(y(s))|| \, ds \leq \lambda \tau d(x,y) \; .$$

We see that for small enough τ, the map ϕ is a contraction.

(iii) With $M = \max\{||f(x(t))|| \mid 0 \leq t \leq 1, \; d(x,c) \leq 1\}$, one has

$$||\phi(c) - c|| = || \int_0^t f(x_0(s)) \, ds|| \leq \int_0^t ||f(x_0(s))|| \, ds \leq M \cdot \tau \; .$$

If $\tau \leq 1$ is small enough, then $M \cdot \tau < (1-\lambda)r$. Using the triangle inequality, we obtain

$$d(\phi(x), c) \leq d(\phi(x), \phi(c)) + d(\phi(c), c) \leq \lambda d(x,c) + M\tau \leq \lambda r + (1-\lambda) r = r$$

proving that ϕ maps $X = \{d(x,c) \leq r\}$ into itself.

(iv) The fixed point ϕ in X obtained by Banach's fixed point theorem is a solution of the differential equation $\dot{x} = f(x)$ with initial value $x(0) = x_0$.

EXAMPLE WITH NO UNIQUE SOLUTION.

The differential equation $\frac{d}{dt}x = \sqrt{x}$ with $x(0) = 0$ has the solution $x(t) = Ct^2/4$ for any C. There are infinitely many solutions with the initial condition $x(0) = 0$. Note that the function $F(x)$ is not differentiable at $t = 0$.

EXAMPLE WITH NO GLOBAL SOLUTION.

The differential equation $\frac{d}{dt}x = x^2$ with initial condition $x(0) = 1$ has the solution $x(t) = 1/(1-t)$. At $t = 1$, the solution has escaped to infinity.

P.S. The photos show Stefan Banach (1892-1945), Emile Picard (1856-1941) and Augustin Cauchy (1789-1857).

DIDFFERENTIAL EQUATIONS IN TWO DIMENSIONS Math118, O. Knill

ABSTRACT. Differential equations in the plane do not show chaotic behavior. An interesting feature in two dimensions are limit cycles and their bifurcation. We look at some examples of such differential equations.

DIFFERENTIAL EQUATIONS. **Ordinary differential equations** are equations for an unknown function $x(t)$ in which which the derivatives with respect to one variable t appears. If derivatives with respect to several variables would occur, one would speak of partial differential equations. By introducing new variables for higher derivatives and possibly for time t, one can always bring it into the form

$$\boxed{\frac{d}{dt}x(t) = f(x(t))}$$

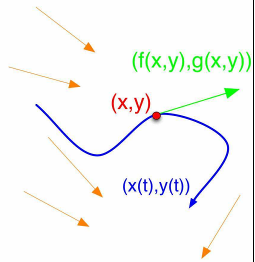
where $x(t)$ is a vector.

EXAMPLE. To write the second order inhomogeneous differential equation $\frac{d^2}{dt^2}x(t) + \frac{d}{dt}x(t) = \sin(t)$ in the above form, introduce $y(t) = \frac{d}{dt}$ and $z(t) = t$. Then $\frac{d}{dt}\begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} y(t) \\ \sin(z(t)) - y(t) \\ 1 \end{bmatrix}$

DIFFERENTIAL EQUATIONS IN THE PLANE. A solution $\vec{x}(t)$ of a differential equation $\frac{d}{dt}\vec{x} = \vec{F}(\vec{x})$ is a vector quantity changing in time. The vector $F(\vec{x}(t))$ is the velocity vector. In two dimensions, we have

$$\begin{aligned} \dot{x}(t) &= f(x,y) \\ \dot{y}(t) &= g(x,y) . \end{aligned}$$

The **vector field** is obtained by attaching a vector $\vec{F}(x,y) = (f(x,y), g(x,y))$ to each point (x,y). Of special importance are **equilibrium points**. These are points, where the velocity is zero.



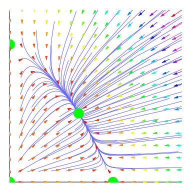
EXAMPLE. COMPETING SPECIES. A population of two species, where both compete for the same food can be modeled by the coupled logistic equations

$$\begin{aligned} \dot{x} &= \alpha x(1 - x/M) - \beta xy \\ \dot{y} &= \gamma y(1 - y/M) - \delta xy . \end{aligned}$$

A specific example is

$$\begin{aligned} \dot{x} &= 2x(1 - x/2) - xy \\ \dot{y} &= 3y(1 - y/3) - 2xy \end{aligned}$$

which has the equilibrium point $(1,1)$ because $(f(1,1), g(1,1)) = 0$. Additionally, one has the equilibrium points $(0,3), (2,0)$ and of course $(0,0)$.

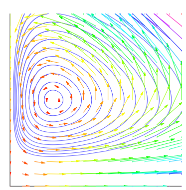


EXAMPLE. PREDATOR-PREY. These systems of the form

$$\begin{aligned} \dot{x} &= \alpha x - \beta xy \\ \dot{y} &= -\gamma y + \delta xy \end{aligned}$$

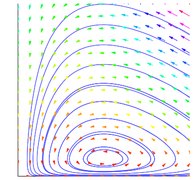
are also known under the name Volterra-Lodka systems. They can describe for example a shark-tuna population. The tuna population $x(t)$ becomes smaller with more sharks. The shark population $y(t)$ grows with more tuna. Historically, Volterra explained so the oscillation of fish populations in the Mediterrian sea. Here is a specific example:

$$\begin{aligned} \dot{x} &= 0.4x - 0.4xy \\ \dot{y} &= -0.1y + 0.2xy \end{aligned}$$



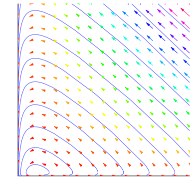
EXAMPLE. AIDS EPIDEMIC. The previous model can also model an epidemic as you can read in detail in Tom's lecture notes. In the interpretation of the epidemic, $x(t)$ is the size of the susceptible population, while $y(t)$ is the size of the infected population. A specific example modeling AIDS is

$$\begin{aligned} \dot{x} &= 0.2x - 0.1xy \\ \dot{y} &= -y + 0.1xy \end{aligned}$$

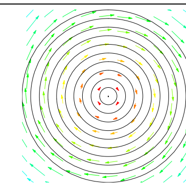


EXAMPLE. EBOLA EPIDEMIC. If the disease kills fast like in the case of ebola, we get a different picture

$$\begin{aligned} \dot{x} &= 0.2x - 0.5xy \\ \dot{y} &= -y + 0.5xy \end{aligned}$$



HARMONIC OSCILLATOR. The system $\dot{x} = y, \dot{y} = -x$ can in vector form $\vec{x} = (x,y)$ be written as $\frac{d}{dt}\vec{x}(t) = A\vec{x}(t)$, with $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. The direction field is always perpendicular to $\vec{x}$ so that by the product differentiation rule $d/dt\vec{x} \cdot \vec{x} = 2\vec{x}' \cdot \vec{x} = 0$ and $|\vec{x}|$ is constant. The solution curves are circles. In the homework, you look at a bit more general case. $\dot{x} = y, \dot{y} = -cx$, where c is a constant.



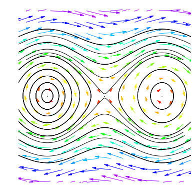
HAMILTONIAN SYSTEMS. If H is a function of two variables, we can look at the system

$$\begin{aligned} \dot{x} &= \partial_y H(x,y) \\ \dot{y} &= -\partial_x H(x,y) \end{aligned}$$

H is called the **energy** or Hamiltonian, x is called the position and y the momentum. Hamiltonian systems preserve energy $H(x,y)$: $\frac{d}{dt}H(x(t),y(t)) = \partial_x H(x,y)\dot{x} + \partial_y H(x,y)\dot{y} = \partial_x H(x,y)\partial_y H(x,y) - \partial_y H(x,y)\partial_x H(x,y) = 0$. The level curves of H are solution curves of the system. The time T maps are integrable. The illustration to the right shows the solution curves for the pendulum $H(x,y) = y^2/2 - \cos(x)$, where

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\sin(x) \end{aligned}$$

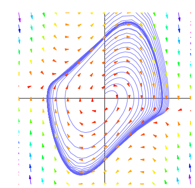
Here x is the angle between the pendulum and y-axes, y is the angular velocity, $\sin(x)$ is the potential.



THE VAN DER POL EQUATION. $\ddot{x} + (x^2 - 1)\dot{x} + x = 0$ appears in electrical engineering, biology or biochemistry. It is an example of a **Lienhard system** differential equations of the form $\ddot{x} + \dot{x}F'(x) + G'(x) = 0$, where $F(x) = x^3/3 - x, g(x) = x$.

$$\begin{aligned} \dot{x} &= y - (x^3/3 - x) \\ \dot{y} &= -x \end{aligned}$$

Lienhard systems often have **limit cycles**, closed solution curves on which trajectories can be attracted to. Lienhard systems are useful for engineers, who need oscillators which are stable under random noise.

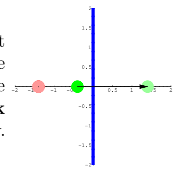


BIFURCATIONS — Math118, O. Knill

ABSTRACT. Equilibrium points can bifurcate. One distinguishes **pitchfork bifurcation** and **blue-sky bifurcation**, which were already knew in the one-dimensional setting. In two dimensions, where limit cycles can occur, it can happen that an equilibrium point produces a limit cycle. This is called the **Hopf bifurcation**.

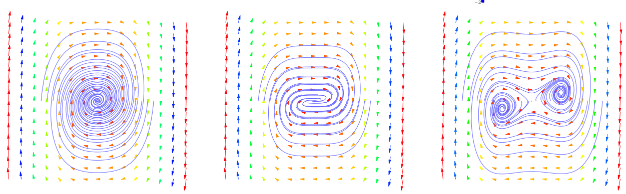
BIFURCATIONS OVERVIEW. If an eigenvalue of the Jacobean DF at an equilibrium point (x_0, y_0) crosses the y-axes, the stability of the equilibrium point changes. As in the discrete case, this is called a **bifurcation**. What possibilties are there? Besides the **pitch-fork** and **blue-sky** bifurcations, we already already in one dimensions, there are now possibilities which are not known in one dimensions. One is called **Hopf bifurcation**, which is the birth of **limit cycles**.

PITCHFORK BIFURCATION. If both eigenvalues were initially in the left half plane and one eigenvalue moves over to the right half plane, then a stable sink becomes a hyperbolic point. This bifurcationis usually associated to the creation of two new stable equilibrium points. This is called a **pitchfork bifurcation**. We know the one-dimensional version of this bifurcation already.

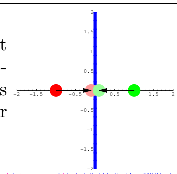


Example: $c = 0$ is a bifurcation parameter for

$$\begin{aligned} \frac{d}{dt}x &= y - 0.3 * x \\ \frac{d}{dt}y &= cx - x^3 \end{aligned}$$

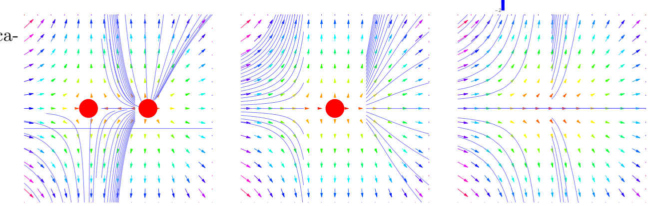


BLUE SKY BIFURCATION. It can happen that a hyperbolic equilibrium point collides with a stable or unstable equilibrium point and disappears. The opposite is also possible. Out of the blue, a parabolic equilibrium appears and splits into two equilibrium points. This is called the **saddle node bifurcation** or **blue-sky bifurcation**.

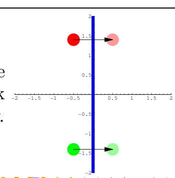


Example: $c = 0$ is a bifurcation parameter for

$$\begin{aligned} \frac{d}{dt}x &= c + x^2 \\ \frac{d}{dt}y &= -y \end{aligned}$$

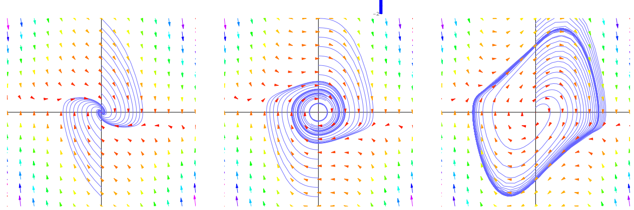


If both eigenvalues are on the left hand side and cross at the same time, we have a **Hopf bifurcation**. In this case, a stable sink becomes an unstable source and a limit cycle will appear nearby.



Example:

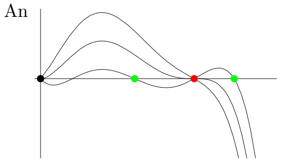
$$\begin{aligned} \frac{d}{dt}x &= y \\ \frac{d}{dt}y &= -x - (x^2 - c)y \end{aligned}$$



MORE BIFURCATIONS WITH LIMIT CYCLES (what follows will not be quizzed). These bifurcations above started with critical points and led to limit cycles. With limit cycles, there are more possibilites:

PITCH-FORK BIFURCATION FOR LIMIT CYCLES.
A stable limit cycles can change stability, become unstable and produce two limit cycles. This is called the **saddle node bifurcation** for limit cycles. An example is given in polar coordinates by

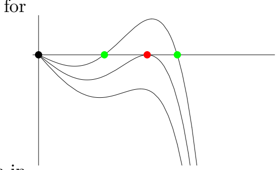
$$\begin{aligned} \frac{d}{dt}r &= r(r(1-r)^3 + c((r-1)^3 + (r-1))) \\ \frac{d}{dt}\theta &= \alpha + r^2 \end{aligned}$$



(It can using the formula $x = r\cos(\theta), y = r\sin(\theta)$ be rewritten as a system in the x, y coordinates.)

SADDLE NODE BIFURCATION FOR LIMIT CYCLES.
The sudden appearance of limit cycles is called **saddle node bifurcation** for limit cycles. An example is given in polar coordinates by

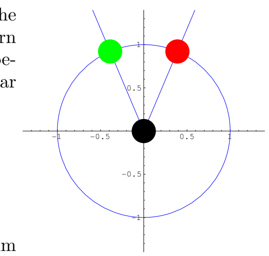
$$\begin{aligned} \frac{d}{dt}r &= cr + r^3 - r^5 \\ \frac{d}{dt}\theta &= \alpha + r^2 \end{aligned}$$



(It can using the formula $x = r\cos(\theta), y = r\sin(\theta)$ be rewritten as a system in the x, y coordinates.)

INFINITE PERIOD BIFURCATION.
A blue-sky bifurcation for equilibrium points can appear on a limit cycle. The limit cycle will become the stable and invariant manifolds of the newly born hyperbolic points. This bifurcation is called **infinite period** bifurcation because the limit cycle period will satisfy $T \to \infty$. An example is given in polar coordinates by

$$\begin{aligned} \frac{d}{dt}r &= r(1 - r^2) \\ \frac{d}{dt}\theta &= c - \sin(\theta) \end{aligned}$$

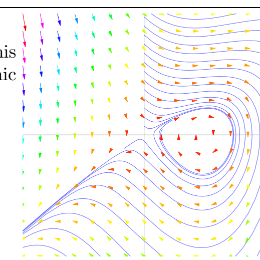


The system has an invariant circle for all c but for $c = 1$, there is an equilibrium point on the circle.

HOMOCLINIC BIFURCATION.
An equilibrium point can collide with a limit cycle and "open" it up. This bifurcation is called a **homoclinic bifurcation**. An example of a homoclinic bifurcation happens for

$$\begin{aligned} \frac{d}{dt}x &= y \\ \frac{d}{dt}y &= cy + x - x^2 + xy \end{aligned}$$

with the parameter $c = -0.86....$



Reversed situations of "supercritical" bifurcations (discussed above) are often called subcritical.

- A stable critical point can collide with two other hyperbolic critical points and become unstable. This is called a **subcritical pitch-fork bifurcation**. An example is $\frac{d}{dt}x = cx + x^3, \frac{d}{dt}y = -y$. This example is often associated to catastrophe like in the example $\frac{d}{dt}x = cx + x^3 - x^5, \frac{d}{dt}y = -y$
- An unstable limit cycle collapses to a stable critical point and becomes an unstable critical point. This is called a **subcritical Hopf bifurcation**.

These situations lead to "catastrophes". The stable equilibrium or cycle "jumps" discontinuously.

LIENHARD SYSTEMS — Math118, O. Knill

ABSTRACT. For a certain class of differential equations called Lienard systems, one can prove the existence of a stable limit cycle. An example is the van der Pol oscillator.

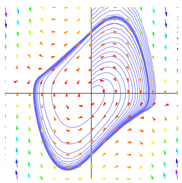
LIENHARD SYSTEMS. A differential equation

$$\frac{d^2}{dt^2}x + F'(x)\frac{d}{dt}x + G'(x) = 0$$

is called a **Lineard system**. With $y = \frac{d}{dt}x + F(x)$, $G'(x) = g(x)$, this is equivalent to

$$\begin{aligned} \frac{d}{dt}x &= y - F(x) \\ \frac{d}{dt}y &= -g(x) \, . \end{aligned}$$

VAN DER POL EQUATION. If $F(x) = c(x^3/3 - x)$ and $G(x) = x^2/2$, we have **van der Pol equation**



$$\boxed{\frac{d^2}{dt^2}x + c(x^2-1)\frac{d}{dt}x + x = 0}$$

Physically, one has a harmonic oscillator $\frac{d^2}{dt^2}x + x = 0$ for $c = 0$. For $c > 0$, some velocity and space dependent force $c(x^2-1)\frac{d}{dt}x$ is added. This force is accelerating the oscillator, if $x^2 < 1$, it is slowing down the oscillator if $x^2 > 1$. For large c, one calls the oscillator a **relaxation oscillator** because the stress accumulated during a slow buildup is relaxed during a sudden discharge.

THEOREM (Lienard) Assume F and g are smooth odd functions such that $g(x) > 0$ for $x > 0$ and such that F has exactly three zeros $0, a, -a$ with $F'(0) < 0$ and $F'(x) \geq 0$ for $x > a$ and $F(x) \to \infty$ for $x \to \infty$. Then the corresponding Lienard system has exacactly one limit cycle and this cycle is stable.

REMARK ON THE FIXED POINT $(0,0)$: Because g is odd with $g(x) > 0$ for $g \geq 0$, we have $g'(0) \geq 0$. The Jacobean matrix

$$\begin{bmatrix} F'(x) & 1 \\ -g'(x) & 0 \end{bmatrix}$$

has the eigenvalues $\lambda_{1,2} = (-F'(x) \pm \sqrt{F'(x)^2 - 4g'(x)})/2$. At the fixed point, the real part of these eigenvalues is positive because by assumption $F'(0) < 0$ and $|\sqrt{F'(x)^2 - 4g'(x)})| \leq |F'(x)|$ since $g'(0) \geq 0$. We see that the fixed point 0 is repelling.

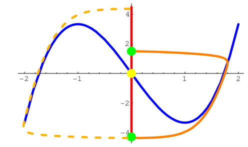
SOME REMARKS. Stable limit cycles appear in ecological, biological as well as mechanical systems. They are relevant because they are in general stable under small changes of the system.

From 1920 to 1950, research on nonlinear oscillations florished. The work was initially motivated by the development of radio and vacuum tube technology, where one realized that many oscillating circuits could be modeled by Lienard systems. This has been applied to many other situations. For example, one has also modeled the periodic firing of nerve cells driven by a constant current using van der Pol type differential equations.

Balthasar Van der Pol (1889-1959) was a Dutch electrical engeneer. He started his investigation on the van der Pol equation in 1926 and also studied versions with periodic forcing term, where chaotic motion can occur.

Lienards theorem was found and published in Russian by **Lienard** in 1958. For the proof of the Lienards theorem, we followed the proof given in the book "Differential equations and Dynamical systems" of Lawrence Perko.
A nice discussion can also be found in the book "Nonlinear dynamics and Chaos" by Steven Strogatz.
For historical facts mentioned in this section, we used "Writing the History of Dynamical Systems: Longe Duree and Revolution, Disciplines and Cultures" by David Aubin and Amy Dahan Dalmedico in Historia Mathematica 29, 2002. One should note also **Mary Cartwright** (1900-1998), who was making important contributions to the theory of nonlinear oscillations and discovered many phenomena later known as chaos (when the oscillator is driven, it becomes chaotic).

PROOF OF LIENHARDS THEOREM.
Draw in the $xy-$ plane the graph of the function $x \to F(x)$. On this graph, the vector field is vertical. It is called a **nullcline**. For $x > 0$ we have $\frac{dy}{dx} < 0$. On the y-axes, the vector field is horizontal because $g(0) = 0$. The y-axes is also a nullcline.



Consider an orbit which starts at $(0, y_0)$ on the positive y axes. It goes to the right because $g(x) > 0$ for $x \geq 0$. Because $g(x) > 0$ for $x > 0$, the orbit also moves down. It has to hit the graph of F. It intersects that nullcline at a point $(x_1, 0)$ with positive vertical velocity and enters the region, where $\frac{dy}{dx} < 0$. It must then go to the left and hit again somewhere the y axes horizontally in some point $(0, y_1) = (0, -S(y_0))$.
Because the differential equations are invariant under the transformation $(x, y) \mapsto (-x, -y)$, we can analyze the fate of the orbit on the left half plane in the same way as on the right plane.

A limit cycle exists if the map $y_0 \to S(y_0)$ has a fixed point. Alternatively, we can express this that the "energy" $H(x,y) = y^2/2 + G(x)$ is the same at $(0, y_0)$ and $(0, y_1)$. The idea of the proof is to determine the energy gain along the orbit and to see that only for one single orbit, the energy is conserved.

Compute

$$\frac{d}{dt}H(x,y) = y\frac{d}{dt}y + g(x)\frac{d}{dt}x = -F(x)g(x)$$

If $F(x(t))$ were positive on the entire trajectory from $(0, y_0)$ to $(0, y_1)$, then $H(0, y_1) - H(0, y_0)$ is positive. It must therefore cross the graph of F at a point, where $F(x) > 0$. The theorem is proven if we can show the following statement about the energy difference

$$\Delta(y_0) = H(0, S(y_0)) - H(0, y_0)$$

depending on the intersection point $(x_1, F(x_1))$ with the null cline.

> If $x_1 \leq a$, then $\Delta(y_0) > 0$. For y_0 such that $x_1 > a$, $\Delta(y_0)$ is a monotonically decreasing function for y_0. and $\Delta(y_0) \to -\infty$ for $y_0 \to \infty$.

As a consequence, there exists then exactly one point y_0, where the energy gain is zero. This point y_0 belongs to a limit cycle. The rest of the proof is devoted to the verification of the above claim.

(i) $\Delta(y) > 0$ **if** y_0 **is such that** $x_1 \leq a$.
Note that $F(x)$ is negative in the interval $[0, a]$. If $x_1 \leq a$, then $x(t) \leq a$ until we hit the y axes again. But since then $F(x(t)) < 0$ and $g(x) > 0$ for $x > 0$, we have $\frac{d}{dt}H(x,y) = -F(x)g(x) > 0$. The energy gain is positive.

(ii) **The monotonicity claim for** $x_1 \geq a$.
Let $A(y_0)$ be the path $(x(0), y(0)) = (0, y_0)$ and $(x(T), y(T)) = (0, y_1)$. From $\frac{d}{dt}H(x,y) = -F(x)g(x)$ we obtain

$$\Delta(H)(y_0) = \int_A -F(x(t))g(x(t))\, dt = -\int_A F(x(y))\, dy = \int_A -\frac{F(x)g(x)}{y - F(x)}\, dx \, .$$

Split the path A into a path A_1 from $(0, y_0)$ to $x(t) = a$, a path A_2 which is the continuation until $x(t) = a$ again and into a path A_3 until $(0, y_1)$. Along A_1 and A_3, we can parametrize the curve by x instead of t, along A_2, we can use the parameter y.

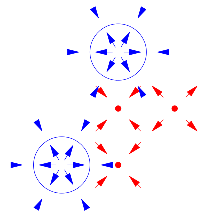
We see that increasing y_0 increases $y(t)$ and so decreases the integral $\Delta_1(H)(y_0) = \int_0^a -\frac{F(x)g(x)}{y-F(x)}\, dx$ along A_1. On A_3 increasing y_0 decreases $y(t)$ which decreases the integral $\Delta_3(H)(y_0) = \int_0^a \frac{F(x)g(x)}{y-F(x)}\, dx$ along along A_3. Along A_2, use y as the variable. Increasing y_0 pushes the path A_2 to the right so that $F(x(t))$ is increasing and the integral $\Delta_2(H)(y_0) = -\int_{y_2}^{y_3} F(x(y))\, dy$ is decreasing. The sum $\Delta(H)(y_0) = \Delta_1(y_0) + \Delta_2(y_0) + \Delta_3(y_0)$ is decreasing in y_0.

(iii) **The limit** $y_0 \to \infty$.
To see that $\Delta(y_0)$ goes to $-\infty$ for $y_0 \to \infty$, we split an orbit into paths B_1, B_2, B_3 in the same way as A_1, A_2, A_3 but where the value of a has been replaced by $a + 1$. The integrals along B_1 and B_3 are bounded by a constant independent of y_0, while the integral along B_2 is bigger or equal to $F(a+1)$ times the y differences of the two points, where $x(t) = a + 1$. This difference goes to $-\infty$ for $y_0 \to \infty$. So, the energy gain along the sum of the paths B_1, B_2, B_3 goes to $-\infty$ for $y_0 \to \infty$.

THE POINCARE BENDIXON THEOREM — Math118, O. Knill

ABSTRACT. The Poincaré-Bendixon theorem tells that the fate of any bounded solution of a differential equation in the is to convergence either to an attractive fixed point or to a limit cycle. This theorem **rules out "chaos" for differential equations in the plane**.

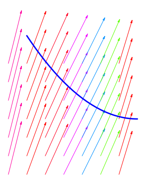
THEOREM (Poincare-Bendixon). Given a differential equation $\frac{d}{dt}x = F(x)$ in the plane. Assume $x(t)$ is an solution curve which stays in a bounded region. Then either $x(t)$ converges for $t \to \infty$ to an equilibrium point where $F(x) = 0$, or it converges to a single periodic cycle.



PRELIMINARIES.

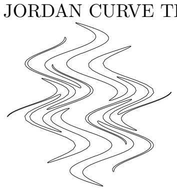
CYCLES, EQUILIBRIA AND CYCLES. Points x, where $F(x) = 0$ are called **equilibrium points** for the differential equation $\frac{dx}{dt}x = F(x)$. If a solution starts at an equilibrium point, it stays at the equilibrium point for ever. If $x(t)$ is a solution curve and $x(t+T) = x(t)$ for some $T > 0$, then the curve is called a **cycle**. Note that we do not include equilibrium points in this definition. The minimal time T for which $x(t+T) = x(T)$ is called the **period** of the cycle.

TRANSVERSE CURVES. A smooth curve $\gamma(s) \in R^2$ is called **transverse** to the vector field $x \mapsto F(x)$ if at every point $x \in \gamma$, the vector $F(x)$ and at least one tangent vector of γ passing through x are linearly independent.

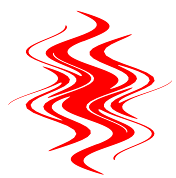


OMEGA LIMIT SET. The **omega limit set** $\omega^+(x_0)$ of an orbit $x(t)$ passing through x_0 is the set of points x, for which there exists a sequence of times t_n such that $x(t_n)$ converges to x. Equivalent is the mathematical statement $\omega^+(x_0) = \bigcap_{s \geq 0} \overline{\{x(t) \mid \; t \geq s\;\}}$, where $\overline{A}$ is the **closure** of a set A. If the ω-limit set of an orbit is a cycle, it is called a **limit cycle**.

JORDAN CURVE THEOREM.



A **Jordan curve** is a simple closed curve in the plane. "Simple" means that the curve should not have selfintersections or be tangent to itself at any point. The **Jordan curve theorem** assures that such a curve devides the plane into two disjoint regions, the "inside" and the "outside". This seemingly elmentary fact is surprisingly hard to prove.



EXAMPLE OF LIMIT CYCLE. The differential equation given in polar coordinates as

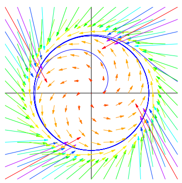
$$\frac{dr}{dt} = r(1-r^2), \frac{d\theta}{dt} = 1$$

is with $x = r\cos(\theta),\; y = r\sin(\theta)$ equivalent to

$$\frac{dx}{dt} = \frac{dr}{dt}\cos(\theta) - r\sin(\theta)\frac{d\theta}{dt} = (1-(x^2+y^2))x - y$$

$$\frac{dy}{dt} = \frac{dr}{dt}\sin(\theta) + r\cos(\theta)\frac{d\theta}{dt} = (1-(x^2+y^2))y + x$$

In this example, all initial conditions away from the origin will converge to the limit cycle.



EXAMPLE OF ATTRACTIVE POINT. The differential equation given in polar coordinates as

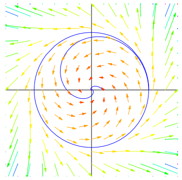
$$\frac{dr}{dt} = r(r^2-1), \frac{d\theta}{dt} = 1$$

is with $x = r\cos(\theta),\; y = r\sin(\theta)$ equivalent to

$$\frac{dx}{dt} = \frac{dr}{dt}\cos(\theta) - r\sin(\theta)\frac{d\theta}{dt} = ((x^2+y^2)-1)x - y$$

$$\frac{dy}{dt} = \frac{dr}{dt}\sin(\theta) + r\cos(\theta)\frac{d\theta}{dt} = ((x^2+y^2)-1)y + x$$

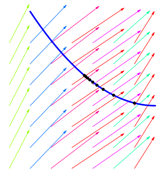
In this example, all initial conditions away from the limit cycle will converge to the origin or to infinity.



PROOF OF THE POINCARE-BENDIXON THEOREM. The aim is to show that if the omega limit set $\omega^+(x_0)$ is nonempty, then it either an equilibrium point or a closed periodic orbit.

(i) There are no equilibrium points on a transverse curve. The vector field f can therefore not reverse direction along the curve.

(ii) Let γ be a transverse curve. If a solution $x(t)$ crosses γ more than once, the successive crossing points form a monotonic sequence on the arc γ.
Proof. Denote by $x(t_1) = \gamma(s_1), x(t_2) = \gamma(s_2)$, the first two crossing times. We can assume that $s_2 \geq s_1$ because if this does not hold, one can reparametrize γ by $s' = 1 - s$ if $s_1 < s_2$. The union of the two smooth arcs $\{x(t) \mid t_1 \leq t \leq t_2\}$ and $\{\gamma(s) \mid s_1 \leq s \leq s_2\}$ is a closed piecewise smooth curve. By Jordan's curve theorem, such a curve divides the plane into two different regions. For $t > t_2$, the solution $x(t)$ stays in one of these regions. For the next crossing $x(t_3) = \gamma(s_3)$ one has therefore $s_3 \geq s_2$.

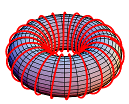


(iii) It follows from (ii) that no more than one point of any transverse arc γ can belong to the ω limit set $\omega^+(x_0)$.

(iv) Given $y_0 \in \omega^+(x_0)$. Because a solution $y(t)$ with $y(0) = y_0$ stays by assumption in a bounded region, the solution $y(t)$ is by the existence theorem for differential equations defined for all times. It stays in $\omega^+(x_0)$ because this set is invariant under the flow. Assume, there exists no stationary point in $\omega^+(x_0)$. There exists then a transverse arc γ passing through y_0. Because $\omega^+(x_0) \cap \gamma$ can have only one intersection and $y(t)$ returns arbitrary close to y_0, the orbit $\{y(t)\}$ through y_0 is a single periodic orbit.

DIFFERENT SURFACES. Does an anlogue of Poincare Bendixon hold also on other two dimensional spaces? The answer depends on the space. On the sphere, the answer is yes, on the torus, there are solutions which are neither asymptotic to a limit cycle or equilibrium point. An example of such a curve is $(t, \alpha t)$ mod 1 which is a solution of the differential equation

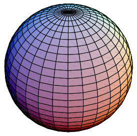
$$\frac{d}{dt}x = 1, \;\; \frac{d}{dt}y = \alpha \; .$$

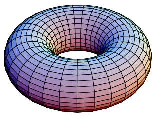


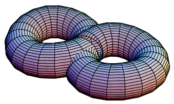
Differential equations of the form

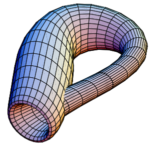
$$\frac{d}{dt}x = F(x,y), \;\; \frac{d}{dt}y = \alpha F(x,y) \; .$$

can even show some weak type of mixing. You explore the question a bit in a homework problem.









BASICS FOR ODES — Math118, O. Knill

ABSTRACT. This is an overview over the stability of equilibrium points of linear differential equations in the plane.

LINEAR SYSTEMS. A linear differential equation in two dimensions has the form

$$\begin{aligned}\frac{d}{dt}x(t) &= ax+by\\ \frac{d}{dt}y(t) &= cx+cy\end{aligned}$$

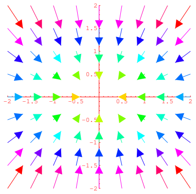
It can be written as $\frac{d}{dt}\vec{x}(t) = A\vec{x}(t)$ with a vector $\vec{x}$ and a matrix A. We denote the eigenvalues of A with λ_1 and λ_2.

If the eigenvalues are different, one can diagonalize A. In the eigenbasis of A, the matrix is $B = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ and the differential equation becomes

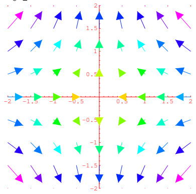
$$\begin{aligned}\frac{d}{dt}x(t) &= \lambda_1 x\\ \frac{d}{dt}y(t) &= \lambda_2 y\end{aligned}$$

with explicit solution $x(t) = e^{\lambda_1 t}x(0), y(t) = e^{\lambda_2 t}y(0)$.

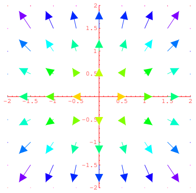
PHASE-PORTRAITS. We plot some vector fields and typical orbits



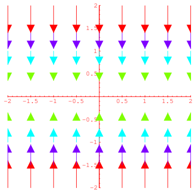
$\lambda_1 < 0$
$\lambda_2 < 0$,
i.e $A = \begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix}$



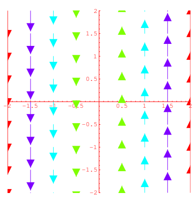
$\lambda_1 < 0$
$\lambda_2 > 0$,
i.e $A = \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$



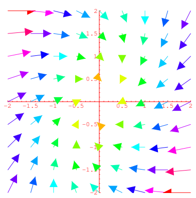
$\lambda_1 > 0$
$\lambda_2 > 0$,
i.e $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$



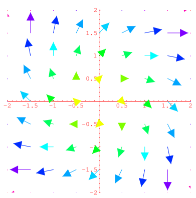
$\lambda_1 = 0$
$\lambda_2 < 0$,
i.e $A = \begin{bmatrix} 0 & 0 \\ 0 & -3 \end{bmatrix}$



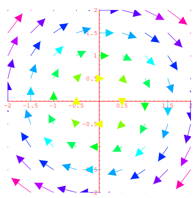
$\lambda_1 = 0$
$\lambda_2 = 0$,
i.e $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$



$\lambda_1 = a+ib, a < 0$
$\lambda_2 = a-ib$,
i.e $A = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}$



$\lambda_1 = a+ib, a > 0$
$\lambda_2 = a-ib$,
i.e $A = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$



$\lambda_1 = ib$
$\lambda_2 = -ib$,
i.e $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

AREA PRESERVATION. A differential equation for which we have solutions for all times defines for each t a map T_t in the plane.

We say a differential equation $\frac{d}{dt}x = F(x)$ is **area-preserving** if each of the time t maps T_t is area preserving.

DIVERGENCE. If F is a vector field, we denote by $\mathrm{div}(F)$ the **divergence** of F. It is in two dimensions, where $F(x,y) = (f(x,y), g(x,y))$ given by the formula $\mathrm{div}(F)(x,y) = f_x(x,y) + g_y(x,y)$.

> A differential equation $\frac{d}{dt}x = F(x)$ is area-preserving if and only if $\mathrm{div}(F)(x,y) = 0$ for all points in the plane.

PROOF. By the change of variable formula $\int\int_{T_t(A)} \; dA = \int\int_A |\det(DT(x)| \; dA$, where $DT(x)$ is the Jacobean matrix of the transformation T at x. Because $T_t(x) = x + tF + O(t^2)$, one has $DT_t = I_2 + tDF + O(t^2)$, where I_2 is the identity matrix. We have $DT_t = \begin{bmatrix} 1+ta & tb \\ tc & 1+td \end{bmatrix} + O(t^2)$ we have $\det(DT_t) = 1 + (a+d)t + O(t^2) = 1 + \mathrm{div}(F)t + O(t^2)$. Therefore $\frac{d}{dt}\int\int_{T_t(A)} \; dA = \int\int_A \frac{d}{dt}|\det(DT(x)| \; dA = \int\int_A \frac{d}{dt}(1 + t\mathrm{div}(F)) \; dA = \int\int_A \mathrm{div}(F)(x) \; dA$. (We could get rid of the absolute value because $1 + t\mathrm{div}(F)$ is positive for small t).

2. PROOF. Define $G(x,y,t) = (f(x,y), g(x,y), 1)$ and a tube like region $\{(x(t), y(t), t) \; (x(0), y(0)) \in A, 0 \leq t \leq \tau\}$ in **space-time**. Applying the **divergence theorem** using $\mathrm{div}(G)(x,y,t) = \mathrm{div}(F)(x(t), y(t))$, using the fact that the flux through the cylindrical walls is zero and the flux throught the bottom is $-\mathrm{area}(A)$ and the flux through the top is $\mathrm{area}(T_\tau(A))$ gives $\mathrm{area}(T_\tau(A)) - \mathrm{area}(A) = \int_0^\tau \int_{T_t(A)} \mathrm{div}(F(x(t), y(t))) \; dAdt$. This elegant proof does not need the coordinate change formula.

DISSIPATIVE SYSTEMS. If $\mathrm{div}(F) < 0$ in a region, then area is shrinking. You will explore some of the consequences of dissipation in the homework. Here just an example:

> PROPOSITION. In a region with $\mathrm{div}(F) < 0$, there are no sources or elliptic equilibrium points.

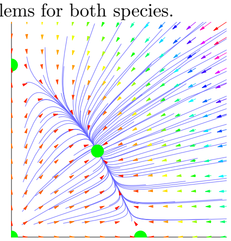
PROOF. If (x_0, y_0) is the equilibrium point, then $\mathrm{div}(F) = \lambda_1 + \lambda_2$. At sources, the real part of both λ_1 and λ_2 are positive. At elliptic equilibrium points, λ_1 and λ_2 are purely imaginary and the sum is 0.

EQUILIBRIUM POINTS. Points, where $F(x,y) = (0,0)$ are called **equilibrium points**. An equilibrium point is called **hyperbolic**, if no eigenvalue has a real part equal to 0. Bifurcations can happen, when an eigenvalue passes through the axes $\mathrm{Re}(\lambda) = 0$. In the hyperbolic case, one can conjugate the system near the equilibrium point to a linear system. This is a continuous version of the Sternberg-Grobman-Hartman theorem.

NULLCLINES. In two dimensions, we can draw the vector field by hand: attaching a vector $(f(x,y), g(x,y))$ at each point (x,y). To find the equilibrium points, it helps to draw the **nullclines** $\{f(x,y) = 0\}, \{g(x,y) = 0\}$. The equilibrium points are located on intersections of nullclines. The eigenvalues of the Jacobeans at equilibrium points allow to draw the vector field near equilibrium points. This information is sometimes enough to draw the vector field **by hand**.

EXAMPLE: COMPETING SPECIES. The system $\dot{x} = x(6 - 2x - y), \dot{y} = y(4 - x - y)$ has the nullclines $x = 0, y = 0, 2x + y = 6, x + y = 5$. There are 4 equilibrium points $(0,0), (3,0), (0,4), (2,2)$. The Jacobian matrix of the system at the point (x_0, y_0) is $\begin{bmatrix} 6 - 4x_0 - y_0 & -x_0 \\ -y_0 & 4 - x_0 - 2y_0 \end{bmatrix}$. Without interaction, the two systems would be logistic systems $\dot{x} = x(6 - 2x), \dot{y} = y(4 - y)$. The additional $-xy$ part is due to the competition. If both x and y become large, then this produce resource problems for both species.

Equilibrium	Jacobean	Eigenvalues	Nature of equilibrium
(0,0)	$\begin{bmatrix} 6 & 0 \\ 0 & 4 \end{bmatrix}$	$\lambda_1 = 6, \; \lambda_2 = 4$	Unstable source
(3,0)	$\begin{bmatrix} -6 & -3 \\ 0 & 1 \end{bmatrix}$	$\lambda_1 = -6, \; \lambda_2 = 1$	Hyperbolic saddle
(0,4)	$\begin{bmatrix} 2 & 0 \\ -4 & -4 \end{bmatrix}$	$\lambda_1 = 2, \; \lambda_2 = -4$	Hyperbolic saddle
(2,2)	$\begin{bmatrix} -4 & -2 \\ -2 & -2 \end{bmatrix}$	$\lambda_i = -3 \pm \sqrt{5}$	Stable sink



DIFFERENTIAL EQUATIONS IN THREE DIMENSIONS Math118, O. Knill

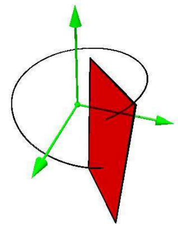
ABSTRACT. Differential equations in space can exhibit more complicated behavior than in the plane. Higher-dimensional systems occur naturally as we will see. Many systems can be studied using a Poincare map.

HOW DO SYSTEMS APPEAR IN THREE DIMENSIONS?

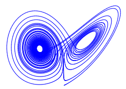
- A second order differential equation $\ddot{x} = f(x, \dot{x}, t)$ can be written with $(x, y, z) = (x, \dot{x}, t)$ as $(\dot{x}, \dot{y}, \dot{z}) = (y, f(x, z), 1)$. Such systems often appear in physcis. The time dependence allows to write the equation in three dimensions.

- A mechanical system of two degrees of freedom defines a flow in four dimensional space. Every coordinate has a position and velocity. Because energy is preserved, the dynamics takes place on a three dimensional energy surface.

POINCARE MAP. Assume we have a differential equation $\frac{d}{dt}x = F(x)$ in space. Given a two-dimensional surface Σ in space, we can start at a point in the plane, wait until the orbit returns back to the plane, hitting it transversely and so define a map from a subset of the plane to the plane. For any surface Σ in space, there is an open subset U, on which the return map T is defined and smooth.



THE LORENTZ SYSTEM. The system has been suggested by Eduard Lorentz in 1963. It is obtained by a truncation of the **Navier Stokes equations**. It gives an approximate description of a horizontal fluid layer heated from below which is itself a model for the earths atmosphere.

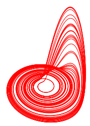


$$\begin{aligned}\dot{x} &= a(y - x)\\ \dot{y} &= cx - xz - y\\ \dot{z} &= xy - bz\end{aligned}$$

For $a = 10, b = 8/3, c = 28$, Lorentz observed a **strange attractor**.

THE ROESSLER SYSTEM. The following system of differential equations in space was found by Otto Rössler in 1976. The system was designed as a model for a strange attractor without any application in mind. It is theoretically interesting because a return map resembles the one dimensional logistic map $f_c(x) = cx(1 - x)$:

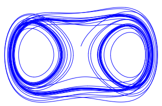
$$\begin{aligned}\dot{x} &= -(y + z)\\ \dot{y} &= x + 0.2y\\ \dot{z} &= 0.2 + xz - cz\end{aligned}$$



It is parametrized by a parameter c. The picture to the right shows an orbit for $c = 5.7$. For parameters in the range $2.5 < c < 10$ one observes a Feigenbaum bifurcation scenario.

