GOLDEN ROTATIONS

OLIVER KNILL

Abstract. These are expanded preparation notes to a talk given on February 23, 2015 at BU. This was the abstract: "We look at Birkhoff sums $S_n(t)/n = \sum_{k=1}^n X_k(t)/n$ with $X_k(t) = g(T^k t)$, where T is the irrational golden rotation and where $g(t) = \cot(\pi t)$. Such sums have been studied by number theorists like Hardy and Littlewood or Sinai and Ulcigrain [41] in the context of the curlicue problem. Birkhoff sums can be visualized if the time interval $[0, n]$ is rescaled so that it displays a graph over the interval $[0, 1]$. While for any L^1-function $g(t)$ and ergodic T, the sum $S_{[nx]}(t)/n$ converges almost everywhere to a linear function Mx by Birkhoff's ergodic theorem, there is an interesting phenomenon for Cauchy distributed random variables, where $g(t) = \cot(\pi t)$. The function $x \to S_[nx]/n$ on $[0, 1]$ converges for $n \to \infty$ to an explicitly given fractal limiting function, if n is restricted to Fibonacci numbers $F(2n)$ and if the start point t is 0. The convergence to the "golden graph" shows a truly self-similar random walk. It explains some observations obtained together with John Lesieutre and Folkert Tangerman, where we summed the anti-derivative G of g, which happens to be the Hilbert transform of a piecewise linear periodic function. Recently an observation of [22] was proven by [40]. Birkhoff sums are relevant in KAM contexts, both in analytic and smooth situations or in Denjoy-Koksma theory which is a refinement of Birkhoff's ergodic theorem for Diophantine irrational rotations. In a probabilistic context, we have a discrete time stochastic process modeling "high risk" situations as hitting a point near the singularity catastrophically changes the sum. Diophantine conditions assure that there is enough time to "recover" from such a catastrophe. There are other connections like with modular functions in number theory or Milnor's theorem telling that the cot function is the unique non constant solution to the Kubert relation $(1/n)\sum_{k=1}^n g((t+k)/n) = g(t)$."

Date: July 31, 2015.

1991 *Mathematics Subject Classification.* Primary: 05C50,81Q10 .

Key words and phrases. Dynamical systems.

1. A very special problem

We look at examples of **Birkhoff sums** of $g(x) = \cot(\pi x)$ over the **golden rotation** $x \to x + \alpha$, where $\alpha = (\sqrt{5} - 1)/2$. This is a distinguished setup as the function $g/2$ is the only non-zero odd function with a constant Fourier transform $\hat{g}/2 = (1, 1, \dots)$ as $g = 2\sum_{k=1}^{\infty} \sin(2\pi kx)$ and the golden ratio α is the only nonzero real number in $[0, 1]$ with a constant continued fraction expansion $\alpha = [1, 1, \dots]$. The later expansion is verified from the defining identity $\alpha = 1 + 1/\alpha$, then plugging in the left hand side into the right. The constant Fourier series comes from expanding the left hand side of $2(1 - \exp(ix))^{-1} = 1 + i\cot(x/2)$ as a geometric series and comparing the imaginary part. As a distribution, the Hilbert transform of g is the Dirac delta $h = \delta_0 - 1$ on the circle because integrating $2\sum_{k=1}^{\infty} \cos(2\pi kx) = -1 + \sum_{k \in Z} e^{2\pi ikx}$ over $[0, 1]$ with a rapidly decreasing test function g gives $-1 + \sum \hat{g}(k)$ which is by the Poisson summation formula $-1 + \sum g(k)$. One can also deduce it from the fact that the anti derivative $\pi G = \log(2 - 2\cos(2\pi x))/2 = \log|1 - e^{2\pi ix}|$ of πg is the Hilbert transform of the piecewise linear function $\pi H = \pi(x - [x] - 1/2) = \arg(1 - e^{2\pi ix})$ which is the anti derivative of the Dirac delta. The exponential of the Birkhoff sum is therefore up to a scaling factor the product $P_n(z) = \prod_{k=1}^{n} (1 - z^k)^{-1}$ whose Taylor coefficients count the number $p(n)$ of partitions of n into maximally n positive summands.

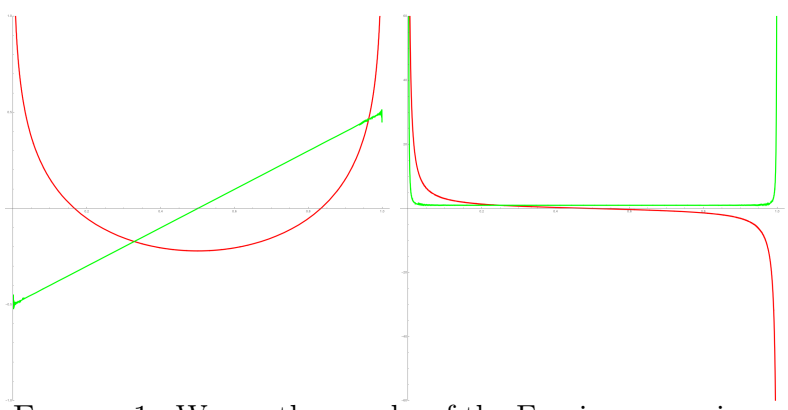

Figure 1. We see the graphs of the Fourier approximations of G, H, defined by $(G + iH) = \log(1 - e^{2\pi ix})/\pi$. $H = (x - [x] - 1/2)$ is piecewise linear and the identity $\pi G(x) = \log(2 - 2\cos(2\pi x))/2$ holds. To the right, we see the derivatives $G' = \cot(\pi x)$ and H' which is the Dirac delta on the circle.

The cot-function appears in different setups and is distinguished in many ways, similarly as the Gaussian functions. It is no surprise that in solid state physics, the **Maryland model** [34, 8] has so many symmetries and explicit formulas. There is some relation with "chaos theory" as iterating the dynamical system $T(x) = \cot(x)$ on the real line produces random numbers: if x_n is an orbit, then the sequence $\operatorname{arccot}(x_n)$ is uniformly distributed on $[-\pi/2, \pi/2]$. When replacing cot with tan, we have a parabolic fixed point $x = 0$ leading to intermittent behavior.

FIGURE 2. Simeon Poissonand Joseph Fourier

2. High risk

For a high risk stochastic process, the variance of the increments is not bounded. We aim to understand sums $\sum_{k=1}^n g(T^k x)$, where T is a measure preserving transformation on a probability space and where $g : X \to \mathbb{R}$ is an observable with a **Cauchy distribution**. In other words, we like to get a grip on the growth $\sum_{k=1}^n X_k$ of identically distributed random variables with zero expectation but in a situation, where the random variables are not necessarily independent. Every Cauchy distributed random variable X_k on a probability space can be realized in the form $X_k(x) = \cot(\pi T^k x)$ with some $T : [0,1] \to [0,1]$. The fact that $\cot(\pi x)$ has a Cauchy distribution follows from the fact that $\arctan'(x)/\pi$ is the Cauchy distribution. While the expectation exists only as a Cauchy principle value, the variance $\int (x^2/\pi)(1+x^2)\, dx$ is infinite, so that we deal with high risk situations. Close encounters to the origin produce large changes in the sum. The Cauchy distributions $(1/\pi)/(1+x^2)$ is special in probability theory as it is the high risk situations analogue of the Gauss function $\exp(-x^2/2)/\sqrt{2\pi}$. Both have central limit theorems because both are invariant when adding independent random variables with that distribution. What happens in such infinite variance stochastic process if correlations are allowed?

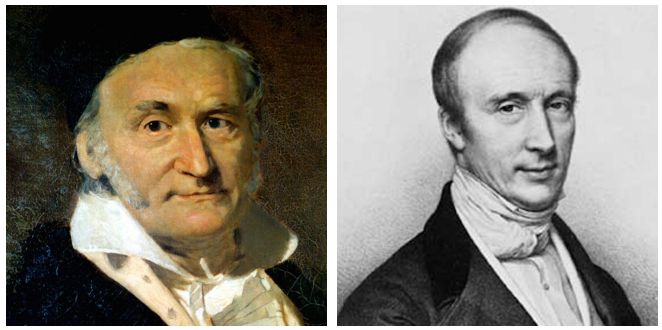

FIGURE 3. Carl Friedrich Gauss August-Louis Cauchy

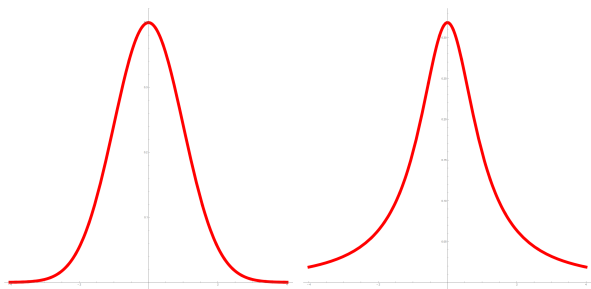

FIGURE 4. The Gaussian and the Cauchy distribution are both special. The Gaussian is an example of a bounded-risk L^1 processes, the Cauchy process in a non-integrable case which means high risk.

3. Birkhoff sums

A function $g : [0,1] \to \mathbb{R}$ and a Lebesgue measure preserving transformation $T : [0,1] \to [0,1]$ defines a sequence of **random variables** $X_k(t) = g(T^k x)$. They form a **discrete stochastic process**, as all the random variables have the same distribution and "time" is the discrete set $k = 1, 2, 3, \dots$ of integers. As the probabilist Joseph Doob noticed first, any discrete process can be realized as such. This shows that part of probability theory can be absorbed within dynamical systems theory. The sum $S_n = \sum_{k=1}^n X_k$ is now a **Birkhoff sum** and S_n/n is a **time average**. In probability, where the X_k are assumed to be independent, we get the **laws of large numbers**. The relation between time averages and space averages is treated with ergodic theorems like the Birkhoff ergodic theorem. Why do we want to study such sums? First of all, it often happens in applications that we see accumulations S_n of quantities when modeling developments like stock

markets, snow fall or capital. Historically, interest in gambling initiated the first steps in probability theory like with Cardano. In a gambling context, X_k represent the winnings or losses in one game and S_n is the total capital, accumulated over time. More fundamentally, such cocycles over a dynamical system allow to understand the underlying dynamical system T, similarly as fibre bundles allow to investigate the structure of the underlying manifold. What is Doob's argument? Given a sequence of random variables $X_k : (\Omega, \mathcal{A}, P) \to \mathbb{R}$ with identical distribution. We can realize them each on the probability space $(\mathbb{R}, \mathcal{B}, \rho)$, where ρ is the law of the random variable X_i. Let $(\Omega^{\mathbb{N}}, \mathcal{A}^{\mathbb{N}}, \rho^{\mathbb{N}})$ and let T be the shift $T(x)_n = x_{n+1}$. Given $\omega \in \Omega$ we have an element $\phi(\omega) = (X_1(\omega), X_2(\omega), \dots)$. The push-forward of P by ϕ onto $(\Omega^N, \mathcal{A}^N)$ is preserved by the shift and $Y_k(x) = x_k = f(T^k x)$ reproduces now the original random variables X_k. By the way, Doob sat in some of Birkhoff's classes at Harvard and according to [36] was thrown out of the one on "aesthetic measures" [5], as he had objected too loudly to some of the methodology of Birkhoff.

