12^2 beautiful mathematical theorems with short proofs

Jörg Neunhäuserer

Contents

1

Introduction

ABSTRACT: We present 12^2 beautiful theorems from almost all areas of mathematics with short proofs, assuming notations and basic results a graduate student will know.

MSC 2000: 00-01

In this notes we present some beautiful theorems of mathematics for the enjoyment of the reader. We include results in almost all areas of mathematics: set theory, topology, geometry, analysis and function theory, number theory, algebra, discrete mathematics, probability theory and dynamical systems. The total number of results we include is 144, fifty of this results are contained in list "The Hundred Greatest Theorems of Mathematics" presented by Paul and Jack Abad in 1999. [1]
Also there is no satisfactory notion of a beautiful mathematical theorem, we think that most well educated fellows, which are interested in mathematics, will agree that all theorems we choose are in fact beautiful. We only include theorems for which we are able to write down a short proof, which means one page at most. Of course our choice of theorems is motivated by our subjective taste. So we excluded theorems from the list not only because we have no short proof but also because the result seems to us not this beautiful. To be more prices we do not know short proofs for more than 30 theorem in the list of Paul and Jack Abad and we do not like around 20 theorem in the list so much. We were able include here more than 20 theorems with short proofs that certainly belong to our Top 100, and are not contained in the list.
Our approach is not self-contained, we include some but not all definitions that we use and we presuppose some elementary results and definitions in geometry, calculus and linear algebra. We think that all graduate students of mathematics or physics and most second year undergraduate students will be familiar with all the material we presuppose. Our proofs are not complectly formalized given main ideas and arguments and not all technical details. In some instances we only sketch the way the prove works. It is a good exercisers to students to fill in the details and perform calculations left the the reader. .

There are many beautiful theorems in mathematics for which we do not have a short and perhaps not even a beautiful proof. Just think about the resent proofs of Fermat's last theorem [26, 24] , the Poincare conjecture [17, 18, 19, 12], the four color theorem [3] or the classification of all finite simple groups, [27]. Such results are not at our focus here. Our aim is to collect knowledge for a good educational background, not discuss topics for specialists. For any mathematician the book may solve as a kind of test, if he or she really has good knowledge of well know short proofs of beautiful theorems. If not, we hope that our proofs are fairly easy to understand and to recognize and may thus increase present knowledge of the mathematical community. In addition we include remarks on related results with references for further reading.
We have no historical predations here, nevertheless the reader will find information in the footnotes who discovered a result. On the other hand the formulation and the proof of the theorems we present is in general not original.
At the end of our notes the reader will find a list of ten conjectures for the evolution of mathematics in the next millennium. Our list is quiet different from the millennium problems of the Claymath institute. We think that our problems are mostly more elementary and thus even harder to solve.

Jörg Neunhäuserer
Goslar, 2013

Standard Notations

We recall some standard notations in mathematics which are usually used in the book.

$\bigcup$	The union of sets
$\bigcap$	The intersection of sets
$\|A\|$	The cardinality of a set
$\sum$	The sum of numbers
$\prod$	The product of numbers
$\mathbb{N}$	The set of natural numbers
$\mathbb{Z}$	The set of integers
$\mathbb{Q}$	The set of rational numbers
$\mathbb{A}$	The set of algebraic numbers
$\mathbb{R}$	The set of real numbers
$\mathbb{C}$	The set of complex numbers
$\|a\|$	The modulus of a number
sin	The sinus function
cos	The cosine function
tan	The tangent function
cot	The cotangent function
log	The natural logarithm
e	The Euler number
π	The Archimedes constant
(x_i)	A sequence or family with index i
$\to$	A limit of a sequence
lim	Limits of sequences or functions
f'	The derivative of a function
$\int f$	The integral of a function

2

Set theory

2.1 Countable sets

We begin the chapter with the definition of countable and uncountable infinite sets.