THE DUFFING SYSTEM $\ddot{x} + b\frac{\dot{x}}{-}x + x^3 - c\cos(t) = 0$ can be written as

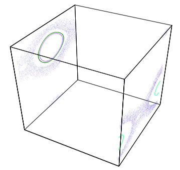
$$\begin{aligned}\dot{x} &= y\\ \dot{y} &= -by - x + x^3 - c\cos(z)\\ \dot{z} &= 1\end{aligned}$$



The Duffing system models a metallic plate between magnets. It is a harmonic oscillator with an additional cubic force, some damping and an external periodic driving force.

THE ABC FLOW. It is a flow with three parameters a, b, c, therefore its name **ABC flow**. An other etymological explanation is that Arnold, Beltrami and Childess worked on this system. Even so the system looks simple, its solutions can be complicated.

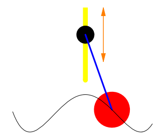
$$\begin{aligned}\dot{x} &= a\sin(z) + c\cos(y)\\ \dot{y} &= b\sin(x) + a\cos(z)\\ \dot{z} &= c\sin(y) + b\cos(x)\end{aligned}$$



FORCED PENDULUM. The differential equation $\ddot{x} = \cos(x) + g\sin(t)$ describes a pendulum which is periodically shaken up and down. The equations

$$\begin{aligned}\dot{x} &= y\\ \dot{y} &= \cos(x) + g\sin(z)\\ \dot{z} &= 1\ .\end{aligned}$$

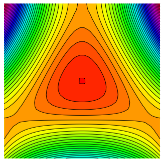
in space have a natural Poincare section $z = 0$.



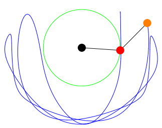
HENON-HEILS SYSTEM. The differential $\frac{dx_i}{dt} = H_{y_i}(x, y)$, $\frac{dy_i}{dt} = -H_{x_i}(x, y)$ with

$$H(x, y) = \frac{1}{2}(y_1^2 + y_2^2 + x_1^2 + x_2^2) + x_1^2 x_2 - \frac{1}{3}x_2^2$$

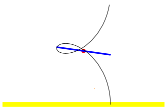
was studied first numerically by Henon and Heils in 1964. Energy surfaces $\{H(x, y) = E\}$ are invariant. For $0 \leq E \leq 1/6$ the surface is bounded and solutions stay bounded. The Poincaré section $\Sigma = \{x_1 = 0\}$ defines an area-preserving map on a subset of the plane.



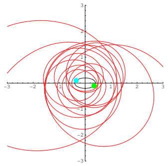
DOUBLE PENDULUM. The double pendulum is described by four variables. Energy conservation defines a differential equation on a three dimensional space. The return map $x = 0$ defines a map on the cylinder. If the gravitational field is zero, the double pendulum is integrable. With gravity $g > 0$, the system is complicated.



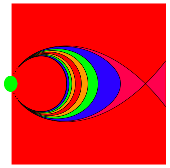
FALLING COIN. A falling coin defines a dynamical system which is often used, to produce random events: you flip a coin or dice and let it hit the ground, where it bounces. Flipping a coin and catching by the hand uses an integrable system. Some people can throw, catch and predict the outcome. If the stick moves in a gravitational field and if there are no impacts, then there is besides energy conservation also momentum conservation: the system becomes integrable. With impact, the system develops chaos.



3 BODY PROBLEM. The restricted three body problem in the plane is the situation, where the third particle is assumed not to influence the two other bodies. By Kepler, the two bodies moves on ellipses and produce a time periodic force on the third body. Therefore, we obtain a differential equation of the form $\frac{d^2}{dt^2}\vec{x}(t) = F(\vec{x}, t)$, where $\vec{x} = (x, y, \dot{x}, \dot{y})$. Energy conservation defines a three dimensional system.



STOERMER PROBLEM. A charged particle in a magnetic dipole field has rotational symmetry and so an angular momentum integral. This allows to reduce the system to a differential equation with four variables. The energy integral defines a flow on a three dimensional space. The system can be studied using a return map. The relevance of the system is the motion of charged particles in the **van Allen belts** and the explanation of the **Aurora Borealis**.

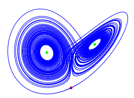


FRACTALS — Math118, O. Knill

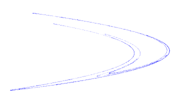
ABSTRACT. In order to define a strange attractor, we have to look at the notion of a "fractal", a set of fractional "dimension". The term fractal had been introduced by Benoit Mandelbrot in the late 70ies. We will see more about fractals later in this course, when we look at complex maps.

STRANGE ATTRACTOR. An **attracting set** of a differential equation $\dot{x} = F(x)$ or map $x \to T(x)$, is called a **strange attractor**, if it has **fractal dimension** (we will define that below), **sensitive dependence on initial conditions** (positive Lyapunov exponent) and which has an **indecomposable physical measure** which means that for almost all initial conditions x_0 and all continuous functions f, the limit $\frac{1}{t}\int_0^t f(T_s(x_0))\, ds$ rsp. $\frac{1}{n}\sum_{k=1}^n f(T^k(x)$ exists and depends only on f and not x_0.

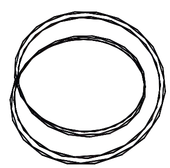
The Lorenz attractor: the dimension is numerically around 2.05 (Doering Gibbon 1995), 2.0627160 (Viswanath, 2002), The in-decomposability (technically called "SRB measure") (Tucker, 2002).



The Henon attractor: the dimension is measures 1.36 (Grebogi, Ott, Yorke, 1987) The in-decomposability had been shown (Benedicks and Carlson, 1991).



The Solenoid: This is a toy attractor, for which all the properties can be proven. It is a strange attractor for a map in space.

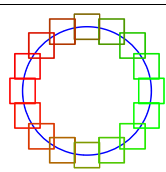


DIMENSION. Let X be a a set in Euclidean space. Define the s-volume of accuracy r of a set X as $h_{s,r}(X) = nr^s$, where n is the smallest number of cubes of side length r needed to cover X. The **s-volume** is the limit $h_s(X) = \lim_{r\to 0} h_{s,r}(X)$. The **box counting dimension** is defined as the limiting value s, where $h_s(X)$ jumps from 0 to infinity.

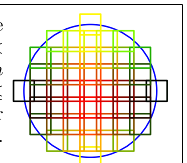
LINE SEGMENT. A **line segment** of length 1 in the plane can be covered with n intervals of length $1/n$ and $h_{s,r}(X) = n(1/n^s)$. For $s < 1$ this converges to 0, for $s > 1$, it converges to infinity. The dimension is 1.

SQUARE. A **square** X of a plane of area 1 in space can be covered with n^2 cubes of length $1/n$ and $h_{s,r}(X) = n^2(1/n^s)$ which converges to 0 for $s < 2$ and diverges for $s > 2$. The dimension is 2.

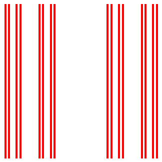
CIRCLE. A **circle** or radius 1 can be covered with with $2\pi n$ squares of length $1/n$ and $h_{s,r}(X) = 2\pi n(1/n^s)$. For $s < 1$ this converges to 0, for $s > 1$, it converges to infinity. The dimension is 1.



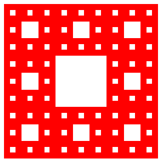
DISC. A **disc** of radius 1 in space can be covered with $\pi n^2/4 < N < \pi n^2$ squares of length $1/n$ and $\pi(n^2/4)/n^2 \leq h_{s,r}(X) \leq \pi n^2/n^s$ which converges to 0 for $s < 2$ and diverges for $s > 2$. The dimension is 2.



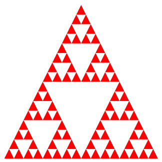
THE CANTOR SET. The **Cantor set** is constructed recursively by dividing the interval $[0,1]$ into 3 equal intervals and cutting away the middle one repeating this procedure with each of the remaining intervals etc. At the k'th stop, we need 2^k intervals of length $1/3^k$ to cover the set. The s-volume $h_{s,3^{-k}}(X)$ of accuracy $1/3^k$ is $2^k/3^{sk}$. It goes to zero if $s < 2/3$ and diverges for $s > \log(2)/\log(3)$.



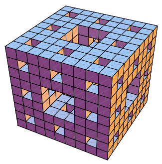
SHIRPINSKI CARPET. The **Shirpinski carpet** is constructed recursively by dividing a square in the plane into 9 equal squares and cutting away the middle one, repeating this procedure with each of the remaining squares etc. At the k'th step, we need 8^k squares of length $1/3^k$ to cover the carpet. The s-volume $h_{s,1/3^k}(X)$ of accuracy $1/3^k$ is $8^k(1/3^k)^s$ which goes to 0 for k approaching infinity if s is smaller than $d = \log(8)/\log(3)$ and diverges for s bigger than d. The dimension of the carpet is $d = \log(8)/\log(3) = 1.893$ a number between 1 and 2. It is a fractal.



SHIRPINSKI GASKET The **Shirpinski gasket** is constructed recursively by dividing a triangle in the plane into 4 equal triangles and cutting away the middle one, repeating this procedure with each of the remaining squares etc. At the k'th step, we need 3^k triangles of side length $1/2^k$ to cover the gasket. The s-volume $h_{s,1/2^k}(X)$ of accuracy $1/2^k$ is $8^k(1/2^k)^s$ which goes to 0 for k approaching infinity if s is smaller than $d = \log(3)/\log(2)$ and diverges for s bigger than d. The dimension of the gasket is $d = \log(3)/\log(2)$, a number between 1 and 2.



MENGER SPONGE. The three-dimensional analogue of the Cantor set in one dimensions and the Shirpinski carpet. One starts with a cube, divides it into 27 pieces, then cuts away the middle third along each axes. It is your task to compute the dimension. Note that the faces of the Menger sponge are decorated by Shirpinski Carpets.



THE PROBLEMS OF THE DEFINITION. If one takes the above definition, then the dimension of the set of rational numbers in the interval $[0,1]$ is equal to 1. A better definition, the **Hausdorff dimension** is needed. We include that definition below but it is a bit more complicated. The problem with the box counting dimension is that the size of the cubes should be allowed to vary. This refinement is similar to the change from the **Riemann integral** to the **Lebesgue integral**. c

HAUSDORFF MEASURE. Let (X,d) be a metric space. Denote by $|A| = \sup_{x,y\in A} d(x,y)$ the **diameter** of a subset A. Define for $\epsilon > 0, s > 0$

$$h^s_\epsilon(A) = \inf_{\mathcal{U}_\epsilon} \sum_{U\in\mathcal{U}_\epsilon} |U|^s ,$$

where $\mathcal{U}_\epsilon$ runs over all countable open covers of A with diameter $< \epsilon$. Such covers are also called **ϵ-covers**. The limit

$$h^s(A) = \lim_{\epsilon\to 0} h^s_\epsilon(A)$$

is called the $s-$ **dimensional Hausdorff measure** of the set A. Note that this limit exists in $[0,\infty]$ (it can be ∞), because $\epsilon \mapsto h^s_\epsilon(A)$ is increasing for $\epsilon \to 0$.

LEMMA: If $h^s(A) < \infty$, then $h^t(A) = 0$ for all $t > s$. Take $\epsilon > 0$ and assume $\{U_j\}_{j\in\mathbf{N}}$ is an open ϵ-cover of A. Then

$$h^t_\epsilon(A) \leq \sum_j |U_j|^t \leq \epsilon^{t-s} \cdot \sum_j |U_j|^s .$$

Taking the infimum over all coverings gives

$$h^t_\epsilon(A) \leq \epsilon^{t-s} \cdot h^s_\epsilon(A) .$$

In the limit $\epsilon \to 0$, we obtain from $h^s(A) < \infty$ that $h^t(A) = 0$.

HAUSDORFF DIMENSION.
Either there exists a number $\dim_H(A) \geq 0$ such that

$$s < \dim_H(A) \quad \Rightarrow \quad h^s(A) = \infty \,,$$
$$s > \dim_H(A) \quad \Rightarrow \quad h^s(A) = 0$$

or for all $s \geq 0$, $h^s(A) = 0$. In the later case, one defines $\dim_H(A) = \infty$.
The number $\dim_H(A) \in [0, \infty]$ is called the **Hausdorff dimension** of A.

FRACTAL. A **fractal** is a subset of a metric space which has finite non-integer Hausdorff dimension.

The Hausdorff dimension is in general difficult to calculate numerically. The central difficulty is to determine the infimum over $\sum_i |U_i|^t$, where $\mathcal{U} = \{U_i\}$ is an ϵ-cover of A. The box-counting dimension simplifies this problem by replacing arbitrary covers by sphere covers and so to replace the terms $|U_i|^t$ by ϵ^t. The prize one has to pay is that one can no more measure all bounded sets like this. In general, the upper and lower limits differ.

UPPER AND LOWER CAPACITY. Given a compact set $A \subset X$. Define for $\epsilon > 0$, $N_\epsilon(A)$ as the smallest number of sets of diameter ϵ which cover A. By compactness, this is finite. Define the **upper capacity**

$$\overline{\dim}_B(A) = \limsup_{\epsilon \to 0} \frac{\log(N_\epsilon(A))}{-\log(\epsilon)}$$

and analogous the **lower capacity** $\underline{\dim}_B(A)$, where lim sup is replaced with lim inf. If the lower and upper capacities coincide, the value $\dim_B(A)$ is called **box counting dimension** of A.

CAPACITY DIMENSION. If the lower and upper capacity are the same, one calls it the **capacity dimension**.

BOX COUNTING DIMENSION. Cover $\mathbf{R}^n$ by closed square boxes of side length 2^{-k}. and let $M_k(A)$ be the number of such boxes which intersect A. Define the box counting dimension

$$\dim_B(A) = \lim_{k \to \infty} \frac{\log(M_k(A))}{\log(2^k)} \,.$$

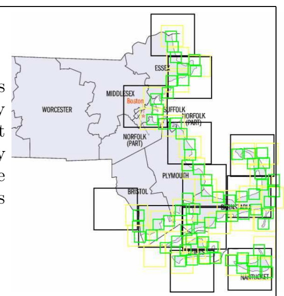
If the capacity dimension exists, then it is equal to the box counting dimension.

PROOF: Any set of diameter 2^{-k} can intersect at most 2^n grid boxes. On the other hand, any box of side 2^{-k} has diameter smaller than 2^{-k+1}. There exists therefore a constant C such that

$$C^{-1} \cdot M_k(A) \leq N_{2^{-k}}(A) \leq C \cdot M_k(A) \,.$$

Therefore
$$\lim_{k \to \infty} \frac{\log(M_k(A))}{\log(2^k)} = \lim_{k \to \infty} \frac{\log(N_{2^{-k}}(A))}{\log(2^k)} \,.$$

SELF-SIMILARITY. The computation of the dimension in the example objects was easy because they are **self-similar**. A part of the object is when suitably scaled equivalent to the object. We we will see more about this when we look at iterated function systems. To measure or estimate the dimension of an arbitrary object, one has to count squares. As an illustration of fractals in nature, one often takes coast lines. A rough estimate of the coast of Massachusetts leads to a dimension 1.3.



HISTORY.

The Cantor set is named after George Cantor (1845-1918), who was putting down the foundations of set theory. Ian Stewart writes in "Does God Play Dice", 1989 p. 121:
"The appropriate object is known as the Cantor set, because it was discovered by Henry Smith in 1875. The founder of set theory, George Cantor, used Smith's invention in 1883. Let's fact it, 'Smith set' isn't very impressive, is it?

The Hausdorff dimension has been introduced in 1919 by **Felix Hausdorff** (1868-1942).

Abram Besicovitch, around 1930, worked out an extensive theory for sets with finite Hausdorff measure.

The name "fractal" had been introduced only much later by Benoit Mandelbrot (1924-) in 1975.

The Sierpinski carpet was studied by **Waclaw Sierpinski** in 1916. He proved that it is universal for all one dimensional compact objects in the plane. This means that if you draw a curve in the plane which is contained in some finite box, however complicated it might be and with how many self-intersections you want, there is always a part of the Shirpinski carpet which is topologically equivalent to this curve.

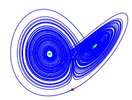
This might not look so surprising but this result is not true for the Shirpinski gasket. The Menger Sponge was studied by **Klaus Menger** in 1926. He showed that it is universal for all one dimensional objects in space. This means whatever complicated curve you draw in space, you find a part of the Menger sponge, which is topologically equivalent to it.

THE LORENZ SYSTEM **Math118, O. Knill**

ABSTRACT. In this lecture, we have a closer look at the Lorenz system.

THE LORENZ SYSTEM. The differential equations

$$\begin{aligned} \dot{x} &= \sigma(y-x) \\ \dot{y} &= rx - y - xz \\ \dot{z} &= xy - bz \ . \end{aligned}$$



are called the Lorenz system. There are three parameters. For $\sigma = 10, r = 28, b = 8/3$, Lorenz discovered in 1963 an interesting long time behavior and an aperiodic "attractor". The picture to the right shows a numerical integration of an orbit for $t \in [0, 40]$.

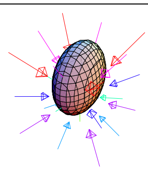
DERIVATION. Lorenz original derivation of these equations are from a model for fluid flow of the atmosphere: a two-dimensional fluid cell is warmed from below and cooled from above and the resulting convective motion is modeled by a partial differential equation. The variables are expanded into an infinite number of modes and all except three of them are put to zero. One calls this a Galerkin approximation. The variable x is proportional to the intensity of convective motion, y is proportional to the temperature difference between ascending and descending currents and z is proportional to the distortion from linearity of the vertical temperature profile.
The parameters $\sigma > 1, r > 0, b > 0$ have a physical interpretation. σ is the Prandl number, the quotient of viscosity and thermal conductivity, r is essentially the temperature difference of the heated layer and b depends on the geometry of the fluid cell.

SYMMETRIES. The equations are invariant under the transformation $S(x, y, z) = (-x, -y, z)$. That means that if $(x(t), y(t), z(t))$ is a solution, then $(-x(t), -y(t), z(t))$ is a solution too.
If $(x_0, y_0, z_0) = (0, 0, z_0)$, then the equations are $\dot{z} = -bz$. Therefore, we stay on the z axes and to the equilibrium point $(0, 0, 0)$.

VOLUME. The Lorenz flow is dissipative: indeed, the divergence of F is negative. The flow contracts volume.

$$\mathrm{div}(F) = -1 - \sigma - b$$

A TRAPPING REGION.



A region Y in space which has the property that if $x(t) \in Y$ then for all $s > t$ also $x(s) \in Y$ is called a **trapping region**. A function, which is nondecreasing along the flow is also called a **Lyapunov function**. Don't confuse this with the **Lyapunov exponent**.

LEMMA. There exists a bounded ellipsoid E which is a trapping region for the Lorenz flow. The time-one map T of the Lorenz flow maps E into the interior of E.

PROOF. We show that the function $V = rx^2 + \sigma y^2 + \sigma(z - 2r)^2$ is a Lyapunov function outside some ellipsoid. Indeed, the time derivative satisfies

$$\dot{V} = -2\sigma(rx^2 + y^2 + bz^2 - 2brz) \ .$$

Define $D = \{\dot{V} \geq 0\ \}$. This is a bounded region. If c the maximum of V in D and $E = \{V \leq c + \epsilon\ \}$ for some $\epsilon > 0$. then E is a region containing D. Outside this ellipsoid E, we have $\dot{V} \leq -\delta$ for some positive δ. With an initial condition $\vec{x}_0$ outside E, the vlaue of $V(x(t))$ decreases and within finite time, the trajectory will enter the ellipsoid E. All trajectories pass inwards through the boundary of E so that a trajectory which is once within E, remains there forever.

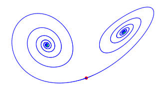
GLOBAL EXISTENCE. Remember that nonlinear differential equations do not necessarily have global solutions like $d/dtx(t) = x^2(t)$. If solutions do not exist for all times, there is a finite τ such that $|x(t)| \to \infty$ for $t \to \tau$.

LEMMA. The Lorenz system has a solution $x(t)$ for all times.

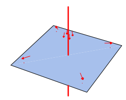
Since we have a trapping region, the Lorenz differential equation exist for all times $t > 0$. If we run time backwards, we have $\dot{V} = 2\sigma(rx^2 + y^2 + bz^2 - 2brz) \leq cV$ for some constant c. Therefore $V(t) \leq V(0)e^{ct}$.

THE ATTRACTING SET. The set $K = \bigcap_{t>0} T_t(E)$ is invariant under the differential equation. It has zero volume and is called the **attracting set** of the Lorenz equations. It contains the unstable manifold of O.

EQUILIBRIUM POINTS. Besides the origin $O = (0, 0, 0$, we have two other equilibrium points. $C^{\pm} = (\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1)$. For $r < 1$, all solutions are attracted to the origin. At $r = 1$, the two equilibrium points appear with a **period doubling bifurcation**. They are stable until some parameter r^*. The picture to the right shows the unstable manifold of the origin for $\sigma = 10, b = 8/3, r = 10$ which end up as part of the stable manifold of the two equilibrium points.

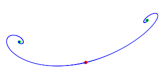


HYPERBOLICITY IN THREE DIMENSIONS. An equilibrium point is called **hyperbolic** if there are no eigenvalues on the imaginary axes. This is quite a wide notion and includes attractive or repelling equilibrium points as well as the possibility to have a one dimensional stable and two dimensional unstable direction or a two dimensional stable and a one dimensional unstable direction.

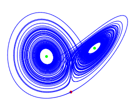


THE JACOBEAN. The Lorenz differential equations $\dot{x} = F(x)$ has the Jacobean $DF(x, y, z) = \begin{bmatrix} -\sigma & \sigma & 0 \\ r - z & -1 & -x \\ y & x & -b \end{bmatrix}$.

THE ORIGIN. At the equilibrium point $(0, 0, 0)$, the Jacobean $DF(0, 0, 0)$ is block diagonal. The eigenvalues are $-b, \frac{-1-s\pm\sqrt{(1-s)^2+4rs}}{2}$. For $r < 1$, where $\sqrt{(1-s)^2 + 4rs} < (1 + s)$, all three eigenvalues are negative. For $r > 1$, we have one positive eigenvalues and two negative eigenvalue. To the positive eigenvalue belongs an unstable manifold which is part of the Lorenz attractor.



THE TWO OTHER POINTS. At the two other equilibrium points, the eigenvalues are the roots of a polynomial of degree 3. For $\sigma > b + 1$ and $1 < r < r^* = (\sigma(\sigma + b + 3)/(\sigma - b - 1)$, all eigenvalues have negative a real part and the two points $C^{\pm}$ are stable. At $r = r^*$, a **Hopf bifurcation** happens. The two stable points $C^{\pm}$ collide each with an unstable cycle and become unstable. For $\sigma = 10, b = 8/3$ we have $r^* = 470/19 = 24.7$.



PERIODIC ORBITS. For large r parameters, the attractor can be single periodic orbit. Known windows are $99.534 < r < 100.795$, $145.96 < r < 166.07$,$214.364 < r < \infty$. Some periodic solutions are knots.

LYAPUNOV EXPONENTS OF DIFFERENTIAL EQUATIONS. If $T_t(x_0) = x_t$ is the time t map defined by the differential equation $\frac{d}{dt}x = F(x)$, then

$$\lambda(F, x) = \lim_{t\to\infty} \frac{1}{t} \log ||DT_t(x)||$$

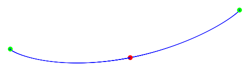
is called the **Lyapunov exponent** of the orbit. It is always ≥ 0. The Lyapunov exponent is for non-periodic orbits only accessible numerically.

THE LORENZ SYSTEM II **Math118, O. Knill**

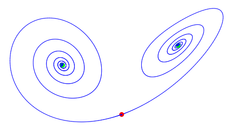
ABSTRACT. This is a continuation of the discussion about the Lorenz system and especially on the r dependence of the attractor.

OVERVIEW OVER BIFURCATIONS. We fix the parameter $\sigma = 10, b = 8/3$.

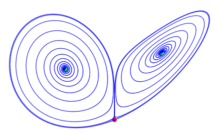
For $0 < r < 1$, the origin is the only equilibrium point and all points attracted to this point (you can find a proof in the book. At $r = 1$, a **pitchfork bifurcation** takes place. The origin becomes unstable and two stable equilibrium points appear. The picture shows the case $r = 1.5$.



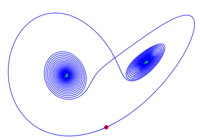
For $1 < r < 13.925$, the unstable manifold of the origin connects to the equilibrium points. The picture shows $r = 10$.



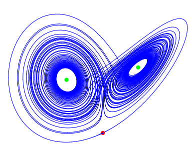
For $r = r_0 = 13.926$, the unstable manifold becomes double asymptotic to the origin.



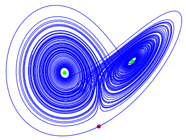
At the parameter r_0, two unstable cycles appear. For $13.926 < r < 24.06$, these cycles come closer to the fixed points $C^{\pm}$. The picture shows the parameter $r = 20$.



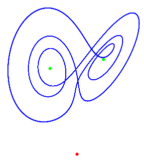
At the parameter $r = r_1 = 24.74 = 470/19$, the unstable cycles collide with the stable equilibrium points and render them unstable. This is called a **subcritical Hopf bifurcation**.



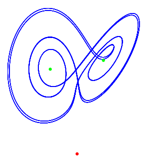
At the parameter $r = 28$, one observes the Lorenz attractor.



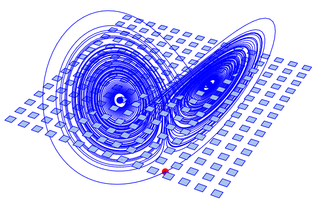
Between $r = 0.99524$ and $r = 100.795$, one observes an infinite series of period doubling bifurcations of stable periodic points (one has to start with the larger value and decrease r). These bifurcations are analogue to the Feigenbaum scenario. The picture shows the parameter $r = 100$.



Here we see the previous stable periodic cycle doubled. The parameter is $r = 99.7$. The period doubling scenario leads to the same Feigenbaum constant as one can see in the one dimensional logistic map family.



RETURN MAP. A good Poincare map is part of the subplane $z = r - 1$. This plane contains the equilibrium point $C^{\pm}$. These points are fixed points of the return map.



HISTORICAL. Lorenz carried out numerical investigations following work of Saltzman (1962). The Lorenz equations can be found in virtually all books on dynamics. We consulted:

- C. Sparrow, "The Lorenz equations: Bifurcations, chaos and strange attractors, Springer Verlag, 1982
- Strogatz, "Nonlinear dynamics and Chaos", Addison Wesley, 1994
- Dynamical systems X, Encyclopadia of Mathematics vol 66, Springer 1988
- Dennis Gulick, Encounters With chaos, Mc Graw-Hill, 1992
- Clark Robinson, Dynamical systems, Stability, Symbolic Dynamics and Chaos, CRC priss, 1995

BILLIARDS I — Math118, O. Knill

ABSTRACT. The billiard dynamical system can be seen as a limiting case of a particle moving in the plane under the influence of a potential V. In the limit, the ODE of three variables becomes a simple map, which still has all the features of differential equations. We discribe the system as an extremization problem, show the existence of periodic orbits and the area-preservation property. We also see that the ellipse is an integrable billiard.

PARTICLE MOTION IN THE PLANE. The motion of a particle in the plane under the influence of a **force** $F(x,y) = (f(x,y), g(x,y) = -\nabla V(x,y)$ is described by the differential equations

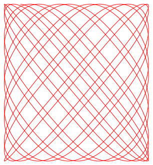
$$\begin{aligned} \frac{d^1}{dt^2} x(t) &= f(x,y) \\ \frac{d^2}{dt^2} y(t) &= g(x,y) \end{aligned}$$

Written as first order system, there are 4 variables x, y, u, v. Energy conservation $H(x,y,u,v) = u^2/2 + v^2/2 + V(x,y) = E$ reduces it to three variables:

$$\begin{aligned} \frac{d}{dt} x &= u \\ \frac{d}{dt} y &= \sqrt{2}\sqrt{E - V(x,y) - u^2/2} \\ \frac{d}{dt} u &= f(x,y) \end{aligned}$$

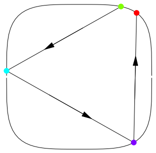
EXAMPLE. For $V(x,y) = x^4 + y^4$, the differential equations are

$$\begin{aligned} \frac{d}{dt} x &= u \\ \frac{d}{dt} y &= \sqrt{2}\sqrt{E - x^4 - y^4 - u^2/2} \\ \frac{d}{dt} u &= -4x^3 \end{aligned}$$

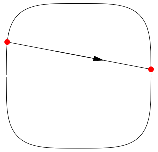


The picture shows an orbit close to a periodic orbit.

THE BILLIARD FLOW. Now, we take a particle in the plane and use a potential V which is zero inside a region G and which is infinite outside G. The mass point will move freely on a straight line until it hits the "wall". There it will reflect, bouncing off using the reflection law "incoming angle"="outgoing angle". The **Birkhoff billiard** is the dynamics of this billiard dynamical system, if the table is convex.

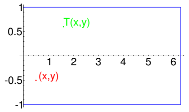


THE BILLIARD MAP. With an initial position s on the boundary, and an angle θ we have new initial position and a new angle. If the boundary of the table is parametrized by $x \in [0,1]$ and the angle by $\theta \in [0,\pi]$, we obtain a map $(s,\theta) \to (s_1, \theta_1)$.



BETTER COORDINATES. If we scale the table such that the table has length 1 and reparametrize the boundary of the table such that x is the **arc length** from some point 0 on the curve to s and take $y = \cos(\theta)$, we obtain a map

$$T : R/Z \times [-1,1],\ T(x,y) = (x_1, y_1)$$

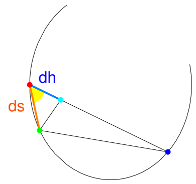


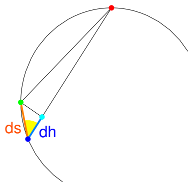
Topologically $R/Z \times [-1,1]$ is an annulus or a cylinder with boundary.

MONOTONE TWIST MAP. One boundary $R/Z \times \{-1\}$ is fixed and the other boundary $R/Z \times \{1\}$ is rotated once. Both boundaries, when the angle is 0 or π consist of fixed points. The map has the **twist property**: $\frac{d}{dy} x_1(x,y) > 0$. We prefer the (x,y) coordinates over the (s,θ) coordinates, because T becomes so area-preserving, as we will see below.

THE LENGTH FUNCTIONAL. Let $h(x_i, x_{i+1})$ denote the Euclidean distance between two points of the table (this is the distance in the plane and not the distance along the boundary). If $x_1, x_2, ..., x_n$ are successive impact points of the trajectory, then $\boxed{\cos(\theta_i) = -h_{x_i}(x_i, x_{i+1}) = h_{x_i}(x_{i-1}, x_i)}$

PROOF: You can see the relation $\cos(\theta) = dh/ds$ by watching the length change $dh = dh(x_i, x_{i+1})$, when x_i is replaced by $x_i + ds$ (first picture). The second formula is seen when observing the length change $dh = dh(x_{i-1}, x_i)$ when x_i is replaced with $x_i + ds$ (second picture).





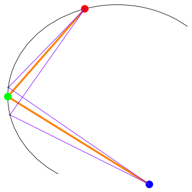
THE EULER EQUATIONS. The billiard map can be described by the equation

$$\boxed{h_{x_i}(x_i, x_{i+1}) + h_{x_i}(x_{i-1}, x_i) = 0}$$

This second order difference equation for the variables x_i is called the **Euler equation** of the billiard system. Given x_0, x_1, we can use these equations to get x_2, then use these equations again to get x_3 etc.

VARIATIONAL PRINCIPLE. If $x_1, x_2, ..., x_n$ is a sequence of impact points of the billiard map and the initial point x_0 and the final point x_{n+1} are fixed, then $x_0, x_1, x_2, ..., x_n$ is a billiard orbit if and only if $(x_1, x_2, .., x_{n-1})$ is a critical point of the function

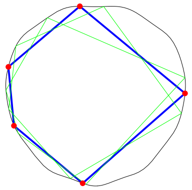
$$\boxed{H(x_1, x_2, ..., x_{n-1}) = \sum_{i=0}^n h(x_i, x_{i+1}).}$$



PROOF: just check that $\nabla H = 0$ gives the Euler equations. In other words, the billiard path extremizes the total length of the path. For $n = 2$, where we extremize $h(x_0, x_1) + h(x_1, x_2)$ we have to find the point x_1 on the table such that the path initiating at x_0 and ending at x_2 and which hits the table at a point x_1 is extremal.
This generalizes the Fermat principle: a light ray reflecting at a curve extremizes the distance to the curve only if in- and out-going angles are the same.

PERIODIC POINTS. A sequence $x_1, x_2, ..., x_n, x_{n=1} = x1$ is a periodic orbit if and only if the total length of the polygon of the impact points is extremal. In other words, we look for critical points of the total length of the closed polygon, which is:

$$\begin{aligned} H(x_1, ..., x_n) &= \sum_{i=1}^n h(x_i, x_{i+1}) \\ &= h(x_1, x_2) + h(x_2, x_3) + ... + h(x_{n-1}, x_n) + h(x_n, x_1) \end{aligned}$$



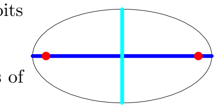
EXISTENCE OF PERIODIC POINTS. Since H is bounded, nonnegative and smooth, we have both a minimum and a maximum. The global minimum is of course when $x_1 = ..., x_n$ are all the same points. The maximum leads to a true periodic point: we have shown

> For a convex smooth billiard table, we find periodic points of minimal period n if n is prime.

PROOF. A continuous function on a bounded and closed subset of R^n has a maximum. The period can not be a factor of n because n was assumed to be prime. You show in a homework that the primality assumption is not necessary.

Example: The long axes and short axes of a convex table are periodic orbits of period 2.

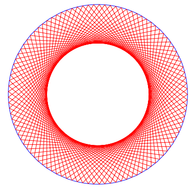
Example: Triangles of maximal total length in the table are billiard orbits of period 3.



BILLIARD IN A CIRCLE. The circle is an example of an integrable billiard. The angle θ and so $F(x,y) = y = \cos(\theta)$ is preserved. The billiard map T on $(R/Z) \times [-1,1]$ is given explicitly by

$$T(x,y) = (x + 2\arccos(y)/(2\pi), y)$$

This is a shear map. On the first coordinate we have a **rational or irrational** rotation.

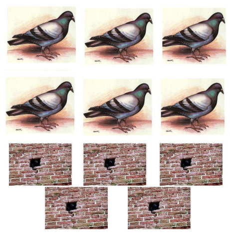


KRONECKER SYSTEM. The dynamical system on the circle obtained by a translation $T(x) = x + \alpha$ mod 1 is called the **Kronecker system**. Let $x_n = [n\alpha] = n\alpha$ mod 1 be the orbit of $T(x) = [x + \alpha]$ on the circle R/Z.

LEMMA. The sequence $x_n = T^n(x_0)$ is dense on $[0,1]$ if α is irrational.

PROOF. Given n divide $[0,1]$ into n equal intervals of length $1/n$. Take an orbit of length $n+1$. By the **pigeon hole principle**, two of these points $0, \alpha, ..., n\alpha$ must be in the same interval and so have distance $< 1/n$. Therefore $\delta = m\alpha = (k-l)\alpha < 1/n$ for some integer m. With an integer N larger than $1/\delta$, the set $\{m\alpha = \delta, 2m\alpha = 2\delta, ..., mN\alpha = N\delta\}$ intersects every interval of length δ at least once. The set $\{x_0, x_1, ..., x_{mN}\}$ intersects every interval of length δ and so every interval of length $1/n$.

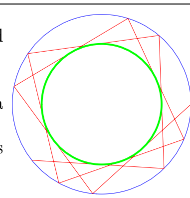
Illustration: 6 pigeons and 5 holes. Two pigeons must be in the same hole.



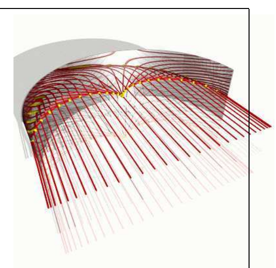
COROLLARY. If (s,θ) is an initial point for the billiard in a circle, then the orbit is periodic if $\theta/(2\pi)$ is rational. The ball will visit arbitrarily close to any given point of the table, if $\theta/(2\pi)$ is irrational.

CAUSTICS. For a billiard curve, one calls a curve a caustic, if the billiard ball, once tangent to that curve, remains tangent after the reflection.

EXAMPLE: For a circular table, every concentric circle inside the table is a caustic. For an ellipse, every confocal ellipse inside the table is a caustic.
EXAMPLE: given a convex curve, we can find a table which has this curve as a caustic using the **string construction**.



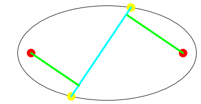
GENERAL CAUSTICS IN OPTICS. Places, where families of light rays focus are called **caustics**. If you take a family of parallel light and reflect it at a circle, then the light rays will focus at a curve which is called the **coffee cup caustic**. If the family of light rays is an orbit of a billiard ball in a table, then caustics might exist or not. In the case of the circle, every orbit produces caustics.



BILLIARD IN AN ELLIPSE.

The billiard in an ellipse is integrable.

PROOF. We find an invariant function $F(x,y)$, which is the product $d_1(x,y), d_2(x,y))$, where $d_i(x,y)$ is the distance of the trajectory to the focal point F_i. You will run a few lines of Mathematica to verify this in class.



BIRKHOFF-PORITSKY CONJECTURE: Is every integrable smooth convex billiard an ellipse? A collaborator of Birkhoff at Harvard with name **Hillel Poritsky** had worked on it and published a paper in 1950, where he made some progress.
The picture shows Poritsky in 1936 at the 42. Summer Meeting of the Mathematical Organizations of America in Cambridge, Massachusetts.

THEOREM. The billiard map is area-preserving.

PROOF. Let $Y \subset T^1 \times [-1,1]$ be disc with boundary C. We show $\int\int_Y \, dydx = \int_Y \, dy'dx'$, where $T(x,y) = (x',y')$, $T^2(x,y) = (x'',y'')$ is the map. (We use primes here not as derivatives Using Greens formula, we get

$$\begin{aligned} \text{Area}(T^{-1}(Y)) &= \int\int_{T^{-1}(Y)} dy\,dx = \int_{T^{-1}(C)} y\,dx = \int_{T^{-1}(C)} h_1(x,x')\,dx \\ &= \int_C h_1(x',x'')\,dx' = \int_C -h_2(x,x')\,dx' = \int_C y'\,dx' = \int\int_Y dy'\,dx' = \text{Area}(Y)\,. \end{aligned}$$

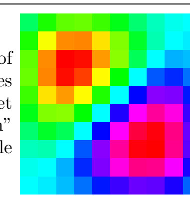
GENERALIZATION. Every map defined by the Euler equations $h_2(x,x') + h_1(x',x'')$ of a smooth generating function $h(x,x')$ is area-preserving in the coordinates $(x,y) = (x, h_1(x,x')$.

EXAMPLE. $h(x,x') = (x'-x)^2/2 + V(x)$ leads to the Euler equation $h_1(x_i,x_{i+1}) + h_2(x_{i-1},x_i) = (x_{i+1} - x_i) + \frac{d}{dx}V(x_i) - (x_i - x_{i-1}) =$ This is the second order difference equation $x_{i+1} - 2x_i + x_{i-1} + V'(x_i) = 0$. Vor $V(x) = c\cos(x)$, this recursion is the **Standard map**. For cubic V, it leads to the Henon map in the plane.

THE JACOBEAN MATRIX. An other proof to show that the map is area-preserving is to compute the Jacobean matrix and to verify that the determinant is 1. We will write down the Jacobean later. An other proof of the area-preservation property is given in proposition 6.4.2 of the textbook.

HISTORY.
Ludwig Boltzmann (1844-1906) studied the hard sphere gas. This is a billiard system.
Emil Artin (1898-1962) looked in 1924 at billiard in the hyperbolic plane. This is of interest in algebra.
Jacques Hadamard (1865-1963) Hedlund-Hopf studied the geodesic flow, which is a generalization of billiards.
George Birkhoff (1884-1944) in 1927, proposed convex billiards as a model for the 3-body problem
Hillel Poritsky in 1950 posed the integrability question.

WHY STUDY BILLIARDS?
It is a beautiful and simple dynamical system featuring all the complexities of more complex systems. It is a limiting case of the geodesic flow and illustrates theorems in topology, geometry or ergodic theory. It is related to Dirichlet spectral problem $\Delta u = \lambda u$ which can be considered the "quantum version" of the billiard problem, where the eigenfunctions describe a quantum particle moving freely in the table with energy λ.



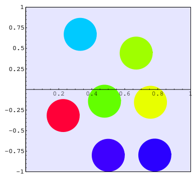
CHAOTIC BILLIARDS — Math118, O. Knill

ABSTRACT. Billiards in tables with negative curvature as well as billiards like the Stadium are chaotic: The Lyapunov exponent is positive. They are actually ergodic: every invariant set of positive measure will have either area 0 or area 1.

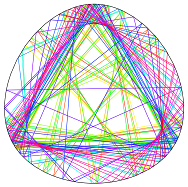
POINCARES RECURRENCE THEOREM. Area preservation allows to make a statement about recurrence of area-preserving map defined on a T invariant subset in the plane. For example, X could be the annulus $R/Z \times [-1,1]$ and T could be a billiard map.

> For every set Y of positive area $|Y|$, there exists n such that $T^n(Y) \cap Y$ has positive area.

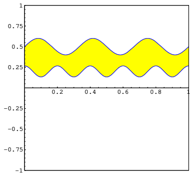
PROOF OF POINCARES THEOREM. Assume there exists a set Y of positive area $m(Y)$ such that $Y_i = T^i(Y)$ satisfies $m(Y_i \cap Y) = 0$ for all $i > 0$. Because $m(Y_i) = m(Y) > 0$ and the total space has finite area, there must exist $0 < i < j$ such that $m(Y_i \cap Y_j) > 0$. (This is a variant of the pigeon hole principle. If you have a cage with finite room and each pigeon needs the same amount of space, only a finite number of pigeons fit). But $m(T^{-i}(Y_i \cap Y_j)) = m(Y \cap Y_{j-i}) > 0$ contradicts that Y and Y_k are disjoint.



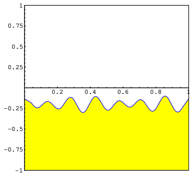
CONSEQUENCE FOR BILLIARDS. Does this mean that if you start shooting from a certain point in a certain direction, there will be times, when the orbit will come back to a similar spot on the table with a similar angle? Not necessarily. For example, if you are on the stable manifold of an unstable periodic point, then the orbit will converge to that periodic orbit. The Poincaré statement is a statement about sets. It assures for example, that if you start shooting from a certain interval on the table in a certain interval of directions, you will come back to that range of initial conditions **with probability 1**.



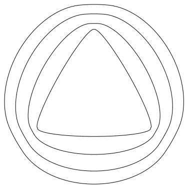
ERGODICITY. Less obvious is the question, whether a given set ever reaches an other set. If all "measurable" invariant subset of the annulus have either area 1 or 0, then the map is called **ergodic**. Measurable is a technical term which assures that the area $\int \int_A 1 \, dxdy$ is defined. Any set which can be defined by a (possibly infinitely) sequence of intersections or unions is measurable.



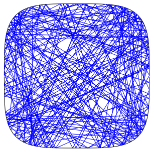
INVARIANT CURVES PREVENT ERGODICITY. If a billiard has an invariant curve which is the graph of a function $\{y = f(x)\}$, then if (x_0, y_0) is below the graph, the entire orbit (x_n, y_n) stays below the graph for all times. The billiard can not be ergodic.



STRING CONSTRUCTION. It had been known since a long times, that if one starts with a convex curve, winds a closed string around it and drags the string around the curve which keeping the string tight, we obtain a table, which has the original curve as a caustic. The picture shows some tables which have a triangle as a caustic. These tables are not ergodic.



GLANCING BILLIARDS. An orbit (x_j, y_j) of a billiard table for which y_j comes arbitrarily close to -1 and arbitrarily close to 1 is called a **glancing billiard orbit**.



> THEOREM. (Birkhoff) There are no invariant curves of T, if and only if there exists a glancing orbit.