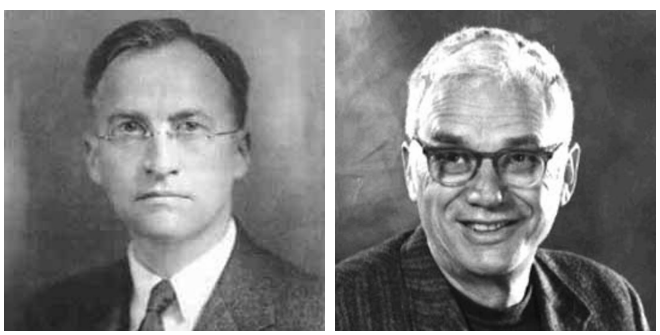

FIGURE 5. George Birkhoff and Joseph Doob

4. JACOBEANS

If $T : [0,1] \to [0,1]$ is a smooth interval map and $g(x) = \log |T'(x)|$ then the growth rate of $S_n = \sum_{k=1}^n X_k$ measures how fast errors propagate. This can be expressed as a Birkhoff sum $\log |(T^n)'(x)| = S_n(x)$ because of the chain rule for functions of one variables. If S_n grows linearly, like for the logistic map $T(x) = 4x(1-x)$ then the derivative $(T^n)'(x)$ of the n-th iterate $T^n = T \circ T \cdots \circ T$ grows exponentially. This means that the system T has **sensitive dependence of initial conditions**. For a measure-preserving map T of a compact smooth manifold M, we can look at the cocycle map $F(x,u) = (Tx, dT(x)u)$ on the compact projective bundle $T_P M$ bundle. It is the fibre bundle where over each point the fibre is the n-dimensional projective space. This is a common

setup in ergodic theory and a point of view taken for example by [25]. One has now a Birkhoff sum with function $g(x,y) = \log(|dT(x)u|/|u|)$. Oseledec's theorem assures that there exist F-invariant measures μ_k on the projective bundle such that the growth rate of S_n is the k'th **Lyapunov exponent**. One would like to know these **Ledrappier measures** μ_k since $\int_{T_P M} g(x,y)\, d\mu_k(x,u)$ is by Birkhoff's ergodic theorem equal to the Lyapunov exponent which leads to the metric entropy. In the case when T is a Bernoulli system, one can actually see that the measures μ_k are a product measure and get results on Lyapunov exponents. This was one of the earliest rigorous results on Lyapunov exponents [14]. In the case when the integral $\int g(x)\, d\mu(x)$ is zero, one can be interested in the sub-exponential growth rate which is related to spectral properties of the spectral measures of the unitary **Koopman operator** $U : f \to f(T)$ on the Hilbert space $L^2(X)$ (see i.e. [17]).

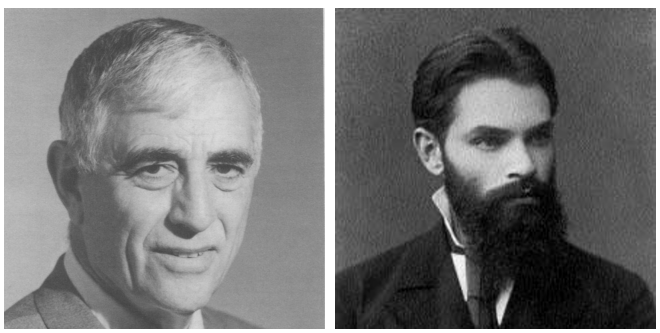

FIGURE 6. Bernard Koopman and Aleksandr Lyapunov

5. Rotation number

The cocycle $F(x,y) = (x+\alpha, A(x)y)$ with $A(x) = \begin{bmatrix} 2-2\cos(x)-E & -1 \\ 1 & 0 \end{bmatrix}$ belongs to the **almost Mathieu operator** $L(x)_n = x_{n+1} - 2x_n + x_{n-1} + 2\cos(x+n\alpha)$ as the time independent second order difference equation $Lx = Ex$ is solved by the first order vector valued equation $\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = A(x) \begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix}$. One knows that the spectrum is a Cantor set [3] of zero measure [24]. The function $\lambda(E) = \log(\det(L-E)) = \mathrm{tr}(\log(L-E)) = \lambda(E) + i\rho(E)$ is Herglotz in the upper half plane $\mathrm{im}(E) > 0$ and remains analytic exactly on the gaps of the spectrum. There are explicit formulas for the derivatives [32]. The real part $\lambda(E)$ is the Lyapunov exponent, the exponential growth rate of the cocycle $A^n(x)$; the imaginary part is the rotation number [33]. For E in a gap,

when viewing $A(x)$ act on the circle T^1, we can see the rotation number as the winding number of the Ledrappier sections [25] which are analytic curves $(x, l^{\pm}(x))$ on the two torus. The homotopy stability of these curves is also called "gap labeling" [4]. For rational α, one can relate the rotation number with the Morse index of the corresponding periodic Jacobi matrix [26]. If the rotation number $\rho(E)$ is constant in a neighborhood of a point E_0, then the complex function $\lambda(E)$ is real analytic and the derivatives can be explicitly computed [32, 33]. If one plots the spectrum of L as a function of α, one obtains the **Hofstadter butterfly**.

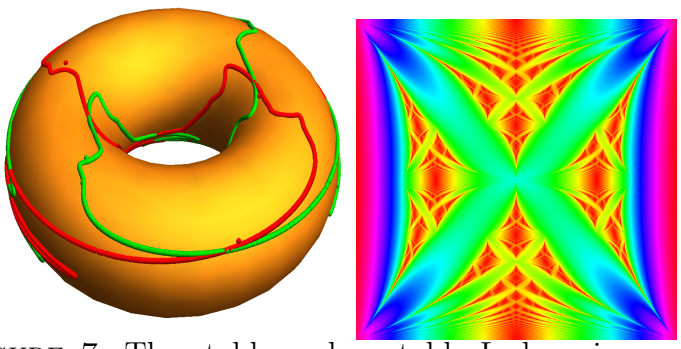

FIGURE 7. The stable and unstable Ledrappier manifolds of $A(x)$ in the almost Mathieu case with golden rotation number. The winding number on the torus is related to the rotation number ρ, which is the Hilbert transform of the Lyapunov exponent λ. The situation is illustrated by the Hofstadter butterfly, which can be visualized by plotting λ as a function of E and α.

6. The entropy problem

For maps in higher dimensions, the growth rate of $\log ||dT^n||$ is no more described by an additive Birkhoff sum but a sub-additive process. The **Standard map** $T_c(x, y) = (2x - y + c\sin(2\pi x), x)$ on the 2-torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ is a prototype problem which remains an enigma. The main open problem is to understand the expected growth of $(T^n)'$, which is the Kolmogorov-Sinai entropy of smooth diffeomorphisms like T. Already in the 1960'ies, Sinai asked whether the entropy of the Standard map is positive. Until now, no value of c is known for which the entropy is known to be positive. Numerically, one measures a lower bound $\log |c|$. I myself tried analytic tools already as an undergraduate: there is a single analytic map T on C^4 which has the property that there are invariant tori X_c on which the map is T_c. The Jacobean

is an analytic matrix valued function. One knows that $\log|A^n(z)|$ is pluri-subharmonic in this case so that averaging over the boundary of a polydisc can be estimated. This is an adaptation of the Herman idea to higher dimensions. But it fails to estimate things as the tori are not polydiscs. The entropy is $\log|\det(L)|$, where L is a random Jacobi operator. One can deform such operators using an isospectral flow [39] and this preserves all spectral data like density of states or Lyapunov exponent [16, 15]. While finite dimensional Toda systems can have be of a scattering or recurrent nature, the ergodic Toda flow does not simplify the situation in the case of the Standard map. The Toda flow is a differential equation for two functions is $\dot{a} = 2a(b - b(T^{-1})$ and $\dot{b} = a^2(T) - a^2$. One can look at piecewise analytic situations and generalize Herman's method by using Jensen formulas. This amounts to estimate the Riesz measures of the subharmonic functions which are spectral arcs of non-selfadjoint operators.

FIGURE 8. Frigyes Riesz and Michael Herman

7. Partition functions

Birkhoff sums are also of interest in number theory. The **partition function** $p(n)$ tells in how many ways a natural number n can be written as a sum of positive integers, has a generating function $\sum_k p_k w^k$ which is the inverse of the Euler function $\prod_{k=1}^n (1-w^k)$. Studying the growth rate of the logarithm of this partial product is a Birkhoff sum over the rotation $x \to x + \alpha$ if $w = \exp(2\pi i\alpha)$. As an illustration on how dynamical methods can shed light on results in number theory, one can deduce from the Gottschalk-Hedlund theorem, and the Hadamard gap theorem as well as Euler's pentagonal number theorem

$$\prod_{k=1}^{\infty}(1-z^k) = 1 - z - z^2 + z^5 + z^7 - z^{12} - z^{15} \cdots = \sum_{n=-\infty}^{\infty} (-1)^n z^{(3n^2-n)/2} ,$$

that the partition function $p(n)$ satisfies $\limsup_n |p(n)|^{1/n} = 1$. In a formal sense, the Pentagonal theorem is the Fourier expansion of the function $\alpha \to e^{S_\infty(\alpha)}$ The fact that $\limsup_n |p(n)|^{1/n} = 1$ is a well known result in additive number theory, but it can be derived using tools from dynamical systems theory.

FIGURE 9. Walter Gottschalk and Gustav Hedlund

8. The upper half plane

If the rotation number α is complex and chosen to be in the upper half plane $\mathrm{Im}(\alpha) > 0$, then the Birkhoff sum $S_n(\alpha) = \log(\prod_{k=1}^n (1 - w^k))$ with $w = e^{2\pi i\alpha}$ converges to an analytic function for $n \to \infty$. Here are some relations: the function

$$f(z) = \prod_{n=1}^{\infty} (1 - z^{2n})$$

is called **elliptic modular function**. The **Dedekind modular η-function** is

$$\eta(\alpha) = z^{1/24} \prod_{n=1}^{\infty} (1 - z^n) \ ,$$

where $z = e^{2\pi i\alpha}$ and $\mathrm{Im}(\alpha) > 0$. It satisfies the functional equation $\eta(A(\tau)) = \epsilon^{1/2}\eta$ for any modular transformation $A(\alpha) = (a\alpha+b)/(c\alpha+d)$ and where $\epsilon = e^{(\pi i(a+d)/12c)-s(d,c)}$ is defined by the **Dedekind sum** $s(h,k) = \sum_{r=1}^{k-1} \frac{r}{k}(\frac{hr}{k} - [\frac{hr}{k}] - 1/2)$. The **modular discriminant** $\Delta(z) = \eta(z)^{24}$ is an example of a **modular form**. Ramanujan conjectured and Deligne proved that the z^p coefficient for prime p has absolute value $\leq 2p^{11/2}$.

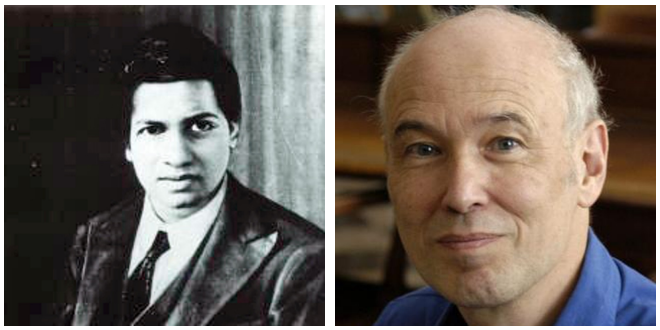

FIGURE 10. Srinivasa Ramanujan and Pierre Deligne

9. Hardy and Littlewood

Ernst Hecke looked at the Dirichlet series $\sum_{n=1}^{\infty} a_n n^{-s}$ with $a_n = g(n\alpha)$ and piecewise linear g. He showed that it has a meromorphic continuation onto the entire plane [11]. Birkhoff sums for $g(x) = \sin(x)^{-1}$ over irrational rotation have been studied in [12]. Hardy and Littlewood showed there that the averaged partial Birkhoff sums S_k/k stay uniformly bounded.