Definition 2.1.1 *A set A is infinite if there is a bijection of A to a proper subset $B \subset A$[1], it is countable infinite, if there is bijective map from A to $\mathbb{N}$. An infinite set is called uncountable if this is not the case.* [2]

If you use an axiomatic set theory you have to suppose the existence of $\mathbb{N}$ (or any inductively ordered set) to make sense of this definition. Using the definition it is easy to prove a general result, that holds for all countable sets.

Theorem 2.1. *Finite products and countable unions of countable sets are countable.*[3]

$\mathfrak{Proof}$. Let A_i for $i = 1 \dots n$ be countable sets with counting maps c_i. Consider the map $C : \prod_{i=1}^{n} A_i \mapsto \mathbb{N}$ given by

$$C((a_1, \dots, a_n)) = \prod_{i=1}^{n} p_i^{c_i(a_i)},$$

where p_i are different primes. Since prime factorization is unique (see theorem 7.2) this map is injective. But each infinite subset of the natural numbers is countable, hence the product is countable as well. Now consider an infinite family of countable sets A_i for $i \in \mathbb{N}$ with counting maps c_i. We have

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} \{c_i^{-1}(n) | n \in \mathbb{N}\} = \{c_i^{-1}(n) | (n, i) \in \mathbb{N}^2\}.$$

This set is countable since $\mathbb{N}^2$ is countable. $\mathfrak{Q.E.D.}$

All students of mathematics are aware of following result, which shows that sets of numbers, that seem to have a different size, have the same size in the set theoretical sense.

[1] This definition is due to the German mathematician Richard Dedekind (1831-1916)
[2] This definition was introduced by German mathematician George Cantor (1845-1918).
[3] The result is due to George Cantor (1845-1918).

Theorem 2.2. *The integers $\mathbb{Z}$, the rationales $\mathbb{Q}$ and the algebraic numbers $\mathbb{A}$ are countable.*[4]

$\mathfrak{Proof}$. The integers are the union of positive and negative numbers, two countable sets, hence they are countable. The map

$$f : \mathbb{Q} \mapsto \mathbb{Z}^2 \qquad f(p/q) = (p, q)$$

is injective and $\mathbb{Z}^2$ is countable, hence $\mathbb{Q}$ is countable as well. There are countable many polynomials of degree n over $\mathbb{Q}$ since $\mathbb{Q}^n$ is countable. Each polynomial has finitely many roots, hence the roots of degree n are countable. $\mathbb{A}$ is countable as the countable union of these roots. $\mathfrak{Q.E.D.}$

2.2 Construction of the real numbers

Now we give a beautiful and simple definition of real numbers, only using the rational numbers $\mathbb{Q}$.

Definition 2.2.1 *A subset $L \subseteq \mathbb{Q}$ is a cut if we have*

$$L \neq \emptyset,\ L \neq \mathbb{Q},$$

$$L \text{ has no largest element},$$

$$x \in L \Rightarrow y \in L\ \forall y < x.$$

The set of real numbers $\mathbb{R}$ is the set of all cuts with order relation given by $\subseteq$, addition given by

$$L + M = \{l + m | l \in L,\ m \in M\}$$

and multiplication given by

$$L \cdot M = \{l \cdot m | l \in L,\ m \in M\ l, m \geq 0\} \cup \{r \in \mathbb{Q} | r < 0\}$$

for positive real numbers K, M, which can be extended to negative numbers in the natural way.[5]

To verify that $\mathbb{R}$, given by this definition is an ordered field, is rather uninteresting, see for instance [**?**] . But the following result is essential for real numbers:

Theorem 2.3. *Every non-empty subset of $\mathbb{R}$, that is bounded from above, has a least upper bound.*

[4] Again this fact was observed by George Cantor (1845-1918).
[5] This is the construction of the German mathematician Richard Dedekind (1831-1916)