PROOF. If there is an invariant curve, there is trivially no glancing orbits because the regions on both sides of the curve are left invariant. Assume now there is no glancing orbit. This means there is an $\epsilon > 0$ such that for all $y_0 < 1 - \epsilon$ we have $y_n > -1 + \epsilon$. Consider the region $Y = \{y < 1 - \epsilon\}$. The set $\bigcup_n T^n(Y)$ is a T-invariant set which does not intersect $\{y > 1 - \epsilon\}$. The boundary of this curve is an invariant curve. (One actually knows that such a curve must be the graph of a Lipshitz continuous function).

THE JACOBEAN. Let κ_i denote the curvature at the impact point and angle θ_i the impact angle and let l_i the length of the path from the impact point x_{i-1} to the impact point x_i. The following formula is well known in geometrical optics and used everywhere in the billiard literature like in the book of Kozlov-Treshchev.

> LEMMA: There are coordinates for which the Jacobean $DT(x_i, y_i)$ of the billiard map has the form
>
> $$B_i = \begin{bmatrix} 1 & 0 \\ -\frac{2\kappa_i}{\sin(\theta_i)} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & l_i \\ 0 & 1 \end{bmatrix}$$

Remark: This is the composition of the Jacobean belonging to the translation and the Jacobean belonging to the reflection at the wall. The value $g_i = \frac{\sin(\phi_i)}{2\kappa_i}$ is the length of the billiard ball in the circle on the normal to the reflection point which is tangent to the table and has radius $1/(2\kappa_i)$.

PROOF OF THE JACOBEAN FORMULA. The formula can be derived geometrically. Instead, we find an algebraic derivation from the Euler equations. It is still a bit messy.

We use the notation h_1, h_{11} for the first and second partial derivative with respect to the first variable and similar h_{12} for the mixed partial derivative. The billiard map $S : \begin{bmatrix} x_i \\ x_{i-1} \end{bmatrix} \to \begin{bmatrix} x_{i+1} \\ x_i \end{bmatrix}$ is equivalent to the second order recursion $h_1(x_i, x_{i+1} + h_2(x_{i-1}, x_i) = 0$. Differentiation of these Euler equation with respect to x_i, x_{i-1} gives $\partial x_{i+1}/\partial x_i = -b_i/a_i, \partial x_{i+1}/\partial x_{i-1} = -a_{i-1}/a_i$, where

$$a_i = h_{12}(x_i, x_{i+1})$$

and

$$b_i = h_{11}(x_i, x_{i+1}) + h_{22}(x_{i-1}, x_i) \; .$$

The Jacobean of S is

$$dS = \begin{bmatrix} -b_i/a_i & -a_{i-1}/a_i \\ 1 & 0 \end{bmatrix} \; .$$

With a first coordinate transformation $F_i = \begin{bmatrix} a_i^{-1} & 0 \\ 0 & 1 \end{bmatrix}$ we can achieve that the determinant is 1:

$$F_i^{-1} dS F_{i-1} = A_i = (a_{i-1})^{-1} \begin{bmatrix} -b_i & -a_{i-1}^2 \\ 1 & 0 \end{bmatrix} \; .$$

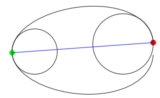
Geometrically, we have

$$a_i = \frac{\sin(\theta_i)\sin(\theta_{i+1})}{l_i}, \; b_i = \sin^2(\theta_i)(\frac{1}{l_i} + \frac{1}{l_{i-1}}) - 2\sin(\theta_i)\kappa_i \; ,$$

where $l_i = h(x_i, x_{i+1})$ are the lengths of the secants, $\theta_i = \theta(x_i, x_{i+1})$ and $\kappa_i = \kappa(x_i)$ are the curvatures at the reflection points. Plugging this in the Jacobean gives with $G_i = \begin{bmatrix} 0 & -\sin(\theta_i) \\ 1/\sin(\theta_i) & \sin(\theta)/l_i \end{bmatrix}$ the new Jacobean

$$G_i^{-1} \cdot A_i \cdot G_{i-1} = \begin{bmatrix} 1 & l_i \\ -\frac{2\kappa_i}{\sin(\theta_i)} & 1 - \frac{2\kappa_i}{\sin(\theta_i)} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{2\kappa_i}{\sin(\theta_i)} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & l_i \\ 0 & 1 \end{bmatrix} \; .$$

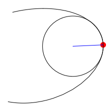
STABILITY OF PERIOD 2 ORBITS. Having the Jacobean given in geometric terms allows to see, whether periodic orbits are stable or not. Inspection of the trace of B_2B_1 (a matrix which is similar to the Jacobean of T^2 and so has the same trace) shows:



LEMMA. Assume ρ_i are the radii of curvature at the impact points. Assume $\rho_1 < \rho_2$. If $l > \rho_1 + \rho_2$ or $\rho_1 < l < \rho_2$, then the periodic orbit of period 2 is hyperbolic. If $l > \rho_2$ or $l < \rho_1 + \rho_2$, it is elliptic.

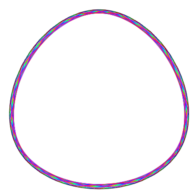
The fastest verification of th lemma is to run a line of Mathematica which gives the trace of the product of the four matrices. For example, the long axis of a non-circular ellipsoid is a hyperbolic periodic point. The short axis is an elliptic periodic point.

CURVATURE. If $r(s)$ is a a curve in the plane parametrized by arc-length, then the curvature $\kappa(t)$ is $|r''(s)|$. If $r(t)$ is the curve given by an arbitrary parameterization, define the unit tangent vector $\vec{T}(t) = \vec{r}\,'(t)/|\vec{r}\,'(t)|$. We get the curvature $\kappa(t) = |\vec{T}'(t)|/|\vec{r}\,'(t)|$. The function $\rho(t) = 1/\kappa(t)$ is called the **radius of curvature**. With the crossed product $(a,b) \times (c,d) = ad - bc$ in two dimensions, we have a more convenient formula $\kappa(t) = \frac{|\vec{r}\,'(t) \times \vec{r}\,''(t)|}{|\vec{r}\,'(t)|^3}$.

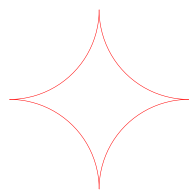


ROLE OF CURVATURE. The curvature of the table plays an important role for the billiard dynamics. Here are some known results:

- Mather has shown that if the table has a flat point, this is a point at which the curvature vanishes like at 4 points of $x^4 + y^4 = 1$, then the billiard map T has no invariant curve at all.
- Lazutkin and Douady have proven using KAM theory that for a smooth billiard table with positive curvature everywhere, there always are "whisper galleries" near the table boundary.



- From Andrea Hubacher (who had obtained this result as an undergraduate student at ETH) is the result that a discontinuity in the curvature of the table does not allow caustics near the boundary. For example, tables obtained by the string construction at a triangle (see homework) do not allow invariant curves near the boundary.
- It is easy to see that billiards for which the table has negative curvature everywhere, the Lyapunov exponent is positive. The Matrices B_i have then positive entries as we will just see.



POSITIVE MATRICES. If we multiply positive matrices with each other, the norm of the product grows exponentially.

LEMMA. If $\det(A(x)) = 1$ for all x and $[A]_{ij}(x) \geq \epsilon > 0$, then the Lyapunov exponent $\lambda(x) = \lim_{n\to\infty} \log ||A(T^{n-1}x)A(T^{k-2}x)\cdots A(x)||$ satisfies $\lambda(A) \geq \frac{1}{2}\log(1+2\epsilon^2)$.

PROOF (Wojtkowski). Define the function F on pairs of vectors by $v = (v_1, v_2) \mapsto F(v) = (v_1 \cdot v_2)^{1/2}$. For a matrix B with determinant 1 satisfying $[B]_{ij}(x) \geq \epsilon$, define $\rho(B) = \inf_{F(v)=1} F(Bv)$.

(i) Given a 2×2-matrix A satisfying $[A]_{ij} \geq \epsilon$. Then $\rho(A) \geq (1+2\epsilon^2)^{1/2}$. Proof: If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $w = (w_1, w_2)$ with $F(w) = (w_1 w_2)^{1/2} = 1$, then $F(Aw) = (aw_1 + bw_2)^{1/2}(cw_1 + dw_2)^{1/2} \geq (ad - bc + 2bc)^{1/2} \geq (1+2\epsilon^2)^{1/2}$.

(ii) $||B|| \geq \rho(B)$. Proof: Take $v = (1,1)$. Then $||A|| \geq \frac{|Av|}{|v|} \geq \frac{F(Av)}{F(v)} \geq \rho(A)$.

(iii) $\rho(AB) = \inf_{F(v)=1} F(ABv) \geq \inf_{F(Bv)=1} \frac{F(ABv)}{F(Bv)} \cdot \inf_{F(v)=1} F(Bv) = \rho(A)\cdot\rho(B)$.

(iv) We get from (ii),(iii),(i) that $\frac{1}{n}\log ||A^n(x)|| \geq \frac{1}{n}\log(\rho(A(T^{n-1}x)\ldots\rho(A(x))) \geq \frac{1}{n}\log((1+2\epsilon^2)^{n/2})$.

CLASSES OF CHAOTIC BILLIARDS. Remember that $g = \frac{\sin(\theta)}{2\kappa}$, and l is the length of the trajectory.

THEOREM (Wojtkowski) Assume, a piecewise smooth convex table has the property that for any pair of points x, x', on the non-flat parts of the curve $2g + 2g' \leq l(x,x')$, with strict inequality on a set of positive measure, then the billiard map T has positive Lyapunov exponents on a set of positive measure.

PROOF. The Jacobian matrix is conjugated to $B_2(x)B_1(x)$. A vector $v = (1, f)$ is mapped by the matrix $B_1(x)$ to the vector $(1, f + l(x))$. This vector is then mapped by $B_2(x)$ to the vector

$$(1 - (f + l(x))/2g(Tx), f + l(x))$$

which is after a rescaling of length equal to the vector

$$(1, \frac{(f+l(x))g(Tx)}{2g(Tx) - f - l(x)} .$$

If we don't care about the length of the vector, the map $v \mapsto B(x)v$ is determined by the map

$$K : f \mapsto f + l \mapsto \frac{1}{1/(f+l) - 1/g(T)} = \frac{(f+l)2g(T)}{2g(T) - f - l} .$$

At each point $x \in X$, we define a basis given by $e_2(x) = (1,0)$ and $e_1(x) = (1, -g(x))$.
Claim: Assume $2g(x) + 2g(Tx) \leq l(x)$ with inequality on a set of positive measure. In this basis, the matrix $B(x)$ is positive and there exists a set of positive measure, where $B(x)_{ij} \geq \epsilon > 0$ for some $\epsilon > 0$ so that we can apply the previous lemma on positive matrices.

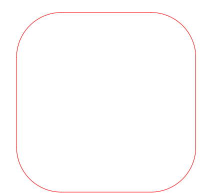
Proof. We have to show that the map K maps the interval $[0, -2g(x)]$ into the interval $[0, -2g(Tx)]$ and into its interior for a set of positive measure because:

$$K(0) = \frac{l(x)2g(Tx)}{2g(Tx) - l(x)} \geq -2g(Tx) .$$

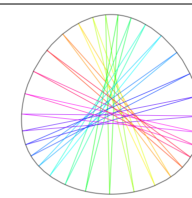
$$K(-2g(x)) = \frac{(-2g(x) + l(x))2g(Tx)}{2g(Tx) + 2g(x) - l(x)} \leq 0 .$$

BUNIMOVICH STADIUM. A famous example is the stadium, where two half circles are joined by straight lines. An other example is the rounded square.

For these billiards, one knows actually much more. They are ergodic and chaotic in the sense of Devaney, a notion we have met earlier in this course. The prove of ergodicity is not so easy. One has to analyze some stable and unstable manifolds and verify that they are dense.



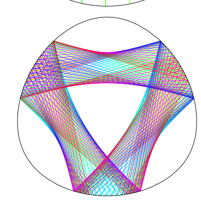
OPEN PROBLEMS. The following problems are open mathematical problems. The first two problems probably go back to Poincaré. The third problem is an old problem in **smooth ergodic theory**. The difficulty of that problem is that for a smooth convex billiards, there are lots of invariant curves and also lots of elliptic periodic orbits consequently, the chaotic regions are mingled well with the stable regions and the techniques described in this handout do not work.



1) Are periodic orbits dense in the annulus for a general smooth Birkhoff billiard?

2) Is the total measure ("area") of the periodic orbits always zero in the annulus? One knows it for period 3 (Rychlik).

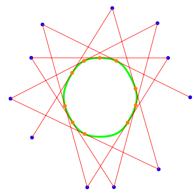
3) Does there exist a smooth convex billiards with positive Lyapunov exponents on a set of positive measure (="area"="probability")?



EXTERIOR BILLIARDS — Math118, O. Knill

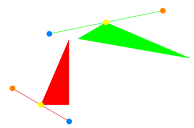
ABSTRACT. We look here briefly at the dynamical system called "exterior billiard". Affine equivalent tables lead to conjugated dynamical systems. One does not know, whether there is a table for which an orbit can escape to infinity nor does not know whether the ellipse is the only smooth convex exterior billiard table for which the dynamics is integrable.

EXTERIOR BILLIARDS. Dual billiards or **exterior billiards** is played outside a convex table γ. Take a point (x, y) outside the table, form the tangent at the table and reflect it at the tangent point (or the mid-point of the interval of intersection). To have no ambiguity with the tangent, γ is oriented counter clockwise. The positive tangent is the tangent at the curve in the same direction.

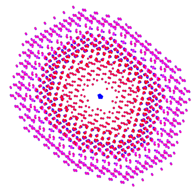


EQUIVALENCE. Assume $S(x) = Ax + v$ is an affine transformation in the plane, where A is a linear transformation and v is a translation vector. Given two tables γ_1, γ_2 such that $S(\gamma_1) = \gamma_2$, then the exterior billiard systems T_{γ_1} and T_{γ_2} are conjugated.

PROOF. Unlike angles, affine transformations preserve ratios and a trajectory of the the exterior billiard at γ_1 is mapped into a trajectory of the table γ_2.

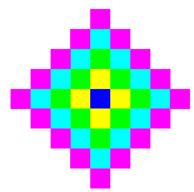


EXAMPLE POLYGONS. Already the case of polygons can be complex. Exterior billiard at a general quadrilateral (=four sided polygon) shows already interesting dynamics. Note that the exterior billiard map is not continuous for polygons. One already does not know whether orbits stay bounded for all quadrilaterals. For regular pentagons, Tabatchnikov was able to compute the Hausdorff dimensions of the closure of some orbits. They are fractals.

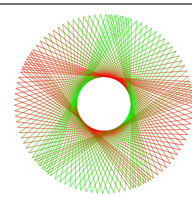


INTEGRABLE PARALLELEPIPED. The exterior billiard at a parallelepiped is integrable.

PROOF. By affine equivalence, it is enough to show this for squares. Check that every orbit is periodic.

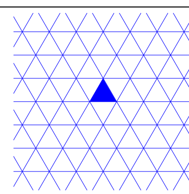


INTEGRABLE ELLIPSE. The exterior billiard at an ellipse is integrable.
PROOF. By affine equivalence, it is enough to show this for circles.



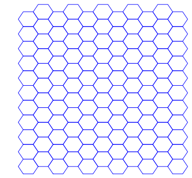
INTEGRABLE TRIANGULAR BILLIARD. The exterior billiard at any triangle is integrable.

PROOF. By affine equivalence, it is enough to show integrability for equilateral triangles. Since every orbit is periodic, we have integrability by a lemma proven earlier. For the hexagon, we have also the property that every orbit is periodic.

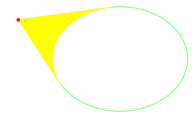


INTEGRABLE HEXAGONAL BILLIARD. The exterior at a regular hexagon is integrable.

PROOF. The key is to see that the successive reflections of the sides of the polygon at the corners of the polygon produces a regular tessellation of the plane.



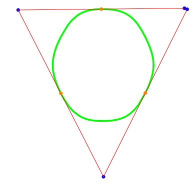
GENERATING FUNCTION. Similar as for billiards, there is a generating function $h(x, x')$ for the exterior billiard. Given two polar angles ϕ, ϕ', draw the tangents with this angle. The function $h(\phi, \phi')$ is the area of the region enclosed by these lines and the curve. We can check that the partial derivative $\frac{\partial}{\partial x} h(\phi, \phi') = -r^2/2$, where r is the distance from the point to the point of tangency. The exterior billiard is area-preserving.



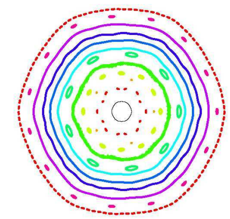
PERIODIC POINTS. By maximizing the functional

$$H(x_1, ..., x_n) = \sum_{k=1}^{n} h(x_i, x_{i+1})$$

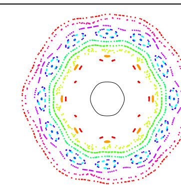
one obtains periodic orbits of the exterior billiard. To say it in words: among all closed polygons for which all sides are tangent to the table, the ones which maximizes the sum of the areas $h(x_i, x_{i+1}$ form a periodic orbit of the dual billiard.



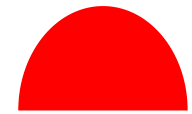
INVARIANT CURVES. For smooth tables, every orbit is bounded. This is a consequence of KAM (Kolmogorov-Arnold-Moser) theory. In that case, there are invariant curves far from the table which enclose the table. A point on this curve will remain on this curve for all times and the dynamics is conjugated to a Kronecker system. A proof of the "invariant curve theorem" is not easy: it requires heavy analytic artillery, modifications of the Newton method or "hard" implicit function theorems. One has to find a smooth invertible map on the circle such that $h(q(x - \alpha), x) + h(x, q(x + \alpha)) = 0$ is satisfied. The irrational rotation number α has to be "far away from rational numbers", one calls this Diophantine. For the story of dual billiards, the proof is even more tricky and has been done by R. Douady.



AN UNSOLVED PROBLEM. Is the ellipse the only smooth convex table for which exterior billiard is integrable?



AN UNSOLVED PROBLEM. Is there a table with an unbounded orbit?
An example of where one does not know the answer, is a semicircle. Tabatchnikov states numerical evidence that for this billiard, there is an unbounded orbit.



HISTORY.
1960. The problem is suggested by B.H. Neumann
1963 The problem is posed by P.Hammer in a list of unsolved problems
1973 In Moser's book "Stable and Random Motion", the stability problem is raised. Some people call exterior billiard also the Moser billiard.
1978 The exterior billiard is also featured in Mosers Intelligencer article "Is the solar system stable".
The photo of Moser to the right had been taken by J. Pöschel in the year 1999, when Moser was lecturing in Edinburgh about twist maps. Moser died in the same year.

FIXED POINT THEOREMS — Math118, O. Knill

ABSTRACT. Fixed point theorems are important in dynamics.

BANACHS FIXED POINT THEOREM.

A contraction T in a complete metric space X has a fixed point.

This theorem can be used for example to prove the existence of solutions to differential equations.

BROWERS FIXED POINT THEOREM.

Every continuous map T from the unit ball $D^n = \{x \in R^n \mid ||x|| \leq 1\}$ onto itself has a fixed point.

SKETCH OF PROOF FOR $n = 2$. If $T(x) \neq x$ for all $x \in D^n$, one can find a continuous map g from D^n to its boundary S^{n-1}: the point $g(x)$ is the intersection of the line through x and $T(x)$ with S^{n-1}. This map is the identity on the boundary. If such a map existed, one could smooth it. We would have a smooth map from the interior of D to S^{n-1}. For most $y \in S^{n-1}$ the set $S^{-1}(y)$ is a curve in D which begins and ends at y. The region it contains must by continuity also be mapped to y and $S^{-1}(y)$ would contain a disc and can not be a curve.

REMARK TO 1D: The Brower fixed point theorem in one dimensions, (D^1 is an interval $[a, b]$) follows from the intermediate value theorem: Since $T(a) \geq a, T(b) \leq b$, the function $g(x) = T(x) - x$ satisfies $g(a) > 0$ and $g(b) < 0$. It must have a root. This root is the fixed point. theorem.

KAKUTANI FIXED POINT THEOREM.

A continuous map T on a compact convex set D in locally convex space X has a fixed point.

(One can relax the condition that T must be a map: it can also be a correspondence for which $T(x)$ is a convex subset of X.) A locally convex set is a vector space in which the topology is given by a sequence of seminorms. An example is $C^\infty(R)$, the space of all infinitely many times differentiable functions.) While von Neumann used Browers fixed point theorem, **John Nash** was among the first to use Kakutani's Fixed Point Theorem in **game theory**, where fixed points can lead to **equilibria**.

POINCARE BIRKHOFF THEOREM.

An area-preserving transformation on the annulus, which moves boundary circles in the opposite directions has at least two distinct fixed points.

Poincare had conjectured this but could no more prove it. The conjecture was therefore called **Poincarés last theorem**. It was George Birkhoff who proved it in 1917.

APPLICATION TO BILLIARDS.

COROLLARY. For a billiard in a smooth convex table, there are at least 2 periodic orbits of type $0 < p/q < 1$ meaning that T^q winds around the table p times.

PROOF. The map T^q leaves one boundary of the annulus $X = T^1 \times [-1, 1]$ fixed, the other boundary is turned around q times. Now define $S(x, y) = (x - 1, y)$ which rotates every point once around. Now, $T^q S^{-p}$ rotates one side of the boundary by $-2\pi p$ and the other side of the boundary by $2\pi(q - p)$. Since the boundary is now turned into different directions, there are fixed points of $T^q S^{-p}$. For such a fixed point $T^q(x, y) = S^p(x, y)$ which is what we call orbit of type $0 < p/q < 1$.

APPLICATIONS TO DUAL BILLIARDS.

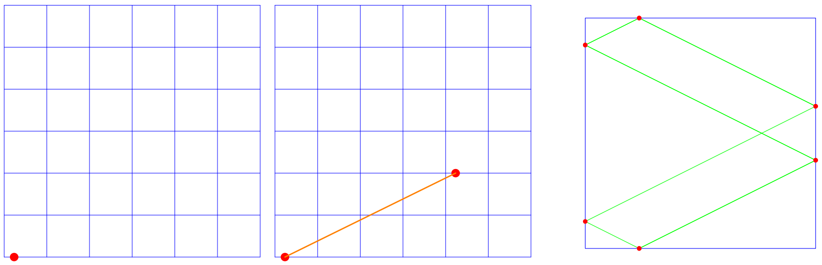
COROLLARY. For exterior billiard at a smooth convex table, there are at least 2 periodic orbits of type $0 < p/q < 1/2$ meaning that T^q winds around the table p times.

Periodic orbits with small rotation numbers p/q are close to the table, periodic orbits with rotation number close to $1/2$ are far away from the table.

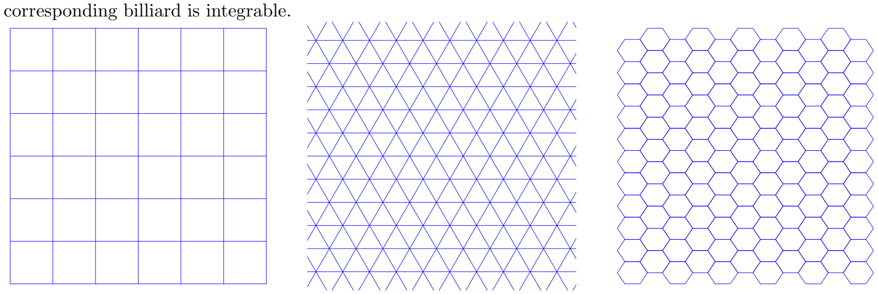
POLYGONAL BILLIARDS **Math118, O. Knill**

ABSTRACT. Billiards in polygons are integrable in the case of rectangles, regular triangles or hexagons.

INTEGRABLE SQUARE. The square and the rectangle are example of an integrable billiard. If θ is the impact angle, then $F(s, \theta) = \sin(2\theta)$ is an integral.



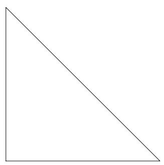
INTEGRABLE POLYGONAL BILLIARDS. If unfolding the polygon produces a tesslation of the plane, the corresponding billiard is integrable.



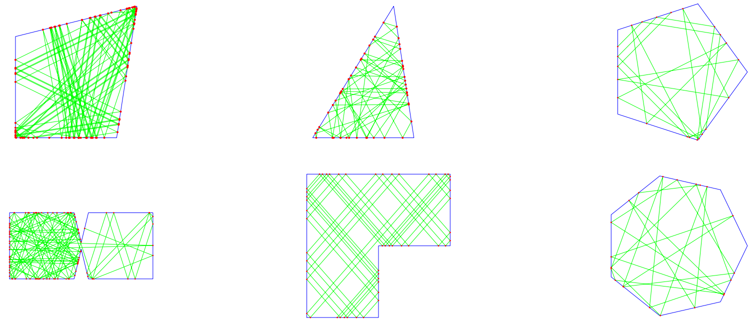
TRIANGULAR BILLIARDS. Even for triangles, the billiard dynamics is complicated. There are many open questions, one of the most astonishing ones is the open problem:

Does every triangular billiard have a periodic orbit?

One can solve the problem for a triangle with a right angle? The answer is easy - if you see it. One can also solve the problem for acute triangles, where the **Fagnano trajectory** connecting the footpoints of the triangles altitudes is a periodic orbit.



LETS PLAY SOME GAMES: Lets mention without proof that the Lyapunov exponent of a polygonal billiard is always zero. The chaos, you obtain with these systems is "weak".



LYAPUNOV EXPONENTS. Because the Jacobean matrix of a billiard is conjugated to $\begin{bmatrix} 1 & l_i \\ -\frac{2\kappa_i}{\sin(\theta_i)} & 1-\frac{2\kappa_i}{\sin(\theta_i)} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{2\kappa_i}{\sin(\theta_i)} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & l_i \\ 0 & 1 \end{bmatrix}$ and the curvature in a polygonal billiard is zero, we have

All Lyapunov exponents are zero in polygonal billiards.

CONNECTIONS WITH OTHER FIELDS. The mathematics of billiards in polygons has relations with other fields like Riemann surfaces, Teichmuller theory and leads to interesting ergodic theory. One knows for example that for a "generic" polygon, the billiard map is ergodic.

CELLULAR AUTOMATA — Math118, O. Knill

ABSTRACT. A shift invariant continuous map on the sequence space A^Z over a finite alphabet A is called a **cellular automaton** or short a CA. These dynamical systems can be considered as discretized cousins of differential equations, for which time, space, as well as the configuration space are discretized.

THE NAME CELLULAR AUTOMATON. Interactions between different scientific fields is always productive. Historically, it seems that cellular automata were introduced in the late 40ies while some applied Mathematicians were dealing with problems from biology. The etymology of the name "CA" could confirm a "bonmot" of Stan Ulam:

> Ask not what mathematics can do for biology.
> Ask what biology can do for Mathematics.

Source: cited from David Campbell, who received his B.A. in chemistry and physics from Harvard in 1966 and worked in nonlinear science. Ulam himself was at Harvard from 1936-1939, eating at Adams house where "the lunches were particularly agreeable" and was also teaching the Math1A here (Source: Ulam: Adventures of a mathematician).
Anyway, it would not surprise if "cellular automaton" had been derived from "cellular spaces" because of mathematical research on biological problems.

SEQUENCE SPACES. Let A be a finite set called the **alphabet** and let A^Z denote the set of all sequences and $\sigma(x)_n = x_{n+1}$ the shift on X. A distance between two sequences is given by $d(x,y) = 1/(n+1)$, where n is the largest number such that $x_i = y_i$ for $|i| \leq n$. Example: Let $A = \{1,2,3,4\}$. For

...	x_{-3}	x_{-2}	x_{-1}	x_0	x_1	x_2	x_3	...
...	1	1	4	3	2	1	1	...

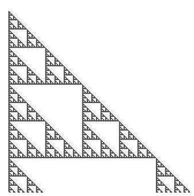
...	y_{-3}	y_{-2}	y_{-1}	y_0	y_1	y_2	y_3	...
...	1	2	3	3	4	1	1	...

we have $d(x,y) = 1/3$, because $x_i = y_i$ if $|i| \geq 3$ but $x_{-2} \neq y_{-2}$.

> LEMMA: X is a **compact metric space** (X,d).

PROOF. To have a metric space, show $d(x,x) = 0, d(x,z) \leq d(x,y) + d(y,z), d(x,y) = d(y,x)$. To have compactness, every sequence $x(k)$ in X must have an accumulation point. That is, there must exist a subsequence $x(k_l)$ in X which converges for $k \to \infty$. See homework.

1D-CELLULAR AUTOMATA. A continuous map T on X which commutes with σ is called a **cellular automaton**. A theorem of Curtis, Hedlund and Lyndon, which we will prove later implies that there is a function ϕ from $A^{2R+1} \to A$ such that $T(x)_i = \phi(x_{i-R}, x_{i-k+1}, ..., x_{i+R})$. The integer R is called the **radius** of the CA. It is assumed that R is the smallest number for which the CA still can be defined like that. One can visualize the dynamics of one dimensional CA by coding each letter in a sequence with a color. The first row is the initial condition. Applying the map gives the second row, etc. Drawing a few iterates produces a **phase space diagram**. The example shows the automaton over the alphabet $\{0,1\}$, where $x_n = x_n + x_{n-1}$ mod 2 and where 0 is black. If initially $x_n(0) = 0$ for $n \neq 0$ and $x_0(0) = 1$, we have an explicit solution formula with binomial coefficients $x_n(t) = \begin{pmatrix} n+t \\ n \end{pmatrix}$ mod(2).



CANTORS DIAGONAL ARGUMENT.

> THEOREM (Cantor) The set $X = A^Z$ is uncountable.

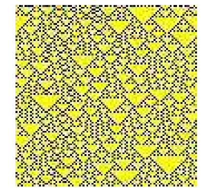
PROOF. If X were countable, one could enumerate all sequences $x(k)$ using integer indices k. Define the "Diagonal" sequence $y_n = (1 + x_n(|n|))$ (here $a+1$ is the next in the alphabet A, or the first element in A, if a was the last). The sequence y is different from any of the sequences $x(k)$ because y and $x(k)$ differ at the k'th entry. The assumption about the enumerability was not possible.

WOLFRAMS NUMBERING OF 1D CA. Any one-dimensional cellular automata with radius 1 and alphabet $\{0,1\}$ can be labeled by a **rule number**. Because there are $2^3 = 8$ possible maps ϕ, we have $2^8 = 256$ possible rules. The **Wolfram number** is $w = \sum_{k=1}^{8} f(k)2^k$, where $y_0 = f(k)$ is the new color for $k = 4x_{-1} + 2x_0 + x_1$.

For example, let $\phi(a,b,c) = a$, then the new middle cell is 1 for the neighborhoods 111, 110, 101, 100 which code the integers 7, 6, 5, 4. So, $f(7) = f(6) = f(5) = f(4) = 1$, and $f(k) = 0$ otherwise. The rule of the automaton is $w = 2^7 + 2^6 + 2^5 + 2^4 = 240$. Indeed, rule 240 is the shift automaton. Let us look at an other example.

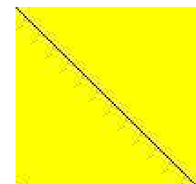
EXAMPLES. The binominal CA discussed above has rule 90. One of the most studied CA is rule 18. Since $18 = 2^4 + 2^1$, which is 10010 to the base 2, we obtain the following function ϕ:

neighborhood (dec)	neighborhood (bin)	new middle cell	factor
7	111	0	128
6	110	0	64
5	101	0	32
4	100	1	16
3	011	0	8
2	010	0	4
1	001	1	2
0	000	0	1



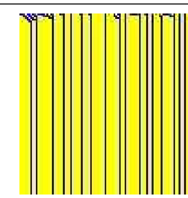
SPEED OF A CA. Every CA has a maximal speed c with which signals can propagate. This means if we take an initial conditions x which is constant outside an interval I, then then $T^k(x)$ will still be constant outside an interval I_k of size $|I_k| \leq |I||2c$

> LEMMA. The speed of a CA is bounded above by the radius R.



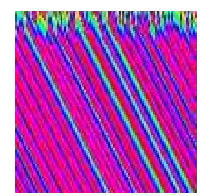
PROOF. Each timestep can change only cells maximally R units to the left or to the right.
Example: The **"Takahashi-Susama Soliton automaton"** is defined on points $x \in \{0,1\}^Z$ for which only finitely many cells are 1. The rule for T is to start from the left and move each 1 to the next 0 position. Since a pack of n adjecent $1's$ moves with speed n, the map T is **not** a cellular automaton.

EXAMPLES.
a) The cellular automaton $T = \sigma^c$ shifting $c \in \mathbf{N}$ entries to the right has the speed c. Since c is also the radius, this shows that the speed can not be faster then the radius R. The **speed ratio** c/R satisfies $c/R \leq 1$.
b) The CA $T(x)_n = (..., a, a, a, a, ...)$ is obtained by a function ϕ which is constant. Every orbit of this automaton is attracted to the fixed point. The speed is zero. The picture to the right shows rule-100 cellular automaton.



POSSIBLE SPEEDS. Note that we can enumerate the set of cellular automata: it is a countable set. Because the set of real numbers in the interval $[0,1]$ is uncountable, we can not obtain all the speeds.

> PROPOSITION. Fix A. For every $0 < a < b < 1$, there is a CA with radius R over the alphabet A for which the speed c satisfies $a \leq c/R \leq b$.



You explore this fact a bit in a homework. The idea is first to use a larger alphabet in order to slow down the motion using internal "color swapping". For different alphabets A, B, a A-automaton can be simulated by a B automaton, possibly changing the radius.

CELLULAR AUTOMATA II — Math118, O. Knill

ABSTRACT. This page contains three mathematical results: the Curtis-Hedlund-Lyndon theorem which says that every continuous, translational invariant map on X is a CA, the proof that σ is chaotic in the sense of Devaney and on a rather technical proof that the topological entropy which we define for CA agrees with the classical topological entropy for general topological dynamical systems.

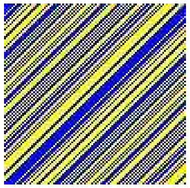
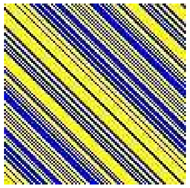
THE CURTIS-HEDLUND-LYNDON THEOREM.

> For every continuous map T on $X = A^Z$ which commutes with σ, there is a finite set $F = \{-R, ..., R\}$ and a map ϕ such that $T(x)_n = \phi(x_{n-R}, ..., x_{n+R})$.

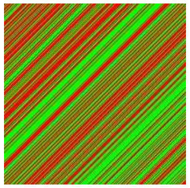
PROOF.
(i) We claim that the map f from X to A defined by $f(x) = T(x)_0$ depends only on $\{x_i, i \in F(x)\}$, where $F(x)$ is some finite set.
Proof: If this were not true, there existed a sequence $x(n_k)$ in X with $n_k \to \infty$ such that $x_l = x(n_k)_l$ for $l \neq n_k$ and $x_l \neq x(n_k)_l$ for $l = n_k$ and $T(x(n_k)) \neq T(x)$. Because $x(n_k) \to x$ for $k \to \infty$, the continuity of T implies that $T(x(n_k)) = T(x)$ eventually because of the finiteness of the alphabet. This is a contradiction to $T(x(n_k)) \neq T(x)$ for all k.

(ii) The set $F(x)$ is independent of x.
Proof. First of all, $x \to m(x)$, where $m(x) = \min(F(x))$ and $x \to M(x)$, where $M(x) = \max(F(x))$ are continuous. This implies that $x_n \to x$ implies $F(x_n) = F(x)$ if $d(x_n, x)$ is close enough. The set $F(x)$ is invariant under the shifts σ by assumption. Assume, there exist two points x, y, where $F(x) \neq F(y)$. We can find z and sequence of translations σ^{n_j} such that $\sigma^{n_j}(z) \to y$ and a sequence of translations m_k such that $\sigma^{m_k}(z) \to y$. We have $F(z) = F(\sigma^{n_j} z)$ and $F(z) = F(\sigma^{m_k} z)$ and so $F(x) = F(y)$.

ISOMORPHIC AUTOMATA. Some of the elementary automata are isomorphic. For example, the parity transformation $P(x)_n = x_{-n}$, then $P^{-1}TP$ is a new elementary automaton with a different number. Also $C(x)_k = (1 - x_k)$ which changes 0 and 1 brings a new automata $C^{-1}TC$. Many of the 256 different rules lead to isomorphic systems. Counting the equivalence classes reduces the number 256 to 88. The pictures to the right show rule 170 and rule 240, the left and right shift.

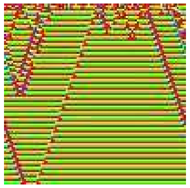
THE "CHAOTIC" SHIFT. The shift map σ is also CA with rule 240.



> CA is chaotic in the sense of Devaney: it has a dense set of periodic points and has a dense orbit.

PROOF. To get a dense orbit, enumerate all finite words w_k and concatenate them together to an infinite sequence y, for $k > 0$. Define $x_k = y_{|k|}$. $T^n(x)$ is dense.
For every x, and every ϵ, there exists a N-periodic sequence y such that $d(x, y) < \epsilon$.

PARTICLES INTERACTIONS. Automata with nearest neighbor interaction and larger alphabets can exhibit already quite interesting behavior. Physicists are intrigued by the similarity to particle physics. Certain configurations travel with some speed, interact and destroy each other like real particles. The picture to the right shows the automaton over the alphabet $Z_p, p = 9$ with $\phi(a, b, c) = a * b * c + 1$. If the CA rule is the "physics" of the "CA micro world", one calls **particles** elements in X which are constant outside some interval and which satisfy $T^n(x) = \sigma^m(x)$. They have speed $v = m/n$. If you are lucky, the interaction of particles produces new particles.



SUBSHIFTS. A closed σ-invariant subset X of A^Z is called a **subshift**. If a subshift X is invariant under a CA map T, we can look at the system (X, T). Examples:

a) If $x = (..., 0, 1, 1, 0, 1, 1, 0, 1, 1, ...)$, then $X = \{x, \sigma(x), \sigma^2(x)\}$ is a subshift. More generally, the set of all M-periodic sequences forms a subshift. Restricting a CA map T onto X means simulating the CA with periodic boundary conditions.

b) Take all sequences with alphabet $\{a, b, c\}$, so that transitions $a- > b- > c- > a$ and $b- > b$ are possible. The space X with words like $(, ..., abcabcabbcabbbcabcabbbbc, ...)$. is an example of a **subshift of finite type**.

c) If T is a cellular automaton map and X is a subshift, then $T(X)$ is a subshift. It is called a **factor** of the original subshift. That is how CA were first introduced by Hedlund.

ATTRACTOR. The image $X_1 = T(X_0)$ of the set of all configurations $X_0 = A^{Z^d}$ is a T invariant subshift. The image $X_2 = T(X_1)$ is invariant too etc. We obtain a nested sequence of subsets $X_0 \supset X_1 \supset X_2....$ The limit $X = \bigcap_k X_k$ is called the **attractor** of the cellular automaton. It is a T-invariant subshift.
EXAMPLES. For the shift σ, the attractor is the entire set A^Z. For the rule 0-automaton, the attractor is a single point.

TOPOLOGICAL ENTROPY OF 1D CA. The topological entropy of a 1D CA is defined as

$$h(T) = \lim_{K \to \infty} \lim_{N \to \infty} \frac{\log(R(N, K))}{N} .$$

where $R(N, K)$ be the number of distinct rectangles of width K and height N which occur in a space-time diagram of T.
The picture to the right shows a rectangle $R(N, K)$ for an automaton, where the attractor is a point. Here $R(N, K)$ depends on K but stays bounded in N. The entropy is zero.

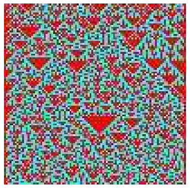
EXAMPLE. The shift $T = \sigma$ has the maximal possible entropy $\log(|A|)$. Take a random sequence x, then $T^n(x)$ will be random sequences too. We have $R(N, K) = |A|^N$.

TOPOLOGICAL ENTROPY IS DIFFICULT TO COMPUTE:

> THEOREM (Hurd,Kari and Culik) Given $\epsilon > 0$. There is no computer algorithm which when given as an input the rule of the CA, the output is the topological entropy up to accuracy ϵ.

The strategy of the proof is to relate the problem of calculating the entropy to the "stopping problem of Turing machines, which is a undecidable problem: there exists no algorithm which takes a Turing machine and decides whether it halts or not.

BOUNDARY CONDITION. If an initial sequence x is periodic, satisfying $x_{i+N} = x_i$ for all i, then $T(x)$ is periodic. We can then watch $x_1, ..., x_N$ and know the entire sequence. In this case, the possible configurations are finite, namely $|A|^N$, where $|A|$ is the cardinality of the alphabet A. The cellular automata map is a map on a finite set X_N.
We can also take fixed boundary conditions, assuming that $x_0 = x_N = 0$. In analogy to PDE's (and CA are in a sense discrete PDE's), one could call this **Dirichlet boundary conditions**.



GROWTH OF LARGEST ATTRACTOR. For a fixed automaton we can look at the size $s(N)$ of the largest attractor on the subshift $X = X_N$ set N periodic sequences. Define the growth rate

$$0 \leq \limsup_N \frac{1}{N} \log(s(N)) \leq \log |A|$$

This growth rate is different from the topological entropy in general: the growth rate of the shift σ is 0, while the topological entropy is $\log |A|$.

GENERAL DEFINITION OF TOPOLOGICAL ENTROPY. The topological entropy of a continuous map T on a compact space X is in general defined as $h(T) = \lim_{\epsilon\to 0} \overline{\lim}_{n\to\infty} \log(M(n,\epsilon))/n$, where $M(n,\epsilon)$ is the minimal number of ϵ-balls in the metric $d_n(x,y) = \max_{0\le i\le n-1} d(T^i x, T^i y)$ which cover X.

The topological entropy of the CA agrees with the general topological entropy:

PROOF. Given two $(N, 2K+1)$-rectangles A, B in the space-time diagram. Enumerate the rows of A and B starting from the bottom with $A_1, \dots, A_N$ and $B_1, \dots, B_N$ and take two elements $x, y \in X$ such that

$$A_j = (T^j(x)_{-K}, \dots T^j(x)_{-1}, T^j(x)_0, T^j(x)_1, \dots, T^j(x)_K) ,$$

$$B_j = (T^j(y)_{-K}, \dots T^j(y)_{-1}, T^j(y)_0, T^j(y)_1, \dots, T^j(y)_K) .$$

Because $A_j = B_j$ if and only if $d(T^j(x), T^j(y)) < 2^{-K}$, we know that $A = B$ implies $d_N(x,y) < 2^{-K}$. On the other hand, if $x, y \in X$ satisfy $d_N(x,y) \ge 2^{-K}$, we have two different rectangles. With

$$M(N, 4\cdot 2^{-k}) \le R(N, 2K+1) \le M(N, 2^{-k}/4) .$$

(i) Left inequality. Take for each $R(N, 2K+1)$ rectangles A a point x such that

$$A_1 = (x_{-K}, \dots x_{-1}, x_0, x_1, \dots, x_K) .$$

This gives a finite set $Y \subset X$ with $R(N, 2K+1)$ points. Every point $x \in X$ has distance $\le 2\cdot 2^{-K}$ to one of the points in Y. The $R(N, 2K+1)$ balls of radius $4\cdot 2^{-k}$ with midpoints in Y cover X. This proves a).

(ii) Right inequality: two different points in Y have distance $\ge 2^{-K}/2$. We need therefore at least $R(N, 2K+1)$ balls of radius $2^{-K}/4$ to cover X.

The two inequalities together give $R(N, 2(K+4)+1)) \ge M(N, 2^{-(K+2)}) \ge R(N, 2K+1)$ so that

$$\lim_{N\to\infty} \frac{\log(R(N, 2(K+4)+1))}{N} \le \lim_{N\to\infty} \frac{\log(M(N, 2^{-(K+2)}))}{N} \le \lim_{N\to\infty} \frac{\log(R(N, 2K+1))}{N} .$$

For $K \to \infty$, the left and right limits converge to the same number. The limit in the middle is the topological entropy.