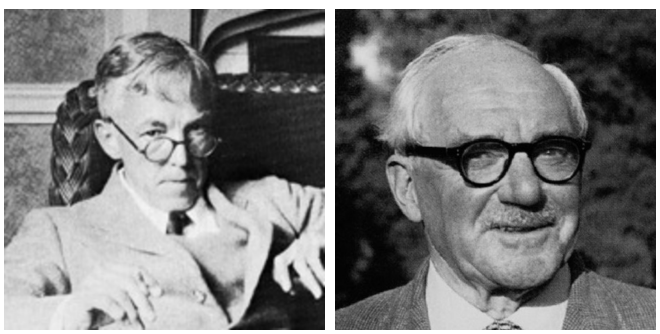

FIGURE 11. Godfrey Hardy and John Littlewood

10. Jacobi Theta

Infinite products of of dynamical type are everywhere. Jacobi's "Aequartro identica ratis abstrura"

$$\prod_{n=1}^{\infty}(1+q^{2n-1})^8 - \prod_{n=1}^{\infty}(1-q^{2n-1})^8 = 16q\prod_{n=1}^{\infty}(1+q^{2n})^8$$

is used in the proof of **Lagrange's theorem** that every positive integer is the sum of four squares. Taking logs, one gets a Birkhoff sum. Of course the question is different. These functions only converge for

$\text{im}(\alpha) > 0$, where the system $z \to qz$ is not area preserving. Relations of the arithmetic nature α, the growth of the Birkhoff sums for real α and the complex functions would not surprise.

Here is a final relation: the classical θ **function**

$$\theta(\alpha) = \sum_{n \in Z} e^{i\pi n^2 \alpha} = \sum_{n \in Z} w^{n^2} = 1 + 2 \sum_{n=1}^{\infty} w^{n^2}$$

with $w = e^{i\pi\alpha}$ is related to the **modular form**

$$\begin{aligned} R(\alpha) = \theta^2(\alpha) &= 1 + 2 \sum_{k=1}^{\infty} g(k\alpha) + g(k\alpha + \pi/2) \\ &= 1 + 2S_\infty(\cot, \pi/4) - 2S_\infty(\cot, -\pi/4) \end{aligned}$$

for $\text{im}(\alpha) > 0$. In other words, the modular form R is up to a constant the sum of two Birkhoff sums which converge for α in the upper half plane.

FIGURE 12. Carl Gustav Jacobi Joseph-Louis Lagrange

11. THE GOLDEN GRAPH

Let q_n be the **Fibonnacci sequence** $q_1 = 1, q_2 = 1, q_3 = 2, q_4 = 3$, etc. We know that $q_{n-1}/q_n \to (\sqrt{5} - 1)/2$ as they are the partial fractions. If we look at a rotation with golden rotation, then the q_n values are the times, when we get closest back to the starting point as then $\alpha q_n - q_{n-1}$ is smallest

Define the **Birkhoff limiting function** of the sum $S_n = \sum_{k=1}^n X_k$ as

$$s(x) = \lim_{n \to \infty} \frac{S_{[q_n x]}}{q_n}$$

for g.

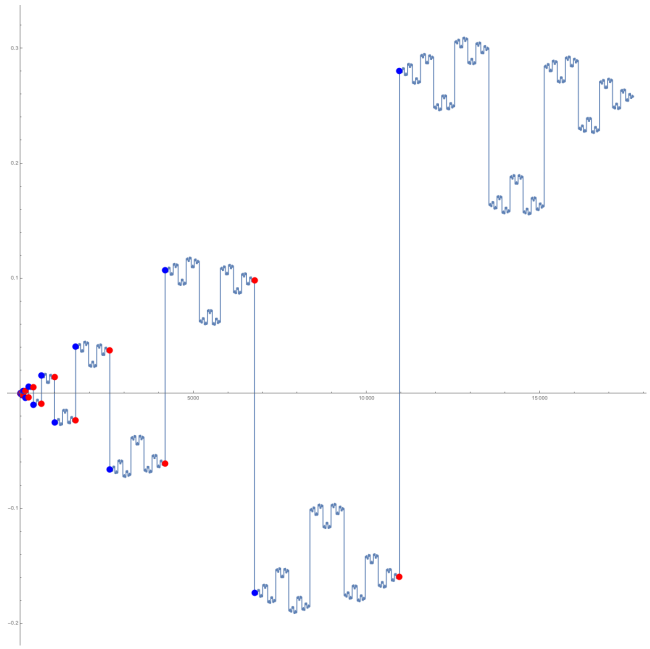

FIGURE 13. The golden graph.

Theorem: The Birkhoff limiting function for $g(x) = \cot(\pi x)$ exists point wise along odd and even subsequences. The graph of $s(x) = \lim_{n\to\infty} s_{2n}(x)$ is **selfsimilar**, as it satisfies $s(\alpha x) = -\alpha s(x)$ and is continuous from the right.

This is exciting as everything is explicit. We could in principle compute S_n/n for arbitrary large numbers like $n = 10^{100}$ even so we are unable to sum up S_n due to lack of computing resources in the universe. The function g can be replaced by any meromorphic function with a single pole, the result does not change.

12. Beta expansion

Numbers can be expanded with respect to any base. Lets take the **binary expansion**, where numbers are written by $x = 0.101000101\ldots$, meaning that $x = \sum_k a_k \alpha^k$ with $\alpha = 1/2$ and $a_k \in \{0,1\}$. Lets look at the graph of the function which assigns to such a x the value $f(x) = \sum (-1)^k a_k$.
This expansion can be done for any $\alpha \in [1/2, 1)$ and called the β expansion. It works in particular for the golden ration, where every x can be written as $\sum_k a_k \alpha^k$ with $a_k \in \{0,1\}$.

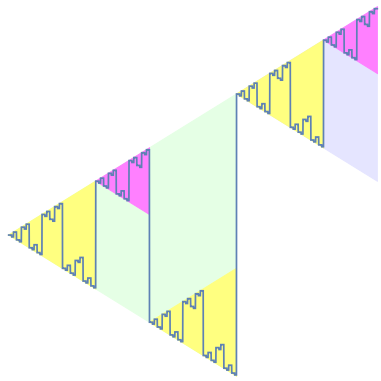

FIGURE 14. The golden graph caricature.

A **caricature for the golden graph** is the function

$$o(x) = \sum_{k=1}^{\infty} (-1)^{k-1} a_k \alpha^k \ ,$$

where $\alpha = (\sqrt{5}-1)/2$ and a_k is the β-expansion of x. The function o has the same symmetry $o(\alpha^n x) = (-\alpha)^n o(x)$ as the golden graph.

13. An explicit formula

The golden graph is a modification of this graph since one can show

$$s(x) = \sum_{k=1}^{\infty} a_k (-\alpha)^k \sigma(y_k) \ ,$$

where $y_0 = 0$, $y_k = \lim_{n\to\infty} [|x_k q_n|\alpha] q_n$ with $x_k = \sum_{i=1}^{k} a_i \alpha^i$.

The function σ is an analytic function which can be constructed:

Theorem: The function $\sigma(y)$ is analytic in y for small $|y|$ with Taylor expansion $\sigma(y) = \sum_{l=0}^{\infty} a_l y^l / l!$, where

$$a_l = \lim_{n\to\infty} \frac{1}{q^{l+1}} \sum_{k=1}^{q-1} k g^{(l)}(\frac{y}{q} + k\alpha)$$

and $q = q_n$ is the n'th Fibonacci number.

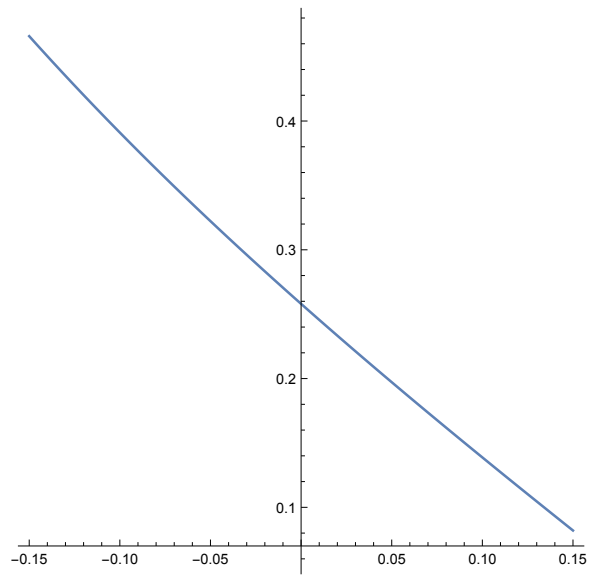

FIGURE 15. The analytic function $\sigma(y)$ determines the function $s(x, y)$ by a β expansion.

14. ZECKENDORFF

The **Zeckendorf representation** of a number is the discrete analogue or dual to a β-expansion of a real number. The picture below illustrates the Zeckendorf representation of $16 = 3 + 5 + 8 = q_2 + q_3 + q_4$ as a sum of Fibonacci numbers as well as the β-expansion of $0.763932\cdots = \alpha^2 + \alpha^3 + \alpha^4$ as a sum of powers of the golden ratio.
Thew expansion is called after the Belgian mathematician Édouard Zeckendorf (1901-1983) who was a dental surgeon, was in prison camps

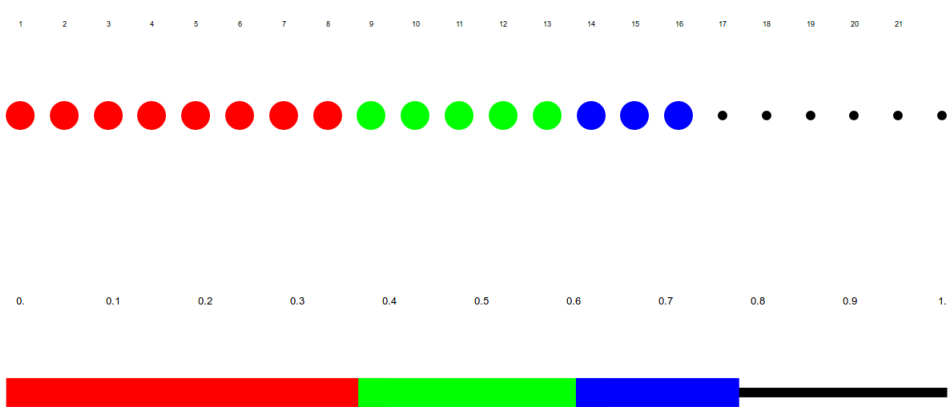

FIGURE 16

during WWII and did mathematics as a hobby. The **Zeckendorf theorem** tells that any integer can be written uniquely as a sum of non-consecutive Fibonacci numbers. The continuous analogue is that every real number $x \in [0,1]$ can be written uniquely as a sum $\sum_k a_k \alpha^k$, where no two consecutive a_k are 1. Here is an illustration of this theorem: for every $x = \sum_i a_i 2^{-i}$ we can define a corresponding Zeckendorf number $F(x) = \sum_i a_{s(i)} \alpha^i$ where $s(i)$ are now nonconsecutive. According to the Zeckendorf theorem, this is a continuous bijection of a compact space and therefore a bijection. It produces a cumulative distribution function $F(x)$ whose derivative $f = F'$ is the Zeckendorf probability density function. It is a singular continuous object.

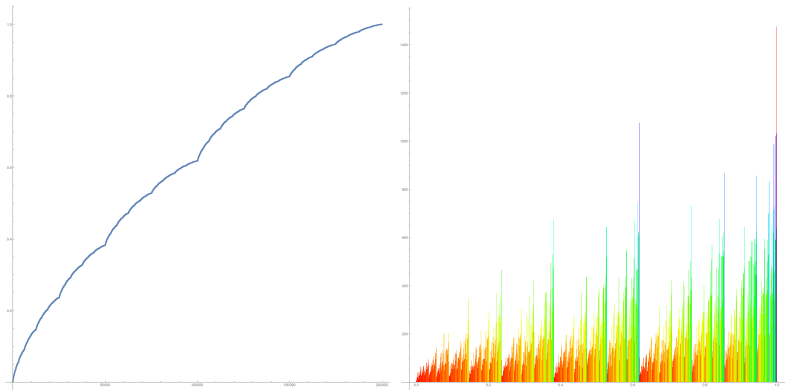

FIGURE 17. The Zeckendorff CDF and PDF.