Proof. Let $\mathbf{A} \subseteq \mathbb{R}$ be bounded by $M \in \mathbb{R}$. For each $L \in \mathbf{A}$ we have $L \subseteq M$. Let

$$\tilde{L} = \bigcup_{L \in \mathbf{A}} L.$$

Obviously $\tilde{L}$ is nonempty and a proper subset of $\mathbb{Q}$. If $\tilde{L}$ would have a largest element l, than $l \in L$ for some $L \in \mathbf{A}$ and l would have to be the largest element of L, a contradiction. If $r \in \tilde{L}$ and $s \leq r$ than $r \in L$ for some $L \in \mathbb{A}$ and $s \in L$ hence $s \in \tilde{L}$. This means that $\tilde{L}$ is a cut; $\tilde{L} \in \mathbb{R}$. Moreover $L \subseteq \tilde{L}$ for all $L \in A$ hence $\tilde{L}$ is an upper bound on $\mathbf{A}$. Also $\tilde{L} \subseteq M$. Since M is an arbitrary upper bound $\tilde{L}$ is the least upper bound. $\mathfrak{Q.E.D.}$

By this construction we obtain the usual representation of real numbers.

Theorem 2.4. *Every real number can be represented by a unique b-adic expansion for $b \geq 2$.*

Proof. Given a real number as a cut $F \subseteq \mathbb{Q}$ let $a_0 = \max\{n \in \mathbb{N} | n \in Q\}$ and define recursively

$$a_n = \max\{a \in \{0, \ldots, b-1\} | a_0 + \sum_{i=1}^{n-1} a_i b^i + a b^n \in F\}.$$

By this we get a representation $a_0.a_1a_2a_3...$ of F, which has no tail of zeros. The other way round, given such a representation, it describes a unique cut by

$$F = \bigcup_{n \in \mathbb{N}} \{q \in \mathbb{Q} | q \leq a_0 + \sum_{i=1}^{n} a_i b^i\}.$$

$\mathfrak{Q.E.D.}$

2.3 Uncountable sets

The following theorem represents a simple and nevertheless deep insight of mankind.

Theorem 2.5. *The real numbers $\mathbb{R}$ are uncountable.*[6]

$\mathfrak{Proof}$. If $\mathbb{R}$ is countable, then $[0, 1]$ is countable as well. Hence there exists a map C from $\mathbb{N}$ onto $[0, 1]$ with

$$C(n) = \sum_{i=1}^{\infty} c_i(n) 10^{-i},$$

[6] Also due to Georg Cantor.

where $c_i(n) \in \{0, 1, \ldots, 9\}$ are the digits in decimal expansion. Now consider a real number

$$x = \sum_{i=1}^{\infty} \bar{c}_i 10^{-i} \in [0, 1]$$

with $\bar{c}_i \neq c_i(i)$. Obviously $C(n) \neq x$ for all $n \in \mathbb{N}$. Hence C is not onto. A contradiction. $\mathfrak{Q.E.D.}$

The next result was quite surprising for mathematicians in the 19th century; in the set theoretical sense there is no difference between euclidian spaces of different dimension.

Theorem 2.6. *There is a bijection between* $\mathbb{R}$ *and* $\mathbb{R}^n$.

$\mathfrak{Proof}$. For simplicity of notation we proof the result for $n = 2$. The general proof uses the same idea. First note that the map $f(x) = \tan(x)$ is a bijection from $(-\pi/2, \pi/2)$ to $\mathbb{R}$ and a linear map from $(-\pi/2, \pi/2)$ to $(0, 1)$ is a bijection. Hence it is enough if we find a bijection between $(0, 1)^2$ and $(0, 1)$. Represent a pair of real numbers $(a, b) \in (0, 1)^2$ in decimal expansion

$$(a, b) = 0.a_1a_2a_3 \ldots, 0.b_1b_2b_3 \ldots).$$

Map such a pair to

$$0.a_1b_1a_2b_2a_3b_3 \ldots.$$

By uniqueness of decimal expansion, using the convention that the expansion has no tail of zeros, this is a bijection. $\mathfrak{Q.E.D.}$

2.4 Cardinalities

Let us go one step further and define the set theoretical size of sets by there cardinality.

Definition 2.4.1 *Two sets* A *and* B *have the same cardinality if there exists a bijection between them. This a an equivalence relation on sets and the equivalence class of* A *is denoted by* $|A|$.

The following theorem shows that there are infinitely many infinite cardinalities.