HIGHER DIMENSIONAL CA

Math118, O. Knill

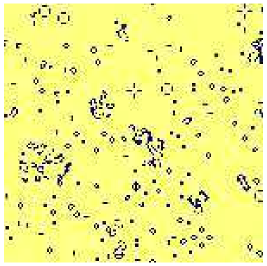
ABSTRACT. We look at some higher dimensional automata like the game of life or lattice gas automata. Note that 2 hours after this lecture, unix time is 1111111111 = Fri, 18 Mar 2005 01:58:31.

HIGHER DIMENSIONAL AUTOMATA. Everything said before can be generalized to higher dimensions. Lets restrict to two dimensions. The space is $X = A^{Z^2}$. It consists of elements $x_{n,m}$, where (n,m) are the coordinates. Define the shifts $\sigma_1(x)_{n,m} = x_{n+1,m}, \sigma_2(x)_{n,m} = x_{n,m+1}$. A continuous map on X which commutes with both σ_i is called a Cellular automaton. We have $T(x)_n = \phi(x_m)$ with $n-m$ in some finite set F. The composition of two CA is a CA. A distance is defined as $d(x,y) = 1/(n+1)$ if $x_k = y_k$ for $|k| \leq n$ and $x_l \neq y_l$ for some $|l| = n$, where $|(i,j)| = |i| + |j|$.

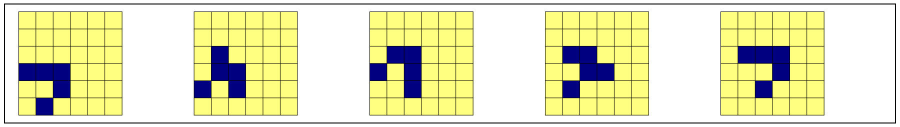
GAME OF LIFE. One of the most famous automaton is **Conways game of life**. A dead cell comes alive if and only if it has three neighbors. A live cell dies if it has less then 2 ore more than 3 neighbors.

SPECIAL SOLUTIONS. A configuration x has **compact support** if there are only finitely many cells which are alive. Examples of solutions with compact support are gliders, stones and blinkers.

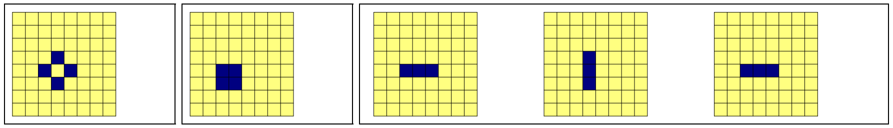
The picture to the right shows life after a random initial condition, after having iterated for 500 iterations.



GLIDERS. Solutions which satisfy $T^n(x) = \sigma^v(x)$ for integer n and $v = (v_1, v_2)$ are called **gliders**. Gliders travel with velocity v/n. If x is a glider, then $T^n(x)$ converges to 0.

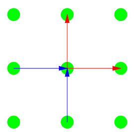


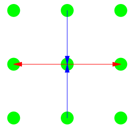
PERIODIC SOLUTIONS. If $T^n(x) = x$, then x is called a periodic solution of T. The left two configurations below show fixed points called "stones". We also see a periodic two orbits called "blinker".

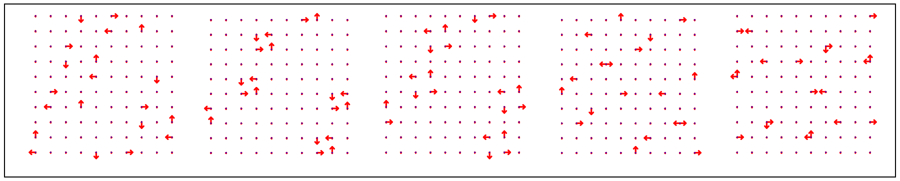


THE HPP MODEL. is a simple deterministic two-dimensional cellular automata designed by Hardy,Pazzis and Pomeau in 1972. Its aim to have a simple toy model to simulate the Navier Stokes equations. The automaton has a color for each of the possible particle configurations. There can be maximally 4 particles at the same spot. One assigns a letter to each of the 16 configurations.

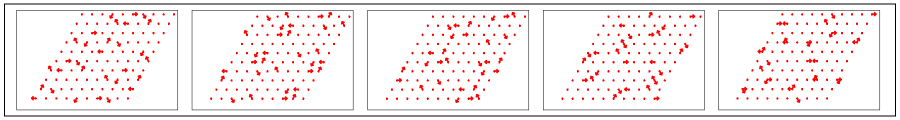
Particles always point away from the origin. Either there is a particle in one of the four directions, or there is not. Once can code each color with a code like $(n,w,s,e) = (1,1,0,1)$ The rules are designed such that particles move freely. For example, if if $x_{n,m} = (0,0,0,1)$ and all other nodes satisfy $x_{i,j} = (0,0,0,0)$, then $x_{n+1,m} = (0,0,0,1)$. A particle has moved from node (n,m) to node $(n+1,m)$. If particles collide with a right angle, they will scatter as if they would pass through each other. If they hit head on, both directions change by 90 degrees.







HEXAGONAL LATTICE GAS CA. Designed by Frisch,Hasslacher and Pomeau in 1985. The rules are designed to conserve particle number and momentum at each vertex. Additionally, there is a random number generator, when particles collide head on. The possible directions in which the particle pair can scatter is chosen randomly. Also this lattice gas automaton conserves particle numbers as well as momentum of the particles.



ATTRACTOR. The image $X_1 = T(X_0)$ of the set of all configurations $X_0 = A^{Z^d}$ is a T invariant subset. The image $X_2 = T(X_1)$ is invariant too etc. We obtain a nested sequence of subsets $X_0 \supset X_1 \supset X_2 ....$ The limit $X = \bigcap_k X_k$ is called the **attractor** of the cellular automaton. It is a closed T-invariant subset and $T(X) = X$.

WHERE DO CA BELONG?

Space	Time	States	Object
Continuous	Continuous	Continuous	Partial differential equations
Continuous	Discrete	Continuous	Maps on function spaces
Discrete	Continuous	Continuous	Coupled differential equations
Discrete	Discrete	Continuous	Coupled map lattices
Discrete	Discrete	Discrete	Cellular automata

PDE Example: $\frac{d}{dt}u(x,t) = f(u(x,t), \frac{d}{dx}u(x,t))$.
Maps on functions: $u(x,t+1) = f(u(x,t))$.
Coupled differential equations: $\frac{d}{dt}u(n,t) = f(u(n-1,t), u(n,t), u(n+1,t))$
Coupled map lattices: $u(n,t+1) = f(u(n-1,t), u(n,t), u(n+1,t))$ with $u(t,x)$ real.
Cellular automata: $u(n,t+1) = f(u(n-1,t), u(n,t), u(n+1,t))$ with $u(t,x)$ finite.

HISTORY. Numerical treatments of ODE's and PDE's leads to CA: Example: the heat equation $u_t = u_{xx}$ leads to a difference equation $u(t+1,x) - u(t) = cu(t,x+1) - 2u(t,x) + u(t,x-1)$ which becomes a CA, when $u(t,x)$ takes finitely many values only. If the PDE is translational invariant, the discretisation is a CA with an alphabet of $1/\epsilon$ elements, if the computing accuracty is ϵ. Difference methods for PDEs were used since a long time, at least since 1920 (L.F. Richardson), and research on it exploded during WW2 and when the first computers appeared (i.e. the first electronic computer ENIAC in 1945). John von Neumann seemed have introduced CA in these years. Ulam claims to have found CAs first in "Adventures of a Mathematician" p.285: "my own simple minded model". 1936 Turing machines are shown to be able to do all computations. A Turing machine with n states and a tape alphabet of k symbols is a special cellular automaton with an alphabet of $n+k$ letters.

1950 Idealized models of biological systems were studied using CA. Ulam and von Neuman called this "nearest neighbor-connected cellular spaces". Source: From Cardinals to Chaos, Ed: Necia Grant Cooper, Cambridge University Press.

1969 Gustav Hedlund considered in the mid 50ies "shift commuting block maps". see "Endomorphisms and automorphisms of the shift dynamical systems" Math. Systems Theory 3, p.320-375, (1969). Hedlund got his PhD at Harvard in 1930.

1970 Conway article on the "game of life" in the Scientific American 223, (October 1970): 120-123. The name CA had already been coined, like in "Essays on cellular automata Ed. Arthur W. Burks, 1970.

2004 MathSciNet shows 3328 papers authored on Cellular automata.

MORE ON CA

ABSTRACT. We add some additional remarks about CA and an open problem.

AUTOMATA ON GRAPHS. Cellular automata can be defined in any dimensions and even on any homogenous graph, where each node looks the same. A popular "two dimensional" example different from the square lattice is the hexagonal lattice. Setting up the CA story on more general graphs is nothing more than changing notation.

A general class of graphs, for which most of the theory goes over are **Cayley graphs** Γ of finitely presented groups like $G = \{a, b \mid a^2 b = ba^2\}$. The graph has nodes for each word in the generators a, b and two nodes v, w are connected, if $va = w$ or $av = w$ or $w = va$ or $w = vb$.
As a metric, one first introduces the geodesic distance in the graph Γ which is the shortest number of steps (applying one of the generators of the group G) to get from one point to the other. Write $|k|$ for the distance to the origin. The distance between two configurations in $X = A^\Gamma$ is still defined as $d(x, y) = 1/(n + 1)$, where $x_k = y_k$ for $|k| \leq n$ and $x_k \neq y_k$ for some k satisfying $|k| = n$.

Hedlunds theorem still applies: a continous map on $X = A^\Gamma$ which is invariant under translations (applying the group G on the Cayley graph) is defined by a local law ϕ.

The proof we have given before applies almost word by word: the continuity of the map T forces a local law. The translational invariance and the fact that the action of the group G on the graph is transitive, implies that the law is the same at every node.

PROBLEMS WITH CA. The discretisation distroys **rotational symmetry**. In the plane, one can make CA more symmetric by using a hexagonal lattice but still, there is no rotational symmetry. Even in the limit when the cells become infinitesimally small, their stucture can be seen from the propagation of solutions.

SURJECTIVITY. Which automata are T are invertible maps on X and so homeomorphisms (every bijective map on a compact space has a continuous inverse). It is also known that an injective CA is surjective. To check injectivity, one actually can restrict to finite configurations. These results had been obtained in the 60ies. Thew fact that injectivity implies surjectivity is called a **"Garden of Eden theorem"**. The from E.F. Moore coined expression "Garden of Eden patterns" is a picturesque name for points in X, which are not in the image of T.

AN OPEN PROBLEM. An automaton T is called transitive, if it has a dense orbit in X. We have seen that the shift is transitive. We also have seen that the shift has a dense set of periodic points. F. Blanchard asks:

> Does every transitive automaton have a dense set of periodic points?

Francois Blanchard writes: "The answer, positive or negative, is a necessary step before on understands the meaning of chaos in the field. " Source: This problem can be found in Michael Misiurewicz list of open problems in dynamical systems (http://www.math.iupui.edu/ mmisiure/open)

THE SEMIGROUP OF CA. If you have a CA T and a CA S defined on the same space X, then $T \circ S$ is a new CA. So, the set of all CA is a semigroup. Historically, this was one of the original ways how CA were introducd because according to Hedlund, cellular automata are just the homomorphism on the category of subshifts. Note that the semigroup of all cellular automata is not commutative.
If you look at the set of all CA which are invertible, then the set of all these cellular automata forms a group. The identity in this group is the trivial CA, where $T(x) = x$.

A CLASS OF REVERSIBLE AUTOMATA. Given an alphabet A and an elementary automaton T defined by a function $\phi : A^3 \to A$ we can define an automaton

$$T(x, y)_i = (y_i + \phi(x_{i-1}, x_i, x_{i+1}), x_i)$$

The map T is now invertible with the inverse $T^{-1}(x, y)_i = (y_i, x_i - \phi(y_{i-1}, y_i, y_{i+1})$. It suffices to look the first coordinate because $y(t) = x(t - 1)$.

This automaton on can actually be written as an automaton on $X = B^Z$, where B is the alphabet $A \times A$. For example, for $A = \{0, 1\}$, the new alphabet B is $\{(0, 0), (1, 0), (0, 1), (1, 1)\}$. The translation if $x_k = (0, 1)$, then this would correspond to $(x_k, y_k) = (0, 1)$ in the original picture.

CA AS MAPS ON SUBSHIFTS. If X is a **subshift** that is a shift invariant subset of A^Z, and T is a CA map, then $T(X)$ is again a subshift. It is called a factor of X. There are some properties of subshifts which stay the same after applying CA maps.

- **Topological transitive**: there exists a dense orbit.
- **Almost periodic = minimal**: every orbit is dense.
- **Uniquely ergodic**: there exists exactly one invariant measure.
- **Strictly ergodic**: minimal and uniquely ergodic.
- **Dense set of periodic orbits**: x periodic orbit: $T^n(x) = x$.
- **Prime**: Every factor of (X, T) is either trivial or isomorphic to X.
- **Totally minimal**: No factor is a finite permutation.
- **Completly positive entropy**: all non trivial factors have positive directional entropy.
- **Zero directional topological entropy**:
- **Topologically strongly mixing**: U, V open. Exists $n \in Z$ such that $U \cap T^n V \neq \emptyset$.
- **Topologically weakly mixing**: $X \times X$ is topologically transitive.
- **Uniquely ergodic, strong mixing**: $\mu(U \cap T^n V) \to \mu(U) \cdot \mu(V)$.
- **Uniquely ergodic, weakly mixing subshifts**: $X \times X$ is ergodic.
- **Sophic**: a factor of a subshift of finite type.
- **Chaotic in the sense of Devaney**: topological transitive and dense set of periodic orbits .

(If one requires additionally that the shift is not periodic, then this property is not invariant. There are shifts which have periodic factors).

> Cellular automata maps can be used to generate new subshifts with given dynamical properties!

Is this useful? It can be. If you have a complex subshift to analyze and if you can show that it is obtained by applying CA maps from a simpler shift, then you have proven that the subshift inherits the properties of the initial subshift.

ABOUT COMPLEXITY. The shift acting on all periodic sequences is not very spectacular. It just rotates a sequence. Every orbit is n periodic. Other cellular automata like rule 30 have complexer behavior when restricted to periodic sequences in the sense that there are longer periodic orbits in that space X of 2^n possible configurations. Note that T can never be transitive on X in the periodic setup because if you start with a constant sequence x, then $T(x)$ is a constant sequence. But orbits can get long.

> The complexity of a dynamical system can depend dramatically on the space, on which it is defined.

> REMINDER: A linear map A like the cat map on R^2 behaves differently then the same map on the torus R^2/Z^2. The map on the torus is complex. However, when restricting the map on the set of rational points $(x, y) \in X$, the map is not complex at all: every orbit is eventually periodic.
>
> REMINDER: The free motion of a particle in the plane is trivial. But when confined to a finite region (a billiard table), the motion can become complex. Then again, restricting this complex motion to some subset can be completely understandable like restrictiion to the invariant curve on which the dynamics is just a translation.

Talking about the complexity of a map or differential equation does not make sense per se. The set X on which one wants to understand the system is important. Complexity is often mentioned in discussions about CA. Like other **buz words**, the word is loaded with many different meanings. One precise mathematical definition is the "computational complexity of a problem" which is a measure on how the number of computations grows with a parameter of the problem.

TURING MACHINES — Math118, O. Knill

ABSTRACT. This is an excursion into a class of dynamical systems called Turing machines. They are remarkable because any computation can be done by Turing machines. Because Turing machines can be realized as subshifts and subshifts are abundant in dynamical systems theory, most dynamical systems like the Henon map would be capable to do any possible computation.

TURING MACHINES. A Turing machine is a dynamical system (Y, T) defined as follows. Define $Y = X \times S = \{0,1\}^Z \times S$, where S is a finite set of states. The set S contains an element 0, which is called the halting state. The set $\{(\dots, 0, \dots)\} \times S$ is called the empty tape. The set X is the space of $0,1$ sequences for which only finitely many 1 are called data. The Turing machine is defined by three maps from finite sets to finite sets.

$f : \{0,1\} \times S \to \{0,1\}$	defines the new letter
$g : \{0,1\} \times S \to S$	defines the new state
$h : \{0,1\} \times S \to \{-1,0,1\}$	decides whether to move the tape to left, right or stay

one can define now a continuous map on the compact metric space Y by

$$T(x, s) = (\sigma^{h(x_0,s)}(\cdots, x_{-2}, x_{-1}, f(x_0, s), x_1, x_2 \cdots), g(x_0, s)) .$$

This dynamical system is called a Turing machine. Note that this is not a CA, since the map does not commute with the shift. But alread John von Neumann noticed that one can find for every Turing machine a CA, which simulates the Turing machine. Note that the set Y is not compact but it is a subset of a compact set.

HALTING STATE The description of a Turing machine is given by a finite amount of information, because the three involved functions map finite sets into finite sets. The set $X \times \{0\}$ is called the halte set. One step of a Turing machine can be described as follows: the Turing machine with tape x and state s moves the tape $h(x, s)$ steps goes into the state s and then writes the entry $f(x, s)$ at the position 0.

CHURCH THESES. Turing showed, that every computation which can be done by known computations can be done by Turing machines. The question of what actually can be computed is probably beyond the scope of mathematics. There is a widely accepted statement called the Church thesis (1934) which tells that everything which can be computed can be computed with a partial recursive function. Such functions can be computed by Turing machines. Everything we know to compute can be computed with partial recursive functions.

TURING MACHINES AS DATA. The set of pairs (T, x) where T is a Turing machine and $x \in X$ is an input data, is countable. We can encode therefore the set of such pairs into data X. Let $TM \subset X$ be the set of all the so obtained pairs (T, x). Denote by H the subset of TM, which consists of halting Turing machines.

DECIDABLE SETS IN TM. A subset Z of TM is called **decidable**, if there exists a Turing machine, which tells after finitely many steps, whether a given $x \in TM$ is in Z or not.

THE HALTE PROBLEM IS NOT DECIDABLE.

THEOREM (Turing) The subset $H \subset TM$ of all halting Turing machines is not decidable.

PROOF. Assume the halting problem is decidable. Then there exists a Turing 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 has as an input an input x and does the following

```
1) Read[x]
2) Define Stop=Halt[(x,x)];
3) While Stop==True repeat Stop:=True.
4) Print[Stop]
```

Now, either (DIAGONAL,DIAGONAL) is in the set H or it is not.

(i) Assume first DIAGONAL is in H. Then the variable *Stop* was *True*, which means that the program DIAGONAL runs for ever. So, Halt[(DIAGONAL,DIAGONAL)]=False, and DIAGONAL is not in H.

(ii) Assume now DIAGONAL is not in H Then, the variable *Stop* becomes *False*, which means that Halt[(DIAGONAL,DIAGONAL)]=true, which implies DIAGONAL is in H.

Since the assumption of the existence of a Turing machine HALT leads to a contradiction, a machine DIAGONAL can not exist. This argument of Turing is very similar to Cantors diagonal argument.

UNIVERSAL TURING MACHINE. Turing also showed the existence of a universal Turing machine. This is a machine which can simulate all Turing machines. The universal Turing machine takes a Turing machine with input (T, x) as input and returns as output, the output of the machine x. What Turing showed 1936 means translated into the dynamical systems language:

The universal Turing machine can be realized as a dynamical system.

Indeed, there exists a compact set X and a continuous transformation T on X, such that for a subset Z of X, (Z, T) can do any computation in Mathematics. This tells us also that there are fundamental limitations, what can be said about dynamical systems in general. There are dynamical systems, so that we can not decide for a given set U and a point x, whether $T^n(x)$ will every enter U or not. Note that all said here about Turing machines is just rephrasing of what Turing knew 70 years ago already in an other language. This has to be said because there is literature which can give the impression that such statements are a new discovery.

BUSY BEAVER. The **busy beaver problem** is the task to construct a Turing machine which has n states not counting the halting state 0 and satisfies the following: The machine starts on the empty tape and should write as many 1 onto the tape as possible before it stops. For $n = 1, 2, 3, 4$, the optimal solutions are known. For $n = 5$, Heiner Marxen has built a Turing machine in 1989 which produces 4098 marks. Its orbit is has length 11'798'826. You find a Mathematica program which simulates this Turing machine on the website.

REDDY'S THEOREM. A **topological dynamical system** is a pair (X, T), where X is a compact metric space and T is a homeomorphism of X. (A homeomorphism is a map which is continuous and invertible and for which the inverse is continuous too). A topological dynamical system is called **expansive**, if there exists $\epsilon > 0$ such that for all $x \neq y \in X$, there exists n such that $d(T^n x, T^n y) \geq \epsilon$. A dynamical system is called **zero dimensional** if X is zero dimensional, that is if there is a basis for X which consists of sets which are both open and closed. (A **basis** is is a set B of subsets such that (i) the empty set is in B, arbitrary unions of sets in B are in B, the intersection of two sets in B is a union of sets in B.)

THEOREM (Reddy) A zero dimensional expansive dynamical system is isomorphic to a subshift.

PROOF (sketch) partition X into n sets X_i which are both closed and open, such that each of the sets has diameter $\leq \epsilon$. An orbit $T^n(x)$ defines a code $y \in A^Z$, where A labels the partition. The expansiveness assures that the encoding is injective.

THE TURING MACHINE AS A SUBSHIFT. We first change the Turing dynamical system to make it expansive. This can be done by a topological trick. The zero-dimensionality is assured already. The abstract theorem of Reddy shows that

COROLLARY. There is a subshift which can simulate the universal Turing machine.

Because a subshift is a subset of the shift and the shift can be realized in a dynamical system with a horse shoe, one obtains

COROLLARY. The map $T(x, y) = (-1.5x^2 - 0.3y, x)$ can simulate any computation.

Proof. An iterate T^m of T contains a horse shoe, on which the dynamics is conjugated to a shift of 2 symbols. The map T^{mk} is on this set conjugated to a shift of 2^k symbols.

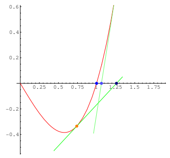
Again, it is important to state that such corollaries are nothing more than climbing onto the shoulders of Turing and other mathematicians working in topological dynamics. While there is nothing original in such statements, it is amusing. It also illustrates that dynamical systems have relations with the foundations of mathematics or what one sometimes calls the "theory of computation".

COMPLEX DYNAMICS

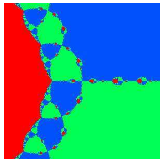
Math118, O. Knill

ABSTRACT. When maps are iterated in the complex plane it leads to interesting dynamics. An example is the Newton method in the complex. We look at some examples and especially show finally that the Ulam map is chaotic. Actually, the interval on which the Ulam map is defined is the Julia set of the corresponding quadratic map.

THE NEWTON METHOD IN THE REAL. The Newton method to find a root of $f(x) = 0$, is to start with a point x_0 and apply the map $T(x) = x - f(x)/f'(x)$. If $T(x) = x$, then $f(x) = 0$. Because $T'(x) = f(x)f''(x)/(f'(x))^2$ is small near $f(x) = 0$, T is a contraction in an interval $[x_0 - \epsilon, x_0 + \epsilon]$ and has a fixed point. The **basin of attraction** of a root x_i are all the points for which $T^n(x) \to x_i$.



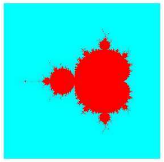
THE NEWTON METHOD IN THE COMPLEX. The Newton method to find a root $f(z) = 0$ can also be done in the complex plane. We start with a point z_0 and apply the map $T(z) = z - f(z)/f'(z)$. If $T(z) = z$, then $f(z) = 0$. Again $T'(z) = f(z)f''(z)/(f'(z))^2$ is small near $f(z) = 0$, the map T is a contraction. The **basin of attraction** of a root x_i are all the points for which $T^n(x) \to x_i$. The picture to the right shows the basins of attractions for each fixed point. Each of this region is the "stable manifold" of the fixed point. The rest is called the **Julia set** of T.



QUADRATIC MAP. The **quadratic map**

$$f_c : z \mapsto z^2 + c$$

with a complex parameter c defines s discrete dynamical system on the complex plane. f_c leaves a set $J_c \subset C$ called **Julia set** and its complement F_c, called the **Fatou set** invariant. The parameter space C is divided into a **Mandelbrot set** M, parameters, where J_c is connected and its complement, where J_c is disconnected.



PARAMETRIZING ALL QUADRATIC MAPS. The quadratic family f_x is not as special as one might think:

LEMMA. A quadratic polynomial $T(z) = az^2 + 2bz + d$ is conjugated by $S(z) = az + b$ to

$$f_c(z) = z^2 + c$$

where $c = ad + b - b^2$.

Proof. Just verify $S^{-1} f_c S(z) = T(z)$.

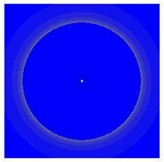
Remark. You show in the homework taht every cubic polynomial $T(z)$ can be conjugated to $f_{a,b}(z) = z^3 - 3a^2 z + b$. The parametrization is chosen so that $-a, a$ are critical points of $f_{a,b}$.
When dealing with maps on the real line, we could also chose the normal form

$$z \mapsto az(1 - z) \; .$$

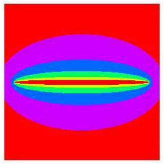
Parametrized like this, the quadratic map is also called the **logistic map**. It maps the interval $[0, 1]$ onto itself. The linear map $S(z) = -az + a/2$ conjugates $z \mapsto az(1 - z)$ to $z \mapsto z^2 + c$, when $c = a/2 - a^2/4$. Especially, the Ulam map is conjugated to f_{-2}.

EXAMPLE. THE SQUARING MAP. Let us look at the map $f(z) = z^2$. If $z = re^{i\theta}$ with $r = |z|$, then $f^n(z) = r^{2^n} e^{i2^n\theta}$. If $r > 1$, then $f^n(z) \to \infty$. If $|r| < 1$, then $f^n(z) \to 0$. If $r = 1$, then $f^n(z) = e^{i2^n\theta}$. On $|z| = 1$, the map is $T(x) = 2x \bmod 1$.

There is a set J on which f is chaotic and the complement F where f is attracted to some attracting fixed point.



EXAMPLE. THE ULAM MAP AS A QUADRATIC MAP. What happens with the Ulam map $f(z) = 4z(1 - z)$ in the complex plane? We have seen that it is conjugated to $f_2(z) = z^2 - 2$. The conjugating map $S(z) = 2 - 4z$ maps the interval $[0, 1]$ to the interval $[-2, 2]$. This interval is invariant and the map T restricted to this interval is the Ulam map.

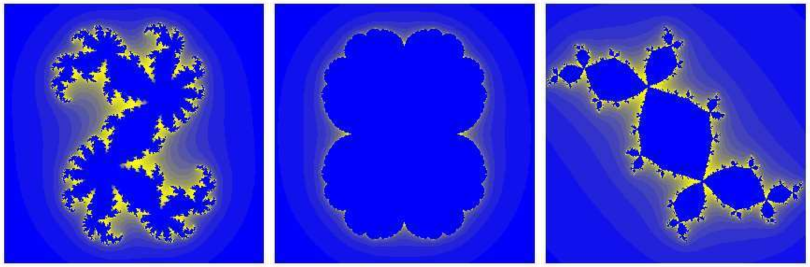


FIXED POINTS. The fixed points of the quadratic map are $z_\pm = (1 \pm \sqrt{1 - 4c})/2$. The value of $f'(z)$ determines the stability. If $|f'(z)| < 1$, then the fixed point is stable, if $|f'(z)| > 1$, it is **unstable**.

Note that when a complex map is written as a real map, then it is not possible that T has a hyperbolic fixed point.

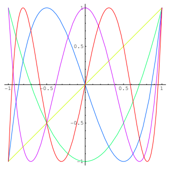
EXAMPLE. $f(z) = z^2 + z + 1$ has the fixed points $i, -i$. Since $f'(i) = 2i$ and $f'(-i) = -2i$, we have $|f'(i)| = 2$ and both fixed points are unstable.

JULIA SETS. Let f be a polynomial. Let P_c denote the set of all points for which $f^n(z)$ stays bounded. This is called the **prisoner set** K (or **filled in Julia set**). The boundary of K is called the **Julia set** J. The complement of J is an open set called the **Fatou set** F of f. It is known that the Julia set is the closure of all repelling periodic points. For the quadratic family, the Julia set is totally disconnected if c is outside the Mandelbrot set and connected, if c is inside the Mandelbrot set.



CHEBYSHEV POLYNOMIALS.

Let $f(z) = 2z^2 - 1$. Because $f(\cos(z)) = 2\cos^2(z) - 1 = \cos(2z)$ we have $f^n(\cos(z)) = \cos(2^n z)$. Actually, the map $S(z) = (z + 1/z)/2$ satisfies $Sz^2 = fS(z)$. In other words, the map S semiconjugates f to the map $g(z) = z^2$ which we have seen above. The conjugating map S maps the unit circle to the interval $[-1, 1]$. This can be used to conjugate the Ulam map to a shift. One generalize this example to the case, where $T_k(z)$ is the Chebychev polynomial $\cos(kz) = T_k(\cos(z))$. (See Homework).



THE ULAM MAP AND THE SHIFT.

The Ulam map $T(x) = 4x(1 - x)$ is chaotic in the sense of Devaney.

Proof. The Ulam map is conjugated to the Chebychev map $C(z) = 2z^2 - 1$. The idea is to use the semiconjugation of the later to $f(z) = z^2$ which is semiconjugated to the shift on $\{0, 1\}^N$. That the later is chaotic in the sense of Devaney had been shown last week in the CA week.

We can find $C(z)$, by forming $\theta = \arccos(z)$ and then get $y = \cos(2\theta)$. if $\arccos(z)/\pi = 0.x_1x_2x_3...$ in binary expansion, then $C(z) = \cos(\pi \cdot 0.x_2x_3x_4...)$.

To find a dense set of periodic points, take a periodic sequence $x \in \{0, 1\}^N$ then $z = \cos(\pi \cdot 0.x_1x_2...)$ is a periodic point of the Ulam map. The map $x \to z$ is continuous and surjectiv. We can find so periodic orbits intersecting each interval $[a, b]$. To show transitivity, take $z = \cos(\pi 0.x_1x_2...)$, a sequence $x \in\in \{0, 1\}^N$ which is transitive (concat an enumeration of all finite words onto each other)

THE MANDELBROT SET — Math118, O. Knill

ABSTRACT. This is a proof a theorem of Douady and Hubbard assuring that the Mandelbrot set is connected. The proof needs some concepts from topology and complex analysis and topology.

BÖTTCHER-FATOU LEMMA.

Assume $f(z) = z^k + a_{k+1}z^{k+1} + \dots$ with $k \geq 2$ is analytic near 0. Define $\phi_n(z) = (f^n(z))^{1/k^n} = z + a_1 z^2 + \dots$. In a neighborhood U of $z = 0$ $\phi = \lim_{n\to\infty} \phi_n(z) : U \to B_r(0)$ satisfies $\phi \circ f \circ \phi^{-1}(z) = z^k$ and $\phi(0) = 0$ and $\phi'(0) = 1$.

PROOF. We show that ϕ_n converges uniformly. The properties $\phi(f(z)) = \phi(z)^k$ as well as $\phi(0) = 0$ and $\phi'(0) = 1$ follow from the assumptions. The function

$$h(z) := \log(\frac{f(z)^{1/k}}{z})$$

with the chosen root $f(z))^{1/k} = z + O(z^2)$ is analytic in a neighborhood U of 0 and there exists a constant C such that $|h(z)| \leq C|z|$ for $z \in U$. U can be chosen so small that $f(U) \subset U$ and $|f(z)| \leq |z|$. We can write $\phi(z)$ as an infinite product

$$\phi(z) = z \cdot \frac{\phi_1(z)}{z} \cdot \frac{\phi_2(z)}{\phi_1(z)} \cdot \frac{\phi_3(z)}{\phi_2(z)} \dots .$$

This product converges, because $\sum_{n=0}^{\infty} \log \frac{\phi_{n+1}(z)}{\phi_n(z)}$ converges absolutely and uniformly for $z \in U$:

$$|\log \frac{\phi_{n+1}(z)}{\phi_n(z)}| = |\log \left[\frac{(f \circ f^n(z))^{1/k}}{f^n(z)}\right]^{1/k^n}| = \frac{1}{k^n} \cdot |h(f^n(z))| \leq \frac{1}{k^n} C \cdot |f^n(z)| \leq \frac{C \cdot |z|}{k^n} .$$

COROLLARY (*). If $c \mapsto f_c(z)$ is a family of analytic maps such that $c \mapsto f_c(z)$ is analytic for fixed z, and c is in a compact subset of C, then the map $(c, z) \mapsto \phi_c(z)$ is analytic in two variables.

PROOF. Use the same estimates as in the previous proof: the maps $(c, z) \mapsto \phi_n(c, z)$ are analytic and the infinite product converges absolutely and uniformly on a neighborhood U of 0.

PROPOSITION The Julia set J_c is a compact nonempty set.

PROOF.
(i) The Julia set is bounded: the Lemma of Boettcher-Fatou implies that every point z with large enough $|z|$ converges to ∞. This means that a whole neighborhood U of z escapes to ∞. In other words, the family $\mathcal{F} = \{f_c^n\}_{n\in N}$ is normal, because every sequence in $\mathcal{F}$ converges to the constant function ∞.

(ii) The Julia set is closed: this follows from the definition, because the Fatou set F_c is open.

(iii) Assume the Julia set were empty. The family $\mathcal{F} = \{f_c^n\}$ would be normal on $\overline{C}$. This means that for any sequence f_n in $\mathcal{F}$, there is a subsequence f_{n_k} converging to an analytic function $f : \overline{C} \to \overline{C}$. Because such a function can have only finitely many zeros and poles, it must be a rational function P/Q, where P, Q are polynomials. If $f_{n_k} \to f$, there are eventually the same number of zeros of f_{n_k} and f. But the number of zeros of f_{n_k} (counted with multiplicity) grows monotonically. This contradiction makes $J_c = \emptyset$ impossible.

COROLLARY. The Julia set J_c is contained in the **filled in Julia** set K_c, the union of J_c and the bounded components of the Fatou set F_c.

PROOF. Because J_c is bounded and f-invariant, every orbit starting in J_c is bounded and belongs by definition to the filled-in Julia set. If a point is in a bounded component of F_c, its forward orbit stays bounded and it belongs to the filled in Julia set. On the other hand, if a point is not in the Julia set or a bounded component of F_c, then it belongs to an unbounded component of the Fatou set F_c.

GREEN FUNCTION. A continuous function $G : C \mapsto R$ is called the potential theoretical **Green function** of a compact set $K \subset C$, if G is **harmonic** outside K, vanishing on K and has the property that $G(z) - \log(z)$ is bounded near $z = \infty$.

The Green function G_c exists for the filled-in Julia set K_c of the polynomial f_c. The map $(z, c) \mapsto G_c(z)$ is continuous.

PROOF. The Boettcher-Fatou lemma assures the existence of the function ϕ_c conjugating f_c with $z \mapsto z^2$ in a neighborhood U_c of ∞. Define for $z \in U_c$

$$G_c(z) = \log |\phi_c(z)| .$$

This function is harmonic in U_c and growing like $\log |z|$ because by Boettcher satisfies $|f_c^n(z)| \geq C|z|^{2^n}$ for some constant C and so

$$G_c(z) = \lim_{n\to\infty} \frac{1}{2^n} \log |f_c^n(z)| .$$

Although G_c is only defined in U_c, there is one and only one extension to all of C which is continuous and satisfies

$$G_c(z) = G_c(f_c(z))/2 . \quad (1)$$

In fact, we define $G_c(z) = 0$ for $z \in K_c$, and $G_c(z) = G(f_c^n(z))/2^n$ otherwise, where n is large enough so that $f_c^n(z) \in U$. We know from this extension that G_c is a smooth real analytic function outside K_c. From the **maximum principle**, we know that $G_c(z) > 0$ for $z \in C \setminus K_c$. We have still to show that G_c is continuous in order to see that it is the Green function. The continuity follows from the stronger statement:

$(z, c) \mapsto G_c(z)$ is jointly continuous.
$G_c^{-1}([0, \epsilon))$ is open in C^2 for all $\epsilon > 0$ if and only if there exists n such that

$$A_n := \{(c, z) \mid G_c(f_c^n(z)) \geq 2^n \epsilon\}$$

is closed $\forall \epsilon > 0$. Given $r > 0$. There exists a ball of radius b which contains all the sets K_c for $|c| \leq r$. For $R \geq G_r(b)$, all the solutions ξ of $G_c(\xi) \geq R$ satisfy $|\xi| \geq b$ if $|c| \leq r$. The set $B = \{(c, \xi) \mid G_c(\xi \geq R\} \cap \{|c| \leq r\}$ is closed. For n large enough, also $A_n \cap \{|c| \leq r\}$ is closed and A_n is closed.

THEOREM (DOUADY-HUBBARD). The Mandelbrot set M is connected.

CORE OF THE PROOF. The Böttcher function $\phi_c(z)$ can be extended to

$$S_c := \{z \mid G_c(z) > G_c(0)\} .$$

Continue defining $\phi_c(z) := \sqrt{\phi_c(z^2 + c)}$ to get ϕ_c having defined in larger and larger regions. This can be done as long as the region $\phi_c^{-1}(\{r\})$ is connected (this assures that the derivative of ϕ_c is not vanishing). Because Equation (1) gives $G_c(c) = 2G_c(0) > G_c(0)$, every c is contained in the set S_c and the map

$$\boxed{\Phi : c \mapsto G_c(c)}$$

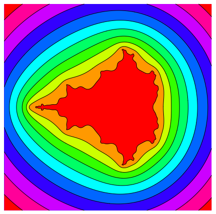
is well defined. It is analytic outside M and can be written as

$$\Phi(z) = \lim_{n\to\infty} [f_c^n(c)]^{1/2^n} .$$

Claim:

$$\Phi : \overline{C} \setminus M \to \overline{C} \setminus \overline{D}$$

is an analytic diffeomorphism, where $\overline{C} = C \cup \{\infty\}$ is the Riemann sphere. (This implies that the complement of M is simply connected in $\overline{C}$, which is equivalent to the fact that M is connected). The picture to the right shows the level curves of the function $\phi_6(c) = [f_c^6(c)]^{1/64}$. The function $(\phi_6(z)$ is already close to the map $\Phi(z)$ in the sense that the level sets give a hint about the shape of the Mandelbrot set.



(1) Φ is analytic outside M. This follows from the Corollary.

(2) For $c_n \to M$, we have $|\Phi(c_n)| \to 1$. Proof. Continuity of the Green function.

(3) The map Φ is proper. (A map is called **proper** if the inverse of any compact set is compact).
Given a compact set $K \subset C \setminus D$. The two compact sets D and and K have positive distance. Assume $\phi^{-1}(K)$ is not compact. Then, there exists a sequence $c_n \in \Phi^{-1}(K)$ with $c_n \to c_0 \in M$ so that $|\Phi(c_n)| \to 1$. This is not possible because $\Phi(c_n) \in K$ is bounded away from D.

(4) The map Φ is open (it maps open sets into open sets). This follows from the fact that Φ is analytic. (This fact is called **open mapping theorem** (see Conway p. 95))

(5) The map Φ maps closed sets into closed sets.
A proper, continuous map $\Phi : X \to Y$ between two locally compact metric spaces X, Y has this property. Proof. Given a closed set $A \subset X$. Take a sequence $\Phi(a_n)$ in $\Phi(A)$ which converges to $b \in Y$. Take a compact neighborhood K of b (use local compactness of Y). Then $\Phi^{-1}(K \cap \overline{\phi(A)})$ is compact and contains almost all a_n. The sequence a_n contains therefore an accumulation point $a \in X$. The continuity implies $\Phi(a_n) \to \Phi(a) = b$ for a subsequence so that $b \in \Phi(K)$. Consequently $\Phi(K)$ is closed.

(6) Φ is surjective.
The image of $\Phi(\overline{C} \setminus M)$ is an open subset of set $\overline{C} \setminus \overline{D}$ because Φ is open. The image of the boundary of M is (use (5)) a closed subset of $\overline{C} \setminus D$ which coincides with the boundary of D because the boxed statement about the the Green function showed $G_c(c) \to 0$ as $c \to M$.

(7) Φ is injective.
Because the map Φ is proper, the inverse image $\phi^{-1}(s)$ of a point s is finite. There exists therefore a curve Γ enclosing all points of $\Phi^{-1}(s)$. Let $\sharp A$ denote the number of elements in A. By the **argument principle** (see Alfors p. 152), we have

$$\sharp(\phi^{-1}(s)) = \frac{1}{2\pi i} \int_\Gamma \frac{\Phi'(z)}{\Phi(z) - s} \, dz$$

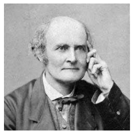
and this number is locally constant. Given $M > 0$, we can find a curve Γ which works simultaneously for all $|s| \leq M$. Because Φ is surjective and $\sharp(\phi^{-1}(\infty)) = 1$, we get that $\sharp(\phi^{-1}(s)) = 1$ for all $z \in C \setminus \overline{D}$ and ϕ is injective.

(8) The map Φ^{-1} exists on $C \setminus D$ and is analytic.
Because an injective, differentiable and open map has a differentiable inverse, (this is called **Goursat's theorem**), the inverse is analytic.

NOTATIONS.

- $f(z)$ is **analytic** in a set U if the derivative $f'(z) = \lim_{w \to 0}(f(z+w) - f(z))/w$ of f exists at every point in U. This means that for $f(z) = f(x+iy) = u(x+iy) + iv(x+iy)$ the partial derivatives $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}$ are all continuous real-valued functions on U. In that case $u(x,y), v(x,y)$ are **harmonic**: $u_{xx} + u_{yy} = 0$.
- $B_r(z) = \{w \mid |z - w| < r\}$ is a neighborhood of z called an **open ball**.
- A sequence of analytic maps f_n **converges uniformly** to f on a compact set $K \subset U$, if $f_n \to f$ in $C(K)$, which means $\max_{x \in K} |f_n(x) - f(x)| \to 0$.
- A family of analytic maps $\mathcal{F}$ on U is called **normal**, if every sequence $f_n \in \mathcal{F}$ has a subsequence which converges uniformly on any compact subset of U. The limit function f does not need to be in $\mathcal{F}$. With respect to the topology of convergence on compact subsets normality is precompactness in this topology: $\mathcal{F}$ is normal, if and only if its closure is compact. The **theorem of Arzela-Ascoli** (see Alfors p. 224) states says that normality of $\mathcal{F}$ is equivalent to the requirement that each f is equicontinuous on every compact set $K \subset U$ and if for every $z \in U$, the set $\{f(z) \mid f \in \mathcal{F}\}$ is bounded. z is part of the **Fatou set** of f, $\{f^n\}_{n \in N}$ is normal in some neighborhood of z. The **Julia set** is the complement of the Fatou set.
- A set is called **locally compact**, if every point has a compact neighborhood. In the plane, a set is compact if and only if it is bounded and closed. A subset is closed, if and only if its complement is open. A subset U is open, if for every point x in U there is a ball $B_r(x)$ which still belongs to U.

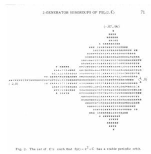
SOME HISTORY:



In 1879, **Arthur Cayley** poses the problem to study the regions in the plane, where the Newton iteration converges to some root.

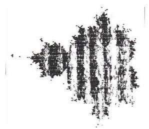
Gaston Julia (1893-1978) and **Pierre Fatou** (1879-1929) both worked already 90 years ago on the iteration of analytic maps. Julia and Fatou sets are called after them. Julia and Fatou were both competed for the 1918 'grand priz' of the academie of sciences and produced similar results. This produced a priority dispute. Julia lost his nose in world war I and had since to wear a leather strap across his face. He had continued with his research in the hospital.

Robert Brooks and Peter Matelski produce in 1978 the first picture of the Mandelbrot set in the context of Kleinian groups. Their paper had the title "The dynamics of 2-generator subgroups of PSL$(2, C)$". The defined $\tilde{M} = \{c \mid f_c$ has a stable periodic orbit $\}$. This set is now called **Brooks-Matelski set** and is now believed to be the interior of the Mandelbrot set M. If the later were locally connected, this would be true: $\text{int}(M) = \tilde{M}$.