FIGURE 18. Leonardo Fibonacci and Edouard Zeckendorff

15. Catalan gold

Problem If $\frac{p_n}{q_n}$ is a **partial fraction** of the **golden ratio** $\alpha = (\sqrt{5} - 1)/2$, then

$$\sqrt{5}(\alpha - \frac{p_n}{q_n}) = \sum_{k=0}^{\infty} \frac{(-1)^{(n+1)(k+1)} c_k}{q_n^{2k+2} 5^k} \ , \tag{1}$$

where $c_k = \frac{(2k)!}{k!(k+1)!}$ are the **Catalan numbers** which have the generating function

$$c(x) = \sum_{k=0}^{\infty} c_k x^k = \frac{2}{1 + \sqrt{1-4x}} \ .$$

For odd n and $p/q = p_n/q_n$, the right hand side of (1) is equal to

$$\sum_{k=0}^{\infty} \frac{c_k}{q^{2k+2} 5^k} = c(\frac{1}{5q^2}) \frac{1}{q^2} = \frac{2\sqrt{5}}{\sqrt{5}q^2 + q\sqrt{5q^2-4}} \ .$$

For even n and $p/q = p_n/q_n$, it is

$$\sum_{k=0}^{\infty} \frac{c_k (-1)^k}{q^{2k+2} 5^k} = c(\frac{-1}{5q^2}) \frac{1}{q^2} = \frac{2\sqrt{5}}{\sqrt{5}q^2 + q\sqrt{5q^2+4}} \ .$$

Claim (1) is now equivalent to the formulas

$$q\alpha - p = \frac{2}{\sqrt{5}q + \sqrt{5q^2-4}}, \quad (p,q) = (q_{2n-1}, q_{2n}) \ ,$$

$$q\alpha - p = \frac{-2}{\sqrt{5}q + \sqrt{5q^2+4}}, \quad (p,q) = (q_{2n}, q_{2n+1}) \ ,$$

which give the n'th **Fibonacci number** q_n from the q_{n-1}. Iterate twice and simplify to generate all even Fibonacci numbers with $q_{2n+2} = T(q_{2n})$, where

$$T(x) = (3y + \sqrt{4+5y^2})/2$$

and $q_{2n+1} = S(q_{2n-1})$ if $S(x) = (3y + \sqrt{-4+5y^2})/2$. Even Fibonacci pairs therefore solve the Diophantine equation $(4+5y^2) = (2x-3y)^2$ which is

$$x^2 + y^2 - 3xy = 1 \ .$$

Similarly, odd Fibonacci pairs solve $x^2 + y^2 - 3xy = -1$.
This quadratic Diophantine equation is the special case $x = 1$ of the **Markoff Diophantine equation**

$$x^2 + y^2 + z^2 = 3xyz$$

(see [10]) for which the singular solutions $(1,1,1)$ and $(1,1,2)$ determine the others. In the tree of Markoff solutions $(1,y,z) = (1, q_{2n-1}, q_{2n+1})$.

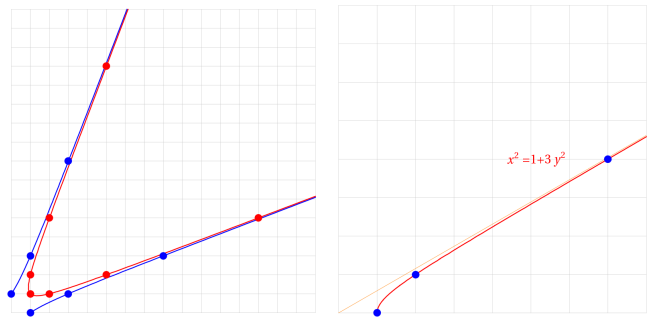

FIGURE 19. The integer lattice points on $x^2+y^2-3xy = 1$ are the even Fibonacci pairs (q_{2n}, q_{2n+2}) and integer lattice points on $x^2+y^2-3xy = -1$ are the odd Fibonacci pairs (q_{2n-1}, q_{2n+1}). The second picture shows the Pell equation.

The fact that the $x^2+y^2-3xy = 1$ has the integer lattice point solutions $(x, y) = (q_{2n}, q_{2n+2})$ implies that the relation $p = q_{2m}$ is **Diophantine**. The later fact was used in 1970 by Matiyasevich to finish the proof of Hilbert's 10th problem on solutions to Diophantine equation and complete the **Davis-Putnam-Robinson-Matiyasevich** theorem.

16. Birkhoff's ergodic theorem

Lets look at bounded random variables $X_n = g(n\alpha)$ defined by like $g(t) = \cos^2(\pi t)$. Because the expectation of the random variables is positive, and by Birkhoff's ergodic theorem $S_n(t)/n \to \int_0^1 g(t)\ dt = const$, we have a linear growth. A functional analytic version of the ergodic theorem is **von Neumann's mean ergodic theorem** which tells that the time averages $\frac{1}{n}\sum_n U^n$ converge in the strong operator topology to the projection onto the eigenspace of $\lambda = 1$. Lets define the **rescaled graph** $s_n(x) = S_{[nx]}/n$ on $[0, 1]$. Birkhoff's ergodic theorem assures that $s_n(x)$ converges to the linear function Mx almost everywhere and Oxtoby refined this in 1952 [30] so that in a strictly ergodic case, we have convergence for all initial points. Birkhoff's theorem holds for any measure-preserving ergodic transformation $T : [0, 1] \to [0, 1]$. Besides extreme cases like $T(x) = x + \alpha$ mod 1 and irrational α, or the Bernoulli process $T(x) = 2x$ mod 1 in which case the random variables X_k are independent, we can also have have more complicated cases like $T(x, y) = (x + \alpha, y + x)$ on the 2-torus and define $g(x, y) = \cos^2(\pi x)$. Also this T could be rewritten as a transformation on $[0, 1]$. The analysis in the last case could be more difficult however

since even the transformation is ergodic, the spectrum is mixed: there are absolutely continuous and discrete components.

FIGURE 20. John Oxtoby and John von Neumann

17. Law of iterated logarithm

If the expectation satisfies $M = \int_0^1 g(t)\, dt = 0$, then we can look at the growth of S_n. By Birkhoff's theorem, the growth rate must be smaller than linear. It is known that it can be arbitrarily close to linear. There is known that there can not be any general law which bounds the growth rate of S_n in the zero mean case except that $S_n = o(n)$ [23]. In the Bernoulli case, where the X_k are IID random variables, where the law of large number has been considered first by Jacob Bernoulli, the growth rate is $\sqrt{t/2}\log\log(t)$. This result is due to Khintchine (1924) and Kolmogorov (1929). Using the functions $s_n(x)$ we can reformulate this as that the graphs of $s_n(x)$ accumulate and have as a subset of the plane an accumulation point which is the modified parabolic region.

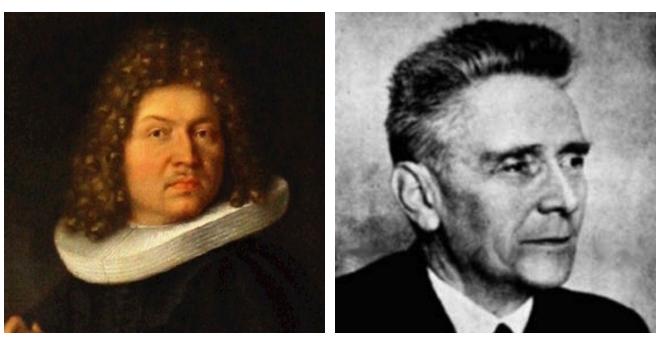

FIGURE 21. Jacob Bernoulli and Aleksandr Khinchin

18. Denjoy-Koksma theory

In the case of an irrational rotation $T(x) = x + \alpha$, the growth depends on Diophantine conditions of the rotation number. If there is a constant

C so that $|p\alpha - q| \leq Cq$, for all p, q and g has bounded variation $V(g) = \sup_P \sum |g(x_{i+1}) - g(x_i)|$, then $S_n \leq \log(n)V(g)$ for all n and there is a sequence of integers q_n for which S_n is bounded. More generally, if α is **Diophantine** of type $r > 1$, meaning $|p\alpha - q| \leq Cq^r$, then

Theorem: $S_n = \sum_{k=1}^n g(k\alpha)$ satisfies $|S_n| \leq Cn^{1-1/r} \log(n) \mathrm{Var}(g)$

Also here, in general, there is no bound on the growth rate for any irrational α, except the trivial bound $S_n \leq nC$. We can imagine taking a real Liouville number extremely close to the rationals leading to very long periods with linear growth, then again long periodic with linear decay then a much longer period with linear growth again etc.
Assume α is Diophantine of type $r > 1$ and g is of bounded variation and $\int_0^1 g(x)\, dx = 0$. Jitomirskaya refined Denjoy-Koksma:

FIGURE 22. Jurjen Koksma and Svetlana Jitomirskaya

19. The KAM Theorem

The Chirikov twist map $T(x, y) = (y - c\sin(\pi x), x)$ on $T^2 = R^2/Z^2$ appears from the recursion $x_{n+1} - 2x_n + x_{n-1} = c\sin(\pi x_n)$ which is a time discretization of the pendulum equation $x'' = c\sin(\pi x)$. The system is also called the **Frenkel-Kontorova model**. When looking at the problem to find invariant curves one looks for a **circle map** $q : \mathbb{T} \to \mathbb{T}$ solving the functional equation

$$F(q) = Lq - V(q) = q(t + \alpha) - 2q(t) + q(t - \alpha) - c\sin(\pi q(t)) = 0 \ .$$

With such a solution the map $\phi : T^1 \to T^2 : t \to (q(t), q(t - \alpha))$ conjugates the irrational rotation on T^1 with the map T restricted to the image of ϕ. **KAM theorem**:

Theorem: If α is Diophantine and $|c|$ is small, then there is a real-analytic solution q of $F(q)(x) = q(x+\alpha) - 2q(x) + q(x-\alpha) - c\sin(\pi q(x)) = 0$.

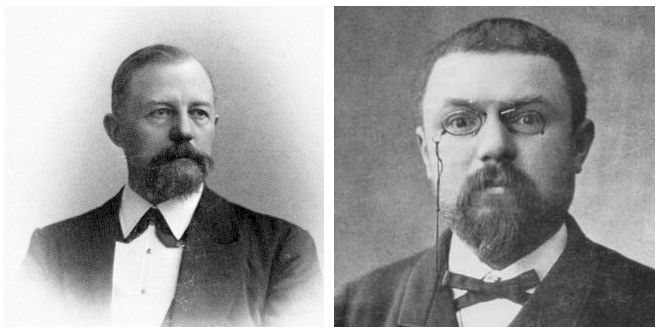

FIGURE 23. Anders Lindstedt and Henri Poincaré

20. KAM DIFFICULTIES

There are some difficulties when trying to solve the functional equation. The operator $L = F'(q)$ can for $c = 0$ be inverted on smooth functions q, as its Fourier transform is diagonal, after the Newton step $q - F(q)F'(q)^{-1}$ the smoothness will have decreased. An other difficulty is that for positive c, the operator $F'(q)$ is in the Fourier picture only a **Toeplitz operator** whose invertibility relies on exponential decay of the Green functions. For $c = 0$, the operator $F'(0)$ is trivially localized and this most likely persists also for small c as the Mathieu story, a simpler Jacobi case, has shown. But we have a deterministic fixed operator and not a probability space of operators for which we have almost all statements. As Kolmogorov-Arnold and Moser have shown, these difficulties can be washed away by combining the Newton step with a smoothing. Even if $F'(q)$ should have a zero eigenvalue, the smoothing will wash the eigenvalue away. John Neuberger's theorem helps but we have still the problem to estimate the norm of $(F'(q))^{-1}$ on a dense set q's in a small neighborhood of $q = 0$ in the Banach space.