Theorem 2.7. A *and the power set* $P(A) = \{B|B \subseteq A\}$ *do not have the same cardinality.*[7]

$\mathfrak{Proof}$. Let $f : A \longmapsto P(A)$ be a function and

$$T := \{x \in A | x \notin f(x)\} \in P(A).$$

[7] Another result of Georg Cantor (1845-1918).

If f is surjective, there exists $t \in A$ such that $f(t) = T$. If $t \in T$. we have by the definition of T $t \notin f(t) = T$. On the other hand if $t \notin T$, we have $t \in T = f(t)$ by the definition. This is a contradiction. $\mathfrak{Q.E.D.}$

Definition 2.8. *The cardinality of A is less or equal to the cardinality of B if there is a injective map from A to B. We write $|A| \trianglelefteq |B|$ for this relation.*

Now we will show that cardinalities are in fact ordered by $\trianglelefteq$. Let us say what an order is:

Definition 2.4.2 *A binary relation $\leq$ is a partial order of a set A, if*

$$a \leq a$$

$$a \leq b \text{ and } b \leq a \Rightarrow a = b$$

$$a \leq b \text{ and } a \leq c \Rightarrow a \leq c.$$

The relation $\leq$ is an order, if in addition

$$a \leq b \vee b \leq a.$$

With this definition we have:

Theorem 2.9. *The relation $\trianglelefteq$ defines an order on cardinalities.* [8]

The proof is simple but we think it was not easy to find.

$\mathfrak{Proof}$. Reflexibility and transitivity of the relation are obvious. We first show anti-symmetry:
If there exists injections $f : A \longmapsto B$ and $g : B \longmapsto A$, then there is a bijection. For a set $S \subseteq A$ let

$$F(S) := A \backslash g(B \backslash f(S))$$

and define

$$A_0 := \bigcap_{i=0}^{\infty} F^i(A).$$

By the rule of De Morgan and the definition of F we have

$$F(\bigcap A_i) = \bigcup F(A_i)$$

and hence $F(A_0) = A_0$, which means $g(B \backslash f(A_0)) = A \backslash A_0$. The map, given by

$$\phi(x) = \begin{cases} f(x) & : \ x \in A_0 \\ g^{-1}(x) & : \ x \notin A_0 \end{cases},$$

[8] This is attributed to Cantor and the German mathematicians Felix Bernstein (1878-1956). and Ernst Schröder (1841-1902)

is a bijection. It remains to show that the order is total:
There is an injective map from A to B or an injective map from B to A. To this end let

$$F := \{f : \bar{A} \longmapsto \bar{B} | f \text{ is a bijection with } \bar{A} \subseteq A \text{ and } \bar{B} \subseteq B\}$$

For each chain $C \subseteq F$ we have $\bigcup C \in F$. Hence by theorem 2.12 below there is an maximal element $(f : \bar{A} \longmapsto \bar{B}) \in F$. We show that $A = \bar{A}$ or $B = \bar{B}$ for this function f. Suppose neither. Then there exists $a \in A \backslash \bar{A}$ and $b \in B \backslash \bar{B}$. But now $f \cup \{(a, b)\}$ would be in F. This is a contradiction to maximality of F. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

As a consequence we have:

Theorem 2.10. $\mathbb{R}$ *and* $P(\mathbb{N})$ *have the same cardinality.*

$\mathfrak{Proof}$. First we see that $|P(\mathbb{N})| = |\{0, 1\}^{\mathbb{N}}|$ since the map $f(A) = (a_i)$ with $a_i = 1$ for $i \in A$ and $a_i = 0$ for $i \notin A$ is a bijection. Furthermore we have $|\mathbb{R}| = |\{0, 1\}^{\mathbb{N}}|$ using binary expansions and anti-symmetry of $\trianglelefteq$. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

The continuum hypotheses, that there is no set $A \subseteq \mathbb{R}$ such that neither $|A| = \mathbb{N}|$ nor $|A| = |\mathbb{R}|$, is independent of the other axioms of set theory, see [10, 6].