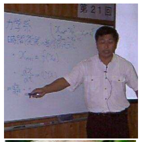
John Hubbard made better pictures of a quite different parameter space arising from Newton's method for cubics. Hubbard was inspired by a question from a calculus student. **Benoit Mandelbrot**, perhaps inspired by Hubbard, made corresponding pictures in 1980 for quadratic polynomials. He conjectured the set M is disconnected because his computer pictures showed "dust" with no connections to the main body of M. It is amusing that the journals editorial staff removed that dust, assuming it was a problem of the printer.

John Milnor writes in his book of 1991: "Although Mandelbrot's statements in this first paper were not completely right, he deserves a great deal of credit for being the first to point out the extremely complicated geometry associated with the parameter space for quadratic maps. His major achievement has been to demonstrate to a very wide audience that such complicated fractal objects play an important role in a number of mathematical sciences."



Adrien Douady and John Hubbard prove the connectivity of M in 1982. This was a mathematical breakthrough. In that paper the name "Mandelbrot set" was introduced. The paper provided a firm foundation for its mathematical study. We followed on this handout their proof. Note that the Mandelbrot set is also **simply connected**, but this is easier to show. Both statments use that a subset of the plane is connected if and only if the complement is simply connected.

Evenso one of the first things which comes in mind, when talking about fractals is the Mandelbrot set. It is not a "fractal": in 1998, **Mitsuhiro Shishikura** has shown that its Hausdorff dimension of M is 2. (M. Shishikura, "The Hausdorff dimension of the boundary of the Mandelbrot set and Julia sets, Annals of Mathematics 147 (1998), 225-267.)

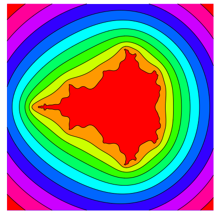


Also for higher dimensional polynomials, one can define Julia and Mandelbrot sets. For cubic polynomials $f_{a,b}(z) = z^3 - 3a^2 z + b$, define the **cubic locus set** $\{(a,b) \in C^2 \mid K_{a,b}$ is connected $\}$, where $K_{a,b}$ is the **prisoner set** $K_{a,b} = \{z \mid f^n_{a,b}(z)$ stays bounded. $\}$. **Bodil Branner** showed around 1985, that the cubic locus set is connected. This generalizes the main result discussed in this handout.

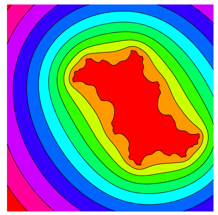
OPEN PROBLEMS. The major open problem is whether the Mandelbrot set is locally connected or not. A subset M of the plane is called **locally connected**, if at every point $x \in M$ if every neighborhood of x contains a neighborhood, in which M is connected. A locally connected set does not need to be connected (two disjoint disks in the plane are locally connected but not connected). A connected set does not need to be locally connected. An example is the union of the graph of $\sin(1/x)$ and the y-axes.

NOTIONS IN COMPLEX DYNAMICS Math118, O. Knill

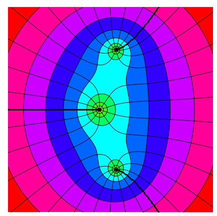
ABSTRACT. This page summarizes some definitions in complex dynamics and gives a brief jumpstart to some notions in complex analysis and topology.



MANDELBROT SET. $f_c(z) = z^2 + c$ is called the quadratic map. It is parametrized by a constant c. The set M of parameter values c for which $f_c^n(c)$ stays bounded. In the homework you seea that $M = \{c, |f_c^n| \leq 2$ for all $n\}$. With $G(c) = \lim_{n\to\infty} \log |(f_c^n(c))^{1/2^n}|$ one can also say $M = \{c \mid G(c) = 0\}$. The level curves of G are **equipotential curves**: if you would charge the Mandelbrot set with a positive charge, $G(z) = c$ is the set of points where the attractive force of an electron to the set is the same. By definition, M is closed. Douady-Hubbard theorem tells it is connected. That M is **simply connected** is much easier to see: it follows from the **maximum principle** that the complement of M is connected.



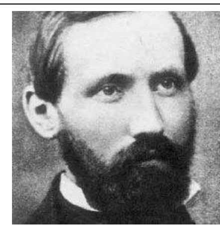
JULIA SET. The set of complex numbers z for which $f_c^n(z)$ stays bounded is called the **filled in Julia set** K_c. It is the set of z for which the function $G_c(z) = \lim_{n\to\infty} \log |(f_c^n(z))^{1/2^n}|$ is zero. Its boundary is called the **Julia set**. The Julia set can be a smooth curve like in the case $c = 0$ or for $c = -2$ but it is in general a complicated fractal. It is known that the Julia set J_c is the closure of the repelling periodic points of f_c. It is also known that f_c restricted to J_c is chaotic in the sense of Devaney. The complement of J_c is called the Fatou set F_c. The bounded components of F_c are called **Fatou components**.



COMPLEX MAPS. A complex map f can be written as a map in the real plane $f(x + iy) = u(x, y) + iv(x, y)$. The derivative at a point z_0 is defined as the complex number

$$a = f'(z) = \lim_{w\to 0}(f(z + w) - f(z))/w .$$

If the derivative exists at each point in a region U and f' is a continuous function in U, the map f is called **analytic** in U.

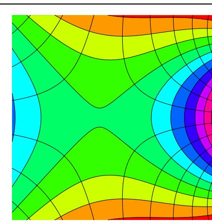


CAUCHY-RIEMANN. Since the linearization of f at z_0 is the map $z \to az$ which is a rotation dilation and the linearization of f is the Jacobean

$$A = \begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix} ,$$

we must have $\boxed{u_x = v_y, u_y = -v_x}$ (A rotation matrix has identical diagonals and antidiagonals of opposite signs and this property is preserved after multiplying the matrix with a constant) . These two equations for u, v are called **Cauchy-Riemann differential equations**.

CONFORMALITY. If $a \neq 0$, then angles are preserved because both rotations and dilations preserve angles. Therefore the rotation dilation $z \to az$ preserves angles. If $f'(z)$ is never zero in a region U, the map f is called **conformal** in U. In that case, it maps U bijectively to $f(V)$ and preserves angles. Angle preservation is useful in cartography or computer graphics.



HARMONICITY. From the Cauchy-Riemann equations follows $u_{xx} + u_{yy} = 0$ and $v_{xx} + v_{yy} = 0$. Therefore, the real and imaginary part of f are **harmonic functions**. The mean value property $\int_{|w-z|=r} u(w(t))\ dt = u(z)$ and $\int_{|w-z|=r} v(w(t))\ dt = v(z)$ for harmonic functions can be written as $\int_{|w-z|=r} f(w(t))\ dt = f(z)$.

TAYLOR FORMULA. Because $df(w(t))/dt = f(x + r\cos(t) + i(y + r\sin(t)) = f'(w)(r\cos(t) + ir\sin(t)) = f'(w)(z - w)$, this can be rewritten as $\int_{|w-z|=r} f'(w(t))dt/(z - w) = f(z)$. This is the **Cauchy integral formula**. Since we can differentiate the left hand side arbitrarily often with respect to z, this proves that an analytic function is arbitrarily often differentiable and $f(w)/(z - w)$ has the n'th derivative $\frac{f(w)}{n!(z-w)^{n+1}}$, we get

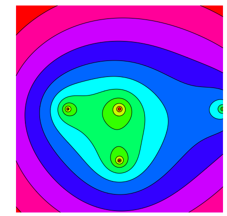
$$f(w) = \sum_n \frac{f^{(n)}(z)(w - z)^n}{n!}$$

which is the familiar **Taylor formula** if f is real.

CAUCHY THEOREM. The Cauchy Riemann equations also prove the **Cauchy formula**. If C is a closed curve in simply connected region U in which f is analytic, then

$$\boxed{\int_C f(z)dz = \int f(z(t))z'(t)\ dt = 0}$$

because the later is the line integral of $F(x, y) = (-v(x, y), u(x, y))$ and **Greens theorem** in multi-variable calculus shows that $\mathrm{curl}(F) = \mathrm{curl}((-v, u)) = (u_x - v_y) = 0$. In other words, the vector-field $F(x, y) = (-v(x + iy), u(x + iy))$ is conservative.



FIXED POINTS. Because the eigenvalues of the rotation dilation A come in complex conjugate pairs, the fixed points or periodic points can not be hyperbolic. Fixed points are either stable sinks, or unstable sources elliptic, conjugated to a rotation. For example, the fixed points of $f(z) = z^2 + c$ are $(1 \pm \sqrt{1 - 4c})/2$ and the linearization at those points is $df(z) = (1 \pm \sqrt{1 - 4c})z$

TOPOLOGY. Here are some topological notions occuring in complex dynamics:
OPEN. A set U in the plane is called **open** if for every point z, there exists $r > 0$ such that $B_r(z) = \{w \mid |w-z| < r\}$ is contained in U. One assumes the empty set to be open. The entire plane is open too.
CLOSED. A set U in the plane is **closed**, if the complement of U is open. The entries plane is closed.
INTERIOR. The **interior** of a set U is the subset of all points z in U for which there exists $r > 0$ such that $B_r(z) \subset U$. If a set is open, then it is equal to its interior.
CLOSURE. The **closure** of a set U is the set of all points which are limit points of sequences in U. It is the complement of the interior of the complement of U. If a set is closed, then U is equal to its closure.
BOUNDARY. The boundary of a set U is the closure of U minus the interior of U. The boundary of a closed set without interior is the set itself.
SIMPLY CONNECTED. A set A is **simply connected**, if every closed curve contained in A can be deformed to a point within A. A simply connected set has no "holes".
CONNECTED. A set A is called **connected** if one can not find two discjoint open sets U, V such that $A \cap U \neq \emptyset$, $A \cap V \neq \emptyset$.

A set A is connected if and only if the complement is simply connected.

To verify that the complement of M is simply connected, one finds a smooth bijection of the complement of the unit disc with the complement of M. The bijection is given by $\Phi(c) = \lim_{n\to\infty}(f_c^n(c))^{1/2^n}$. The Mandelbrot set M is connected as well as simply connected. The Julia sets J_c are connected, if c is in M.
COMPACT. A subset of the complex plane is called **compact** if it is closed and bounded. A sequence in a compact set always has accumulation points. The Mandelbrot set as well as the Julia sets are examples of compact sets.
PERFECT SETS. A subset J in the complex plane is **perfect** if it is closed and every point z in J is accumulation point of points in $S \setminus z$. Perfect sets contain no isolated points.
NOWHERE DENSE. A subset J in the complex plane is **nowhere dense** if the interior of its closure is empty. A Julia set J_c is nowhere dense if c is outside the Mandelbrot set.
CANTOR SET. A perfect nowhere dense set is also called a **Cantor set**. An example is the **Cantor middle set**. A Julia set J_c is a Cantor set if c is outside the Mandelbrot set.

THE BERNOULLI SHIFT — Math118, O. Knill

ABSTRACT. When equipped with an invariant measure, which is the area measure when representing it as the Baker map, the shift is called the Bernoulli shift. It produces independent random variables.

A SHIFT INVARIANT MEASURE. We have defined a map S from the unit square $Y = [0,1) \times [0,1)$ to the sequence space $X = \{0,1\}^Z$ by

$$S(u,v)_n = \begin{cases} 0 & u_n < 1/2 \\ 1 & u_n \geq 1/2 \end{cases}$$

if $T^n(u,v) = (u_n, v_n)$ is the orbit of the Baker map. This was called **symbolic dynamics**. We can use the map S to measure subsets in X by requiring that it preserves the measure: the left half of the square of area 1/2 is mapped into the set of sequences x which satisfy $x_0 = 0$, the right half of the square of area 1/2 is mapped into $\{x \mid x_0 = 1\}$. The set $\{x_0 = 0, x_1 = 1\}$ in X corresponds to the lower left quarter of the square which has area 1/4.

THE BERNOULLI MEASURE. The space X can be equipped with a shift invariant probability measure P. In that case, we say $P[U]$ is the measure or the probability of U. We can define $P[U]$ as the area of $S^{-1}(U)$ in the square. We know then that

$$P[x_{n+1} = f_1, ..., x_{n+m} = f_m] = 2^{-m} .$$

This measure is called a **Bernoulli measure**. It is **invariant under the shift**. for any subset U of X, then $P[\sigma(U)] = P[U]$.

> If U is a subset of the square and S is the map conjugating the Baker map to the shift, then $P[S(U)]$ is the area of U.

RANDOM VARIABLE. A **random variable** is a (continuous) function from X to R. Examples of random variables are $X_k(x) = x_k$. Two random variables are called **independent** if $\mathrm{P}[\{Y = a, Z = b\}] = \mathrm{P}[\{Y = a\}]\mathrm{P}[\{Z = b\}]$ for any choice a, b.

> The random variables $X_k = x_k$ in the Bernoulli shift are independent.

PROOF. $\mathrm{P}[X_k = a, X_l = b] = 1/4$ for any choice of a, b. This is the same as $\mathrm{P}[Y = a]\mathrm{P}[Z = b] = (1/2) \cdot (1/2)$.

In other words, one can use the Bernoulli shift or the Baker map to produce **random numbers**.
This is not a very practical way to produce random numbers: lets look at the first coordinates, when applying the Baker map, we have $T^n(x) = 2^n x \bmod 1$. If we start with a a rational number, then $T^n(x)$ will be attracted by a periodic orbit like for example $1/3, 2/3, 1/3, ...$ For a practical generation of random numbers other maps are better suited.

EXPECTATION. The **expectation** of a random variable which takes finitely many values $f_1, ..., f_m$ is

$$E[Y] = P[Y = f_1]f_1 + ... + P[Y = f_n]f_n$$

Two random variables Y, Z are called **uncorrelated** if $E[YZ] = E[Y]E[Z]$. Two independent random variables are automatically uncorrelated.

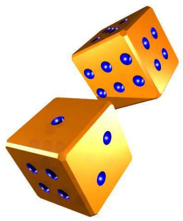
EXAMPLE. $A = \{\text{head}, \text{tail}\}$ models throwing a coin The random variable

$$X(x) = \left[\begin{array}{ll} 3 & x_0 = \text{head} \\ 5 & x_1 = \text{tail} \end{array} \right.$$

has the expectation

$$E[X] = P[X = 3]3 + P[X = 5]5 = 3/2 + 5/2 = 4 .$$

EXAMPLE. Consider the shift over the alphabet $A = \{1, 2, 3, ..., 6\}$. The random variables $X_1, X_2, ...$ simulate the outcomes of a dice event. If $X_3 = 5$, then the third dice rolling produced a 5. These random variables are uncorrelated and independent.



THE LAW OF LARGE NUMBERS. The law of large numbers tells that if X_k are independent random variables with the same distribution, then

$$\frac{1}{n} \sum_{k=1}^{n} X_k$$

converges to the common expectation $E[X_k]$ for almost all experiments.

EXAMPLES. In the dice case, we have for almost all sequences x, that $\frac{1}{n} \sum_{k=1}^{n} X_k \to 7/2$.

OTHER MEASURES. The set X can be equipped with other measures. Assume the letter $x_k = 1$ should have probability p and $x_k = 0$ should have probability $1 - p$. In that case, the probability $\mathrm{P}[x_1 = a_1, ..., x_n = a_n]$ is $\binom{n}{k} p^k (1-p)^{n-k}$, where k is the number of times, $a_i = 1$. Knowing the probability of all these events defines the invariant measure. All theses measures are called Bernoulli measures.

MARKOV CHAINS. Often, one does not know the invariant measure, but one knows the **conditional probabilities**: $P[x_{n+1} = a \mid x_n = b] = M_{ab}$. In words, the probability that $x_{n+1} = a$ under the condition $x_n = b$ is P_{ab}. The matrix M_{ab} is called a **Markov matrix**. It has the property that the sum of coefficients in each column is equal to 1. The matrix M is a $n \times n$ matrix, if the alphabet A has n elements. You have seen examples of the following fact in linear algebra:

> The eigenvector $p = (p_1, ..., p_n)$ to the eigenvalue q of the matrix M normalized so that the $p_1 + ... + p_n = 1$ defines a Bernoulli probability measure on X.

EXAMPLES.
a) If $M_{ij} = 1/2$ for all i, j, we have the Bernoulli shift.
b) If $M = \begin{bmatrix} 1/2 & 2/3 \\ 1/2 & 1/3 \end{bmatrix}$, we can read off the probabilities p that $x_n = 1$ and $1 - p$ that $x_n = 0$ by computing the eigenvector v of M to the eigenvalue 1 and normalizing it, so that the sum of its entries is 1.
c) If $M = \begin{bmatrix} 1/3 & 1 \\ 2/3 & 0 \end{bmatrix}$, we obtain a measure supported on the Fibonnacci shift introduced above. The transitions 11 is not possible.

MEASURES ON SUBSHIFTS OF FINITE TYPE. If we use a Markov matrix for which $M_{ab} = 0$ if ab is a forbidden word, then we obtain an invariant measure for the subhift of finite type by inducitively determining the probability of the cylinder sets $P[\{x_0 = a_0, ..., x_n = a_n\}]$ using the **Bayes formula** $P[A|B] = P[A \cap B]/P[B]$. Subshifts of finite type have a lot of invariant measures. Markov matrices provide a possibility to define such measures.

MEASURES ON SUBSHIFTS. Every subshift X has an invariant measure. It can be obtained by averaging along an orbit. This averaging does not converge in general, but there is a subsequence, along which the limit sets $P[A] = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} 1_{T^k(x) \in A}$

UNIQUELY ERGODIC SUBSHIFTS. If there is only one shift-invariant measure then the subshift is called **uniquely ergodic**. An example are Sturmean sequences, which are obtained by doing symbolic dynamics on using a half open I and an irrational rotation on the circle. The There is only one invariant measure, because also the irrational rotation on the circle has only one invariant measure.

ERGODIC THEORY. The part of dynamical systems, which deals with invariant measures of a map or dynamical system is called **ergodic theory**. It has close relations to probability theory. The law of large numbers we mentioned here has a generalization which is called **Birkhoffs ergodic theorem**.

SHIFTS IN QUADRATIC AND STANDARD MAP — Math118, O. Knill

ABSTRACT. We look on this page at an analytic proof that there is an invariant shift embedded in some Hénon maps, Standard maps or quadratic maps. The proof uses the **implicit function theorem** and is based on an idea of Aubry and Abramovici called **anti-integrable limit**.

THEOREM OF DEVANEY-NITECKI. Fix $b \neq 0$. For large enough c, the Hénon map $H : (x, y) \mapsto (x^2 - c - by, x)$ has an invariant set K such that T restricted to K is conjugated to the shift

$$S = (..., x_{-1}, x_0, x_1, x_2, ...) \to (..., x_0, x_1, x_2, x_3, ...)$$

on all sequences with two symbols.

PROOF. With the new parameter $a = 1/\sqrt{c}$ and the new coordinates $q = x \cdot a, p = y \cdot a$, the map becomes

$$T(q,p) \mapsto (\frac{q^2-1}{a} - bp, q)$$

and is equivalent to the recurrence

$$a \cdot q_{n+1} + a \cdot b \cdot q_{n-1} = q_n^2 - 1 \; .$$

We look for sequences $q_n = q(S^n x)$, where S is the shift on the space of all sequence $X = \{-1, 1\}^Z$ and where q is a continuous map from X to R. We have to solve

$$a \cdot q(Sx) + a \cdot b \cdot q(S^{-1}x) - (q(x)^2 - 1) = 0 \; .$$

With the map $F : R \times C(X) \to C(X)$ defined by

$$F(a, q)(x) = a \cdot q(Sx) + a \cdot b \cdot q(S^{-1}x) - (q(x)^2 - 1)$$

this equation can be rewritten as $F(a, q) = 0$. The partial derivative $F_q(a, q)$ is

$$F_q(a, q)u = a(u(S) + b \cdot u(S^{-1})) - 2q \cdot u \; .$$

The map $F(0, q) : C(X) \to C(X)$ has the property that every function $q \in C(X)$ with values in $\{-1, 1\}$ is a solution of $F(0, q) = 0$. We take for such a solution the map $q(x) = x_0$.

The derivative $F_q(0, q)$ is the linear map

$$(F_q(0, q)u = -2q \cdot u$$

which is invertible because q is bounded away from 0.

By the implicit function theorem, there exists a solution $a \mapsto q_a = G(a)$ satisfying $F(a, q_a) = 0$ for small a. Define $\phi_a : X \to R^2$ by

$$\phi_a(x) = (q(x), q(S^{-1}x)) \; .$$

The map ϕ_a is continuous, because q and T are continuous.
Using $F(a, q) = 0$, we check that

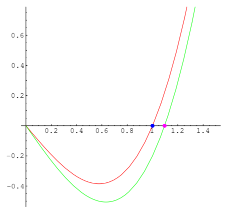
$$\begin{aligned} \phi_a \circ T(x) &= (q(Sx), q(x))) = (\frac{(q(x)^2 - 1)}{a} - b \cdot q(S^{-1}x), q(x)) \\ &= T(q(x), q(S^{-1}x)) = T \circ \phi_a(x) \end{aligned}$$

for all $x \in X$.

The map is injective because if two points x, y are mapped into the same point in R^2 then the fact that $q_a(x)$ is near $q_0(x) = x_0$ implies $x_0 = y_0$. The conjugation $\phi_a \circ S^n(x) = T^n \circ \phi_a(x)$ gives us $T^n(x) = T^n(y)$ and so $x_n = y_n$ for all n.

ϕ has a continuous inverse because every bijective map from a compact space to a compact space has a continuous inverse. The map is indeed a homeomorphism from X to a closed subset $K = \phi(X) \subset R^2$.

THE IMPLICIT FUNCTION THEOREM. Given a family $q \to F(a, q)$ of maps, parametrized by a parameter a. If $F(0, q_0) = 0$ and $F'(0, q_0) \neq 0$, then there exists a continuous function q in some interval I such that $F(a, q(a)) = 0$ for $a \in I$.

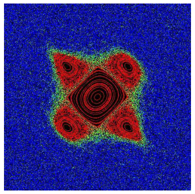


PROOF. The Newton map $T_a(q) = q - F(a, q)/F'(a, q)$ has as a stable fixed point which is the root $q(a)$. This fixed point exists for small a and changes continuously with a.
This proof works also in infinite dimensional spaces, in which it is possible to differentiate. An example is the space $C(X)$ of continuous functions on a compact set X. Example: let $F(f) = f^3 + 5f$. The function F maps a continuous function to a continuous unction. One has $F'(f)g = (3f^2 + 5)g$. Example: let $F(f) = f(x^2)$. Because this is a a linear map in f, we have $F'(f)g(x) = f(x^2)g(x)$.

HORSE SHOES IN THE STANDARD MAP. For large enough c, the Standard map $T : (x, y) \mapsto (2x + c\sin(x) - y, x)$ has an invariant set K such that T restricted to K is conjugated to the shift

$$S = (..., x_{-1}, x_0, x_1, x_2, ...) \to (..., x_0, x_1, x_2, x_3, ...)$$

on all sequences with two symbols.



PROOF. If $T^n(q, p) = (q_n, p_n)$ is an orbit of the Standard map, then $p_n = q_{n-1}$ and so $q_{n+1} - 2q_n + q_{n-1} + c\sin(x_n) = 0$. With $\epsilon = 1/c$, this means

$$\epsilon(q_{n+1} - 2q_n + q_{n-1}) + \sin(x_n) = 0$$

Let X be all $\{0, 1\}$ sequences. Consider the space of all continuous functions q from X to $[0, 2\pi]$.
If we find a solution q to the equation

$$F(\epsilon, q) = \epsilon(q(\sigma x) - 2q(x) + q(\sigma^{-1}x)) + \sin(q(x)) = 0$$

then q is a conjugation from (X, σ) to $(q(X), T)$ showing that we can find a shift similar as the horse shoe construction does.

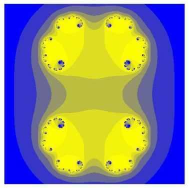
(i) There is a solution for $\epsilon = 0$: Just take $q(x) = \pi x_0$. Because $\sin(0) = \sin(\pi) = 0$, the equation $\sin(q(x)) = 0$ is satisfied.

(ii) In order to have a solution for small ϵ, we compute the derivative of $L = F_q(0, q) = \cos(q)$ and see whether it is invertible. Indeed, since $L = \cos(q(x)) = \pm 1$, we can invert L, the inverse is actually equal to L. (Note that F has as an argument a function q and the the derivative $F_q(a, q) = \lim_{u \to q}(F(a, q + u) - F(a, q))/u$ is defined with respect to the function q. It was computed in the same way as derivatives with respect to real parameters.)
(iii) The implicit function theorem now assures that we can find for small ϵ a function q_ϵ which satisfies $F(\epsilon, q_\epsilon) = 0$. This function q_ϵ conjugates the shift with the standard map T_c restricted to the set $K = q_\epsilon(X)$. Since $\epsilon = 1/c$, this conjugation works for large enough c.

JULIA SETS. The same construction works also for the map $f(z) = a(z^2 - 1)$. We look for a function $q \in C(X, C)$ such that $q(\sigma) - a(z^2 + 1) = 0$. With $\epsilon = 1/a$, this is

$$F(\epsilon, q) = \epsilon q(\sigma) - (z^2 - 1) = 0 \; .$$

For $\epsilon = 0$, the function $q(x) = (2x_0 - 1)$ is a solution. The derivative $L = F_q(0, q) = 2q$ is invertible. We have solutions for small ϵ, which corresponds to large a. Actually, the image $q(X)$ is just the Julia set of f.



SUMMARY. The anti-integrable limit construction allows to get embedded shifts in a purely **analytic** way using the **implicit function theorem**. In comparison, the construction of a **horse shoe** is a **geometric** construction. Finding a **generating partition** is a more **combinatorial** task. The shift brings different areas of mathematics together.

SYMBOLIC DYNAMICS **Math118, O. Knill**

ABSTRACT. We have seen shifts as cellular automata, in a horse-shoes or in Julia set. We look at this dynamical system a bit closer.

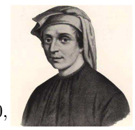
THE SHIFT. Given a finite alphabet A, define $X = A^N$ and $\sigma(x)_n = x_{n+1}$. This dynamical system is called the **one sided shift**. The shift on A^Z is called the **two sided shift**. While the later is invertible, the first is not.

SUBSHIFTS. The shift restricted to a closed shift-invariant subset X of A^Z is called a **subshift**.

EXAMPLE. Let $T(x) = x + \alpha \bmod 1$ and $Y = [0, a)$ and interval. Look at all sequences obtained by taking a point x and defining $x_n = 1_Y(x + n\alpha)$, where $1_Y(x) = \begin{cases} 1 & x \in Y \\ 0 & x \notin Y \end{cases}$. That is

$$x_n = \begin{cases} 1 & (x_0 + n\alpha) \bmod 1 \in Y \\ 0 & (x_0 + n\alpha) \bmod 1 \notin Y \end{cases} .$$

Lets assume for example, $Y = [0, 1/2)$ and $\alpha = \sqrt{2}$. With the starting point $x = 0$, we obtain the sequence $\{x_0, x_1, x_2, ...\} = \{1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, ...\}$. The image X of the map S is a closed subset of the sequences. Every orbit of the shift σ in X is dense.

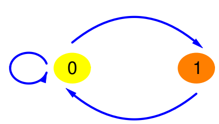


A particular interesting case is $\alpha = (\sqrt{5} - 1)/2$ and $F = [0, \alpha)$. If $x_0 = 1$, then $x_1 = 0$. If $x_0 = 0$, then $x_1 = 1, x_2 = 0$. One can obtain the sequence also by applying the substitution rule $1 \to 0, 0 \to 10$. A sequence obtained like this is called a **Fibonacci sequence**. Here is part of the sequence: $..., 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0,$

EXAMPLE: SUBSHIFTS OF FINITE TYPE. Given a finite set of words K over an alphabet A. The set X of all sequences, in which the words of K do **not** appear, is called a **subshift of finite type**.
The **language** of X is the set of all words which occur in sequences of X. There is a finite set of words which can build up any sequence $x \in X$ and such that the forbidden words determine which words can be adjacent and which not. We can define a **directed graph** (V, E), which has as vertices these words and where an arrow goes from one word to an other if these words can be glued together. One says that the graph represents the subshift.

EXAMPLE. Assume $K = \{00, 111\}$ are the forbidden words, then a sequence can be ...0101101011010110101010011011.... We can get any sequence by gluing together words $w_1 = 01, w_2 = 11$ and $w_3 = 10$. The combinations $w_1 \to w_1, w_1 \to w_2, w_2 \to w_1, w_3 \to w_1, w_3 \to w_2, w_3 \to w_3$ are possible.

EXAMPLE. Let $K = \{11\}$, then X consists of all sequences, where no double 11 occur. The language of X is $\{0, 1, 00, 10, 01, 10, 000, 001, 010, 100, 101, 0000, 0001, ...\}$. With the set $V = \{00, 01, 10\}$ of words one can build any sequence. The gluing $00 \to 01, 01 \to 01, 00 \to 00$. $10 \to 10, 10 \to 01$ are possible, while the gluing $01 \to 10$ is not possible.



SOFIC SHIFTS. If X is a subshift of finite type and T is a cellular automaton map, then $T(X)$ is called a **sofic shift**.

Sophic shifts produce **regular languages**, languages accepted by finite state automata, but they are in general no more of finite type. The next example shows this.

EXAMPLE. The **even shift** is the set of all $x \in \{0, 1\}^Z$ so that between any two 1, there is an even number of 0's. The even shift is not a subshift of finite type, but it is a sofic shift. Start with the subshift of finite type, with the forbidden word 00. Take the elementary CA which gives only 0 for $1, 0, 1$ and $0, 1, 0$ and $0, 1, 1$. For example, $x =$ ...0111101011110110111011111101111111111... is mapped to $y =$ ...0000111001001100.... The image of this cellular automaton consists of all sequences for which 0 occurs only in blocks containing an even number.

IRREDUCIBLE SHIFTS. A subshift is called **irreducible** if the language $B(X)$ has the property if v, w are words in $B(X)$, then there is a word u in $B(X)$ such that vuw is also in $B(X)$.

PROPOSITION. A subshift (X, T) is irreducible if and only if it is transitive.

PROOF. Assume the subshift is irreducible. We show that for every n, there is an orbit which comes $1/n$ close to any point in X. To do so, make a list of all words $w_0, ..., w_L$ of length $2n + 1$ which appear in X. By assumption we can fill in words $v_1, ..., v_L$ such that $w_0 v_1 w_1 v_1 .... v_L w_L$ is part of a sequence $x \in X$. Now, $T^n(x)$ comes $1/n$ close to any point in X.
On the other hand, if (X, T) is transitive, there is a point x such that $T^n(x)$ is dense. Given two words u, w which in the language of X, there exists n such that $(T^n(x)_1 .... T^n(x)_k) = u$ and m such that $(T^m(x)_1 .... T^m(x)_l) = w$. The word v between u and w in the sequence x is the one we need to prove irreducibility.

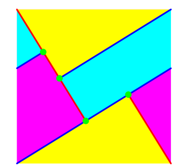
OVERVIEW. class of all subshifts $\supset$ class of all sofic shifts $\supset$ class of all shifts of finite type
CA leave the class of sofic shifts invariant because the composition of two CA is again a cellular automaton.

MINIMAL SHIFTS. A subshift (X, σ) is called **minimal**, if every orbit is dense. Note that minimal shifts can not have periodic points unless it is periodic itself.

EXAMPLE: STURMEAN SEQUENCES. **Sturmean sequences** $x_n = 1_A(t + n\alpha)$, where α is irrational and A is an interval on the circle are minimal because the irrational rotation on the circle is minimal and the symbolic map S is continuous and invertible. Because every orbit $T^n(x)$ of the irrational rotation is dense, also the corresponding orbit $S(T^n(x))$ is dense.

EXAMPLE: The full shift as well as subshifts of finite type are **not** minimal. They have many periodic orbits.

SYMBOLIC DYNAMICS. The basic construction of symbolic dynamics for a given dynamical system (Y, T) is to find a **partition** of the set X into subsets $A_0, ..., A_{n-1}$. Every point x is then assigned a sequence where $x_k = a$ if $T^k(x) \in A_a$. This **generating partition** defines a map S from Y to $X = \{0, ..., n-1\}^N$ if T is not invertible, or to $X = \{1, ..., n\}^Z$ if T is invertible. The map S conjugates (Y, T) to the subshift $(\overline{S(Y)}, \sigma)$. The map S is continuous, but it is in general neither injective nor surjective. In the homework, you deal with a a partition in case of the cat map. It is called **Markov partition**.

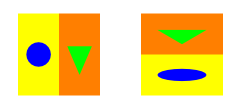


EXAMPLE. Let $T(y) = y + \alpha$ a rotation on the circle $Y = R/Z$. With $A_0 = [0, 1/2), A_1 = [1/2, 1)$, the sequence $x = S(y)$ is called a **Sturmean sequence**. The map S is a continuous map from the circle to the sequence space. But the image is not the entire space. For example, it does not contain any periodic sequences.

THE BAKER TRANSFORMATION. The baker transformation is a map on the square $Y = [0, 1) \times [0, 1)$: The map preserves area and is invertible

$$T(u, v) = \begin{cases} (2u, v/2) & , 0 \leq u < 1/2 \\ (2u - 1, (v + 1)/2) & , 1/2 \leq u \leq 1 \end{cases} \qquad T^{-1}(u, v) = \begin{cases} (u/2, 2v) & , 0 \leq v < 1/2 \\ ((u + 1)/2, 2v - 1) & , 1/2 \leq v \leq 1 \end{cases}$$

The inverse is obtained by switching u and v, applying T and switching u and v again. Now take the **generating partition** $A_0 = \{u \in [0, 1/2)\}, A_1 = \{u \in [1/2, 1]\}$. The symbolic dynamics of a point (u, v) defines a sequence $x \in \{0, 1\}^Z$.



THEOREM. The map S is an invertible map from the square Y to $X = S(Y)$ and $\sigma \circ S(u, v) = S \circ T(u, v)$. For any given sequence x in the image $S(Y)$, we can get back $(u, v) = S^{-1} x$ with

$$u = \sum_{k=0}^{\infty} x_k 2^{-k-1}, v = \sum_{k=-\infty}^{-1} x_k 2^k .$$

EXAMPLES:

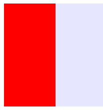
$...x_{-2}x_{-1}, x_0x_1x_2x_3...$	(u, v)
...0000, 10000..	$(1/2, 0)$
...0001, 00000..	$(0, 1/2)$
...0000, 01110..	$(7/16, 0)$
...0000, 11100..	$(7/8, 0)$
...0001, 11000..	$(3/4, 1/2)$
...0011, 10000..	$(1/2, 3/4)$
...0111, 00000..	$(0, 7/8)$
...1110, 00000..	$(0, 7/16)$

Remark: While S is injective, it is not surjective. (The point (....0000000, 0111111...) is not reached, but represented by (....0000000, 1000000...) $\sim (1/2, 0)$ While the map S^{-1} is continuous, T and S are both not.

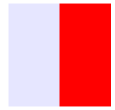
I: The binary expansion of u is $u = 0.x_0x_1x_2....$

$$u = \sum_{i=0}^{\infty} x_k 2^{-k-1} .$$

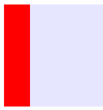
$x_0 = 0$ means that $u \in [0, 1/2)$. Note that $u = 1/2$ gives $x_0 = 1$.



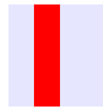
$x_0 = 1$ means that $u \in [1/2, 1)$.



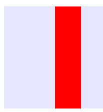
$x_0 = 0, x_1 = 0$ means $u \in [0, 1/2)$ and $2u \in [0, 1/2)$ which is equivalent to $u \in [0, 1/4)$.



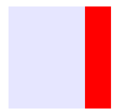
$x_0 = 0, x_1 = 1$ means $u \in [0, 1/2)$ and $2u \in [1/2, 1)$ which is equivalent to $u \in [1/4, 1/2)$.



$x_0 = 1, x_1 = 0$ means $u \in [0, 1/2)$ and $2u \in [0, 1/2)$ which is equivalent to $u \in [1/2, 3/4)$.



$x_0 = 1, x_1 = 1$ means $u \in [0, 1/2)$ and $2u \in [1/2, 1)$ which is equivalent to $u \in [3/4, 1)$.

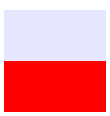


In general, fixing $x_0, ..., x_{n-1}$ determines in which of the 2^n intervals $[(k-1)/2^n, k/2^n)$ the coordinate u is.

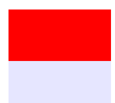
II: The binary expansion of v is $v = 0.x_{-1}x_{-2}x_{-3}....$

$$v = \sum_{k=-\infty}^{-1} x_k 2^k .$$

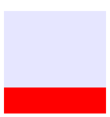
$x_{-1} = 0$ means that $v \in [0, 1/2)$. T maps the left half of the square to the lower half of the square so that T^{-1} maps the lower half of the square to the left half.



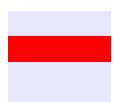
$x_{-1} = 1$ means that $v \in [1/2, 1)$.



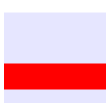
$x_{-1} = 0, x_{-2} = 0$ means $v \in [0, 1/2)$ and $2v \in [0, 1/2)$ which is equivalent to $v \in [0, 1/4)$.



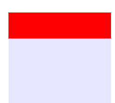
$x_{-1} = 0, x_{-2} = 1$ means $v \in [0, 1/2)$ and $2v \in [1/2, 1)$ which is equivalent to $v \in [1/4, 1/2)$.



$x_{-1} = 1, x_{-2} = 0$ means $v \in [0, 1/2)$ and $2v \in [0, 1/2)$ which is equivalent to $v \in [1/2, 3/4)$.



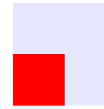
$x_{-1} = 1, x_{-2} = 1$ means $v \in [0, 1/2)$ and $2v \in [1/2, 1)$ which is equivalent to $v \in [3/4, 1)$.



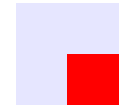
Fixing $x_{-1}, ..., x_{-n}$ determines in which of the 2^n intervals $[(k-1)/2^n, k/2^n)$ the coordinate v is.

III: Combination of part I and Part II

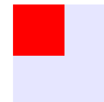
$x_{-1} = 0, x_0 = 0$ means $u \in [0, 1/2)$ and $v \in [0, 1/2)$.



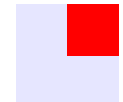
$x_{-1} = 0, x_0 = 1$ means $u \in [0, 1/2)$ and $v \in [1/2, 1)$.



$x_{-1} = 1, x_0 = 0$ means $u \in [1/2, 1)$ and $v \in [0, 1/2)$.



$x_{-1} = 1, x_0 = 1$ means $u \in [1/2, 1)$ and $v \in [1/2, 1)$.

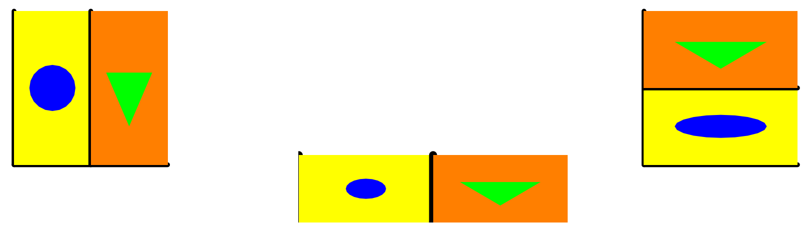


Fixing $x_{-m}, ..x_0, ..x_n$ determines in which of the 2^{n+m+1} rectangles $[(k-1)/2^n, k/2^n) \times [(l-1)/2^{m-1}, l/2^{m-1})$ the coordinate (u, v) is.

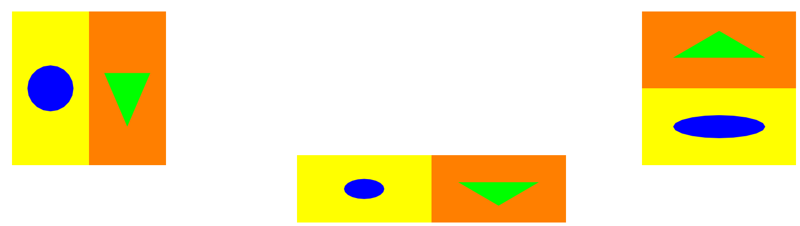
IV: Symmetry

We know $u = 0.x_0x_1x_2x_3x_4....$ Because replacing T and T^{-1} corresponds to switching u with v and replacing the partition $A_0, A-1$ with $B_0 = \{v < 1/2\}, B_1 = \{v \geq 1/2\}$, the itinery y with respect to the new partition gives $v = 0.y_0y_1y_2y_3y_4...$ Because $T(A_0) = B_0$, we have $v = 0.x_{-1}x_{-2}x_{-3}....$

BAKER MAP. In the baker map, the second rectangle is translated straight onto the first rectangle.



FAT HORSE SHOE MAP. The symbolic dynamics of the horse shoe is similar except that the second rectangle is turn around by 180 degrees. In the horse shoe, the stretching is stronger. There was a set K which never leaves the rectangle (the horse shoe is kind of a "Julia set").



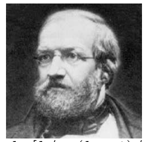
TWO REMARKS. The baker map can also be conjugated to the **right shift** $\sigma x_n = x_{n-1}$. If we take the same generating partition A_0, A_1, then the formulas for S^{-1} become $u = \sum_{k=-\infty}^{0} x_n 2^{-n-1}, v = \sum_{k=1}^{n} x_n 2^{-n}$. In many treatments of symbolic dynamics of the Baker transformations, one neglects things of area zero. In that case, it does not matter, what boundary we take for the generating partition. If we want the symbolic dynamics to work for **every** point in the square Y, then we remove the right and upper boundaries in all rectangles which appear as we have done that here.

APPROXIMATION OF NUMBERS **Math118, O. Knill**

ABSTRACT. The approximation of real numbers by rational numbers is a special and solvable case of solving the logarithm problem in dynamical systems.

DIRICHLET THEOREM. Let $x \in [0,1]$ be a real number in and $n > 1$ be an integer. There exist integers p and $1 \leq q \leq n$ $1 \leq p \leq n$ such that

$$|x - \frac{p}{q}| \leq \frac{1}{qn} .$$



PROOF. The Pigeonhole principle shows that at least one of the n intervals $[k/n, (k+1)/n]$ in $[0,1]$ contains two elements of the set $\{0, \{x\}, \{2x\}, ...\{nx\}\}$, where $\{kx\}$ is the fractional part of kx. So $|(kx - lx) + p| \leq 1/n$ for some integer p and $q = k - l < n$. Division through q gives $|x + p/q| \leq 1/(nq)$.

APPROXIMATION. For any irrational x, there are infinitely many p/q such that $|x - p/q| \leq 1/q^2$.

PROOF. If x is rational, $q = 0$ is possible and the result is not true. If x is irrational, then $k - l = 0$ is not possible and $q > 1$. Now $|x + p/q| \leq 1/(nq) \leq 1/q^2$.

CONTINUED FRACTION EXPANSION. We have seen the same result using continued fraction expansion p_n/q_n because

$$|x + p_n/q_n| \leq 1/(q_{n-1}q_n) \leq 1/q_n^2$$

There is a huge difference between this result and the above result

The pigeonhole principle is **not constructive**. It does not tell you what p/q is. The continued fraction expansion is **constructive**. You can determine p/q efficiently. The Dirichlet method needs n computations to determine the approximation, the continued fraction method essentially $\log(n)$.

THE LOG PROBLEM IN DYNAMICAL SYSTEMS.

Given a point x and a set I. At which time does the orbit of x enter I. For differential equations, we want to solve $T^t(x) = y$ up to some error, for maps, we want to solve $T^n(x) = y$ up to some error.

EXAMPLES.