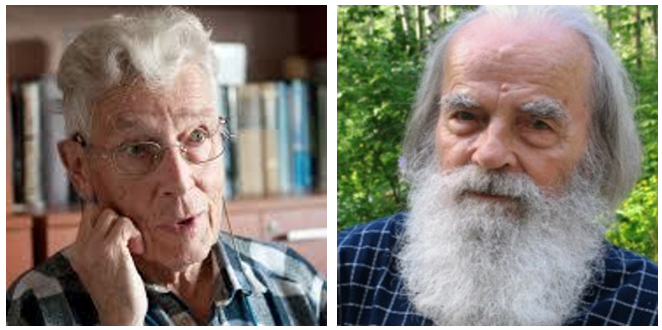

FIGURE 24. Michael Henon and Boris Chirikov

21. Variational problem

This is a variational problem as the equation for q is the Euler equation to a variational problem given by the Percival functional $\mathcal{L}(q) = \int_0^1 (q(t+\alpha) - q(t))^2/2 + c\cos(\pi q(t))/\pi \; dt$. **Aubry-Mather theory** shows that this variational problem has solutions for all c but then q is no more smooth. If L were invertible, we could invoke the implicit function theorem and assure that solutions persist. This can be used in hyperbolic situations, where it is also called the **anti-integrable limit** of Aubry-Abramovici [2] which applies for large c.

FIGURE 25. Serge Aubry and John Mather

22. Small Divisors

For $c = 0$, a solution to $F(q)(x) = q(x+\alpha) - 2q(x) + q(x-\alpha) = 0$ is $q_0(x) = x$. To study the functional derivative $L = F_q(0,q) = q(x+\alpha) - 2q(x) + q(x-\alpha)$, we look at the equation in a Fourier basis. For $q(x) = \sum_{n\neq 0} c_n e^{inx}$, we have

$$Lq = \sum_n c_n(e^{in\alpha} - 2 + e^{-in\alpha})e^{inx} = \sum_n 2c_n(\cos(n\alpha) - 1)e^{inx}$$

so that

$$L^{-1}q = \sum_{n=1}^{\infty} \frac{c_n}{2(\cos(n\alpha)-1)} e^{inx} \ .$$

We see the appearance of **small divisors** $V_n = 2(\cos(n\alpha)-1)$. The operator $K = -L^{-1}$ is diagonal and a finite dimensional approximation K_n satisfies $\log(\det(K_n)) = \sum_{k=1}^{n} \log(2-2\cos(k\alpha))$. We started with this Birkhoff sum.

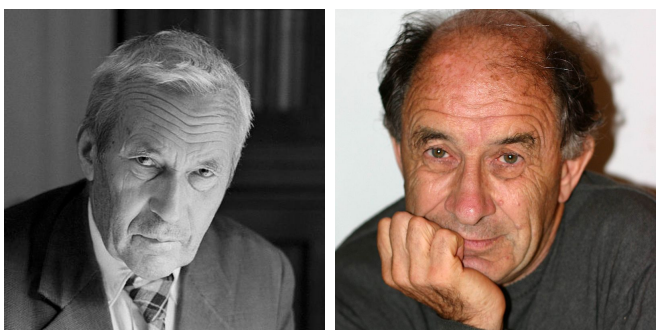

FIGURE 26. Andrey Kolmogorov and Vladimir Arnold

23. NEUBERGERS THEOREM

In 2007, John Neuberger told me about his implicit function theorem [28]. The smooth version goes as follows: let F be a smooth map from a Banach space X onto itself. Assume that X is compactly embedded in a second Banach space Y. Assume there is a dense set W in $B_r(q_0)$, so that for $q \in W$, there is a $h \in B_r(0)$ so that $F'(q)h = -F(q_0)$. Then there is a $q \in B_r(q_0)$ with $F(q) = 0$.

In one dimension and $q_0 = 0$, Neuberger's implicit function theorem applies to the situation that if $h(x) = -f(0)/f'(x)$ has absolute value $< r$ for a dense is $|x| < r$, then there is a root of f near 0. In other words, if $|f'(x)| > r|f(0)|$ for all $|x| < r$, then we have a root. Assume $r = 1$, and $|f'(x)| > |f(0)|$ for all $|x| < 1$, then we have a root.

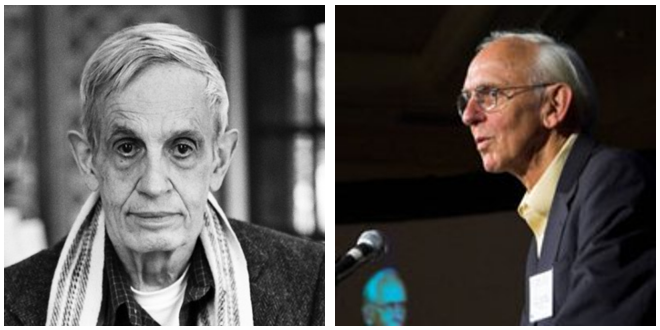

FIGURE 27. John Nash and John Neuberger

24. NEUBERGER FOR KAM?

In the summer of 2008, John Lesieutre explored in a PRISE project [21] whether the theorem is strong enough to prove the twist map theorem. The goal looks simple. Take the golden ratio α. Find for small c a function q close to $q(x) = x$ such that

$$F(q) = q(t+\alpha) - 2q(t) + q(t-\alpha) - c\sin(\pi q(t)) = 0 \ .$$

This could be written as a fixed point problem $q = G(q) = (q(t+\alpha) + q(t-\alpha) - c\sin(\pi q(t)))/2$. but no fixed point theorem bites as G is not a contraction nor leaves a convex set invariant.
Neuberger's theorem can be compared with Newtons method, where we replace q with $q+h$, where $h = -F'(q)^{-1}F(q)$ or $-F(q) = F'(q)h$. The theorem assumes that the iteration step can be done on a dense set. The conclusion is that there is a fixed point. Neuberger's theorem implies the classical implicit function theorem If $F(q_0, 0) = 0$ and $F_q(q_0, 0)$ is invertible, then for any small enough $|\epsilon|$, there exists $q \in B_r(q_0)$ such that $F(q, \epsilon) = 0$.

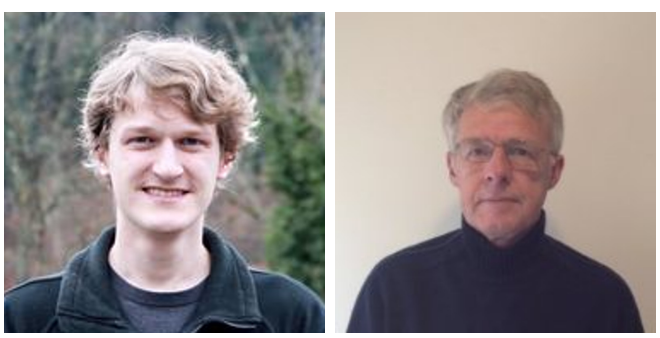

FIGURE 28. John Lesieutre and Volkert Tangerman

25. Green functions

What is needed to prove the KAM theorem? The Hessian matrix L at a critical point of the Percival functional is a Toeplitz matrix, a bounded operator L_{nm} such that $L_{n,n+k}$ decays exponentially with $k \to \infty$. For $q = 0$, the operator L is diagonal and it is invertible on the subspace of $l^2(N)$ which exponentially decay. For nonzero q, there is a heavy machinery of Green functions which shows that the operator is still invertible in general. There are two difficulties: the theory of almost periodic matrices assures the invertibility only for almost all x. The second difficulty is that we need to have a Banach space in which we can estimate $F'(q)$ on a dense set. Both are not unreasonable because the operators have in general pure point spectrum and because one has shown in Jacobi cases that the Lyapunov exponent (the decay rate) is continuous in some situations. But the difficulties look not easy as the theory of the almost Mathieu operator shows. In the KAM case, the operators are Toeplitz operators. But Neuberger's theorem would be highly intuitive and practical to compute q as it essentially boils down to a Newton step. There is an additional idea in the Nash-Moser implicit function theory approach: after each iteration step, a smoothing operation is performed. The reason is that inverting L brings us into a larger Banach space. Controlling these iterations is what makes KAM difficult. Toeplitz matrices of this type have been studied in [6]. One difficulty of applying this however that we need to know the situation for a specific operator and not for almost all parameters in a probability space.

Figure 29. Otto Toeplitz and Jean Bourgain

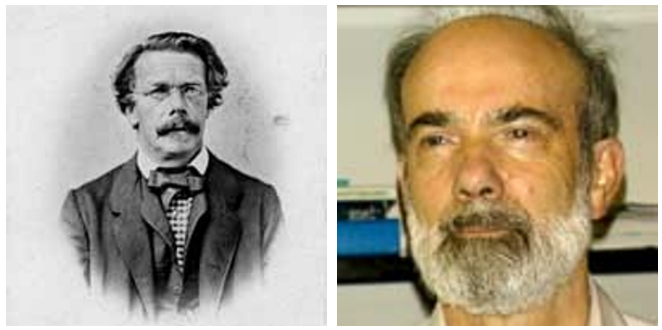

FIGURE 30. Ludwig Otto Hesse (1811-1874) and Ian Percival

26. Determinants

The work of summer of 2008 was not enough. We could not use Neuberger's theorem to prove the KAM theorem yet. Either Neubergers theorem is not strong enough, or we were not and the later is well possible. Neuberger's theorem only requires that we can invert $F'(q)$ on a dense set of q''s. This looks like a perfect fit for KAM as we can invert F' in the Diopantine case for a dense set of real analytic functions. Here is the difficulty: while for $c = 0$, the operator F' is a Jacobi matrix for $c > 0$, the operator is **Toeplitz matrix**

$$K = \begin{bmatrix} \dots & c_1 & c_2 & c_3 & c_4 & c_5 & \dots \\ c_1 & V_1 & c_1 & c_2 & c_3 & c_4 & \dots \\ c_2 & c_1 & V_2 & c_1 & c_2 & c_3 & \dots \\ c_3 & c_2 & c_1 & V_3 & c_1 & c_2 & \dots \\ c_4 & c_3 & c_2 & c_1 & V_4 & c_1 & \dots \\ c_5 & c_4 & c_3 & c_2 & c_1 & V_5 & \dots \end{bmatrix}$$

with $V_k = 2(\cos(k\alpha)-1)$ where the $c_i = \hat{\phi}(i)$ for a real analytic function ϕ for $i > 1$ decay exponentially. One studies such problem using Green functions. One has to establish the exponential decay of the Green function. This boils down to determinants. In the case $c = 0$, after truncating the matrix, we get the determinant $\prod_{k=1}^{n} |2\cos(2\pi k\alpha) - 2|$. Taking logarithms leads us to the function $g(x) = \log|2\cos(2\pi x) - 2|$. We had to study Birkhoff sums. What is needed to prove KAM? We have to be able to invert L on a dense set. Enough would be to show exponential decay of Green functions for $K(q)$ for a dense set of q. This would even lead to a nice numerical Newton method to find the solution q of the variational problem. There might be analytic tools like [7], but it does not look easy.

27. Holomorphic dynamics

Here is a motivation from [22]. Consider the nonlinear complex dynamical system $T(z,w) = (cz, w(1-z))$ in C^2, where $c = \exp(2\pi i\alpha)$ and α is the golden mean. This is one of the simplest quadratic systems which can be written down in $\mathbb{C}^2$. How does the orbit behave on the invariant cylinder $\{|z| = 1\} \times \mathbb{C}$ starting at $(c,1)$? We have $T^n(z,w) = (z_n, w_n) = (c^n z, w(1-z)(1-cz)\dots(1-c^{n-1}z))$ and $\log|w_n| = \sum_{k=1}^n \log|1-e^{2\pi ik\alpha}| = \frac{S_n}{2}$ for $(w_0,z_0) = (c,1)$ because $2\log|1-e^{ix}| = \log(2-2\cos(x))$. The study of the global behavior of the holomorphic map T in C^2 leads to the Birkhoff sum over the golden circle on a subset because for $r = |z| < 1$, where $g_r(x) = \log|1-re^{ix}|$ is real analytic and the Birkhoff sum converges by Gottschalk-Hedlund. It follows that for $r < 1$ the orbits have the graph of a function $A : \{|z| = r\ \} \to \mathbb{C}$ as an attractor. For $r = |z| > 1$, we have $|w_n| \to \infty$. So, all the nontrivial dynamics of the quadratic map happens on the subset $\{|z| = 1\ \} \times \mathbb{C}$.