At the end of the section we present a result showing again how "big" the continuum is:

Theorem 2.11. *The set of* $C(\mathbb{R}, \mathbb{R})$ *of continuous functions* $f : \mathbb{R} \longmapsto \mathbb{R}$ *has cardinality* $|\mathbb{R}|$.

$\mathfrak{Proof}$. We have

$$|\mathbb{R}^{\mathbb{Q}}| = |\mathbb{R}^{\mathbb{N}}| = |\{0, 1\}^{\mathbb{N}\times\mathbb{N}}| = |\{0, 1\}^{\mathbb{N}}| = |\mathbb{R}|$$

using $|\mathbb{Q}| = |\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$. But a continuous function on $\mathbb{R}$ is determined by its values on the rational numbers. Hence $|C(\mathbb{R}, \mathbb{R})| \trianglelefteq |\mathbb{R}^{\mathbb{Q}}| = |\mathbb{R}|$. On the other hand the constant functions are continuous hence $|C(\mathbb{R}, \mathbb{R})| \trianglerighteq |\mathbb{R}|$. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

2.5 The axiom of choice with some consequences

The following axiom of set theory is the foundation for most part of modern mathematics:

Axiom of choice: For any family of sets $(A_i)_{i\in I}$ with index set I there is a function

$$f : I \longmapsto \bigcup_{i\in I} A_i$$

with $f(i) \in A_i$ for all $i \in I$.[9]

The axiom seems to be self-evident, its consequences are not at all obvious. We will use the axiom to find strong results on well-ordered sets. Therefore a definition:

Definition 2.5.1 *An order is a well-ordering, if every subset of A has a minimal element.*

With this definition we obtain the following beautiful and strong result:

Theorem 2.12. *Every non-empty partially ordered set, in which every chain (i.e. ordered subset) has an upper bound, contains at least one maximal element.*[10]

$\mathfrak{Proof}$. Let A be non-empty and partially ordered. Assume A has no maximal element. Then especially the upper bound of of any chain is not maximal. We proof that under this assumption there is an unbounded chain. This is a contradiction.
Let W be the set of all well-ordered chains in A, then by the axiom of choice there is a function

$$c : W \longmapsto A$$

with $C(K) \in A \backslash K$ and $c(K) > K$. Now we call a chain K in A a special chain if K is well ordered and

$$\forall x \in K : C(K_x) = x \text{ where } K_x = \{y \in A | y < x\}.$$

We prove that for two special chain K, L either $K = L$ or $K = L_x$ or $L = K_x$ for some $x \in A$, this implies that the union of special chains is a special chain. Assume that $K \neq L$ and $K_x \neq L$ for all $x \in A$. Since K is well ordered $K \backslash L$ has a minimal element k. Since L is well ordered $L \backslash K_k$ has a minimal element l. Now $K_k = L_l$ and hence $k = l$, which prove our claim.
Let $\bar{K}$ be the union of all special chains. This is a special chain and especially a chain and hence bounded by some $a \in A$. By assumption a is not in A. Hence there is another $b \in A$ with $b > a$. Thus $b \notin \bar{K}$. On the other hand $b \cup \bar{K}$ is a special chain. Hence $b \cup \bar{K} \subseteq K$ which implies $b \in \bar{K}$. A contradiction. $\mathfrak{Q.E.D.}$

One surprising application of Zorn's lemma is the following

Theorem 2.13. *Every set has a well ordering.*[11]

$\mathfrak{Proof}$. Let $\mathbf{C}$ be the set of all well ordered subsets of a set with the partial order

$$(C_1, <_1) < (C_2, <_2) :\Leftrightarrow C_1 \subseteq C_2 \text{ and } <_1 \subseteq <_2 .$$

Every ordered subset of $\mathbf{C}$ has an upper bound by the union of elements of the ordered set. Hence by Zorn's lemma there is a maximum, say $(M, <)$ of $\mathbf{C}$. Assume $M \neq A$. Then there is a $x \in A \backslash M$. But $M \cup \{x\}$ with the definition $M < x$ is well