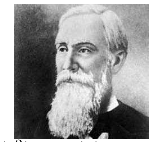
- If $T(x) = x + \alpha$ mod1 and $x = 0$ is a real and $y = 0$. Determining $T^q(t) = 0$ is the problem to find n such that $|q\alpha - p| = y$ for some integer p. In other words, we want to find close solutions of $|\alpha - p/q| = 0$. The continued fraction expansion gives such values.
- The differential equation $\dot{x} = ax$ has the solution $T^t(x) = e^t x(0)$. To solve $T^t(x) = a^t = y$, we have $t = \log_a(y)$. Computation of the real logarithm is a special case of the dynamical logarithm problem.
- Given an prime number p and an integer a, we have a map $T(x) = ax \bmod p$ on the set $X = \{1, ...p-1\}$. For given x and y, to compute n such that $T^n(x) = y$ is called the **discrete logarithm problem** in number theory. Logarithms are called **indices** in number theory. For a composite $n = pq$, if you could solve $a^k = 1 \bmod n$ we could find p. For example $5^4 = 1 \bmod 15$ so that $\gcd(4+1, 15) = 5$ is a factor. **The discrete log problem is harder then factoring**.
- If $T^t(x)$ is the evolution of the weather and x is the current meteorological condition and y is a severe storm, determining t such that $T^t(x)$ is close to y is a an example of a dynamical logarithm problem.
- If $T^t(x)$ is the position of an asteroid relatively to the earth and $y = 0$, then $T^t(x) = y$ determines the time it takes until the asteroid has an impact. It is an example of a dynamical logarithm problem.
- If T is the cellular automaton realization of a Turing machine, x is the initial condition with the empty tape and y is the "halt" state, then $T^n(x) = y$ determines how long it takes until the Turing machine halts. It is an example of a dynamical logarithm problem.

HURWITZ THEOREM. For any irrational x, there are infinitely many p/q such that $|x - \frac{p}{q}| \leq \frac{1}{\sqrt{5}q^2}$.

PROOF (Borel) One of the consecutive continued fraction convergent $p_{n-1}/q_{n-1}, p_n/q_n, p_{n+1}/q_{n+1}$ satisfies this bound. This is not so difficult to prove but could be part of a project.

This result can not be improved. The golden ratio satisfies this bound. There is an interesting story attached. If one takes away the bad example (the golden ratio) and all numbers which can be obtained by applying a modular transformation $T(x) = (ax+b)/(cx+d)$ with integers a, b, c, d satisfying $ad - bc = 1$, then the bound $\sqrt{5}$ can be improved to $\sqrt{8}$ which is the best possible bound attained by the **silver ratio** $\sqrt{2} + 1$.

SOLVING THE LOG PROBLEM FOR IRRATIONAL ROTATION. The following theorem solves the dynamical log problem for irrational rotations on the circle. Given two points on the circle, we can **construct** integers q_n such hat $T^n(x) = x + q_n\alpha$ is close to y.

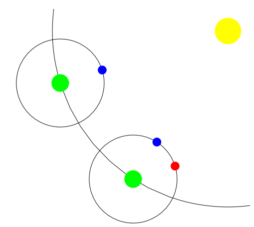


TCHEBYCHEV THEOREM. Assume x is irrational with periodic approximation p_n/q_n. Assume y is real. For every n, there exists $k \leq q_n$ such that $\{y + kx\} < 3/q_n$.

PROOF. Because $|x - p_n/q_n| \leq 1/(q_n q_{n-1})$, we can write $x = p_n/q_n + \delta/(q_n^2)$ with $|\delta| < 1$, where p_n/q_n are the periodic approximations of α.

Choose an integer t with $|q_n x - t| \leq 1/2$ so that $y = t/q_n + \delta'/(2q_n)$, $|\delta'| \leq 1$. Find k, l satisfying $q_n/2 \leq k \leq 3q_n/2$ with $p_n k - q_n l = t$. Then $|xk - l - y| = p_n k/q_n + \delta k/(q_n^2) - l - t/q_n - \delta'/(2q_n)| = |k\delta/q_n^2 - \delta'/(2q_n)| < k/(q_n q_n) + 1/(2q_n)$. Because $k < 3q_n/2$, the right hand side is $\leq 3/q_n$.

ECLIPSES AND PERIODIC APPROXIMATION. A **synodic month** is defined as the period of time between two new moons. It is $\alpha = 29.530588853$ days. The **draconic month** is the period of time of the moon to return to the same node. It is $\beta = 27.212220817$ days. Intersections between the path of the moon and the sun are called **ascending and descending nodes**. Such an intersection is called a solar eclipse. This appears in a period of a bit more then 18 years = 6580 days which is called one Saros cycle). This cycle and others are obtained from the continued fraction expansion of α/β. It is said that Thales using the Saros cycle to predict the solar eclipse of 585 B.C. The next big eclipse will happen May 26, 2021. Source: http://www.websters-online-dictionary.org/definition/english/mo/month.html



The Eclipse cycles can be explained using the continued fraction expansion (see homework).

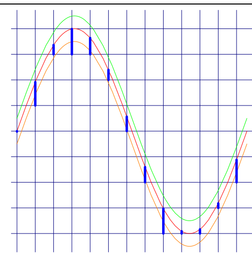
cycle	eclipse	synodic	draconic
fortnight	14.77	0.5	0.543
month	29.53	1	1.085
semester	177.18	6	6.511
lunar year	354.37	12	13.022
octon	1387.94	47	51.004
tritos	3986.63	135	146.501
saros	6585.32	223	241.999
Metonic cycle	6939.69	235	255.021
inex	10571.95	358	388.500
exeligmos	19755.96	669	725.996
Hipparchos	126007.02	4267	4630.531
Babylonian	161177.95	5458	5922.999

See http://www.phys.uu.nl/ vgent/calendar/eclipsecycles.htm for more details.

LATTICE POINT PROBLEMS — Math118, O. Knill

ABSTRACT. Finding lattice points close to curves leads to problems in dynamical systems theory.

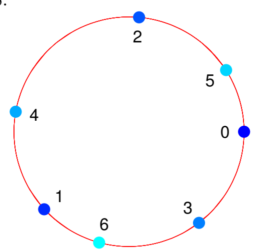
CURVES AND DYNAMICAL SYSTEMS. A curve $r(t) = (t, f(t))$ in the plane defines a sequence of points $x_n = f(n) \bmod 1 = f(n) - [f(n)]$ on the circle $T = R/Z$ and so a dynamical system $T : X \to X$, where X is the closure of all the translates of sequences $x = x_n$ and T is the shift.

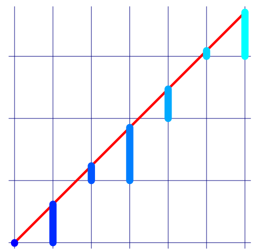


More generally, with the vectors $\vec{x}_n = (x_n, x_{n-1}, ..., x_{n-d})$, we can define a map $T(\vec{x}) = (x_{n+1}, x_n, ..., x_{n-d+1})$ on the d-dimensional torus $T^d = R^d/Z^d$. (For curves in space, there is a map on a higher dimensional torus, for two dimensional surfaces, time becomes two dimensional).

EXAMPLE STURMIAN SEQUENCES.

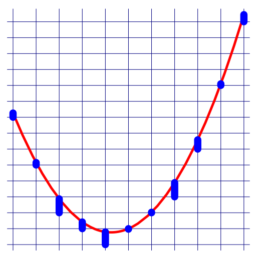
If $r(t) = (t, \alpha t)$ is a line in the plane with slope α, then $x_n = \alpha n \bmod 1$ and $\vec{x} = (..., n\alpha, ...)$ is called a **Sturmian sequence**. The map T is a rotation on the circle. It is a prototype of what one calls an **integrable system**, systems in which one can for example solve the dynamical logarithm problem.





EXAMPLE: PARABOLIC SEQUENCES.

For the parabola $r(t) = (t, \gamma + \alpha t + \beta t^2)$ we obtain the sequence $x_n = \gamma + n\alpha + n^2\beta \bmod 1$. it leads to a measure preserving transformation on the two dimensional torus $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 2\alpha \\ 0 \end{bmatrix} = A\vec{x} + \vec{b}$

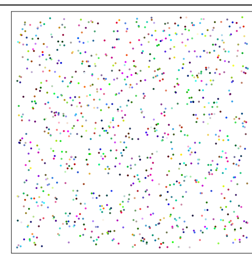


POLYNOMIALS If $p(x)$ is a polynomial of degree n, define $p_n(x) = p(x), p_{n-1}(x) = p_n(x+1) - p_n(x), p_{n-2} = p_{n-1}(x+1) - p_{n-1}(x), .., p_0(x) = \alpha$. Each p_i is a polynomial of degree i. If $T(x_1, x_2, ..., x_n) = (x_1 + \alpha, x + 2 + x_1, ..., x_n + x_{n-1})$, then $T(p_1(n), p_2(n), ..., p_d(n)) = (p_1(n+1), ..., p_d(n+1)) = (p_1(n) + \alpha, p_2(n) + p_1(n), ..., p_d(n) + p_{d-1}(n))$.

QUADATIC CASE: $p_2(x) = \gamma + \beta x + \alpha x^2, p_1(x) = p_2(x+1) - p_2(x) = \alpha + \beta + 2\alpha x, p_0(x) = p_1(x+1) - p_1(x) = 2\alpha$. We have a map $T(x, y) = (x + 2\alpha, x + y)$.

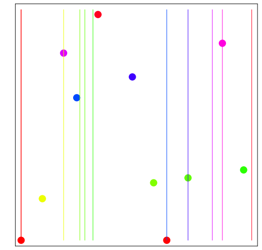
WEAK CHAOS IN PARABOLIC SEQUENCES.

The map $T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2\alpha \\ x + y \end{bmatrix}$ has zero Lyapunov exponent $\frac{1}{n}\lim(\log ||dT^n||)$. There is no sensitive dependence on initial conditions. If α is is irrational, then the map has only one invariant measure, the area. The map is also minimal: every orbit is dense. It is not chaotic in the sense of Devaney. It does not have even one single periodic orbit. The map T is an example of a system exhibiting a "weak type of chaos". There is no hyperbolicity present like in the cat map. Still, a single orbit covers the torus densely.



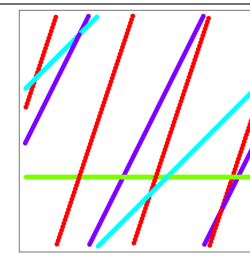
THE INTEGRABLE FACTOR IN PARABOLIC SEQUENCES.

If we look at the lines $y = const$, then these lines are tossed around in a regular way by the dynamics.



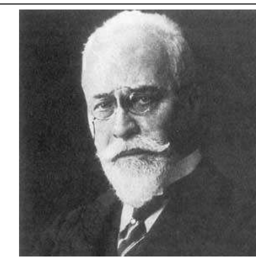
SOME DECAY OF CORRELATIONS.

The system also has mild chaotic behavior. A curve $y = const$ experiences a shear. Lets take a random variable $f(x, y) = f(x)$ which is independent of y. The random variables $f, f(T), ... f(T^n), ...$ show some decay of correlations $\int_{T^2} (f(T^n(x, y)) f(x, y) - f(x, y)^2) \, dxdy \to 0$ as time progresses.



WHY CONSTRUCT LATTICE POINTS CLOSE TO CURVES?

0) The problem is relevant in **cryptology**.
1) Estimating points close to curves is a problem in the **matric theory of Diophantine approximation**.
2) Finding points close enough to algebraic curves like $z = \sqrt{p(x)}$ lead to actual rational points on the manifold solving **Diophantine equations**.
3) Estimating lattice points in regions is a problem in the **geometry of numbers**, a field founded by Hermann Minkowski.
4) It relates to recurrence problems for classes of **dynamical systems**. It is a source for new type of dynamical systems.



CRYPTOLOGICAL APPLICATION: FACTORING INTEGERS. Given an integer $n = pq$ which is the product of two prime factors p, q, we want to find numbers y such that $y^2 = O(n^\alpha) \bmod (n)$, with α as small as possible. One way to do that is to look at numbers $[\sqrt{nx}]^2 \bmod n$. More generally, one can look at integer points (x, y) close to the curve $y^2 = np(x)$. As closer we are to the curve, as smaller $y^2 - np(x) = a$ is. Any algorithm which would find a of the order $O(n^\alpha)$ would with $\alpha < 1/2$ improve the speed of the current factorization algorithms.

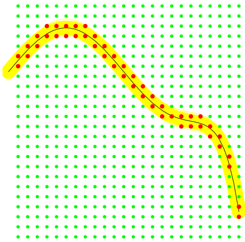
FACTORING ALGORITHMS. Some of the best factoring algorithms for a composite number $n = pq$ are based on an idea of Fermat: find x such $x^2 \bmod n$ is a small square y^2, then $x^2 - y^2$ is a multiple of n and $gcd(x - y, n)$ likely a factor of n. Example of algorithms are the **Morrison Brillard algorithm**, the **quadratic sieve** or the **number field sieve**. These methods allow to construct x for which y is of the order $\sqrt{n}$. A method to construct numbers x with $x^2 \bmod n$ of the order $n^{1/2-\epsilon}$ for some $\epsilon > 0$ would improve factorization methods.

EXAMPLE: PELL'S EQUATION. With $p(x, y) = x^2$, the curve $y^2 - nx^2 = 1$ is a hyperbola with asymptotes $y = \pm\sqrt{n}x$. The equation $y^2 = 1 + nx^2$ is called **Pells equation** or **Brouncker equation**. Integer points close to the line $y = \sqrt{n}x$ can be found using the continued fraction algorithm: if $\sqrt{n} \sim y_j/x_j$, then $y_j^2 - nx_j^2 = a$ and $y_j^2 = a \bmod n$. Because $\sqrt{n} = y/x + C/x^2$ we have $x\sqrt{n} - y = C/x$ and $x^2 n - y^2 = (x\sqrt{n} + y)C/x = C\sqrt{n} + Cy/x \sim 2C\sqrt{n}$. Here $\theta = 1/2$.

EXAMPLE PARABOLA. $p_n(x, y) = 2n + x$. The curve $y^2 = np_n(x)$ is a parabola. The tangent at $(x, y) = (0, \sqrt{2n^2 + 1})$ to the curve has slope $n/\sqrt{8n^2 + 1}$. The Diophantine error is $O(1/x)$. The nonlinearity error $y''(0)x^2 \sim x^2 n^2/y^3$. We have $y = O(n)$. In order that $1/x = n^2 x^2/y^3$, we must have $x = n^{1/3}$. The error is then $y/x = n^{2/3}$ so that $\alpha = 2/3$. If we could get rid of the quadratic or cubic erros, α would get smaller.

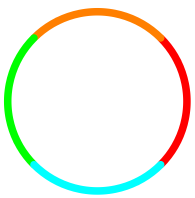
POINTS CLOSE TO A CURVE. The following result is a contribution to the **geometry of numbers**.

THEOREM. For every $0 \leq \delta < 1/3$ and every three times differentiable curve of finite length, there exists a positive constant C depending only on the curve, such that the number $M(n,\delta)$ of $1/n$-lattice points in a $1/n^{1+\delta}$ neighborhood of the curve satisfies $M(n,\delta)/n^{1-\delta} \to C$ for $n \to \infty$.

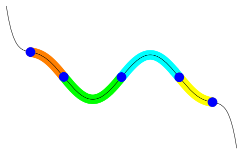


Remarks: if the curve is not a line, the constant C is positive. The constant can change under rotations of the curve, but does not change under translation of the curve.

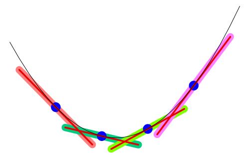
OUTLINE OF THE PROOF.



Cut the curve so that each piece is a graph



Cut the curve to have line segments or curves with nonzero curvature



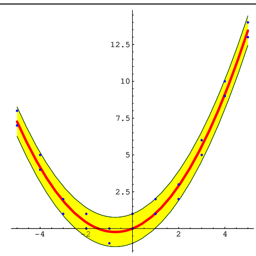
Approximate the curve by a polygon with Diophantine slopes

Remark. The polygon pieces have to be large enough so that continued fraction algorithm finds lattice points. On the other hand, the pieces have to be small enough to get a small nonlinearity error. A compromise is possible for $\delta < 1/3$. This bound $1/3$ is a limitation of the method. Results in the **metric theory of Diophantine approximation** indicate that $\delta < 1/2$ should be possible. Numerical experiments suggest that one can go even higher. An approximation by polynomials of higher degree could also put the bound higher. But then the proof no more be **constructive**.

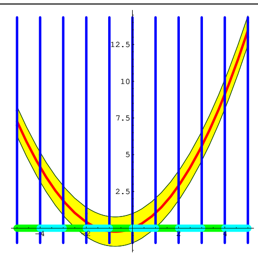
After cutting the curve into pieces, we can reformulate the theorem as follows:

THEOREM (Same result reduced to graph) Given a curve which is the graph of a smooth function $f(t)$ such that $f''(t) \geq \epsilon > 0$ on $[0,1]$. If $M(n,\delta)$ is the number of $1/n$-lattice points between $f(t) - 1/n^{1+\delta}$ and $f(t) + 1/n^{1+\delta}$. Then there is a constant C such that

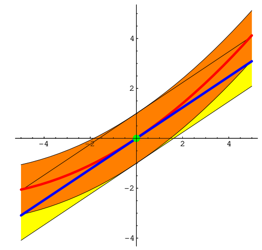
$$\frac{M(n,\delta)}{n^{1-\delta}} \to C \; .$$



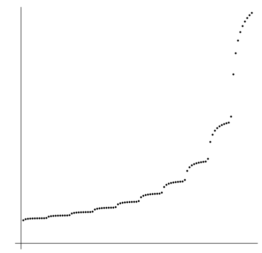
PROOF part (0). Let $[a,b] = f'[0,1]$ be the interval of possible slopes $f'(t)$ of f on $[0,1]$. Choose and fix a number $\delta < \theta < 1/3$ and call $\epsilon = 1/3 - \theta$. Let K be the maximum of $f''(x)$ on the interval $[0,1]$. For every n, divide the interval $[a,b]$ into $r(n,\theta) = [n^{1-\theta}]$ intervals I_k, called **small intervals**. The number of $1/n$-intervals in each of these intervals I_k is $[n^\theta]$. Call $M_k(n,\delta)$ the number of $1/n$ lattice points in the parallelepiped J_k above the interval I_k between $f(t) - 1/n^{1+\delta}$ and $f(t) + 1/n^{1+\delta}$.



RROOF part (i) (Nonlinear error) On one of the small intervals, the discrepancy of the curve to a tangent line is bounded above by $K/n^{2-2\theta} < K/n^{1+\theta+\epsilon}$. This uses Taylors formula $f(x+s) \in [f(x) + f'(x)s - Ks^2, f(x) + f'(x)s + Ks^2]$. It follows that if $M_{k,x}(n,\delta)$ denotes the number of lattice points in a $n^{-(1+\delta)}$ neighborhood J_k of a line segment at x above the interval I_k, then $(M_{k,x}(n,\delta) - M_k(n,\delta))/n^{1-\delta} \to 0$.



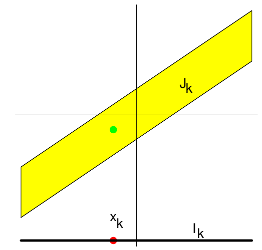
PROOF part (ii) (Sufficiently many strongly Diophantine slopes). Let $h(n,\delta)$ denote the number of intervals I_k, in which we can find x_k for which the slope $f'(x_k) = [a_0; a_1, a_2, ...]$ satisfies $a_i \leq \sqrt{r(n,\delta)}$. Then $h(n,\delta)/r(n,\delta) \to 1$ for $n \to \infty$.
Reformulation: the set of all numbers $y = [u, v, a_1, a_2, ...]$ with $u, v \leq M$ is $1/M^2$ dense on a set $Y_M \subset [0,1]$ with $|Y_M| \to 1$. A new reformulation: the set $\{f(u,v,x) = 1/u + 1/(v+x) = (v+x)/(u(v+x)+1) \mid u, v \leq M \ \}$ for $x \in [0,1]$ is $1/M^2$ dense on a set Y_M which has asymptotically full measure 1. This is a multivariable calculus problem: for $u, v \geq \sqrt{M}$, the distance from one point to the next is of the order $1/M^2$ because $f_v(u,v,x) = 1/(1+u(v+x))^2$.



PROOF part iii) (Reformulation for a line segment). Each of the $h(n,\theta)$ parallelograms J_k above I_k has slope α_k, thickness $n^{-1-\delta}$ and contains $[n^\theta]$ lattice units. In a scale, where the lattice size is 1, we have the following problem:

> Estimate the number of lattice points in a parallelogram J_k of length $[n^\theta]$ and thickness $n^{-\delta}$ for which the continued fraction expansion of the slope $\alpha_k = f'(x_k) = \alpha_k = [a_1, a_2, ...]$, with $a_i < n^\delta$.

The answer is that that there are n^ϵ lattice points.



PROOF part iv) (Number of lattice points in a Diophantine parallelogram J_k). There exists $c_k(n), d_k(n)$ such that the line segment J_k contains at least $[c_k(n)n^\epsilon]$ lattice points and maximally $[c_k(n)n^\epsilon]$ lattice points. Furthermore, $c_k(n) \to 1$ and $d_k(n) \to 1$ uniformly in k. There is a more general result of Schmidt and which even gives the error term.

PROOF part v) (Putting things together) The total number $M_{k,x_k}(n,\delta)$ of lattice points is between $c(n)h(n,\delta)n^\epsilon$ and $d(n)h(n,\delta)n^\epsilon$. Because of ii), we know it is between $c(n)r(n,\delta)n^\epsilon = c(n)n^{1-\delta}$ and $d(n)r(n,\delta)n^\epsilon = c(n)n^{1-\delta}$. Dividing by $n^{1-\delta}$ and using $c(n), d(n) \to 1$, we get the result.

AN OPEN PROBLEM. There is an efficient method to solve the dynamical logarithm problem for the map $T(x) = x + \alpha$: the continued fraction expansion gave an efficient method to find lattice points close to a line.

> Is there an efficient way to solve the dynamical logarithm problem for
>
> $$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2\alpha \\ x + y \end{bmatrix} \; .$$
>
> on the torus. A concrete problem: for $\alpha = \pi$, find n such that $T^n(0.5, 0.5)$ is within distance 10^{-1000} of $(0,0)$.

Geometrically, we look for an efficient method to find lattice points close to the parabola $y = \alpha x^2 + \beta x + \gamma$ with irrational α. Of course, we could just list all numbers $[\alpha n^2 + \beta n + \gamma]$ and see which one is close, this is not practical. While we can find in a few thousand computation steps an integer n such that $[\alpha n]$ is smaller than 10^{-1000} (it is a [P] problem) more then 10^{100} computations seem needed in the parabolic case (is it a [NP] problem?). Note that the big bang happend about 10^{17} seconds ago.

STRICT ERGODICITY (* not treated in class) Math118, O. Knill

ABSTRACT. The irrational rotation on the circle is a minimal uniquely ergodic system. Other systems occuring in number theory have the same property.

ERGODICITY. A map T is ergodic if for every function $f(x) = \sum_{n \in Z} c_n e^{inx}$ with finite $\sum_n c_n^2$, the condition $f(T) = f$ implies $f = \text{const}$.

THEOREM. For irrational α, the map $T(x) = x + \alpha$ is ergodic.

PROOF. Comparing Fourier coefficients of $f(T)$ and f gives $e^{in\alpha} c_n = c_n$ so that $c_n = 0$ unless $n = 0$.

UNIQUE ERGODICITY. A continuous transformation T on a compact topological space is called **uniquely ergodic** if there is only one invariant measure μ of T.

KRONECKER-WEYL THEOREM. The only measure which is invariant under an irrational rotation is the length measure dx.

PROOF. A measure μ is a linear map from the space of all continuous functions $C(X)$ to R given by $\mu(f) = \int f(x)\, d\mu(x)$. If μ is T invariant, then $\mu(f(T)) = \mu(f)$ and by linearity $\mu(\frac{1}{n}\sum_{k=1}^n f(T^k)) = \mu(f)$. Because for $f(x) = e^{ikx}$, we have

$$\frac{1}{n}\sum_{k=1}^n f(T^k) = \frac{1}{n}\sum_{k=1}^n e^{ij(x+k\alpha)} = \frac{e^{ijx}}{n}\frac{(1-e^{ijn\alpha})}{(1-e^{i\alpha})} \to 0$$

also for any $f = \sum_k e^{ikx}$ we have $\mu(f) = \mu(\frac{1}{n}\sum_{k=1}^n f(T^k)) \to c_0$ for $n \to \infty$ which implies $\mu(f) = c_0 = \int f(x)\, dx$.

MINIMALITY. A map T is called **minimal**, if every orbit of T is dense.

THEOREM. The irrational rotation on the circle is minimal.

PROOF. This follows in a constructive way from Chebychevs theorem. For every x and y and $\epsilon > 0$, there exists n such that $|x + n\alpha - y| < \epsilon$.

STRICT ERGODICITY. A map is called **strictly ergodic**, if it is both minimal and uniquely ergodic.

COROLLARY. The irrational rotation on the circle is strictly ergodic.

HIGHER DIMENSIONAL GENERALIZATION. The above statements go through word by word for a rotation $T(x) = x + \alpha$ with vectors $\alpha = (\alpha_1, ..., \alpha_d)$ for which $n \cdot \alpha = n_1\alpha_1 + ... + n_d\alpha_d) = 0$ implies $n = 0$. We call such vectors **irrational**. Functions of several variables have a Fourier expansion too: $f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\cdot x}$, where $n = (n_1, ..., n_d)$ runs over all lattice points in Z^d.

COROLLARY. The irrational translation on the torus $T^d = R^d/Z^d$ is strictly ergodic.

PROOF. We have shown both minimality as well as unique ergodicity.

THEOREM (FURSTENBERG) If α is irrational and $b_{ij} \in \mathbf{Z}, 1 \leq j < i \leq d$ real with $b_{i,i-1} \neq 0$. Then $T(x_1, \dots, x_d) = (x_1 + \alpha, x_2 + b_{21}x_1, ..., x_d + b_{d1}x_1 + ... + b_{d,d-1}x_{d-1})$ defines a uniquely ergodic system on $\mathbf{T}^d$.

It can be written as $\vec{x} \mapsto A\vec{x} + e_1\alpha$, where $A = \begin{bmatrix} 1 & 0 & . & 0 \\ b_{21} & 1 & . & 0 \\ b_{31} & b_{32} & . & . \\ . & . & . & . \\ b_{d1} & . & . & 1 \end{bmatrix}$.

PROOF. Fourier theory shows that T is ergodic: $f(T) = \sum_n c_n e^{in\cdot T(x)}$ with $n \cdot T(x) = (n_1, ..., n_d) \cdot (x_1 + \alpha, x_2 + b_{21}x_1, ..., x_d + b_{d1}x_1 + ... + b_{d,d-1}x_{d-1}) = n_1\alpha + An \cdot x$. Comparing Fourier coefficients gives $c_{An} = c_n e^{2\pi i n_1 \alpha}$ which implies $n = (n_1, 0, ..., 0)$ and therefore $c_n = c_n e^{2\pi i n_1\alpha}$ which implies that $c_n = 0$ unless $n = 0$.

Unique ergodicity is shown with induction to d. We know it for $d = 1$, where the system is an irrational rotation. To prove the result in dimension d, write $T(\vec{x}, x_d) = (S(\vec{x}), x_d + A \cdot \vec{x})$. Note that S does not depend on x_d. By induction, S is uniquely ergodic on T^{d-1}. Given invariant measure μ for T, the projection of μ on $\mathbf{T}^{d-1}$ is S-invariant. and by induction assumption the volume measure $dx_1...dx_{d-1}$.

Because T commutes with $R(\vec{x}, y) = (\vec{x}, y + \beta)$, a T invariant measure must also be R_β invariant for every β. By Birkhoffs ergodic theorem, we know that μ almost all points $x = (\vec{x}, x_d)$ are **generic** in the sens that $\mu(f) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k x)$. Assume $x = (\vec{x}, y)$ is generic. Then also $(\vec{x}, y + \beta)$ is generic.

A uniquely ergodic system on the torus which preserves the volume measure $dx_1...dx_n$ is automatically minimal: if there were an orbit x which were not dense, then its closure Y would be a T invariant set which is not the entire torus. This set would carry an other invariant measure.

ILLUSTRATION. Lets see this in the case $T(x, y) \to (x, x + y) \to (x + \alpha, x + y)$. When projecting onto the first coordinate, we have the uniquely ergodic map $x \to x + \alpha$. The key is that the map T commutes with $R(x, y) = (x, y + \beta)$:

$$T(R(x,y)) = T(x, y+\beta) = (x+\alpha, x+y+\beta), R(T(x,y)) = R(x+\alpha, x+y) = (x+\alpha, x+y+\beta) .$$

If (x_n, y_n) is an orbit, then the distribution of x_n on the first coordinate is the measure dx. Assume two different points $(x, y), (x, y + \beta)$ with irrational β produce measures $\mu(x, y), \mu(x, y + \beta)$ which must coincide.

APPLICATION: Let $p(x)$ be polynomial of degree n. Define $p_n(x) = p(x), p_{n-1} = p_n(x + 1) - p_n(x), p_{n-2} = p_{n-1}(x + 1) - p_{n-1}(x), .., p_0(x) = \alpha$. Each p_i is a polynomial of degree i. With

$$T_p = \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 + \alpha \\ x_2 + x_1 \\ ... \\ x_n + x_{n-1} \end{bmatrix}$$

we have $T_p(p_1(n), p_2(n), ..., p_d(n)) = (p_1(n + 1), ..., p_d(n + 1))$.

CORROLLARY. If $p = a_n x^n + ... + a_1 x + a_0$ is a polynomial of degree n and assume a_n is irrational, then T_p is a uniquely ergodic transformation on the n dimensional torus which preserves the volume $\mu = dx_1...dx_n$.

QUESTION. Are polynomials the only functions f for which one can describe $f(n)$ mod 1 by a finite dimensional system?

EXAMPLES.
1) For $f(x) = \sqrt{x}$. What dynamical system does $\sqrt{n}$ mod 1 generate?
2) Does $f(x) = \exp(x)$ generate an infinite dimensional system?
3) If $f(x)$ is a k-periodic function, then $f(n)$ is periodic too For $f(x) = \sin(2\pi x\alpha)$ with irrational α, then $f(n)$ is an almost periodic sequence. The system on the torus $(x, y) \to (x + \alpha), \sin(2\pi x))$ allows to read of $f(n)$ in one coordinate.
4) For rational functions like $f(x) = x/(1 + x^2)$, the system has a fixed point which attracts all points.

OTHER STRICTLY ERGODIC SYSTEMS. Any factor of a strictly ergodic system is strictly ergodic. This applies to symbolic dynamics.

Doing symbolic dynamics with a strictly ergodic system produces strictly ergodic subshifts. Let $A_1, A_2, .., A_n$ be a partition of T^d into subsets, define $S(x)_n = k$ if $T^n(x) \in A_k$. This defines a subshift which is strictly ergodic.

EXAMPLES. Sturmean sequences $x_n = 1_A(x + n\alpha)$ and especially the Fibonacci sequence are uniquely ergodic subshifts. Applying cellular automata maps on such subshifts generates new subshifts which are strictly ergodic. CA maps preserve both minimality as well as unique ergodicity.

NUMBERS AND DYNAMICAL SYSTEMS — Math118, O. Knill

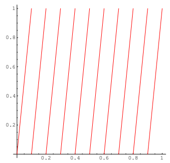
ABSTRACT. Numbers can be represented in various ways. In many cases, the representation of real numbers can be seen as a construction in symbolic dynamics.

REPRESENTATIONS OF REAL NUMBERS.

Given a finite generating partition $A_0, A_1, ..., A_n$ of the interval $[0,1)$, define $f(y) = i$ if $y \in A_i$ and a map $T : [0,1) \to [0,1)$ we can look at the orbit of a point y and define the sequence $x_n = f(T^n y)$.
We are interested in cases, where the sequence x_n determines x for all x. If T is a piecewise smooth expanding map, then this is the case.

Many representations of numbers as sequences of a finite symbols is described by symbolic dynamics.

DECIMAL EXPANSION. Let $T(x) = 10x$ and $f(x) = [10x]$ where $[r]$ is the **integer part** of r. Let $A_0, A_1, ...A_9$ be the intervals defined by $A_k = \{f(x) = k\}$.
This is the decimal expansion of x. From the sequence a_j, we can reconstruct $x = \sum_{j=1}^{\infty} a_j 10^{-j}$.



CONTINUED FRACTION EXPANSION. Take $T(y) = 1/y$ mod 1 and $f(y) = [1/y]$. For a point y, define the sequence $a_n = f(T^n(y))$. It is called the **continued fraction expansion** of y. if y is a rational number, then

$$y = [a_0; a_1, ...a_n] = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + ... \frac{1}{1+a_n}}} = p_n/q_n$$

If y is an irrational number, then

$$y = [a_0; a_1, ...a_n, ...] = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + ... \frac{1}{1+a_n} + ...}}$$

EXAMPLES. $\sqrt{2} = [1; 2, 2, 2, 2, 2, ...]$. Since $1/(2+x) = x$ has the solution $\sqrt{2} - 1$.
$(\sqrt{5} - 1)/2 = [1; 1, 1, 1, 1, 1, ...]$. Since $1/(1+x) = x$ has the solution $(\sqrt{5} - 1)/2$.
$5/7 = [0; 1, 2, 2]$

PARTIAL QUOTIENTS. The **partial quotients** p_n/q_n satisfy the recursion $p_n = a_n p_{n-1} + p_{n-2}$, $q_n = a_n q_{n-1} + q_{n-2}$ with the initial conditions $p_{-1} = 1, p_0 = a_0, q_{-1} = 0, q_0 = 1$ so that $p_0/q_0 = a_0, p_1/q_1 = a_0 + 1/a_1 = (a_0 a_1 + 1)/a_1$.

CONVERGENCE ESTIMATES. One can write the second order recursion as a first order recursion $\begin{bmatrix} p_n \\ p_{n-1} \end{bmatrix} = \begin{bmatrix} a_n & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p_{n-1} \\ p_{n-2} \end{bmatrix}$. In the product of matrices $A^n = A_n ... A_0 = \begin{bmatrix} p_n & q_n \\ p_{n-1} & q_{n-1} \end{bmatrix}$ each matrix $A_k = \begin{bmatrix} a_k & 1 \\ 1 & 0 \end{bmatrix}$ has determinant (-1). The product has therefore the determinant $(-1)^n$. This gives the important identity

$$\boxed{p_{n-1} q_n - p_n q_{n-1} = (-1)^n}$$

which implies $p_{n-1}/q_{n-1} - p_n/q_n = (-1)^n/(q_n q_{n-1})$. Since $q_n \geq q_{n-1} = 1$, we have $q_n \geq n$ and $|p_{n-1}/q_{n-1} - p_n/q_n| \leq (-1)^n/n^2$ so that p_n/q_n is a Cauchy sequence. Because p_n/q_n is alternatively below and above x (look at the images of the basis vectors of A_k), we have even the bound

$$\boxed{|x - \frac{p_n}{q_n}| < \frac{1}{q_n q_{n+1}} < \frac{1}{q_n^2}}$$

SOLVING LINEAR EQUATIONS. Given a, b, c, how do we solve $ax + by = c$ for integers x, y?
Solution: we can solve $p_{n-1} q_n - p_n q_{n-1} = (-1)^n$ by making the continued fraction expansion of p_n/q_n then multiply the result with $(-1)^n c$.

EXPANSION OF PI. To find the continued fraction expansion of $x = \pi$: $\pi = 3 + 1/(7 + ...)$, look at the orbit of $x = 0.141592653...$ under the map $T(x) = \{1/x\}$ and see in which intervals they fall.

T[x_] := Mod[1/x, 1]; S = NestList[T, Pi − 3, 10]; f[x_] := Floor[1/x]; Map[f, S]

Mathematica has already built in the continued fraction expansion as a basic function:

ContinuedFraction[Pi, 10]

The result is $\pi = [3; 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, 2, 1, 1, 2, 2, 2, 2, 1, 84, 2, 1, 1, 15, 3, 13, 1, 4, ...]$. Continued fraction expansion of π has been computed up to 10^8 terms. One can use partial quotients like

$$\pi \sim [3; 7] = 22/7 = 3.14286$$

$$\pi \sim [3; 7, 15] = 333/106 = 3.14151$$

$$\pi \sim [3; 7, 15, 1] = 355/113 = 3.14159$$

to approximate π with rational numbers. Mathematica has the reconstruction of a number from the continued fraction built in too:
FromContinuedFraction[3, 2, 1]

KHINCHIN CONSTANT. If $[a_0; a_1, a_2, ...]$ is the continued fraction expansion of a number, then the limit $(a_1 a_2 a_3 ... a_n)^{1/n}$ exists for almost all irrational numbers. The limit is called **Khinchins constant**. Numerical experiments indicate that this limit is obtained for π but one does not know.

β-EXPANSION. A generalization of the decimal or expansion with respect to an integer base is the **beta expansion**. For any given real number $\beta > 1$, define the map $T(x) = \beta x$ and $f(x) = [\beta x]$. One has still $x = \sum_{i=1}^{\infty} a_i \beta^{-i}$ however, the transformation is no more so easy to understand as in the integer case. For example, T_β does not preserve the length measure dx any more in general.

PERIODIC POINTS. As in any dynamical system, also for dynamical systems which define number, periodic points are important. Examples:

- **Rational points** are eventually periodic points of the decimal expansion.
- quadratic irrationals are eventually periodic points of the continued fraction expansion.
- Numbers which lead to eventually periodic orbits of the β-expansion are called **beta numbers**.

The determination whether an orbit is eventually periodic or not is nontrivial. For example it is unknown whether $\pi + e$ is rational. In other words, one does not know whether the shift on $X_{\pi+e,10}$ is eventually periodic.

BETA NUMBERS. An interesting question is for which real numbers β and $x = 1$, the attractor is a periodic orbit. If this is the case, then β is called a **beta number**. Examples are **Pisot numbers**, algebraic integers $\beta > 1$ for which all conjugates β^σ have norm $|\beta^\sigma < 1$ besides the identity. The positive root of $x^3 - x - 1 = 0$ is known to be the smallest Pisot number. If $|\beta^\sigma| \leq 1$ for any embedding and β is not a Pisot number, it is called a **Salem number**.

NORMALITY. If every word of length k in the decimal expansion of π appears with probability 10^{-k}, then π is **normal**. One does not know whether this is true. **Normality** results are hard to get. And normality with respect to one base does not mean normality with respect to an other base. Normality is a statement with respect a specific shift invariant measure and If a number is normal with respect to all bases is called **absolutely normal**. A well studied open problem is

Is π normal with respect to any base or even absolutely normal?

STRANGE SINGULARITIES AND ORBITS Math118, O. Knill

ABSTRACT. Non-collision singularities are possible in the Newtonian n-body problem by careful construction. Also the construction of special solutions to the n-body problems is an art.

PAINLEVES CONJECTURE: Painlevé asked in his Stockholm lectures of 1895: for $n > 3$, do there exist solutions of the Newtonian n-body problem with singularities that are not due to collisions?

HISTORY. Zeipels theorem showed that singularities of the Newtonian n-body problem are either collisions or configurations for which particles escape to infinity in finite time. Poincaré seems have considered this question already, even so he never wrote it down. Painlevé gave Poincarée credit for having asked that some $x_i(t)$ might go to infinity or oscillate wildly like $\sin(1/(t-\tau)$ as t converges to the singularity. Painlevé himself proved that non-collision singularities do not exist for the three body problem. Painlevés question whether non-collision singularities can occur, stayed open until Jeff Xia constructed non-collision singularities in 1992. (By the way, Xia was at Harvard from 1988-1990, so some of the final polishing of this paper could have been done here). Another mathematician, Joseph Gerver, had also been in the race but considered a planar approach, where the number of particles is large. John Mather and Richard Mc Gehee had already in 1974 shown that particles can escape to infinity but their construction on the one dimensional line and binary collisions were allowed. While it is known that for four bodies, non-collision singularities have measure zero, one does not know whether they exist. There is a construction of a planar 4 body situation of Gerver from 2003 which suggests that the answer could be yes.

A THEOREM OF PAINLEVE.

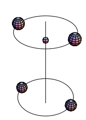
THEOREM (1897) There are no non-collision singularities in the three body problem.

PROOF. The Lagrange-Jacobi equation is $\ddot{I} = U + 2H$, where $I = \sum_{j=1}^{3} m_j r_j^2$ is the **moment of inertia**. I is a measure of diameter of the triangle defined by the positions of the three particles. These equations imply that whenever two particles come close, the triangle they span has to become large. By the triangle inequality, two sides of the triangle are then large. The Sundman-van Zeipel lemma assured that $I(t) \to I^*$ for $t \to \tau$ with $I^* = \infty$ if there is a non-collision singularity. Assuming $I(t) \to \infty$ for $t \to \tau$ we have $I(\ddot{t}_k) \to \infty$ for some sequence of times $t_k \to \tau$ which means $U(t_k) \to \infty$. This implies that two of the three particles must come close to each other. In the same time, the third "lonely" particle has to be far away from these two particles because $I(t) \to \infty$. Because the acceleration of the lonely particle and the center of mass of the binary both stay bounded for $t \to \tau$, these positions converge to a definite finite value for $t \to \tau$. The collision assumption means that the binary system collides for $t = \tau$ but at a finite distance from the third particle. Consequently $I(\tau) = I^* < \infty$, which is in direct contradiction to the assumption $I^* = \infty$.

THEOREM OF XIA.

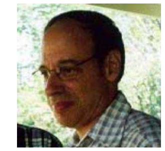
Non-collision singularities exist in the Newtonian 5 body problem. There are initial conditions for the Newtonian 5 body problem in which the bodies escape to infinity in finite time.

BASIC IDEA. The setup is to add a second binary solar system to the Sitnikov system. The planet moving on the z-axes visits alternatively the two binary systems. The timing is done in such a way that the planet will bounce back accelerated after visiting one of the systems. The energy is drawn from the potential energy of the two binary systems which move closer and closer together. The four suns have all the same mass. The upper and lower "solar systems" have opposite angular momentum and their "Kepler orbits" are highly eccentric.

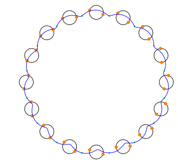


THEOREM OF GERVER. Joseph Gerver proved a theorem for the planar case:

THEOREM. For large n, non-collision singularities exist for the planar n-body problem

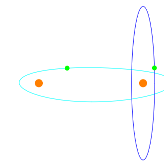


BASIC IDEA. There are 3N bodies in the plane. The configurations are symmetric with respect to rotations by $2\pi/N$. There are N binary systems in which all suns have the same mass. There are N planets which move from one pair to the other. The successive time spans, which the planets need to jump from one to the next system forms a sequence Δ_k with the property that $\sum_k \Delta_k < \infty$.



GERVERS SUGGESTION: Are there planar four body configurations in which particles escape to infinity in finite time?

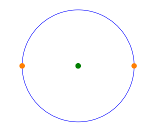
Gervers model contains two planetary systems: there are two suns S_1, S_2 with large mass and two planets P_1, P_2 with small mass. Planet P_2 circles sun S_2 in an elliptical orbit. Planet P_1 circles around Sun S_1 and visits the second planetary system, where it alternatively gains angular momentum and energy.



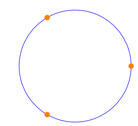
SPECIAL SOLUTIONS. An interesting research topic is the search for special solutions of the 3 body problem.

EQUILIBRIUM SOLUTIONS. Whenever we studied differential equations, we were interested in **equilibrium solutions**, stationary solutions. Are there equilibrium solutions for the Newtonian n body problem? The answer is no: from $\ddot{x}_k = 0$ would imply that $U_{x_k} = 0$ and from Eulers theorem on homogeneous functions that $-U = \sum_{j=1}^{n} x_k U_{x_k} = 0$. But the potential U is clearly positive everywhere.

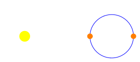
EULERS SOLUTIONS (1767) Euler was the first who found special solutions to the three body problem. In these solutions, the three bodies rotate on circles but remain on a line. The Euler solution and the Lagrange solution below are the only solutions for which the particles move uniformly along circular orbits in a fixed plane.



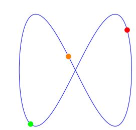
LAGRANGIAN SOLUTIONS (1772). The three bodies are on an equilateral triangle. This system appears in nature: the Trojan asteroids together with Jupiter and the sun essentially move according to this. Lagrange, who found this solution did not think this has any significance in astronomy.



HILLS SOLUTIONS. These are configurations resembling the Earth-Moon-Sun system. Two bodies move closely around each other while both of them circle a third body.