28. Experiments with Tangermann

The function $g(x) = \log(2\sin(\pi x))$ has the property that $G' = \pi\cot(\pi x)$. We noticed that $S_{q_n}(\alpha)$ converges to a finite nonzero limit for $n \to \infty$ and that $S_k(\alpha)/\log(k)$ takes values in the interval $[0,2]$ and has a **limiting distribution** in $[0,2]$.

We only realized in [18] that $G(x) = \log(2-2\cos(2\pi x))/2$ is the Hilbert transform of the piecewise linear function $H(x) = x - [x] - 1/2$. It follows from $H(x) = \pi(x-[x]-1/2) = \arg(1-e^{2\pi ix})$ and $G(x) = \log(2-2\cos(2\pi x))/2 = \log|2\sin(\pi x)| = \log|1-e^{2\pi ix}|$. By Denjoy-Koksma, we know now that the Birkhoff sum of G grows like $C\log(g)$ if α is of constant type.

Figure 31. Ernst Hecke and David Hilbert

29. The Verschueren Mestel paper

[40] prove the observation obtained in [22]. The Birkhoff sum $S_n(\alpha)$ is the logarithm of the product $P_n(\alpha) = \prod_i k = 1^n 2\sin(\pi k\alpha)$. This product has been considered already by Sudler in 1964 [37] but looked at the maximal growth $|S_n| = \sup_\alpha |S_n(\alpha)|$ which Sudler showed that $|S_n|/n$ has a limit for any α. Sudler seemed have been motivated in particular by the theory of partitions as the Taylor expansion of the infinite product $\prod_{k=1}^\infty (1-x^k)$ is given by the Euler Pentagonal Number theorem as $\sum_{m=-\infty}^\infty (-1)^m x^{3m^2-m)/2}$. Lubensky showed in 1983 that $S_n(\alpha)/n \to 0$ almost everywhere. This follows from the fact that the Hilbert transform leaves the L^2-norm invariant and because by Denjoy-Koksma, the sum $\overline{S}_n/\log(n)$ stays bounded.

As the Verschueren-Mestel paper mentions, the study of the sin product goes back to Theodore Motzkin, Culbreth Sudler or Freiman Halberstam. Verschueren-Mestel show that P_{F_n} converges to some constant and that $S_n(\omega)/n$ stays in some interval.

Figure 32. Theodore Motzkin and Ben Mestel

30. Siegel theorem

The **Poincaré-Siegel theorem** in complex dynamics tells that if $f(z)$ is an analytic function with a fixed point $z = 0$ and $f'(0) = \lambda = \exp(2\pi i\alpha)$ with Diophantine α, then there exists $u(z) = z + q(z)$ such that $f(u(z)) = u(\lambda z)$ holds in a disc of radius ϵ around 0. The conformal map $u(z) = z + q(z)$ conjugates f to its linearization at 0. The result applies to the function function $f(z) = \lambda z + z^2/2$ for example. The Siegel disc at 0 is the maximal region on which on still has a conjugation to a rotation.

A key is the following: given a function $f(z) = \lambda z + g(z)$ which is analytic in the unit disc. For small $\epsilon > 0$, the Schröder equation $\lambda z + g(z + q(z)) = q(\lambda z)$ has a solution q which is analytic in the disc of radius ϵ. To solve $F(q) = q(\lambda z) - \lambda q(z) - g(z + q(z)) = 0$, Take the Banach space X of all analytic functions q on $D(2\epsilon)$ with sup norm satisfying $q(0) = q'(0) = 0$. It is compactly embedded in the Banach space Y of analytic functions on $D(\epsilon)$ by the **Arzela-Ascoli theorem**. As the origin in X, we take the function $q_0(z) = 0$. Let W the dense set of all polynomials, $v(z) = \sum_{n=2}^N v_n z^n$ in X. With

$$Lu = F'(q)u = u(\lambda z) - \lambda u(z) - g'(z + q(z))u(z) \ ,$$

we have to solve $(Lu)(z) = -F(0) = -g(z) = \sum_{n=1}^\infty g_n z^n$. With $g'(z + q(z)) = \sum_{n=1}^\infty c_n z^n$. We have

$$(Lu)_n = \lambda^n u_n - \lambda u_n - \sum_{k+l=n} c_k u_l \ .$$

In this Taylor basis,

$$L = \begin{bmatrix} V_1 & 0 & 0 & 0 & 0 & 0 \\ c_1 & V_2 & 0 & 0 & 0 & 0 \\ c_2 & c_1 & V_3 & 0 & 0 & 0 \\ c_3 & c_2 & c_1 & V_4 & c_1 & 0 \\ c_4 & c_3 & c_2 & c_1 & V_5 & 0 \\ c_5 & c_4 & c_3 & c_2 & c_3 & \dots \end{bmatrix} \ ,$$

where the diagonal matrix D has entries $V_n \geq C/n^2$. If $q \in W$, the side diagonals decay arbitrarily fast. The side diagonal of L^{-1} decays like ϵ^n. We find $h \in X$ such that $Lh = g$.

FIGURE 33. Carl Ludwig Siegel and Ernst Schroeder

31. Some calculus

Lets again take $g(x) = \cot(\pi x)$ and let G be the anti derivative of g so that $2\pi G(x) = \log(2 - 2\cos(2\pi x)) = 2\log(2\sin(\pi x)) = 2\log|1 -$

$e^{2\pi i x}|$. The **Euler's product formula** for the sinc-function $\text{sinc}(\pi x) = \frac{\sin(\pi x)}{x\pi} = \prod_{k\neq 0} \frac{x-k}{k}$ gives with $2\sin(\pi x) = \exp(\pi G(x))$ the Euler's formula for $g(x) = G'(x)$. The identity $2(1-\exp(ix))^{-1} = 1 + i\frac{\sin(x)}{1-\cos(x)} = 1 + i\cot(\frac{x}{2})$ relates the cotangent Birkhoff sum with a Birkhoff sum studied by Sinai and Ulcigrai [41].

FIGURE 34. Yakov Sinai and Corinna Ulcigrai

32. KUBERT RELATIONS

The fact that $\cot(\pi x)$ satisfies the identity $(1/n)\sum_{k=1}^n g(t+k/n) = g(t)$ seems have been discovered independently by many. I also had been excited to discover this experimentally and called it a solution to the **Birkhoff renormalization equation** [20]. I had wondered whether the cot function is besides the constants the only solutions but could not prove it. Jeff Lagarias has pointed out to me that the relation is called **Kubert relation** (named after Daniel Kubert (1947-2010) and that John Milnor has proven their uniqueness in 1983 [27]. Kubert worked first alone and later with Serge Lang on the functional equation

$$\frac{1}{n^s}\sum_{k=1}^n g(t+\frac{k}{n}) = g(t) \ .$$

For $s = 1$, $\cot(\pi x)$ the only odd function up to a multiplication with a constant and the constant function 1 spans all the even functions satisfying the relation. For $s = 0$, the identity is solved by $\log(2\sin(\pi x))$ which is the cyclotomic identity as well as its Hilbert transform $x - 1/2$. For $s = -2$ we have $\csc^2(\pi x)$ as well as the symmetric Hurwitz zeta function $\zeta_2(x) - \zeta_2(1-x)$. Milnor shows in [27] that all solutions can be obtained by taking anti-derivatives or derivatives. They are all interesting functions like Bernoulli polynomials. The story of the Kubert relations shows why the special functions are so interesting for Birkhoff

sums. If we take a rational α, then the sum is explicitly solved. In some sense, the Birkhoff sums become integrable.

FIGURE 35. Adolf Hurwitz and Serge Lang

33. Growth of the determinant

If α is Diophantine of bounded type and we take the Birkhoff sum for G, then there exists a constant C such that for almost all x, we have $S_n \leq C \log(m)$. The reason is that $G(x) = \log(2 - 2\cos(2\pi x))/2$ is the Hilbert transform of $H(x) = x - [x] - 1/2$ because $H(x) = \pi(x - [x] - 1/2) = \arg(1 - e^{2\pi i x})$ and $G(x) = \log(2 - 2\cos(2\pi x))/2 = \log|2\sin(\pi x)| = \log|1 - e^{2\pi i x}|$. Now use the Denjoy-Koksma theory.

We can also see that on the level of Fourier transform as $H(x) = -\sum_{n=1}^{\infty} \sin(2\pi n x)/n$ and $G(x) = -\sum_{n=1}^{\infty} \cos(2\pi n x)/n$ so that $G + iH = \sum_n e^{2\pi i n x}/n$. This relates with the polylogarithm $L(z,s) = \sum_{n=1}^{\infty} z^n/n^s$, an example of a random zeta function.

While $G'(x) = \pi \cot(\pi x) = \pi g(x)$, the derivative of H is only defined as a distribution. There is no corresponding Hilbert dual result therefore for the cot function. And indeed, there is logarithmic bound for the Birkhoff sum of cot.

34. Dirichlet series

Given a stochastic process X_k, one can look at the **random Dirichlet series**

$$\zeta(s) = \sum_k X_k e^{-\lambda_k s}$$

In the case $\lambda_k = k$ it produces the **random Taylor series**

$$\sum_k X_k z^k$$

with $z =^{-s}$ in the case $\lambda_k = \log(k)$ we get the **random zeta functions** $\sum_k X_k/k^s$.
The case when g has zero mean is the interesting case as $X = M$ just adds a standard Riemann zeta function. In [21], we looked at the case when X_k were obtained from an irrational rotation. In the zeta function case we proved that if α is Diophantine and g is real analytic, then the random zeta function has an analytic continuation onto the entire complex plane.

FIGURE 36. Peter Dirichlet and Brook Taylor

35. ZETA FUNCTIONS

There are various zeta functions [38]. The major versions either use spectral properties or periodic orbits. The former is a quantum version, the later a classical version. They can be related or mixed. For manifolds for example, the length spectrum of periodic orbits is linked to the spectrum of the Laplacian. For a subshift f of finite type defined by a matrix A one has the **Bowen-Lanford formula** $\exp(\sum_n z^n \mathrm{Fix}(f^n)/n) = \det(1 - zA)^{-1}$. The right hand side is **spectral**, the left hand side **dynamical**. For graphs, there is a spectral version $\sum_k \lambda_k^{-s}$ and a orbit version, the **Ihara zeta function** $\prod_p (1 - z^{|p|})$, where p runs through all prime paths of length $|p|$. The classical Riemann zeta function can be seen as the spectral zeta function of the Dirac operator. The "Golden key" formula of Euler $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$ shows a connection with "primes". The Euler golden key relates already a quantum with a classical concept. This is already present in determinants. The formula $\log(\det(\exp(-sA))) = \sum_k \lambda_k^{-s}$ relates a spectral property on the right with a path integral over closed simple graphs.