[9] The axiom was first formulated by the German mathematician Ernst Zermelo (1871-1953).
[10] This was proved by the American-German mathematician Max Zorn (1906-1993).
[11] This theorem is due to Zermelo.

ordered again. A contradiction to maximality. Hence $M = A$ is well ordered.$\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

The theorem is in fact very strange, if we try to imagine a well ordering of the reals. Another nice application of Zorn's lemma can be found in the foundation of linear algebra. We assume some basic notions in this field to prove:

Theorem 2.14. *Every vector space has a basis.*

$\mathfrak{Proof}$. The set Ξ of all linear independent subsets of a vector space V is partially ordered by inclusion $\subseteq$. Let $\mathbf{C}$ be a chain in Ξ, then $\mathfrak{C} = \bigcup_{C \in \mathbf{C}} C$ is an upper bound of this chain in Ξ. By the Zorn's lemma there is a maximal element B in Ξ. For every $v \in V$ is set $B \cup \{v\}$ is not linear independent, hence v is a linear combination of elements in B, and B is a basis of V. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

All the consequences of the axiom of choice presented here are in fact equivalent to the axiom, see [9]. At the end of the section we introduce another nice fundamental set theoretical concept.

Definition 2.5.2 *A filter F on a set X is a subset of $P(X)$ such that*

$$X \in F, \quad \emptyset \notin F$$

$$A, B \in F \Rightarrow A \cap B \in F$$

$$A \in F, A \subseteq B \Rightarrow B \in F.$$

A filter F is maximal, if there is no other filter G on X with $F \subset G$ and a filter is an ultra filter, if either $A \in F$ or $X \backslash A \in F$ for all $A \subseteq X$.

The existence of an ultrafilter is a further consequence of the axiom of choice.

Theorem 2.15. *Every filter can be extended to an ultra filter.*[12]

Let F be a filter on X and $\mathfrak{F}$ be the set of all filters on X that contain F. $\mathfrak{F}$ is partially ordered by inclusion. If $\mathfrak{C}$ is a chain in $\mathfrak{F}$, the set $\bigcup \mathfrak{C}$ is upper bound of $\mathfrak{C}$ in $\mathfrak{F}$. Hence by Zorn's lemma there is a maximal filter that contains F. It remains to show that a maximal filter is an ultrafilter. Assume that a filter F is not ultra. Let $Y \subseteq X$ be such that neither $Y \in F$ nor $X \backslash Y \in F$. Let $G = F \cup \{Y\}$. This set has the intersection property and may obviously be extended to a filter. Hence F is not maximal. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

2.6 The Cauchy functional equation

This section is devoted to a strange and remarkable consequence of the axiom of choice. We define:

[12] This theorem is due to Polish logician and mathematician Alfred Tarski (1901-1983).

Definition 2.6.1 *A function $f : \mathbb{R} \longmapsto \mathbb{R}$ is additive, if it satisfies the Cauchy equation*

$$f(x+y) = f(x) + f(y)$$

for all $x, y \in \mathbb{R}$

Assuming the notion of continuity it is nice to see that:

Theorem 2.16. *Every continuous additive function is linear.*[13]

$\mathfrak{Proof}$. By additivity of f we have

$$qf(p/q) = f(p) = pf(1) \Rightarrow f(p/q) = \frac{p}{q} f(1)$$

or all $p, q \in \mathbb{N}$. Moreover $f(1) = f(0) + f(1)$ hence $f(0) = 0$ and

$$f(x) = f(0) - f(-x) = -f(-x).$$

We thus see that f is linear on $\mathbb{Q}$ and by continuity linear on $\mathbb{R}$. $\mathfrak{Q.E.D.}$

In fact it is enough to assume that the function is Lebesgue measurable, which implies continuity for additive functions, see [**?**]. On the other hand by the axiom of choice we have the following theorem.