MOORE CHOREOGRAPHIES. Three bodies of equal mass follow each other on a figure eight type orbit. These solutions have been discovered by Cristopher Moore in 1993 through computer calculations.



LITERATURE. A vivid account on the history of non-collision singularities also containing many annectotes about the discovery is the book "Celestial Encounters" by Florin Diaco and Philp Holmes. The article "Off to infinity in Finite Time" by Donald Saari and Jeff Xia gives a nice summary. For a suggestion, how a four body noncollision singularity might work, see Joseph Gervers article "Non collision Singularities: Do four bodies Suffice?".

N-BODY PROBLEMS — Math118, O. Knill

ABSTRACT. The Newtonian n-body problem influenced the development of mathematics at several occasions. For example, it was the catalisator for the development of calculus or topology. In this section, we look at general facts about the n-body problem like existence of solutions and the nature of singularities like Zeipel theorem which distinguishes collision and noncollision singularities by the convergence of the moment of inertia.

NEWTON EQUATIONS.

Celestial mechanics is the study of the **Newtonian n-body problem**, the study of the differential equations

$$\boxed{m_j \ddot{x}_j = -G \sum_{i \neq j} \frac{m_j m_i (x_j - x_i)}{|x_i - x_j|^3}}$$

The vectors x_j are the positions of the bodies with mass m_j and G is the gravitational constant. If the initial positions and velocities of the bodies are known, then the equations determine the position of the bodies at later times as long as solutions exist. The **phase space** of the system is the 6n-dimensional space $M \times \mathbf{R}^{3n}$, where $M = \mathbf{R}^{3n} \setminus \Delta$ with the **collision set**

$$\Delta = \bigcup_{i \neq j} \Delta_{ij} = \bigcup_{i \neq j} \{ x \in \mathbf{R}^{3n} \mid x_i = x_j \} .$$

HAMILTON EQUATIONS FOR THE N-BODY PROBLEM.

A point (x, y) with $x = (x_1, \dots, x_n), y = (y_1, ..., y_n)$ in the phase space encodes both the positions x_i and the momenta $y_i = m_i \dot{x}_i$ of the bodies. The function

$$H(x, y) = \sum_{j=1}^n \frac{y_j^2}{2m_j} - U(x), \quad U(x) = G \sum_{i<j} \frac{m_i m_j}{|x_i - x_j|}$$

on the phase space is the **energy** of the particle system. One calls it the **Hamiltonian**. The Newton equations can be rewritten as **Hamilton equations**

$$\boxed{\dot{x}_j = \nabla_{y_j} H(x, y), \quad \dot{y}_j = -\nabla_{x_j} H(x, y) .}$$

10 CLASSICAL INTEGRALS. An **integral of motion** of a Hamiltonian system is a quantity which is conserved along the orbits.

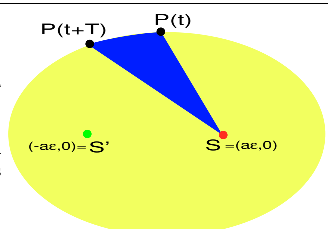
The n-body problem in three dimensions has the **10 classical integrals** of motion:

a) The **total momentum** $Y = \sum_{i=1}^n y_i$.
b) If $Y = 0$, the **position** $C = \sum_{i=1}^n m_i x_i$ of the center of mass.
c) The **total angular momentum** $L = \sum_i x_i \times y_i$.
d) The **total energy** H.

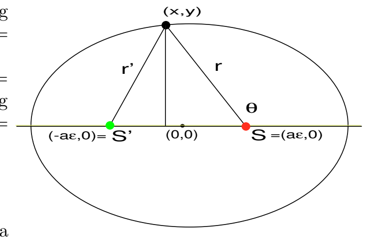
Proofs.
a) Every term in the sum $\dot{Y}$ appears twice but with opposite sign.
b) From $\dot{Y} = C$ follows that if $Y = 0$ then C is constant.
c) $\dot{L} = \sum_{i=1}^n \dot{x}_i \times y_i + \sum_i x_i \times \dot{y}_i$. The second sum is zero because $x_i \times (x_i - x_j) = -x_i \times x_j$ and because each term in the remaining sum appears twice with opposite sign.
d) $\dot{H} = \sum_{i=1}^n H_{x_i} \dot{x}_i + H_{y_i} \dot{y}_i = \sum_{i=1}^n H_{x_i} H_{y_i} - H_{y_i} H_{x_i} = 0$.

THE 2 BODY PROBLEM. After a change of coordinates, one can assume that the center of mass $C = m_1 x_1 + m_2 x_2$ is at the origin. If $q = x_1 - x_2$, then $\ddot{q} = \ddot{x}_1 - \ddot{x}_2 = m_2 G (x_2 - x_1)/|x_2 - x_1|^3 - m_1 G (x_1 - x_2)/|x_2 - x_1|^3 = -(m_1 + m_2) G q / |q|^3$. This is a 1-body problem for a particle with position q and mass $m = m_1 + m_2$ moving in a central field. The angular momentum $L = m x \times \dot{x}$ and the energy $2E/m = \dot{x}^2 + G/|x|$ are conserved quantities.

THE 2. KEPLER LAW. Because $\ddot{x}$ is parallel to x, we get $\dot{L} = 0$. From the conservation of L follows that the vector x stays in a plane, where we can us polar coordinates $x = (r\cos(\theta), r\sin(\theta))$. The constant quantity $L = mr^2\dot{\theta}$ can be interpreted as df/dt, where f is the area swept over by the vector x. We have derived the "area law", Keplers second law: "the radius vector x passes the same area in the same time."



THE 1. KEPLER LAW. An ellipse with focal points $S' = (-a\epsilon, 0), S = (a\epsilon, 0)$ is the set of points (x, y) whose distances r' and r to S' and S satisfy $r' + r = 2a$. The number ϵ is called the **eccentricity**. From $(2a - r)^2 = r'^2 = r^2 \sin^2(\theta) + (2a\epsilon + r\cos(\theta))^2$, we obtain $r = a(1-\epsilon^2)/(1+\epsilon\cos(\theta))$, the **polar form** of the ellipse. Differentiation of this with respect to time, using $\dot{\theta} = L/(mr^2)$ leads to $\dot{r} = a(1-\epsilon^2)\sin(\theta)(1+\epsilon\cos(\theta)^{-2} L/(mr^2) = L\epsilon\sin(\theta)/(ma(1-\epsilon^2))$ and $\ddot{r} = L^2\epsilon\cos(\theta)/(m^2 a(1-\epsilon^2) r^2)$.
With $n = x/r$, one has $\ddot{x} = (\ddot{x} \cdot n)n$ and $x = nr$ gives $\dot{x} = \dot{n}r + n\dot{r}, \ddot{x} = \ddot{n}r + 2\dot{n}\dot{r} + n\ddot{r}$ so that $\ddot{x} \cdot n = \ddot{n} \cdot nr + 2\dot{n} \cdot n\dot{r} + \ddot{r}$. Using $n \cdot n = 1 \Rightarrow n \cdot \dot{n} = 0, \ddot{n} \cdot n + \dot{n} \cdot \dot{n} = 0$ and $\dot{n} \cdot \dot{n} = \dot{\theta}^2 n^\perp \cdot n^\perp = \dot{\theta}^2 = L^2/(mr^2)^2$, we have



$$\boxed{\ddot{x} \cdot n = \ddot{r} - L^2/(m^2 r^4)}$$

With $1/r^4 = (1/r)(1/r^3) = (1+\epsilon)/(r^3 a(1-\epsilon^2))$ and the formula for $\ddot{r}$ we get $\ddot{x} \cdot n = L^2\epsilon\cos(\theta)/(m^2 a(1-\epsilon^2) r^2) - (L^2/(m^2))(1+\epsilon)/(r^3 a(1-\epsilon^2)) = -L^2/(a(1-\epsilon^2) r^2$ so that $\ddot{x} = (\ddot{x} \cdot n)n = -L^2/(a(1-\epsilon^2)) x/r^3 = -Gx/r^3$.

THE 3. KEPLER LAW. If T is the period of the orbit, the the third Kepler law states that T^2/a^3 is constant. Indeed, if $f(t)$ is the area swept by the radial vector from time 0 to time t, then $\dot{f}(t) = L$ implies that the area of the ellipse $\pi a^2 \sqrt{1-\epsilon^2}$ is equal to LT. From $T = \pi a^2 \sqrt{1-\epsilon^2}/L$, we get

$$\boxed{T^2/a^3 = \pi^2 a(1-\epsilon^2)/L^2 = \pi^2/G}$$

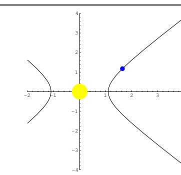
The third Kepler law allows to determine the gravitational constant G from the period and the geometry of the ellipse.

EXAMPLE. A Mars year is 1.88 earth years. How much longer is the length of the major semiaxes of the Mars orbit than the semiaxes of the earth orbit?
Answer: we know $T_{mars}^2/r_{mars}^3 = T_{earth}^2/r_{earth}^3$ so that $r_{mars} = r_{earth}(T_{mars}/T_{earth})^{2/3} = r_{earth} 1.88^{2/3} = 1.523....$ Mars is about one and a half times further away from the sun then the earth.

REMARKS

- To derive the first Kepler law starting with the ellipse is easier than taking off from the differential equations. The later approach is possible but the steps are harder to motivate.
- All Kepler laws crucially depend on the conservation of L.

CULOMB CASE. The case $\epsilon > 1$ corresponds to a negative G, where particles repel each other. The third Kepler law does then no more apply and the curve "ellipse" will be a "hyperbola" in the first law. The second law is unchanged. In this **Coulomb** case of the n-body problem, the total energy is always positive.



OTHER POTENTIALS.

If the interaction potential can be changed to $\ddot{x} = -Gx/r^{\alpha}$, where α is an integer. We have seen the case $\alpha = 3$. For other α, the first Kepler law still applies. Formula $\dot{\theta} = L/(mr^2)$ still applies. Also the derivation of the formula for $\ddot{x} \cdot n = \ddot{r} - L^2/(m^2r^3)$ is still valid. The left hand side is $-G/r^{\alpha-1}$ which leads to the ordinary differential equation

$$\boxed{\ddot{r} = -G/r^{\alpha-1} + L^2/(m^2r^3)} \qquad (*)$$

for $r(t)$. Knowing $r(t)$ gives then $\theta(t)$ from $\dot{\theta} = L/(mr^2)$. The global behavior depends on the constants G, L. The case $\alpha = 4$ corresponds to the natural Newton interaction in 4 dimensions.
You show in the homework:

In four dimensional space, planetary motion is unstable.

$\alpha = 3$ is the Kepler case with elliptic stable motion.
The case $\alpha = 2$ can physically be realized two massive parallel lines. (The general evolution of two rigid attracting lines in three dimensions is more complicated and form a special case of an interaction of two tops.)

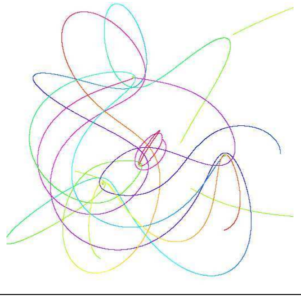
The case $\alpha = 1$ can be realized by the motion of two massive parallel planes. Such planes attract each other with constant force independent of the distance. The equation of motion $\ddot{x} = -G\text{sign}(x)x/|x|$. The three body problem in this case is already interesting. In the case $\alpha = 0$, each coordinate moves according to the harmonic oscillator.

A theorem of Bertant states that only for $\alpha = 3$ (the Kepler case) and $\alpha = 0$ (the harmonic oscillator), all bounded orbits are periodic.

MORE REMARKS.

- To derive the first Kepler law starting with the ellipse is easier than taking off from the differential equations. The later approach is possible but the steps are harder to motivate.
- All Kepler laws crucially depend on the conservation of L.
- In $d \geq 2$-dimensions, one would take the potential $U(x) = \sum_{i<j} \frac{m_i m_j}{|x_i - x_j|^{d-2}}$. In $d = 2$, the natural potential is $U(x) = G\sum_{i<j} m_i m_j \log|x_i - x_j|$.
- A natural regularisation of the singular potential is obtained by replacing the force by $G \cdot (|x|^2 + \epsilon)^{-d/2}$. In that case, one does not have to exclude the collision set Δ.
- The phase space of the system is called with the fancy name **cotangent bundle** of M. Such terminology is not necessary when we deal with particles moving in the open region M of an Euclidean space. However, if one would describe Newtonian particles on surfaces like the sphere or tori, where the interaction potential had to be modified, then the fancier notation is justified. We could for example look at the natural n body problem on a torus or the sphere.
- One would need $(6n - 1)$ integrals of motion to solve the n-body problem explicitly. The 10 classical integrals are not enough to find explicit solutions if $n > 2$. The first mathematical proof of this fact was given by Poincaré in a special case of the three body problem using new qualitative methods.

THE THREE BODY PROBLEM. With 3 or more bodies, the problem becomes chaotic. On the right hand side, you see an orbit computed with the n-body solver "xstar". We will look at the restricted three body problem later in more detail and see in a special situation, the Sitnikov case, that chaos can occur.

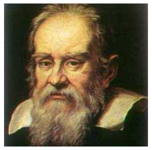


SOME HISTORY.

Aristoteles (384-322 BC) First model of solar system: planets as well the sun move around earth on perfect circles.

Claudius Ptolemaeus (78-150 AC) extended Hipparchus's system of epicycles to explain geocentric theory. Introduced 80 epicycles to explain the motions of sun, moon and 5 planets.

Galileo Galei (1564-1642) discovers Jupiter moons, suns pots etc. Famous for his fight for a Copernican theory with the inquisition. Mathematical work on moments and center of gravity.



Johannes Kepler (1571-1630) builds on the observations of Tycho Brahe. He finds the first and second Kepler law in 1609, the third in 1619.

Joseph-Louis Lagrange (1736-1813) Worked on the 3-body problem, the motion of the moon, and perturbations of comet orbits by the planets as well as the stability of the solar system.

Pierre-Simon Laplace (1749-1827) Investigated the inclination of planetary orbits, studied of planets were perturbed by their moons and the stability of the solar system.

Jean Le Rond d'Alembert (1717-1783) Improved Newton's definition of force in his Trait de dynamique published in 1743. This also contains d'Alembert's principle of mechanics.

George Birkhoff (1884-1944) Tools from probability theory statistical mechanics lead to ergodic theory. An example is Birkhoffs ergodic theorem. Poincare-Birkhoff fixed point theorem.

Jürgen Moser (1928-1999) The "M" in KAM theory. Book with Siegel in Celestial mechanics. Mosers contribution to KAM is the twist map theorem. Worked also on integrable n-body problems.

Hipparchus (190-120 BC) had a moon theory built on epicycles. Still an earth centered system.

Nicolas Copernicus (1473-1543) introduced a heliocentric system as well as secondary epicycles. This is a first step towards perturbation theory (which later would be seen as the Fourier approximation of real motion).

Tycho Brahe (1546-1601) revolutionized astronomy with new instruments and observations. For practical reasons, he used both heliocentric and earth centric coordinate systems.

Isaac Newton (1643-1727) Put celestial mechanics on a solid mathematical foundation and developed calculus simultaneously with Leibniz. Derivation of Keplers laws from basic principles.

Leonard Euler (1707-1783) wrote a 775 page work on the motion of the moon. He won several prizes from the Paris Académie des Sciences in the area of celestial mechanics.

Siméon Denis Poisson (1781-1840) who had Laplace and Lagrange as teachers published in 1808 work on the perturbations of the planets. He used series expansions to derive approximations.

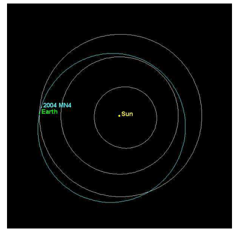
With **Henry Poincaré** (1854-1912) at the end of the 19'th century, the n-body problem was studied with new geometric and topological methods.

Andrey Kolmogorov (1903-1987) The beginning of KAM-theory, which is named after Kolmogorov, Arnold and Moser. Kolmogorov also put probability theory on a solid foundation and worked on a theory of turbulence.

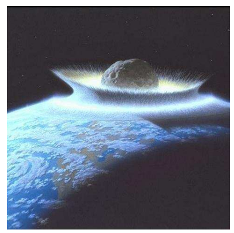
Vladimir Arnold (1937-) Progress on stability questions with perturbative methods (KAM). The concept of **Arnold diffusion** demonstrates a mechanism for instability.

ABSTRACT. The Newtonian 3 body problem can exhibit chaos. The simplest situation is when the third body moves in the time dependent potential of a binary system but itself does not influence the motion of the binary system. A first example is the **Sitnikov problem**, where one can establish the existence of a **horse shoe** which leads to a in general **chaotic calendar** for inhabitants of the Sitnikov planet. An other example is the circular planar restricted three body problem which leads to cases, where one has an area preserving map on a region with finite area. It is also a historically important example because some results in ergodic theory like Poincare recurrence and topology like fixed point theorems were developed with the three body problem in mind.

RESTRICTED THREE BODY PROBLEMS. The **restricted 3-body problem** deals with the situation, where one of the three bodies has a neglectable mass, and moves under the influence of the two other bodies which evolve according to Keplers law. Lets call here the two heavy bodies the **double star binary system** and the third body the **planet**.



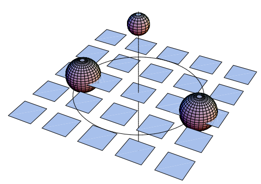
ASTEROID 2004 MN4 IMPACT RISK? In December 2004, Asteroid 2004 MN4 was given a 1/233 chance, then a 1/38 chance to hit the earth in April 13, 2029. Despite numerological support for bad luck like 2+0+2+9=13 and 1+3=4=shi also means "death" in Japanese, subsequent observations have shown that there will be no impact in 2029. It will pass by the Earth at a distance of between 15'000 and 25'000 miles, about a tenth of the distance between the Earth and the Moon and be so close that it can e seen with the naked eye. The change of orbit might put 2004 of a collision course in 2034, 2035 or 2036. One will know more in 2029.



SITNIKOV PROBLEM. The **Sitnikov problem** deals with the situation, where the double star system moves in the xy-plane and the planet is on the z-axes. Both stars have equal mass m normalized to $m = 1/2$ and move on elliptic orbits, where the center of mass is at rest. The third body has no mass. Its z coordinate satisfies the **Sitnikov differential equation**

$$\frac{d^2 z}{dt^2} = -\frac{z}{(z^2 + r(t)^2)^{3/2}} ,$$

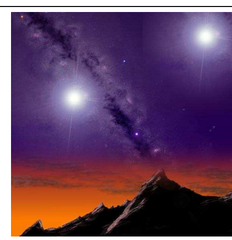
where $r(t)$ is the distance of a sun to the origin at time t. By normalizing time, we can assume that $r(t)$ has period 2π. For small values of the eccentricity ϵ of the ellipse, one has $r(t) = \frac{1}{2}(1 - \epsilon \cos(t)) + O(\epsilon^2)$.



SITNIKOV YEAR. A **Sitnikov year** is the time it takes to return to the xy-plane, the summer position on Sitnikov planet. Winter is when the planet has the maximal distance to the stars. The inhabitants on "Sitnikov" know to measure time and count the number of **Sitnikov days** in one Sitnikov year k by

$$s_k = [(t_{k+1} - tk)/2\pi] .$$

Far away from the double star system, a winter day could look as in the picture to the right.



A CHAOTIC CALENDAR.

THEOREM (Sitnikov-Moser) For sufficiently small eccentricity $\epsilon > 0$, there exists an integer m such that for any sequence $s_1, s_2, ...$ of of integers $s_k \geq m$, there exists a solution of the Sitnikov differential equation for which year k has s_k days.

REMARKS. One can also allow $s_k = \infty$ in which case, the planet would escape for ever, or the solar binary system could capture an orbit which stays bounded for ever. The proof of the theorem relies on the horse shoe construction and is robust. The result therefore holds also for planets with small positive mass. The result can be shown to be true for all $0 < \epsilon < 1$ except a discrete set of values.

Most orbits in this dynamical system go to infinity. It is not quite clear what the **filled in Julia set** is, the points which stay bounded for all times. Sitnikov-Moser theorem constructs a Cantor set of points which stay bounded for ever. It is not excluded that there are some stable elliptic periodic points. Numerical experiments suggest that such stable periodic points exist but I have not seen a proof. The stability problem is in nature similar to the one for the quadratic Henon map in the plane and depends on subtle Diophantine properties which have to be satisfied for the periodic points. We expect for most parameter values ϵ a set of positive area stays bounded. This could be good news for Sitnikov inhabitants.

The bad news is that these regions might be very small and a small disturbance - for example by an asteroid - could free the Sitnikov planet and send its inhabitants to a deadly eternal winter ride. One of the last pictures taken from that escaping planet could look as the picture above.

TO THE PROOF (Moser 1973).

Look at the **Poincare return map** to the plane with polar coordinates $(r, \phi) = (|v|, t)$, where v is the velocity of the planet and $t \; mod \; 2\pi$ is the time given by the suns clock. $t = 0$ corresponds to the moments, when the suns are closest to the z axes. The return map is defined in a simple closed region D_0. Outside this region, the orbit escapes. Here is an outline of the proof. The details are quite technical and can be found in Mosers book.

(0) The return map T_ϵ maps D_0 into $D_1 = \rho(D_0)$, where ρ is the reflection $(v, t) \to (v, -t)$. The map T_ϵ is area preserving: the area element $2v dv dt = dE \; dt$ is preserved.

(i) For small enough ϵ, the boundaries of D_0 and D_1 are smooth curves which intersect transversely. The proof of this fact is done by writing the right hand side of the Sitnikov equations as a power series in ϵ and neglecting ϵ^2 and larger terms. This computation from perturbation theory allows to establish that the angle between the boundary curves becomes nonzero.

(ii) For $\epsilon = 0$, the map T_0 is integrable and of the form

$$T_0 \begin{bmatrix} v \\ t \end{bmatrix} = \begin{bmatrix} v \\ t + f(v) \end{bmatrix}$$

where $f(v) \to \infty$ if $v \to 2$. The differential equation is in this case

$$\ddot{z} = \frac{-z}{(z^2 + 1/4)^{3/2}} .$$

This is an integrable system: indeed, the energy

$$E = \frac{1}{2}\dot{z}^2 - \frac{1}{\sqrt{z^2 + 1/4}} \geq -2$$

is conserved and the map leaves its level curves of E invariant. The origin is a fixed point, each circle gets rotated and the rotation becomes faster and faster until the boundary $E = 0$ is reached. In physical terms, this means that if we start with a larger initial velocity, it takes longer to return.

(ii) There are horse shoes arbitrarily close to the boundary of D_0. This is a consequence of i and ii and will be explained in class. (needs a good picture)

PLANAR CIRCULAR THREE BODY PROBLEM. The **planar restricted 3-body problem** deals the situation, where one of the three bodies has neglectable mass, but moves under the influence of two other bodies which evolve along circles according to Keplers law. An example is the motion of the moon in the influence of the earth and sun. A second example is the motion of an asteroid under the influence of the sun and Jupiter, the second largest body in our solar system. An other example is the motion of a planet in a binary star system.

ROTATING COORDINATE SYSTEM. Assume $\vec{y} = R(\omega t)\vec{x}$, where $R(\alpha)$ is a rotation in the plane with angle α. We can write $R(\omega t) = e^{A\omega t}$, where $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

LEMMA. In the rotating coordinate system

$$\frac{d^2}{dt^2}\vec{y} = R\frac{d^2}{dt^2}\vec{x} + 2A\omega R\frac{d}{dt}\vec{x} - R\omega^2\vec{x}$$

one observes additionally to the **rotated forces** also a **centrifugal force** and a velocity dependent **Coriolis forces**.

PROOF. Differentiating twice the identity $y = Rx$ using $\dot{R} = \omega AR$ gives $\dot{y} = \dot{R}x + R\dot{x} = \omega AR\vec{x}$ and $\ddot{y} = \omega^2A^2R\vec{x} + 2AR\dot{x} + R\ddot{x}$. Because $A^2 = -1$, this gives the equation in the lemma. The same calculation in coordinates: $\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \end{bmatrix} = R\begin{bmatrix} \dot{x}_1 - \omega x_2 \\ \dot{x}_2 + \omega x_1 \end{bmatrix}$ and $\begin{bmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{bmatrix} = R\begin{bmatrix} \ddot{x}_1 - \omega^2 x_1 - 2\omega x_2 \\ \ddot{x}_2 - \omega^2 x_2 + 2\omega x_1 \end{bmatrix}$, where $R = \begin{bmatrix} \cos(\omega t) & -\sin(\omega t) \\ \sin(\omega t) & \cos(\omega t) \end{bmatrix}$. Remark. The same computation can be done in three dimensions, where both the centrifugal and Coriolis forces can be expressed using cross products.

THE EQUATIONS OF THE PLANAR CIRCULAR 3-BODY PROBLEM. Two stars of mass $m_1 = \mu, m_2 = 1-\mu$ move on circular orbits along their center of mass. Going into a rotating **inertial coordinate system** (Keplers 3. law implies from zero eccentricity uniform rotation), in which the stars are fixed at the points $(1-\mu, 0), (-\mu, 0)$, the equations of motion become

$$\frac{d}{dt}x_k = E_{y_k}, \frac{d}{dt}y_k = -E_{x_k} ,$$

where $E = \frac{1}{2}(\dot{y}_1^2 + \dot{y}_2^2) + 2x_2y_1 - 2x_1y_2 - \frac{\mu}{r_1} - \frac{(1-\mu)}{r_2}$ is the Hamilton function. Here $r = \sqrt{x_1^2 + x_2^2}$ is the distance of the planet to the origin, $r_1 = \sqrt{(x_1 + \mu - 1)^2 + x_2^2}$ and $r_2 = \sqrt{(x_1 - \mu)^2 + x_2^2}$ are the distances from the planet to the two stars. We can decompose $E = (\dot{x}_1^2 + \dot{x}_2^2)/2 - U(x_1, x_2)$ with $U = \frac{1}{2}r^2 + \frac{\mu}{r_1} + \frac{(1-\mu)}{r_2}$. The function E is called the **Jacobi integral**. It contains $\frac{1}{2}r^2$ called **centrifugal potential** and $\dot{x}_1^2 + \dot{x}_2^2$, the **Coriolis potential**. How did we get that? The Newton equations in the rotating coordinate system are according to the previous lemma:

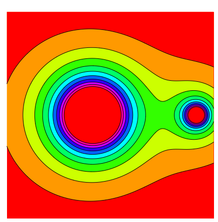
$$\begin{aligned} \ddot{x}_1 - 2\dot{x}_2 &= \tfrac{\partial}{\partial x_1}U \\ \ddot{x}_2 + 2\dot{x}_1 &= \tfrac{\partial}{\partial x_2}U \end{aligned}$$

After multiplying the first equation with $\dot{x}_1$ and the second with $\dot{x}_2$, addition gives $\dot{x}_1\ddot{x}_1 + \dot{x}_2\ddot{x}_2 = \frac{\partial}{\partial x_1}U\dot{x}_1 + \frac{\partial}{\partial x_2}U = \dot{U}$ so that $E = (\dot{x}_1^2 + \dot{x}_2^2)/2 - U$ is conserved. Introducing $y_1 = \dot{x}_1 - x_2, y_2 = x_1 + \dot{x}_2$ leads to the Hamilton equations at the top of this box.

What is the deal? We started with the Newton equations $\ddot{y}_i = \frac{\partial}{\partial x_1}W$ and ended up with a system looking more complicated. But it is not! In the original coordinates, the potential W is time dependent! Especially, there was no energy conservation. Going into the rotating coordinate system led us to a Hamiltonian system with a preserved quantity, the Jacobi integral.

HILLS REGION. Assume $E = c_1$ and $c < c_1$. The regions $U(x_1, x_2) = c$ bound regions in the (x_1, x_2) plane called **Hills regions**.

LEMMA. If (x_1, x_2) is in a Hills region $U \geq c$, then $(x_1(t), x_2(t)$ is in the Hills region for all times.



For large c, these regions consist of three parts. Two in the neighborhood of the two stars (satellite bound by one of the bodies) and one far away (asteriod encircling both). They define an allowed region in which the planet can stay. A large c corresponds to the case, where one is either close to one of the stars with large gravitational potential or very far away, with large centrifugal potential.

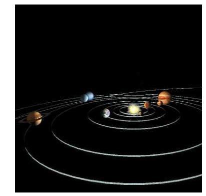
RECURRENCE. The energy surfaces $E = c$ are invariant as are the sets $\{(x_1, x_2) \mid a \leq -E(x_1, x_2) \leq b\}$ for $a < b$. If $c < c_1$, then

$$G = \frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2) - E > c_1 > c .$$

So, (x_1, x_2) stays in a bounded region. Also (x_1, x_2, y_1, y_2) stays in a bounded set. The differential equation preserves the four dimensional volume. When normalizing the volume to 1, we obtain a probability space. The time 1 map is a measure preserving map on that space and Poincares recurrence theorem applies.
There is a subtlety with this argument which has to be mentioned: Not all solutions in the finite region have a global solution. There are initial conditions, in which the planet crashes into one of the suns but these cases can be shown to have zero volume.

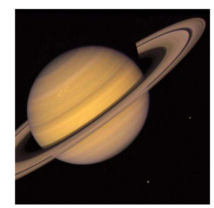
CHAOS IN THE SOLAR SYSTEM. Chaos in the solar system has been measured at different places:

1) The solar system itself is weakly chaotic. The Lyapunov exponent has been measured to be very small $2.8 \cdot 10^{-15}$. For Pluto the Lyapunov exponent had been measured $7 \cdot 10^{-16}$. Numerical experiments have also been done with other parameters. The heliocentric distance for outer planets would behave much more erratically, if the sun would have 1/3 less of its current mass, suggesting that some of the outer planets like Neptune or Uranus would escape in such a case. For our solar system, it looks as if one can not predict the trajectory of the earth for time periods exceeding 100 Million years. More precisely, the uncertainty of 1 km in the initial condition could lead to an uncertainty of the order of 1 astronomical unit in 100 Million years. Numerical simulations of the solar system have been done for time intervals reaching 35 billion years.



2) Many **comets and asteroids** in the solar system have irregular orbits. Numerical experiments have been done for example in the case of the asteroid **Chiron**. To measure sensitive dependence on initial conditions, one starts integrating with various close initial conditions and looks at the outcome. Chiron will undergo several close approaches to planets. One estimates a 1/8 chance that Chiron will eventually leave the solar system. Other objects have an other fate. The comet **Shoemaker-Levy 9** had a spectacular impact with Jupiter in July 1994 after having been disrupted by a close Jupiter approach in 1992.

3) The tumbling of Saturns little moon **Hyperion**. Most satellites in the solar system are in synchronous rotation, keeping one face towards the planet. Hyperion has an irregular shape and is known to tumble erratically in its orbit. The **Cassini spacecraft** will fly past this moon later this year, on September 26, 2005. The Lyapunov exponent of the irregular tumbling motion has been been measured to be of the order 10^{-7}.



4) The motion of charged particles in a magnetic dipole field has been shown to be chaotic. Brown has constructed a horse shoe for the return map. The dynamics can be reduced to a relatively simple Hamiltonian system

$$H(q_1, q_2, p_1, p_2) = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}(\frac{1}{q_1} - \frac{q_1}{(q_1^2 + q_2^2)^{3/2}})^2$$

called the **Stoermer problem**. The dynamics of charged particles in the **van Allen belts** can explain the **aurora Borealis**.

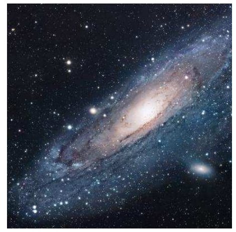
For the Lyapunov exponent data on this box, we the sources:
P. Gaspard: "Chaos Scattering and Statistical mechanics", 1998
I. Peterson: "Newtons Clock: Chaos in the solar system", 1993
C.D. Murray and S.F. Dermott: Solar system dynamics", 2001
D. Goroff: Editorial introduction article in "New Methods of Celestial Mechanics by H. Poincare".
K. Zyczkowski "On the stability of the Solar system".
For the planar 3 body problem, we followed Siegel-Moser. Sitnikovs problem is treated in detail in Mosers 1973 book.

SINGULARITIES **Math118, O. Knill**

ABSTRACT. Singularities for the n-body problem can occur when bodies collide or when bodies escape to infinity. A theorem of van Zeipel shows that these are the two onlyi possibilities.

OPEN PROBLEM. Lets start with a major open problem in celestial mechanics.

Is it true that the Newtonian n-body problem has a full measure set of initial conditions, for which the solutions exist for all times?



COLLISIONS.

If $x(t) \to \Delta$ for $t \to \tau$, then $x(\tau)$ is called a **collision singularity**. Collisions can already occur in the 2-body problem, if the total angular momentum of the two bodies is zero. Analysing collision singularities involving more than two bodies helps to understand what happens when particles move close to such collision configurations. It is known that initial conditions leading to collisions are rare in the n-body problem. Noncollision singularities in which particles escape to infinity in finite time exist already for the 5-body problem.

Our galaxy and M31, the Andromeda galaxy, form a relatively isolated system known as the **local group**. The center of mass of M31 approaches the center of mass our galaxy with a velocity of 119 km/s. In about 10^{10} years, these galaxies are likely to collide. Such a collision would have dramatic consequences for both systems. Nevertheless, even a direct encounter would probably not lead to any collision of stars.

EXISTENCE OF SOLUTIONS.

For every point (x, y) in phase space, there exists $\tau = \tau(x, y)$ such that for $t \in [0, \tau(x)]$ the Newtonian n-body equations have a unique solution (x^t, y^t). Moreover, if K is a closed and bounded subset in the phase space, then there exists $\delta > 0$ such that (x^t, y^t) is outside K for $t \in [\tau - \delta, \tau]$.

(i) The first statement follows from a general existence theorem for differential equation $\dot{x} = f(x)$ on a subset M of Euclidean space. The function $\dot{x} = f(x)$ is Lipshitz continous on a bounded open set in M.

(ii) For any compact (closed and bounded) set K, there is a time $\tau_K = \min_{x \in K} \tau(x) > 0$ such that for all initial conditions $x \in K$, a solution exists in the time interval $[0, \tau_K)$.
Therefore, if $x(t)$ exists in the interval $[0, \tau)$ and the solution can not be extended beyond τ, then for $t \in (\tau - \tau_K, \tau]$, $x(t)$ is outside K.

SINGULARITIES. A point $(x, y) \in (T^*M)^n$ is called a **singularity** if $\tau(x, y) < \infty$. A singularity is called a **collision** if there exists $x \in \Delta$ such that $x^t \to x$. A singularity which is not a collision is a **pseudo collision** or a **non collision singularity**.

The existence theorem shows that if a singularity is approached, then the some velocities become unbounded. It is not possible that positions become unbounded but velocities stay bounded.

PAINLEVE THEOREM. If (x, y) is a singularity, then $|U(x^t)| \to \infty$ for $t \to \tau(x, y)$. In other words, the minimal distance between two particles goes to zero. This result holds in any dimensions and for any potential $U = u(|x|)$ satisfying $u(r) \to \infty$ for $r \to 0$ and such that $u \in C^2([\epsilon, \infty))$ for every $\epsilon > 0$.

PROOF. Assume the contrary: there exists $\delta > 0$ such that $\min_{i \neq j} |x_i^t - x_j^t| \geq \delta$ for $t \in [0, \tau)$. We want to show that τ is not maximal.

(i) The differential equation $\dot{x} = f(x)$ with $|f| \leq M$ in $B_r(x_0)$ and $f \in C^1$ has a solution x^t with $x^0 = x_0$, as long as $|t| \leq r/M$. The piece of orbit $\{x^t\}_{t \in [0, r/M]}$ is contained in $B_r(x_0)$.

Proof. See the proof of the Cauchy-Piccard existence theorem.

(ii) There exists M such that $|\nabla_x U| \leq M$ for $x \notin B_r(x^0)$.

Proof. We have $0 \leq -U \leq C/\rho$, where C is a constant depending only on n and the masses m_j. Therefore, we have $|\nabla_x U| \leq C/\rho^2$.

(iii) There exists M such that $|y_j| \leq M$.

Proof. This follows from the decomposition of the energy $H = K + U$ and the boundedness of U. $\sum_{j=1}^d y_j^2 / 2m_j \leq H + 2M^2/d$.

(iv) For t arbitrarily close to $\tau(x, y)$, we can extend the solution for the time interval $[0, r/2M]$.
Proof. Using $(ii), (iii)$, we can apply (i).

MOMENT OF INERTIA. The number $I(x) = \sum_{i=1}^n m_i |x_i|^2$ the **moment of inertia** of the configuration.

LAGRANGE-JACOBI FORMULA. $\frac{1}{2}\ddot{I}(x^t) = U(x^t) + 2H(x^t, y^t) = T(y^t) + H(x^t, y^t)$, where $H(x, y) = T(y) - U(x)$ is decomposition of the energy into kinetic and potential energy.

PROOF. From $\frac{1}{2}\dot{I} = \sum_{j=1}^d m_j (x_j, \dot{x}_j)$, we get

$$\begin{aligned} \frac{1}{2}\ddot{I} &= \sum_{j=1}^d m_j (x_j, \ddot{x}_j) + 2T = \sum_{j=1}^d (x_j, -\nabla_{x_j} U(x)) + 2T \\ &= U + 2T = -U + 2H = T + H \; . \end{aligned}$$

We have used that U is homogeneous of degree -1: $U(\lambda x) = \lambda^{-1} U(x)$ which gives with the Euler identity $(x, \nabla_x U) = -U$.

REMARK TO 4D. Interesting is the analoguous case in $n = 4$, where U is homogeneous of degree -2. Then $\frac{1}{2}\ddot{I} = 2H$ is constant. This shows that we have in the case of a negative initial energy $H < 0$ always collapse in finite time and that solutions can stay bounded only on the energy surface $H = 0$. You have this fact in the case of the Kepler problem in four dimension.

SUNDMAN-VAN ZEIPEL LEMMA If (x, y) is a singularity, there exists $I^* = I(x^{\tau(x,y)}) \in [0, \infty]$ such that $I(x^t) \to I^*$ for $t \to \tau(x, y)$. The same relation holds for potentials for which $x \cdot \nabla_x U(x) + U(x)$ is globally bounded.

PROOF. From the Lagrange formula and the theorem of Painlevé, we see that $\ddot{I} > 0$ for t near $\tau(x, y)$. This implies that $\dot{I}$ is monotonically increasing and one can assume that $\dot{I}$ is always positive or always negative in the interval $[t, \tau]$ because one could else, if it changes sign, make the interval smaller. The positive function I is therefore monotonic and has a limit.

VAN ZEIPEL's THEOREM. This is a heavy theorem. Even so the proof had been simplified considerably by McGehee, its not possible to hide that this is a relatively deep result:

THEOREM. If (x, y) is a singularity, then $I(x^{\tau(x,y)}) < \infty$ if and only if (x, y) is a collision. In other words, $I(x^{\tau(x,y)}) = \infty$ if and only if (x, y) is a pseudo-collision.

The proof follows closely McGehee's 1986 paper.

PROOF (i) Clusters. Denote by ω a partition of the set $N = \{1, \dots, n\}$. For $\mu \subset N$, define $\Delta_\mu = \{x \in \mathbf{R}^{3n} \mid i,j \in \mu \Rightarrow x_i = x_j\}$ and $\Delta_\omega = \bigcap_{\eta \in \omega} \Delta_\mu$.

PROOF (ii) New scalar product. Consider the scalar product in $\mathbf{R}^{3n}$ by $< x, x' >= \sum_j m_j(x_j, x'_j)$, where $(\cdot,\cdot)$ is the standard scalar product in $\mathbf{R}^3$. The norm $||x||$ of x in this scalar product allows to rewrite the moment of inertia as $I(x) = ||x||^2$.

PROOF (iii) Orthogonal decomposition. Define for $\mu \subset N$ the linear map $\mathbf{R}^{3n} \to \mathbf{R}^3$

$$x \mapsto \pi_\omega x = c_\mu x = \sum_{i \in \mu} m_i x_i / \sum_{i \in \mu} m_i$$

and the linear map π_ω from $\mathbf{R}^{3n}$ to $\mathbf{R}^{3n}$. We have $(\pi_\omega x)_i = \pi_\mu x$ if $i \in \mu$. This is an orthogonal projection with range Δ_ω and kernal $\Gamma_\omega = \{\sum_{j \in \mu} m_j x_j = 0 \; \forall \mu \in \omega\}$. Denote by $\Pi_\omega = Id - \pi_\omega$ the orthogonal projection onto Γ_ω. Write $x = \pi_\omega x + \Pi_\omega(x) = z + w$.

PROOF (iv) Moment of inertia. Define $I_\omega(x) = ||\pi_\omega x||^2 = \sum_{\mu \in \omega} (\sum_{j \in \mu} m_j) |c_\mu x|^2$. Denote by J_μ the moment of inertia of a subsystem having particles $j \in \mu$ and by $J_\omega = \sum_{\mu \in \omega} J_\mu$ the sum of these moment of inertias. The equation

$$||x||^2 = ||\pi_\omega x||^2 + ||\Pi_\omega x||^2 = I_\omega(x) + J_\omega(x)$$

means that the total moment of inertia is the sum of the moment of inertias of the subsystems and the fictious system obtained from the center of masses of the subsystems.

RROOF (v) Potential energy. Define $U_{ij}(x) = \frac{1}{2} \frac{m_i m_j}{|x_i - x_j|}$ for $i \neq j$ and $U_{ij}(x) = 0$ if $i = j$. Let $V_\mu(x) = \sum_{i,j \in \mu} U_{ij}$ be the potential energy of the subsystem μ and $V_\omega(x) = \sum_{\mu \in \omega} V_\eta(x)$ the sum of the potential energies of the subsystems of a partition ω. Define $U_{\mu\nu}(x) = \sum_{i \in \mu, j \in \nu} U_{ij}(x)$ if $\mu \cap \nu = \emptyset$ and $U_{\mu\nu}(x) = 0$ else. The potential energy due to the interaction of the subsystems is $U_\omega = \sum_{\mu,\nu \in \omega} U_{\mu\nu}$. The total potential energy $U(x)$ can be written as

$$U(x) = U_\omega(x) + V_\omega(x) \; .$$

PROOF (vi) Dynamics. For $z \in \Delta_\mu$, we have $V_\omega(x + z) = V_\omega(x)$ which gives $V_\omega(x + \pi_\omega y) = V_\omega(x)$ for all $y \in \mathbf{R}^{2n}$. Differentiation of this with respect to y and putting $y = 0$ gives $\nabla V_\omega(x)\pi_\omega = 0$. Because π_ω is orthogonal, we have therefore $\pi_\omega \nabla V_\omega(x) = 0$. Applying the projection π_ω on $\ddot{x} = \nabla U(x) = \nabla U_\omega + \nabla V_\omega$ gives

$$\ddot{w} = \pi_\omega \ddot{x} = \pi_\omega \nabla U_\omega \; ,$$

from which we derive

$$\frac{d^2}{dt^2} I_\omega(x) = \frac{d^2}{dt^2} < \pi_\omega x, \pi_\omega x >= 2 < \pi_\omega \dot{x}, \pi_\omega \dot{x} > +2 < \pi_\omega x, \pi_\omega \nabla U_\omega(x) > \; .$$

PROOF (vii) Statement of the goal: We assume that $I(x^t) \to I^* < \infty$ and show that x^t converges.

RROOF (viii) The collision set Δ^*. By assumption on the theorem, the set

$$\Delta^* = \bigcap_{t < \tau(x,y)} \overline{O(t, t^*)} \subset \Delta$$

with $O(a,b) = \{x^t\}_{t \in (a,b)}$ is nonempty and compact. For each partition ω define $\Delta^*_\omega = \Delta^* \cap \Delta_\omega$. From the partitions ω with Δ^*_ω we choose a partition with minimal cardinality and fix this partition for the rest of the proof.