FIGURE 37. Rufus Bowen and Oscar Lanford III

36. Classical and quantum

If A is a $n \times n$ matrix, we can look at the function

$$\det(1 + xA) = \sum_k x^k \mathrm{tr}(\Lambda^k A)$$

related to the characteristic function of A. The expression $\mathrm{tr}(\Lambda^k A)$ is a count over "path integrals" which are here periodic orbits. Using the Taylor series $\log(1 + x) = x - x^2/2 + x^3/3 - x^4/4$ .... we have

$$\det(1 + xA) = \exp(\mathrm{tr}(\log(1 + xA))) = \exp(\sum_{n=1}^{\infty} (-1)^{n+1} x^n \mathrm{tr}(A^n)/n) \ .$$

If A is the Laplacian of a graph, then $\det(1 + xL)$ is for $x = 1$ the number of **spanning forests** by the **Chebotarev-Shamis** theorem [31, 19]. Since $\mathrm{tr}(A^n)$ counts the number of closed paths of length n in the graph, where a loop leads to a penalty factor $-d(x)$ with degree $d(x)$, one can see that it is a **path integral** and $\det(1 + xA)$ as a generating function for the closed paths. In dynamical system theory, where $f : M \to M$ is a map, the **Artin-Mazur zeta function** is defined as $\exp(\sum_{n=1}^{\infty} x^n |\mathrm{Fix}(f^n)|)/n$. Ruelle combined the **Fredholm determinant** and **Artin-Mazur Zeta function** to the **Ruelle zeta function**

$$\exp(\sum_{n \geq 1} x^n \sum_{p \in \mathrm{Fix}(f^n)} \mathrm{tr}(A^n(p))) \ ,$$

where $A^n(p) = A(f^{n-1}p) \cdots A(p)$ is the cocycle matrix product of a matrix-valued function $A : M \to M(n, R)$ over the dynamical system $f : M \to M$. If M is the one point space, then this reduces to the Fredholm determinant. If A is the 1×1 matrix 1, then this is the Artin-Mazur zeta function. The spectral version is obtained by looking

at the positive eigenvalues $\lambda_1, \dots, \lambda_n$ of A and defining a zeta function

$$\exp(\zeta_A(s)) = \exp(\sum_k \lambda_k^{-s}) = \exp(\mathrm{tr}(A^{-s})) = \det(\exp(A^{-s})) \ ,$$

where A^{-s} is understood by diagonalization and restricting to the complement of the zero eigenspace. The concept of Zeta function naturally combines the concept of determinants, closed loops and rooted spanning forests. It is at the heart of understanding the relation between quantum and classical properties.

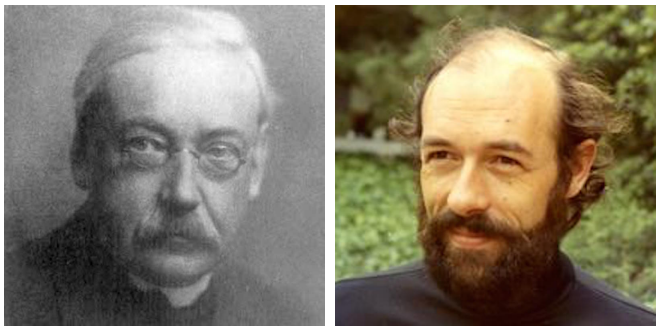

FIGURE 38. Erik Fredholm and David Ruelle

FIGURE 39. Michael Artin and Barry Mazur

37. Baby Riemann hypothesis

The zeta function of a geometric object with exterior derivative d is the $\sum_{\lambda>0} \lambda^{-s}$ where λ runs through the positive eigenvalues of $D = d + d^*$. (The negative eigenvalues are a mirror and do not lead to additional information, just argument ambiguity when defining the $-s$'th power) This definition works for manifolds or graphs. For the circle for example, where $D = -i\frac{d}{dx}$ has the eigenvalues n and eigenfunctions e^{inx}, we get the classical Riemann zeta function.

In the case of a circular graph $\zeta_n(2s)$ agrees with the Zeta function of the Laplacian

$$L = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{bmatrix}$$

We have $\zeta_n(s) = \sum_{k=1}^{n-1} 2^{-s} \sin^{-s}(\pi \frac{k}{n})$. It is a Birkhoff sum $\sum_{k=1}^{n-1} g(\pi k/n)$ with the complex function $g_s = (2\sin(x))^{-s}$. We proved in [20] that the roots of $\zeta_n(s)$ converge to the line $\mathrm{Re}(z) = 1$.

Theorem: The roots of $\zeta_n(s)$ converge to the line $Re(s) = 1$.

This means that the zeta function of the Laplacian converges to the line $\mathrm{Re}(z) = 1/2$. This result has nothing to do with the Riemann Hypothesis as we deal with concrete analytic functions and not with the Riemann zeta function which needs analytic continuation to access the function values on the critical axes.

FIGURE 40. Pierre-Simon Laplace Bernhard Riemann,

38. The Smith determinant

In February 2015, Omar Antolin showed me a proof of a rediscovery of Juan Jose Alia Conzalez that the determinant of the matrix $A_{ij} = \gcd(i,j)$ has determinant $\prod_{i=1}^n \phi(i)$. The determinant $A_{ij}(s) = \gcd(i,j)^s$ has also an explicit formula $\zeta_n(s) = \prod_{k=1}^n k^s \prod_{p|k}(1-1/p^s)$, a **Jordan totient function**. [35]. The Euler golden key $\zeta(s) = \sum_p (1-$

$1/p^s)^{-1}$ reminds about zeta functions. The roots of $\zeta(s) = \det(A_{ij}(s))$ as a function of s are on the imaginary axes because $(1 - e^{-s\log(p)}$ has roots at $2\pi ki/\log(p)$. The determinant is called the **Smith determinant**, named after Henry J. S. Smith, a remarkable mathematician who also is known for the **Smith normal form** of a matrix as well as the discoverer of the Cantor set. The Jordan totient function can not only be studied as a function of s. Because it is a determinant, one can look at the eigenvalue distribution of the matrices $A_{ij}(s)$. The distribution looks pretty regular. Since for complex s, the operators are no more self adjoint, one needs to look at the spectrum in the complex plane. Now one can look at $\log(\zeta_n(s))$ which is a Birkhoff sum of $g(k) = s\log(k) + \sum_{p|k} \log(1 - 1/p^s)$, only that $g(k)$ is not dynamically generated by a transformation. But there is still some almost periodicity coming in. The roots of $\zeta(s)$ on the imaginary axes are a union of almost periodic sets. Camille Jordan from the Jordan totient function is also remembered because of the Jordan curve theorem and Jordan normal form but not Gauss-Jordan elimination.

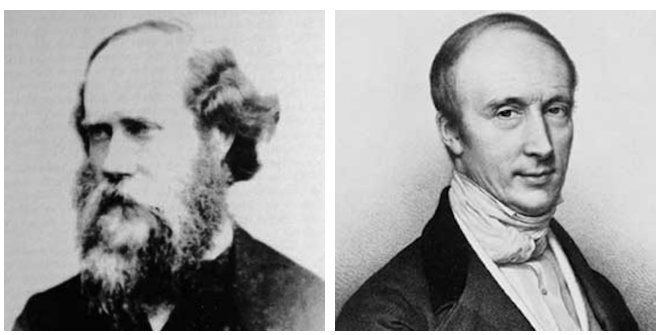

FIGURE 41. Henry Smith and Camille Jordan

39. Baconians and Cartesians

The topic of golden rotations is on one side very concrete and special, on the other hand touches on very general topics like correlated stochastic processes with infinite variance. Both the topic of correlated stochastic processes as well as the study of high risk = infinite variance situations will certainly both become larger fields of probability theory. In the rest, I allow me to comment on the dichotomy of "example" versus "generality" which is present in the story of "golden rotation".

In a foreword of [29], Freeman Dyson indicated, how the views of **Francis Bacon** and **René Descartes** produced both a polarizing tension

as well as a cross fertilization in science. According to Dyson, the Baconians are travelers, exploring science using examples and collecting samples, while the Cartesians stay at home and deduce the truth using axioms and pure thought. In mathematics, the pendulum between Baconian and Cartesian domination regularly swings for and back. Georg Cantor and Alexander Grothendieck were Cartesian mathematician, Henry Poincaré or Paul Erdös were Baconians. Dyson points out that French mathematics was mostly dominated by Cartesians while English mathematics adopted mostly the Baconian point of view, but Dyson also makes clear that this cliché is not universal: Marie Curie was a Baconian, while Isaac Newton was at heart a Cartesian.

FIGURE 42. René Descartes and Francis Bacon

40. Rotating tops and Fourier basis

Having pursued my own experimental Baconian attempts using computer explorations in high school I turned heavily to Cartesian views first in college, reading Bourbaki and Wittgenstein, admiring the categorical, general approaches. Some of my teachers corrected this drift: **Eugene Trubowitz** told me after I presented in his seminar a monodromy theorem in a categorical way the following advise: "Don't look at the problem with a telescope!", in a mechanics exam, I explained, following an appendix of [1], the motion of the dynamics of the **n-dimensional top**. This required to look at tangent bundles of cotangent bundles of Lie groups, Fröhlich drew the Eiffel tower onto the blackboard, placed a top on it and asked: "how does it move?" and let me go with the words: "Herr Knill, the Bourbaki times are over".

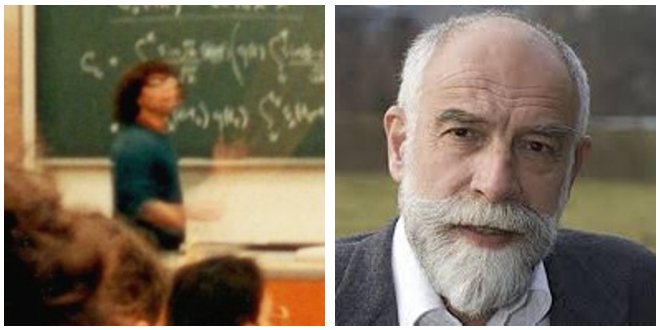

FIGURE 43. Eugene Trubowitz and Jürg Fröhlich

The Baconian-Cartesian encounter happened during a final exam, where Corneliu Constantinescu, a true incarnation of Descartes who taught with extreme clarity and Jürgen Moser, a reborn Francis Bacon who disliked abstraction for the sake of abstraction, examined me both orally. All three together were joined in an office, Constantinescu started to examined me in real analysis, asked me for a proof of the Radon-Nykodin theorem. I could do that very well and prove the theorem in all details like a machine. Constantinescu would interrupt if I use a lemma and ask me to prove the lemma, and so on. I knew every word of the lectures by heart and had no problem reproducing the proofs. When it was Moser's turn to examine me in functional analysis, he started me with the question: "What is the spectrum of the Fourier transform?" I was stunned because this question had never come up, nor had I ever thought about it. I could have produced the proof of the spectral theorem for normal operators but to look at the map $T : f \to \hat{f}$ itself as an operator had not occurred to me before. I think I could get a partial answer like that there are Gaussian eigenfunctions to the eigenvalue 1, but I did not get through stating the entire spectrum. The answer is $\sigma(G) = \{1, -1, i, -i\}$ since T is unitary and $T^4 = 1$ and explicit Hermite eigenfunctions can be written down.

FIGURE 44. Corneliu Constantinescu and Jürgen Moser

41. An intriguing question

During an other final examinations I once again was examined orally by a pair of Baconian and Descartian examiners. These final exams were important as it was still possible and happened regularly that students would fail the final exams and have to leave ETH without a degree. I had taken several logic courses from **Ernst Specker** and **Hans Läuchli**. These two examiners liked to sit comfortably on a couch, while the student was peppered and grilled on the blackboard. Specker was clearly a Baconian mathematician while Läuchli was more of a Descartian. I had learned also single and multivariable calculus from Läuchli, but was examined by him in **Non-standard analysis** on that occasion, a course I had taken from him.

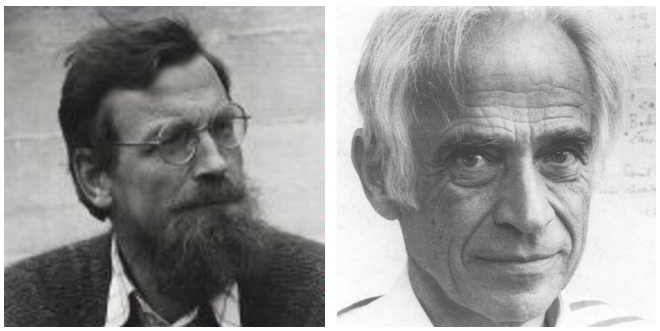

FIGURE 45. Hans Läuchli and Ernst Specker

The exam started with Specker, who would test me in logic. His question completely stunned me: "Herr Knill, verzelled Sie üs öppis" ("Mr. Knill, tell us something!") It is a startling examination question. Well, I had been excited about Goedel's incompleteness theorem and started to explain that theorem. After about 10 minutes, Specker turned to

his younger colleague and asked slowly in Swiss German: "Du Hans, weisch Du was dä Herr Knill doo macht?" ("Hans, do you know what Mr. Knill is doing there?"). Läuchli shook his had and replied equally in Swiss German: "No, I have no idea". Startled, I struggled through the rest but got saved by non-standard analysis, then my favorite topic. I guess that the initial part might also have just been staged for them to get out some entertainment from these exams. Specker was known for his special humor like announcing a talk about the "tractatus" of Wittgenstein, but then arrive with no intention to talk himself but handpick some of his unprepared non-math colleagues in the audience to have them discuss Wittgenstein in a circle surrounded by a delighted audience of people, who were spared the reaping to this rather cruel academic **"hunger game"**. Both Specker and Läuchli were great teachers. I learned two semester linear algebra from Specker and two semesters calculus from Läuchli. The lectures of Läuchli were extremely clear and precise, while Specker would not hesitate to improvise and interact or play with the class. Läuchli was brilliant and Specker was inspiring. A Specker quote reported by his student and friend Läuchli [13] brings it to the point: "Teaching mathematics to good students is like telling fairy tales to children. A world of its own is unveiled. Those who enter it can explore it further and even add to it." I appreciate today both the exposure of Descartian and Baconian approaches both in learning mathematics as well as well as in teaching it.

Figure 46. Kurt Gödel and Ludwig Wittgenstein

42. Generality versus examples

Cartesians use an axiomatic, deductive approach, Baconians like to look at cases and experiment. The beauty of simple examples of dynamical systems like the iteration of quadratic maps, playing billiards in planar convex tables or studying concrete systems like the Hénon map or the study operators like the Mathieu operator are typical Baconian

approaches. While this was popular 20-30 years ago, the Cartesian abstract point of view is now in full swing similarly when Julia and Fatou made the first steps in chaotic complex dynamics. We again live in a "neo Bourbaki" time. Clashes between the approaches have occurred again and again: the Cartesian Cantor was in opposition to Poincaré, a Baconian.

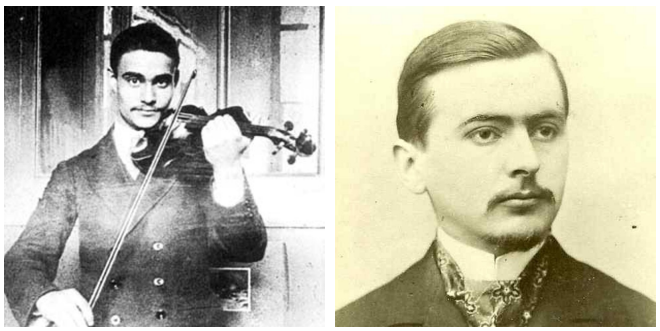

FIGURE 47. Gaston Julia and Pierre Fatou

In some theoretical physics areas like in string theory, the deductive way is completely detached from any experiments or even the outlook to ever be able see the structures in experiments. It can be seen as an invasion of Descartian ideas into physics. On the other hand, the field of experimental mathematics has entered mathematics and conquered part of mathematics. Many mathematicians including myself do lots of experiments while many physicists study very abstract mathematical structures. As Dirac already pointed out, the polarization of Baconian to Descartian is kind is a caricature also and often it is difficult to "classify" a mathematician. The topic has also been addressed in [9], where the cultures are addressed as "theory builders" and "problem solvers". Its hard to tell for example, in which category Gauss or Euler or Kolmogorov belong to, as they developed both theories in pure thought process but also experimented and played a lot with concrete structures and model problems.

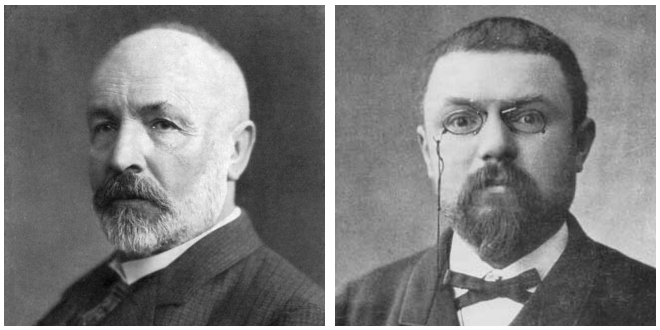

FIGURE 48. Georg Cantor and Henri Poincaré

In dynamics, a Cartesian approach is to look at the structure of dynamical systems and to investigate globally what happens. One can look for example, what happens generically in the class of all dynamical systems. There is generic ergodicity for example. Much of Smale's work is Descartian, as he outlined the general structure of dynamical systems, while much of Milnor's work is more of a Baconian nature, dealing with well chosen examples and problems. While also Smale worked with concrete examples (the horse shoe illustrates this), it is a generic feature of dynamical systems with topological entropy, exotic spheres, counter examples to the Hauptvermutung, or the work on one-dimensional dynamics which are very special in the work of Milnor. But good examples can make up an entire theory: iterating the quadratic map covers most essential features of polynomial map, iterating the Hénon map reveal much about the structure of dissipative or conservative maps, the Lorentz system provides lessons for general dynamical systems etc. Remarkably, both started in topology and moved to dynamical systems. Both approaches are valuable and also in the class room, each has advantages: the Baconian point of view can be more inspiring and motivating, while the Cartesian point of view is more organized and clear. It is the Swiss in me who does not like to take sides but stay neutral. I believe that any student should become exposed to both type of mathematics and mathematicians, as both approaches have their advantages both in research as well as in the classroom.

FIGURE 49. Stephen Smale and John Milnor

References

[1] V.I. Arnold. *Mathematical Methods of classical mechanics*. Springer Verlag, New York, second edition, 1980.

[2] S. Aubry and G.Abramovici. Chaotic trajectories in the Standard map. the concept of anti-integrability. *Physica D*, 43:199–219, 1990.

[3] A. Avila and S. Jitomirskaya. The ten martini problem. *Annals of Mathematics*, 170:303–342, 2009.

[4] J. Bellissard, A.Bovier, and J.-M.Ghez. Gap labelling theorems for one dimensional discrete Schrödinger operators. *Rev. Math. Phys*, 4:1–37, 1992.

[5] G.D. Birkhoff. *Aesthetic measure*. Harvard University Press, 1933.

[6] J. Bourgain. *Green's function estimates for lattice Schrödinger operators and applications*, volume 158 of *Annals of Mathematics Studies*. Princeton Univ. Press, Princeton, NJ, 2005.

[7] J. Bourgain and S. Jitomirskaya. Continuity of the Lyapunov exponent for quasiperiodic operators with analytic potential. *J. Statist. Phys.*, 108(5-6):1203–1218, 2002. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays.

[8] H.L. Cycon, R.G.Froese, W.Kirsch, and B.Simon. *Schrödinger Operators—with Application to Quantum Mechanics and Global Geometry*. Springer-Verlag, 1987.

[9] W.T. Gowers. Two cultures of mathematics. In P. Lax V.Arnold, M. Atiyah and B. Mazur, editors, *Mathematics: Frontiers and Perspectives*, 2000.

[10] R. K. Guy. *Unsolved Problems in Number Theory*. Springer, Berlin, 3 edition, 2004.

[11] G. H. Hardy and J. E. Littlewood. Some problems of diophantine approximation: The analytic character of the sum of a dirichlet' s series considered by hecke. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 3 (1), 1924.

[12] G. H. Hardy and J. E. Littlewood. Some problems of Diophantine approximation: a series of cosecants. *Bulletin of the Calcutta Mathematica Society*, 20(3):251–266, 1930.

[13] Gerhard Jäger, Hans Läuechli, Bruno Scarpellini, and Volker Strassen, editors. *Ernst Sepcker, Selecta*. Birkhäuser, Basel-Boston-Berlin, 1990.

[14] H. Kesten. Symmetric random walks on groups. *Transactions of the AMS*, 92:336–356, 1959.

[15] O. Knill. Factorisation of random Jacobi operators and Bäcklund transformations. *Communications in Mathematical Physics*, 151:589–605, 1993.

[16] O. Knill. Isospectral deformations of random Jacobi operators. *Communications in Mathematical Physics*, 151:403–426, 1993.

[17] O. Knill. Singular continuous spectrum and quantitive rates of weakly mixing. *Discrete and continuous dynamical systems*, 4:33–42, 1998.

[18] O. Knill. Selfsimilarity in the birkhoff sum of the cotangent function. http://arxiv.org/abs/1206.5458, 2012.

[19] O. Knill. A Cauchy-Binet theorem for Pseudo determinants. *Linear Algebra and its Applications*, 459:522–547, 2014.

[20] O. Knill. The zeta function for circular graphs. http://arxiv.org/abs/1312.4239, December 2013.

[21] O. Knill and J. Lesieutre. Analytic continuation of Dirichlet series with almost periodic coefficients. *Complex Analysis and Operator Theory*, 6(1):237–255, 2012.

[22] O. Knill and F. Tangerman. Selfsimilarity and growth in Birkhoff sums for the golden rotation. *Nonlinearity*, 21, 2011.

[23] U. Krengel. *Ergodic Theorems*, volume 6 of *De Gruyter Studies in Mathematics*. Walter de Gruyter, Berlin, 1985.

[24] Y. Last. Zero measure spectrum for the almost Mathieu operator. *Communications of Mathematical Physics*, 164:421–432, 1994.

[25] F. Ledrappier. Positivity of the exponents for stationary sequences of matrices. In *Lyapunov exponents (Bremen, 1984)*, volume 1186 of *Lecture Notes in Math.*, pages 56–73. Springer Verlag, Berlin, 1986.

[26] John N. Mather. Amount of rotation about a point and the Morse index. *Comm. Math. Phys.*, 94(2):141–153, 1984.

[27] J. Milnor. On polylogarithms, Hurwitz zeta functions and the Kubert identities. *Enseignement Mathématique*, 29:281–322, 1983.

[28] J. Neuberger. The continuous Newton's method, inverse functions, and Nash-Moser. *Amer. Math. Monthly*, 114(5):432–437, 2007.

[29] P. Odifreddi. *The Mathematical Century*. Princeton University Press, 2004.

[30] J. Oxtoby. Ergodic sets. *Bull. Amer. Math. Soc.*, 58:116–136, 1952.

[31] P.Chebotarev and E. Shamis. Matrix forest theorems. arXiv:0602575, 2006.

[32] D. Ruelle. Analyticity properties of the characteristic exponents of random matrix products. *Advances of Mathematics*, 32:68–80, 1979.

[33] D. Ruelle. Rotation numbers for diffeomorphisms and flows. *Ann. Inst. Henri Poincaré Phys. Théor.*, 42:109–115, 1985.

[34] B. Simon. Almost periodic Schrödinger operators IV. the Maryland model. *Annals of Physics*, 159:157–183, 1985.

[35] H.J.S. Smith. On the Value of a Certain Arithmetical Determinant. *Proc. London Math. Soc.*, S1-7(1):208, 1876.

[36] J.L. Snell. A conversation with joe doob. *Statistical Science*, 12(4), 1997. http://www.dartmouth.edu/ chance/Doob/conversation.html.

[37] C. Sudler. An estimate for a restricted partition function. *Quart. J. Math. Oxford Ser. (2)*, 15:1–10, 1964.

[38] A. Terras. *Zeta functions of Graphs*, volume 128 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 2011.
[39] H. Toda. *Theory of nonlinear lattices*. Springer-Verlag, Berlin, 1981.
[40] P. Verschueren and B. Mestel. On the growth of sudler's sine product at the golden rotation number. http://arxiv.org/abs/1411.2252, 2014.
[41] Ya.G.Sinai and C. Ulcigrai. A limit theorem for Birkhoff sums of non-integrable functions over rotations. In *Geometric and probabilistic structures in dynamics*, volume 469 of *Contemp. Math.*, pages 317–340. Amer. Math. Soc., Providence, RI, 2008.

Department of Mathematics, Harvard University, Cambridge, MA, 02138