PROOF (ix) Bound the force in a neighborhood G of Δ^*. Since Δ^* is compact we can find an open neighborhood G of Δ^* and a constant M such that

$$||\nabla U_\omega||, | < \pi_\omega x, \nabla U_\omega(x) > | \leq M \; .$$

PROOF (x) If Δ^* is a subset of Δ_ω, then x^t converges.
If $\Delta^* \subset \Delta_\omega$, then $z^t = \pi_\omega x^t$ converges for $t \in \tau(x,y)$. There exists t_2 such that $x^t \in G$ for $t \in (t_2, \tau(x,y))$. From $\ddot{w} = \pi_\omega \nabla U_\omega(x)$ and the bound in (ix), we get $||\ddot{w}|| \leq M$ for $t \in (t_2, \tau(x,y))$. It follows that w^t approaches a limit w^* for $t \to \tau(x,y)$. Hence $x^t = w^t + z^t \to w^* + 0$ converges.

PROOF (xi) The situation that Δ^* is a not a subset of Δ_ω is not possible. Assume Δ^* is a not a subset of Δ_ω. In claim (ix) below, we will derive a contradiction and so finish the proof of the theorem.

PROOF (xii) Definition of a compact set $K_\sigma \subset \mathbf{R}^{3n}$. Choose a bounded open subset B of Δ_ω such that $\Delta^*_\omega \subset B \subset \overline{B} \subset \Delta_\omega \subset G$. Let D_σ denote an open ball of radius σ in the linear space Γ_ω. Define the compact set

$$K_\sigma = \overline{B} \times \overline{D}_\sigma \; .$$

Since the boundary δB of B is compact and B does not intersect Δ^*, there exists σ_0 and $t_0 < \tau$ such that $O([t_0, \tau)) \cap \overline{D}_{\sigma_0} \times \delta B = \emptyset$. We can choose σ_0 so small that additionally $K_{\sigma_0} \subset G$.
Since by our assumption, Δ^* is not a subset of Δ_ω, there exists $0 < \sigma < \sigma_0$ such that for infinitely many values of t close to t^*, we have $x^t \notin K_\sigma$. Choose and fix σ with this property.

PROOF (xiii) Definition of a time t_1. Chose t_1 so small that

$$|I(x^t) - I^*| \leq \frac{\sigma^2}{12}, \; \forall t_1 \leq t < t^* \; .$$

PROOF (xiv) Definition of a time interval $I = [a,b]$ with some properties. There exists an interval $I = [a,b]$ such that
1) $O(I) \subset K_\sigma$ 2) $||\Pi_\omega x^a|| = ||\Pi_\omega x^b|| = \sigma^2$
3) $\min_{t \in [a,b]} ||\Pi_\omega x^t|| < \sigma^2/2$ 4) $b - a < \sigma/\sqrt{3M}$.
Proof. Because x^t comes arbitrarily often arbitrarily close to Δ^*_ω, x^t must enter and leave K_σ infinitely many often. 1) is therefore no problem for intervals arbitrarily close to τ. 3) can be met for intervals arbitrarily close to τ because x^t comes arbitrarily often arbitrarily close to Δ^*_ω, where $||\Pi_\omega x^t|| = 0$. 2) is clear because if x^t enters K_σ it can not enter through $\overline{D}_\sigma \times \delta B$ and must therefore enter through $\delta \overline{D}_\sigma \times B$.

PROOF (xv) Let $s \in (a,b)$ be such that $I_\omega(x^t)$ is maximal. Remember that $I(x^t) = I_\omega(x^t) + J_\omega(x^t)$ converges for $t \to \tau$ so that a maximum exists from (vii) and (vi).

PROOF (xvi) Derive a contradiction. From $\min_{t \in [a,b]} ||\Pi_\omega x^t|| < \frac{\sigma^2}{2}$ and $|I(x^t) - I^*| < \sigma^2/12, t_1 \leq t < t^*$ we obtain

$$I_\omega(x^s) > I^* - \frac{7\sigma^2}{12} \; .$$

From $||\Pi_\omega x^a|| = ||\Pi_\omega x^b|| = \sigma^2$ and $|I(x^t) - I^*| < \sigma^2/12, t_1 \leq t < t^*$ we obtain

$$I_\omega(x^b) < I^* - \frac{11\sigma^2}{12}$$

so that

$$I_\omega(x^s) - I_\omega(x^b) > \frac{\sigma^2}{3} \; .$$

On the other hand we have for $t \in [a,b]$

$$\frac{d^2}{dt^2} I_\omega(x) = 2 < \pi_\omega \dot{x}, \pi_\omega \dot{x} > +2 < \pi_\omega x, \pi_\omega \nabla U_\omega(x) > \geq 2 < \pi_\omega x, \pi_\omega \nabla U_\omega(x) > \geq 2M \; .$$

Because s is a local maxium of $I_\omega(x^t)$, we know that

$$I_\omega(x^s) - I_\omega(x^b) \leq M(b-s)^2 < \frac{\sigma^2}{3} \; ,$$

where the last inequality uses 4) from (ivx).

REMARK. Edvard Hugo von Zeipel (1873-1959) was a Swedish astronomer. Van Zeipel's result holds for every potential $U(x)$ for which one can prove a Sundman-Van Zeipel lemma. For the Newton potential in four dimensions, where $\ddot{I} = const$, we know trivially that I^* exists in $(0, \infty)$. It follows that for the graviatational Newton potential in four dimensions, there are only collision singularities. For negative energy they have full measure.

SHORTEST PATHS IN THE PLANE — Math118, O. Knill

ABSTRACT. The minimization of the arc-length while connecting two points in the plane has been studied by Archimedes already. It can also be solved, if the arc length is generalized. It leads to differential equations.

PLANE. Given two points P, Q in the plane. What is the path connecting P with Q which minimizes the length? While everybody knows that the straight line solves this problem, how does one prove this?

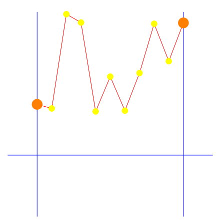
CONNECTING POINTS IN THE PLANE. Let $f(x)$ be a graph over the interval $[a,b]$ such that $P=(a,f(a))$ and $Q=(b,f(b))$. The length of this graph is

$$I(f) = \int_a^b \sqrt{1+f'(x)^2}\, dx \, .$$

Which function f minimizes that? We could look at paths connecting points (x_i,y_i) with $(x_0,y_0)=P$ and $(x_n,y_n)=Q$ using $f_i(x)=f(a)(x-x_i)+(x-x_i)(y_{i+1}-y_i)/(x_{i+1}-x_i)$, to connect neighboring points. The length of such a graph is by Pythagoras $I(y_1,...,y_{n-1})=\sum_{i=0}^{n-1}\sqrt{(x_{i+1}-x_i)^2+(y_{i+1}-y_i)^2}=\sum_{i=0}^{n-1} l_i$. To minimize this, the gradient of I must vanish. Because the partial derivative with respect to y_i is $(y_i-y_{i-1})/l_i-(y_{i+1}-y_i)/l_i=\sin(\alpha_i)-\sin(\alpha_{i+1})$, all the slopes of the polygonal graph must agree and the line has to be a straight line. We have verified

LEMMA. Among all polygonal graphs connecting P and Q, the straight line has minimal length.

One can also see by the triangle inequality that any corner in the graph can be shortened. A polygon which is not a straight line can be shortened by a definite amount. For any given differentiable function f, we can approximate the graph of f by piecewise linear graphs of g_n so that the length differences ϵ_n of the f and g graphs goes to zero. If there was a f for which the length were by an amount $\delta>0$ smaller than the length of the straight line, we could approximate that function f with a polygon g_n for which $\epsilon_n<\delta$ and have a polygon with smaller length contradicting the lemma. We have now shown:



THEOREM (Archimedes). Among all differentiable functions whose graph connects two points P and Q in the plane, the straight line minimizes the length.

Remark: this proof seems oblivious, since it can be shot down with mathematical cannon called "calculus of variations". Besides the fact that it is always nice to avoid heavy artillery, if not needed, the Archimedes proof has an advantage: it goes through also in a larger class of rectifiable functions which do not need to be differentiable like Snells refraction example below. The discretization approach also generalizes to inhomogeneous media, where it gives a numerical method. Remarkably, the proof does not need the notion of "derivative" at all, if one defines "rectifiable curves", as curves for which the lengths of the polygonal approximations converges and replaces $\nabla I=0$ with the triangle inequality.

INHOMOGENEOUS MEDIUM. Lets assume that we are in a medium, where it is difficult to travel at some places and hard at others. If we replace the length by the work

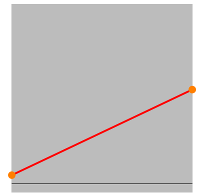
$$I_g(f)=\int_a^b g(x,f(x))\sqrt{1+f'(x)^2}\, dx$$

and again ask for the problem to find the most efficient path connecting two points P and Q, the result will critically depend on the function $g(x,y)$. There will be no more unique solutions. Lets discretize the problem again: we have to minimize $I(y_1,...,y_{n-1})=\sum_{i=0}^{n-1} g(x_i,y_i)l_i$. Setting the partial derivatives with respect to y_i equal to zero shows that $g_y(x_i,y_i)l_i+g(x_i,y_i)(y_{i+1}-y_i)/l_i=C$ is constant. This allows to compute recursively the slope

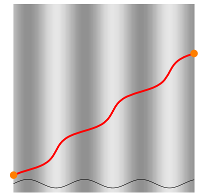
$$\sin(\alpha_i)=(C-g_y(x_i,y_i))/g(x_i,y_i)\, .$$

The constant C is obtained by the requirement that P and Q are connected.

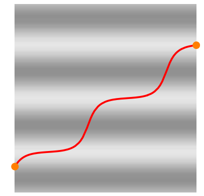
EXAMPLES. The following examples were obtained by numerically solving for the shortest path connecting two given points.



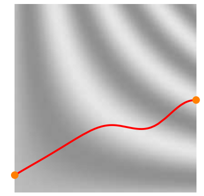
Flat medium. The shortest connection between two points is a line.



Rippled medium. The path prefers to stay in the bright regions, where traveling is easy.



The inhomogeneity is vertical. Again, the particle prefers to stay in the bright regions.



A more general optimization problem: again the path tries to avoid staying too long in the dark area.

EULER-LAGRANGE EQUATIONS. Let $F(t,x,p)$ be a function of three variables. We look at the **variational problem** to extremize

$$I(\gamma)=\int_a^b F(t,x(t),\dot{x}(t))\, dt$$

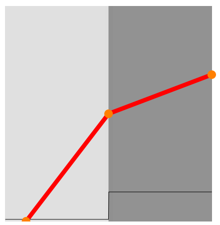
among all smooth paths γ connecting $x(a)$ with $x(b)$. If $t\mapsto h(t)$ is an other path, then $(I(\gamma+h)-I(\gamma))=D_hIh+O(h^2)$ for $h\to 0$ defines a "directional derivative" D_hI called here the **first variation**. By linearizing F, we know that $I(\gamma+h)-I(\gamma)=\int_a^b F_x(t,x,\dot{x})+F_{\dot{x}}(t,x,\dot{x})\, dth+O(h^2)=\int_a^b F_x(t,x,\dot{x})-\frac{d}{dt}F_{\dot{x}}(t,x,\dot{x})\, dth+O(h^2)$. The first variation is zero if $\boxed{F_x(t,x,p)=\frac{d}{dt}F_p(t,x,p)}$ for all t. These are the **Euler-Lagrange equations**.

INHOMOGENOUS PLANE. If $\gamma: t\mapsto (t,x(t))$ is a curve in the plane, we can look at $\int_a^b F(t,x,\dot{x})\, dt=\int_a^b \sqrt{1+\dot{x}(t)^2}\, dt$. The Euler equations show that $\dot{x}/\sqrt{1+\dot{x}(t)^2}$ is time independent. Therefore $\ddot{x}$ is constant and consequently, the optimal curve is a straight line. In the inhomogeneous case, the Euler-Lagrange equations for $F(t,x,\dot{x})=g(t)\sqrt{1+\dot{x}^2}$ are $0=\frac{d}{dt}(\frac{\dot{x}(t)g(t)}{\sqrt{1+\dot{x}(t)^2}})$. This proves

SNELLS THEOREM. $g(t)\dot{x}/\sqrt{1+\dot{x}^2}=g(t)\sin(\alpha(x))$ is constant, where $\alpha(x)$ is the angle the curve makes with the x axes.

SNELLS LAW. A limiting situation is when the medium has two densities like air and water. In this situation, the Euler-Lagrange equations do not help. But the Archimedes approach still works. If $g=u$ on the left hand side and $g=v$ on the right hand side, then $\sin(\alpha_i)=\sin(\alpha_{i+1})$ as before in the left or the right region and $u(y_i-y_{i-1})/l_i-v(y_{i+1}-y_i)/l_i=u\sin(\alpha_i)-v\sin(\alpha_{i+1}=0$ at the boundary. Therefore, the shortest path is a line with angle α on the left hand side and angle β on the right hand side and

$$\boxed{u\sin(\alpha)=v\sin(\beta).}$$



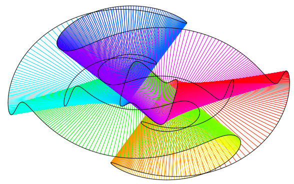
This is called **Snells law** named after **Willebrord Snel**, who had discovered this refraction law. Descartes and Fermats thought about this too. Their dispute about this is described in Nahins book "When least is best". For a more general density distribution Archimedes proof also gives that $g(t)\sin(\alpha(x))$ is constant. **Archimedes proof is more powerful: it leads to a result for nonsmooth $g(t)$.**

AN INITIAL VALUE PROBLEM. With the assumption that a particle moves without an influence of an external force and minimize the action, we are lead to a dynamical system. Start at a point P and a direction v. The extremization requirement leads to a **Newton law**, which is a differential equation of the form $\ddot{x}=f(x,\dot{x},t)$. One can actually derive all of Newtons law from a minimization principle. Extremization of action is one of the most important principles in physics: Newton equations, Maxwell equations, Einstein equations can be derived like this.

WAVE FRONTS AND CAUSTICS **Math118, O. Knill**

ABSTRACT. Wave fronts which start at a point evolve and break at caustics. Given a metric in Euclidean space, the wave fronts form a one-parameter family of piecewise smooth surfaces.

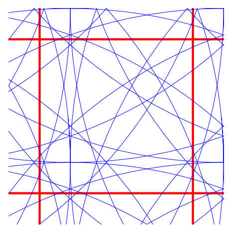
WAVE FRONTS AND CAUSTICS. The set of points reached at time t from a given point x form the **wave front** $K_x(t)$ of x. If the geodesics starts with an initial velocity $(\cos(\phi), \sin(\phi)$, it reaches at time t the point $K_x(t, \phi)$. A **conjugate point** of x is a point $K_x(t, \phi)$, for which $DK_x(t, \phi)$ has zero determinant. The set C_x of all conjugate points $K(t, \phi)$ form the **caustic** of x. The caustic of a curve $\phi \mapsto r(\phi)$ in the plane is defined as the set K_γ of points for which $DK_\gamma(t, \phi)$ has zero determinant, where $K_\gamma(t, \phi)$ is the point reached when we start at $r(\phi)$ in the normal direction $n(\phi)$. Given a closed compact surface and a point P. How does the wave front $K_(t)$ look like? Does it become dense on the surface?



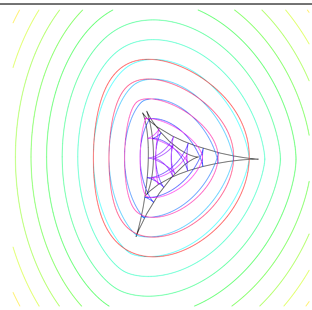
EXAMPLES.

FLAT TORUS. On the flat torus, the wave front $K_x(t)$ becomes dense on the surface for every point x. The caustic is empty. The picture to the right shows the wave front on the flat torus at time 3.

ROUND SPHERE. The wave front $K_x(t)$ is a circle or a point at all times. In the case of the flat torus, the caustic is empty, in the case of the sphere, the caustic C_x consists of two points, x and the antipole $S(x)$ of x.



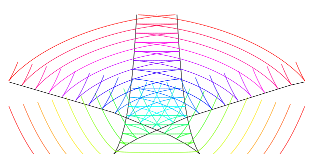
CAUSTIC FLAT CASE. Let $\gamma : r(\phi) = (x(t), y(t)$ be a curve in the flat plane and let $n(\phi) = (-y'(\phi), x'(\phi)$ be the normal vector to the curve and $\rho(\phi) = 1/||n(\phi)|| = 1/||r'||$. Then $K_\gamma(t, \phi) = r(\phi) + tn(\phi)\rho(\phi) = (x(\phi) - ty(\phi)\rho(\phi, y(\phi) + tx(\phi)\rho(\phi)$ so that $DK_\gamma(t, \phi) = \begin{bmatrix} n(\phi)\rho(\phi) & r'(\phi) + tn'(\phi)\rho(\phi) + tn(\phi)\rho'(\phi) \end{bmatrix} = \begin{bmatrix} n(\phi)\rho(\phi) & r'(\phi) + tn'(\phi)\rho(\phi) \end{bmatrix} = 1/\rho + t\rho^2(\phi) \begin{bmatrix} n(\phi) & n'(\phi) \end{bmatrix}$ using $\det(\vec{a}, \vec{b} + \vec{a}) = \det(\vec{a}, \vec{b})$. The caustic of the curve γ is called the **evolute** of the curve.



EXAMPLE: Locally, we can represent a plane curve as a graph $(x, f(x))$. The wave front $W(t, x) = (x, f(x)) + t(-f'(x), 1)/\sqrt{1 + f'(x)^2}$ has the caustic

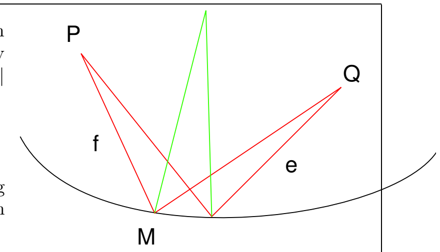
$$\{(t, x) = (1 + f'(x)^2)^{3/2}/f''(x), x)\} .$$

For example, for $f(x) = x^2$, we have $\{W((1 + 4x^2)^{3/2}/2, x)\} = \{(-4x^3, 1/2 + 3x^2)\}$ which is essentially the graph of $y = x^{2/3}$. For $f(x) = x^4$, we have $\{W((1+16x^6)^{3/2}/(12x^2), x)\} = \{(2x/3 - 16x^7/3, 7x^4/3 + x^{-2}/12)\}$.



THE MIRROR EQUATION. If P and Q are successive points on a caustic for a geodesic ray which is reflected at the boundary point M with curvature κ and impact angle θ, then $f = |P - M|$ and $e = |Q - M|$ satisfy

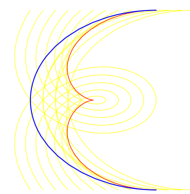
$$\boxed{\frac{1}{f} + \frac{1}{e} = \frac{2\kappa}{\sin(\theta)}}$$

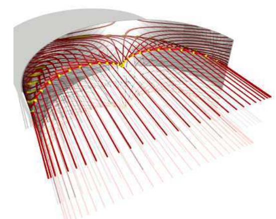


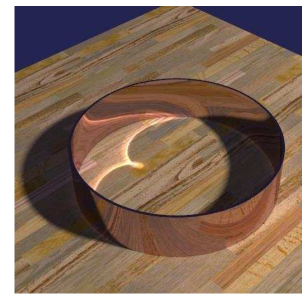
PROOF. The change of the incoming angle $d\theta_1$a and the outgoing ray $d\theta_2$ is related by $d\theta_2 = 2d\theta - d\theta_1$. The claim follows from $d\theta = 1/\rho = \kappa, d\theta_1 = \sin(\theta)/f, d\theta_2 = \sin(\theta)/e$.

Interpretation: If $P = x$ is a point, then Q is a point of the differential geometrical caustic C_x of the point x. If you light a flashlight at P, then the point Q will be a focal point, where the light density is strong.

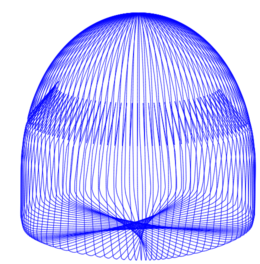
THE COFFEE CUP CAUSTIC. If $r(t) = (-\sin(t), \cos(t))$ is the boundary of the cup and light enters in the direction $(-1, 0)$, then the impact angle θ is just t. The curvature $\kappa(t)$ is 1. Parallel light coming from the right focuses at infinity so that $1/f = 0$. The light which leaves into the direction $(\cos(2t), \sin(2t))$ focuses after reflection at a distance $e = \sin(\theta)/(2\kappa) = \sin(\theta)/2$. The caustic is therefore parameterized by $(-\sin(t), \cos(t)) - (\cos(2t), \sin(2t))\sin(t)/2 = (-\sin(t) + \cos(2t)\sin(t)/2, \cos(t) + \sin(2t)\sin(t)/2)$. Image credit for the picture to the right: Henrik Wann Jensen 1996.



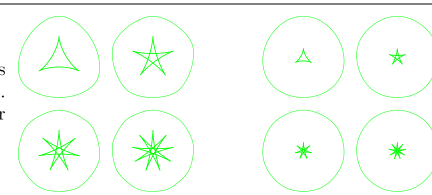




CAUSTICS OF BILLIARDS. The word "caustic" has different meaning in billiards and in differential geometry. Caustics can be defined for any family of light rays. In differential geometry, one looks at all the light rays which are emitted at one spot or all light rays emitted orthogonally to a given curve. If we look at all the light rays emitted from a point x in a billiard table, we will see caustics too. The differential geometrical C_x will be dense however in general. In billiards, we have looked at the caustic of a family of rays which correspond to billiard trajectories on an invariant curve. However, there are some cases, where there is a direct connection between differential geometrical caustics and caustics of billiards. We can deform a sphere in such a way that the caustic of a point on the sphere is the caustic of a special billiard table. We have used this construction once to find metrics on spheres for which the caustics is nowhere differentiable.



CAUSTICS OF BILLIARDS. Caustics of billiards can be quite complicated. To the right, we see some examples for billiards in tables of equal thickness.

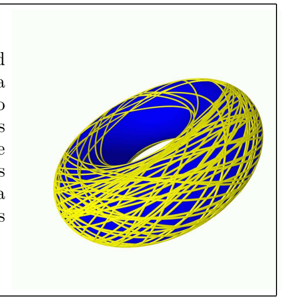


GEODESICS — Math118, O. Knill

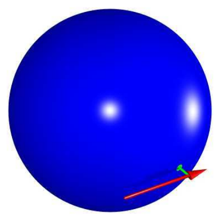
ABSTRACT. Light moves on shortest paths. The corresponding dynamical system is called the **geodesic flow**. We will see examples of geodesic flows which are integrable like the flow on a surface of revolution. This is an introduction to geodesic flows without Riemannian geometry which allows to go straight to the essential math without too much formalism.

ARCHIMEDES THEOREM. We have seen that the shortest distance between two points in Euclidean space is the line. We have proven this in the case of the plane without use of derivatives. This "Archimedes proof" can be generalized to higher dimensional Euclidean spaces too.

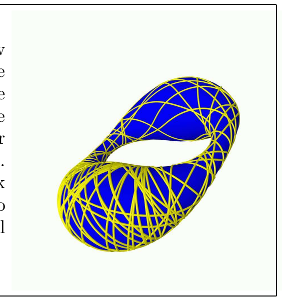
DEFINITION. Given a smooth surface in space, a point P on the surface and initial tangent velocity vector v. Define a path on the surface by letting a particle move freely in space under the influence of a force perpendicular to the surface in such a way that the particle stays on the surface. This defines a path on the surface called **geodesic flow**. This dynamical system can be described using differential equations too. However, for many of the examples considered here, we can work with the intuitive notion. If the surface has a boundary, then we have a **surface billiard**. In that case, we assume the mass point bounces off the boundary according to the usual billiard law.



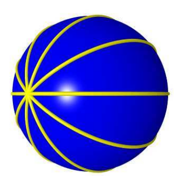
The force $F(x, v)$ perpendicular to the surface at the point x to the direction v can be computed by intersecting the plane spanned by the unit normal vector $\vec{n}$ and the vector v with the surface, leading to a curve with a **radius of curvature** r. Applying the centrifugal force $F(x, v) = |v|^2 n/r$ assures that the particle stays on the surface. The number $\kappa(x, v) = 1/r(x, v)$ is called the **sectional curvature** at the point in the direction v.



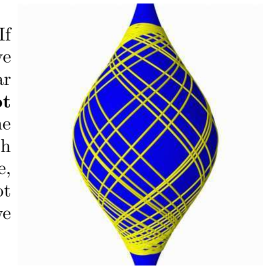
MOTIVATION. The numerical method, we used to compute the geodesic flow on some of the pictures on this page is a mechanical one. We constrain the free motion onto the surface. Given a surface X in space we look at the free evolution of the particle subject to a strong force which pulls the particle to the surface. That force is always perpendicular to the surface and so perpendicular to the velocity of the particle. Especially, it does not accelerate the particle. Do a free evolution in space for some time dt, then projection the vector back onto the surface. $X(u, v) \to X(u, v) + V \to X(u_1, v_1)$ This method is so efficient and simple, that we have let the ray-tracing program (Povray) do all the computation for the pictures.



EXAMPLE: GEODESICS ON THE SPHERE.
On a sphere, the mass-point is at any time subject to a force which goes through the center of the sphere. Angular momentum conservation $\frac{d}{dt}L = \frac{d}{dt}x \times v = 0$ implies that the particle stays on a plane spanned by the normal vector and the initial vector v. The geodesic curve is the intersection of the plane with the sphere: it is a grand circle. The plane can be seen as a limiting case of the sphere, when the radius goes to infinity. A particle which initially is on a plane and has a velocity tangent to the plane stays on the plane without any need of constraint. The geodesic curves consist of lines.

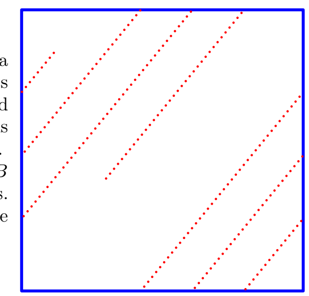


EXAMPLE: GEODESICS ON SURFACE OF REVOLUTION. If ϕ is the angle between a longitudinal line and the geodesic curve and r is the distance from the axes of rotation, then the angular momentum $L = r\sin(\phi)$ is conserved. It is called the **Clairot** integral. Examples of surfaces of revolution are the cylinder, the cone or the torus. If we write the torus as part of the plane with a space dependent metric which depends only on one coordinate, we have a geodesic flow on a surface of revolution. The Clairot integral $r\sin(\phi)$ is the analogue of Snells integral $g(x)\sin(\alpha)$ we have seen before.



METRIC AND DISTANCE. Consider a two-dimensional parametrized surface $(u, v) \mapsto r(u, v)$. At a point $(u, v, r(u, v))$, we have the tangent vectors $dx = r_u du, dy = r_v dv$ The distance element $ds = \sqrt{dx \cdot dx + dy \cdot dy}$ satisfies $ds^2 = (r_u du + r_v dv)^2 = r_u \cdot r_u du du + r_u \cdot r_v du dv + r_v \cdot r_u dv du + r_v \cdot r_v dv dv$. With $g = \begin{bmatrix} r_u \cdot r_u & r_u \cdot r_v \\ r_v \cdot r_u & r_v \cdot r_v \end{bmatrix}$, this can be written as $ds^2 = (du, dv) \cdot g(du, dv)$. A new dot product $< a, b >= a \cdot gb$ and length $||a|| = \sqrt{< a, a >}$ allows to write the length of a curve as $\int_a^b ||r'(t)||\, dt$. Riemanns view is to start with a two dimensional surface M and a symmetric matrix at each point $g_{ij}(x, y)$ defined so that both eigenvalues of g are positive everywhere. The pair (M, g) defines a **Riemannian manifold**. One can measure distances on it without referring to the ambient space in which the surface is embedded.

EXAMPLE: GEODESICS ON THE FLAT TORUS. Because a region in a flat torus can be seen as a region in the plane, geodesics on the flat torus are made of lines. With $g_{ij} = 1$ if $i = j$ and $g_{ij} = 0$ if $i \neq j$ as in the case of the plane, the differential equations for the geodesics are $\ddot{x}^k = \Gamma_{ij}^k \dot{x}^i \dot{x}^j = 0$. There is no acceleration. The fact that the shortest connections between two points A, B on the flat plane are straight lines can be seen in different ways. The straight line gives a distance between the two points as we have seen before in the plane.



EXAMPLE: HILLY REGION. Let $r(u, v) = (u, v, f(u, v))$ be a parameterization of the graph of f. The metric is $g(u, v) = \begin{bmatrix} r_u \cdot r_u & r_u \cdot r_v \\ r_v \cdot r_u & r_v \cdot r_v \end{bmatrix} = \begin{bmatrix} 1 + f_u^2 & f_u f_v \\ f_u f_v & 1 + f_v^2 \end{bmatrix}$. So, if $r(u(t), v(t))$ is a curve on the surface, we can calculate its length. We should get the same result as if we would compute the length of the curve $r(t) = (u(t), v(t), f(u(t), v(t)))$ in three dimensional flat space. But with the internal formalism, it is possible to compute the length without using the third dimension.

CONNECTION. When minimizing the length of a curve, we have to find the Euler Lagrange equations. This involves differentiating the metric g further. The **Christoffel symbols** are defined as

$$\Gamma_{ijk} = \frac{1}{2}[\frac{\partial}{\partial x^i} g_{jk}(x) + \frac{\partial}{\partial x^j} g_{ik}(x) - \frac{\partial}{\partial x^k} g_{ij}(x)] .$$

For a parametrized surface, this is

$$\Gamma_{111} = r_{uu} \cdot r_u,\ \Gamma_{112} = r_{uu} \cdot r_v \qquad \Gamma_{211} = r_{vu} \cdot r_u,\ \Gamma_{212} = r_{vu} \cdot r_v$$
$$\Gamma_{121} = r_{uv} \cdot r_u,\ \Gamma_{122} = r_{uv} \cdot r_v \qquad \Gamma_{221} = r_{vv} \cdot r_u,\ \Gamma_{222} = r_{vv} \cdot r_v$$

FREE MOTION ON A SURFACE. A particle of momentum p has the Lagrangian $F(t,x,p) = \frac{1}{2} g_{ij}(x) p^i p^j$. We use **Einstein summation convention** to automatically sum over pairs of lower and upper indices. We want to minimize $I(x) = \int_a^b F(t,x,\dot{x}) dt = \int_{t_1}^{t_2} g_{ij}(x)\dot{x}^i \dot{x}^j \; dt$ With $F_{p_k} = g_{ki} p^i$ and $F_{x_k} = \frac{1}{2} \frac{\partial}{\partial x^k} g_{ij}(x) p^i p^j$ and the identities $\frac{1}{2} \frac{\partial}{\partial x^j} g_{ik}(x) \dot{x}^i \dot{x}^j = \frac{1}{2} \frac{\partial}{\partial x^i} g_{jk}(x) \dot{x}^i \dot{x}^j$ $g_{ki} \ddot{x}^i = -\Gamma_{ijk} \dot{x}^i \dot{x}^j$ and the definitions $g^{ij} = g_{ij}^{-1}$, $\Gamma^k_{ij} := g^{lk} \Gamma_{ijl}$ this can be written as

$$\boxed{\ddot{x}^k = -\Gamma^k_{ij} \dot{x}^i \dot{x}^j}$$

Because F is time independent, $H(p) = p^k F_{p^k} - F = p^k g_{ki} p^i - F = 2F - F = F(p)$ is constant along the orbit.

GEODESICS ON A SURFACE With $G(t,x,p) = \sqrt{g_{ij}(x) p^i p^j} = \sqrt{2F}$, the functional $I(\gamma) = \int_{t_1}^{t_2} \sqrt{g_{ij}(x) \dot{x}^i \dot{x}^j} \; dt$ is the **arc length** of γ. The Euler-Lagrange equations $\frac{d}{dt} G_{p^i} = G_{x^i}$ can using the previous function F be written as $\frac{d}{dt} \frac{F_{p^i}}{\sqrt{2F}} = \frac{F_{x^i}}{\sqrt{2F}}$ Which means $\frac{d}{dt} F_{p^i} = F_{x^i}$ because $\frac{d}{dt} F = 0$. Even so we got the same equations as for the free motion, they are not equivalent: a reparametrization of time $t \mapsto \tau(t)$ leaves only the first equation invariant and not the second. The distinguished parameterization for the extremal solution is proportional to the arc length. The relation between the two variational problems for energy and arc length is a special case of the **Maupertius principle**.

EXAMPLE: GEODESICS ON THE HYPERBOLIC PLANE. This is an example, where the surface is not given as an embedded surface in R^3. Instead, we assume that the distance on the upper half plane H is given by the formula

$$L(\gamma) = \int_a^b \frac{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}}{y(t)} \; dt \; .$$

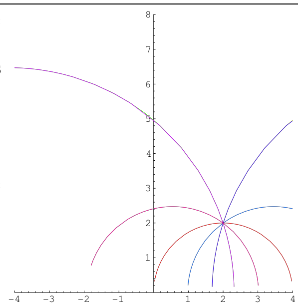
THEOREM. On the hyperbolic plane, geodesics between two points P, Q is the circle through P, Q which hits the x axes in right angles.

PROOF. For $P = (x,a), Q = (y,b)$, the distance is $d(P,Q) = \int_a^b y'(t)/y(t) \; dt = |\log(b/a)|$. The geodesic connection is a line. Now see H as part of the complex plane and note that **Moebius transformation**

$$T(z) = \frac{(az+b)}{(cz+d)}$$

with $ad - bc = 1$ maps circles to circles or lines is an isometry: $d(P,Q) = d(T(P), T(Q))$. Indeed, the two formulas $\mathrm{Im}(T(z)) = \mathrm{Im}(z)/|cz+d|^2$ and $d/dt T(z(t)) = z'(t)/|cz+d|^2$ imply

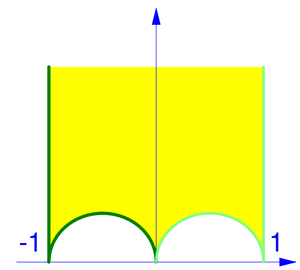
$$\int_a^b \frac{d/dt T(z(t))}{\mathrm{Im}(T(z(t)))} \; dt = \int_a^b \frac{z'(t)}{\mathrm{Im}(z(t))} \; dt \; .$$



To see that a Moebius transformation preserves circles, note that one can write T as a composition $T = T_2 I T_1$, where $T_1(z) = cz+d, T_2(z) = a/c + (ad-bc)z/c$ and where $I(z) = 1/z$ is the inversion at the unit circle. Because all three transformations preserve circles also A circle through the origin is maped into a line. If a,b,c,d are real, then T maps the upper half plane onto itself.

CHAOTIC GEODESIC FLOW. We have seen that the cat map $T(x,y) = (2x+y, x+y)$ is integrable and harmless on the plane. You have computed in a homework an integral, a function $F(x,y)$ which is invariant under T. When projecting the map onto the torus R^2/Z^2, then chaos happens. We have seen that the map allows a description by a symbolic dynamical system. Especially, it is chaotic in the sense of Devaney. A similar thing happens when we look at the geodesic flow on the upper half plane H. The orbits are circles. Even so you have sensitive dependence on initial conditions (as you can see in the picture above that if you start with different direction from the same point, the trajectories separate fast). We can do the analogue of the torus construction on the hyperbolic plane: take a discrete subgroup Γ of the group of all Möbius transformations.

For example Γ could be the subgroup of Möbius transformations with integer entries. It is called the **modular group**. An other subgroup is the **modular group lambda** Λ of all transformation $T(z) = (az+b)/(cz+d)$, where a,d are odd integers and b,d are even integers. The equivalent region to the square in the case of the torus is the **fundamental region** H/Λ which is displayed to the right. Billiard trajectories move on circles, when hitting the the boundary z of the region they enter at an other place $\gamma(z)$ similar than Pacman does for the torus. The corresponding flow is chaotic for any known notion of chaos.



THE DOUGHNUT. The rotationally symmetric torus in space is parameterized by

$r(u,v) = ((a + b\cos(2\pi v))\cos(2\pi u), (a + b\cos(2\pi v))\sin(2\pi u), b \sin$

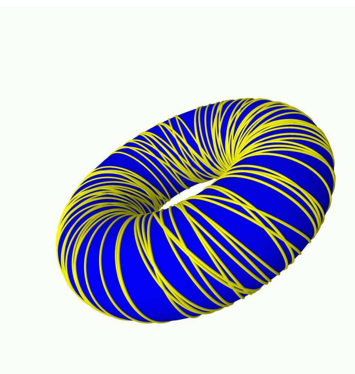
where $0 < b < a$. The metric is

$$\begin{aligned} g_{11} &= 4\pi^2 (a + b\cos(2\pi v))^2 = 4\pi^2 r^2 \\ g_{22} &= 4\pi^2 b^2 \\ g_{12} &= g_{21} = 0 \end{aligned}$$

so that length of a curve is measured with the formula

$$\int_a^b 4\pi^2 (r(u(t), v(t)))^2 \dot{u}^2 + b^2 \dot{v}^2) \; dt \; .$$

The circles $v = 0, v = 1/2$ are geodesics as are all the circles $u = u_0$. The surface is rotationally symmetric and one has the Clairot integral.



HOPF-RYNOV THEOREM ETC. The geodesic flow is defined for all times for closed complete surfaces without boundary. On every point on the surface and in any direction, there exists exactly one geodesic curve. Every geodesic subsegment of a geodesic curve is a geodesic curve. The shortest path between two points on the surface is a geodesic. But as the sphere shows, not every geodesic is the shortest path (you might go into the wrong direction on the grand circle). If two points are close enough, then the shortest geodesic connecting the two points is the shortest curve.

REMARKS. It is not custom to **define** the geodesic flow by constraining the free flow to the surface. But it is a useful fact and used for proving the integrability of the geodesic flow on the ellipsoid. The construction works in general: the **Nash embedding theorem** assures that any Riemannian surface can be embedded isometrically in an Euclidean space.

CONCLUSIONS (preliminary) — Math118, O. Knill

ABSTRACT. We summarize the main points of this course and add some didactical comments.

DYNAMICAL SYSTEMS. While the notion of dynamical systems can be defined in much greater generally, all dynamical systems considered here were either given by a map T on space X or by a differential equation $\dot{x} = f(x)$.

MATHEMATICAL STRUCTURES. The space X can carry different structures. It can be **topological**, **measure theoretical**, **combinatorial**, **geometrical** or **analytical**. Stressing the topological structure leads to topological dynamics, using an invariant measure reaches out to **probability theory** or **ergodic theory**, the geometrical structure is involved when dealing with differentiable functions and subject to **differential geometry**. Combinatorial structures come into play, when doing symbolic dynamics, when dealing with complexity or counting issues. The analytic structure is involved when the map can be extended to the complex, crossing the boundary to **complex analysis**, **algebraic geometry** or **potential theory**.

Topic	Examples	Key points
dynamical systems	semigroup action	the subject has relations with virtually any field of mathematics
1D dynamics	quadratic map, Ulam map	periodic points and their bifurcations, conjugation, Lyapunov exponents
2D dynamics	Henon map, Standard map	horse shoe construction, stability of periodic points, stable and unstable manifolds, Jacobean
2D differential equations	van der Pool equation, linear systems	Poincare-Bendixon
3D differential equations	Lorentz system	Poincare return map, Hopf bifurcation, Lyapunov function, fractals
billiards	polygons, ellipse, stadium	variational principle to construct orbits, effect for chaos
cellular automata	elementary 1D automata, life, lattice gases	topology of sequence space, attractor special solutions
complex dynamics	quadratic maps	Newton method, stability of periodic points, conjugation to normal form
symbolic dynamics	Baker map, full shift, Fibonacci shift, even shift	graphs from forbidden words, symbolic dynamics in general system
dynamics in number theory	irrational rotation, maps on finite sets	continued fraction expansion, dynamic logarithm problem, dynamical systems from curves
celestial mechanics	Kepler, Sitnikov, restricted planar 3-body problem	integrals, horse shoe construction, rotating coordinate systems
geodesic flow	plane, sphere, surfaces of revolution	surface billiards, integrals, caustics, calculus of variations

Some of main points I wanted to make in this course:

- Even deterministic systems lead to unpredictable or uncomputable problems.
- Some systems allow explicit solutions, other systems remain mysterious.
- The history of dynamical systems often sits at the heart of the history of mathematics or science.
- The subject has connections with many other fields of mathematics.
- Dynamical systems theory has many applications.
- There are many open problems left in the area of dynamical systems.

WHAT DID WE LEAVE OUT? First of all, each of the topics could be extended to a full course. There are also important fields, which have not been touched at all: partial differential equations and systems in fluid dynamics in particular, systems with higher dimensional time as they appear in statistical mechanics, dynamical systems of algebraic origin. A large area for dynamics is also **game theory** or the theory of **neural networks**. Then there are problems of **statistical flavor** which deals with the problem to find the laws of the dynamical system from data. A particular case in statistics is to recover the space X the transformation T as well as the measure μ which produces the data. An other untouched area is **artificial intelligence**, where dynamical systems play a role too, especially in inverse problems. Finally, there are quantum versions of many dynamical systems considered so far. For billiards or surface billiards, the quantum problem is the study of the Laplacian on the surface with Dirichlet boundary conditions. The eigenvectors of the Laplacian v_n in the limit $n \to \infty$ have connections with the billiard or geodesic flow on the surface. **Quantum dynamical systems** can be obtained reformulating things first on a **function space**. For a topological dynamical system (X, T), consider the space $\tilde{X} = C(X)$ of all continuous functions on X. The map T induces a linear map $\tilde{T}$ on $\tilde{X}$ by $\tilde{(T)}(f)(x) = f(T(x))$. Allowing more general spaces $\tilde{X}$ C^* algebras allows the study of quantum versions. Also measure theoretical systems (X, T, μ) can be reformulated in function space. Instead of$\tilde{T}(f) = f(T)$ on all bounded measurable functions, consider the dynamics of a general unitary operator or more generally an automorphism on a von Neuman algebra. Also geometric structures have been "quantized" leading to a subject called "noncommutative geometry". Lets mention the topic of **perturbation theory**, which is used for example to prove the persistence of stable motion (KAM) or the existence of homoclinic points (Melnikov theory). Finally, there is **spectral theory**, the study of the unitary operator $U_t f = f(T_t)$ on $L^2(X, \mu)$ for a map or flow T preserving a measure μ.

DIDACTICS. We have covered a lot of material in this course. To avoid being shallow, examples were chosen at the heart of the subject. One could teach this course with the material from the first or second week, but in more depth. That would make sense too. I personally think that in a time where knowledge is accumulated at a tremendous speed, it makes sense to be trained also in the process of acquiring a lot of knowledge in a short time. Equally important is the ability to solve not so straightforward problems and to find creative solutions.

KNOWLEDGE VERSUS CREATIVITY. I can notice more and more that results are published which have been found a long time ago. Also referees often don't know about entire areas which would be relevant. Even special areas of dynamical system theory have fragmented and specialists know only part of it. It is relatively easy to be creative, when ignoring knowledge. It is much harder to find new results in the context of what is known. The right balance has to be found. In a first stage of research, avoiding the literature might be a good idea since too much information can be deadly for creative work. But after having figured out a way to solve the problem, looking up the literature is a necessity to face the possibility that a result has been proven already, maybe a special case of a much more general result. In that "library stage", a lot of information has to be processed in a short time. In a time, where patent offices pass sometimes requests which have a long time been "prior art" and in the public domain, some effort to pass some of the information which is available in books, in databases or papers to the brain has been made. Fortunately, technology softens some of the need to know vast amount of information. Still, most information is not online, nor in text books, not even in recent papers. The challenge is to balance two different but equally important things:

Acquire: process, absorb and learn information

Inquire: question, generate new ideas and solutions

HOMEWORK: having graded the homework myself, I can assess that most homework questions seemed have been just at the right level of difficulty. Some students have spent a lot of time cracking some of the homework problems so that increasing the difficulty level would not make a lot of sense. I think most of the homework problems could not be solved without spending a few hours each week. The act of grading was for me an additional valuable resource to gauge the progress of the class and adapt the difficulty of the lectures. Notes were written typically just before each lecture so that an adaptation of speed was possible.

QUIZZES: the weekly quizzes tested knowledge and presence in the classroom. They also served as a tool to gauge, how the information have been absorbed during lecture.

PROJECTS: are on the way. Assessment about them will be added later here.