Theorem 2.17. *There are uncountable many nonlinear additive functions.*[14]

$\mathfrak{Proof}$. Consider the real numbers $\mathbb{R}$ as a vector space over the field of rational numbers $\mathbb{Q}$. By theorem 2.12 there is a basis B of this vector space. This means that any $x \in \mathbb{R}$ has a unique representation of the form

$$x = \sum_{i=1}^{n} r_i b_i$$

with $\{b_1, \ldots, b_n\} \subseteq B$ and $r_i(x) \in \mathbb{Q}$. Let $f : B \longmapsto \mathbb{R}$ be an arbitrary map and define a extension $f : \mathbb{R} \longmapsto \mathbb{R}$ by

$$f(x) := \sum_{i=1}^{n} r_i f(b_i).$$

On the one hand an obvious calculation shows that f is additive. On the other hand f is linear if and only f is linear on B. But there are uncountable many functions $f : B \longmapsto \mathbb{R}$ that are not linear. $\mathfrak{Q.E.D.}$

[13] This is due to the French mathematician August Lois Cauchy (1889-1857).
[14] This was proved by the German mathematician Georg Hamel (1877-1954).

2.7 Ordinal Numbers and Goodstein sequences

Definition 2.7.1 *An ordinal number is a well order set α with $x = \{y \in \alpha | y < x\}$ for all $x \in \alpha$. We define an order on the ordinal numbers by $\alpha < \beta$ if $\alpha \in \beta$.*[15]

Theorem 2.18. *The ordinals are well ordered by $<$.*

$\mathfrak{Proof}$. Let A be a set of ordinals than $\min(A) = \bigcap A$ is well ordered since the elements of A are well ordered. Moreover A is minimal since $\min(A) \in \alpha$ for all $\alpha \in A$. $\mathfrak{Q.E.D}$.

Definition 2.7.2 *Construct ordinals by*

$$\alpha + 1 = \alpha \cup \{\alpha\}$$

for each ordinal α and

$$\beta = \bigcup_{\alpha \in A} \alpha$$

for a countable set of ordinals A. Especially

$$0 = \emptyset,\ n + 1 = n \cup \{n\},\ \omega = \bigcup_{i=0}^{\infty} i,\ (n+1)\omega = \bigcup_{i=0}^{\infty} n\omega + i$$

$$\omega^{n+1} = \bigcup_{i=0}^{\infty} i\omega^n,\ \omega^\omega = \bigcup_{i=0}^{\infty} \omega^i, \ldots, \epsilon, \ldots$$

$$\begin{aligned} 0 < 1 < 2 < \cdots < \omega < \omega + 1 < \omega + 2 < \cdots < \omega + \omega \\ = 2\omega < 2\omega + 1 < \cdots < 3\omega < \cdots < \cdots < \omega\omega \\ = \omega^2 < \cdots < \omega^3 < \cdots < \omega^\omega < \cdots < \omega^{\omega^\omega} < \cdots \omega^{\omega^{\omega \ldots}} = \epsilon \cdots \end{aligned}$$

Fig. 2.1. Ordinal numbers

The theory of ordinals has an elementary and beautiful consequence.

Definition 2.7.3 *Given a number $N \in \mathbb{N}$ we define the Goodstein sequence recursively. $G_2(N) = N$. If $G_n(N)$ is given, write $G_n(N)$ in hereditary base n notation. Replace n by $n+1$ and subtract one to obtain $G_{n+1}(N)$. Consider for example $N = 3$:*

$$G_2 = 3 = 2^{2^0} + 2^0 \to G_3 = 3^{3^0} + 3^0 - 1 = 3^{3^0} \to G_4 = 4^{4^0} - 1 = 4^0 + 4^0 + 4^0$$

$$\to G_5 = 5^0 + 5^0 \to G_6 = 6^0 \to G_7 = 0.$$

[15] Ordinal numbers were introduced by Cantor. . The definition given here is due to a Hungarian-American mathematician John von Neumann (1903-1957).

Theorem 2.19. *All Goodstein sequences eventually terminate with zero,*

$$\forall N \in \mathbb{N}\ \exists n \in \mathbb{N}\ :\ G_n(N) = 0.$$[16]

$\mathfrak{Proof}$. We construct a parallel sequence of ordinal numbers not smaller than a given Goodstein sequence by recursion. If $G_n(N)$ is given in hereditary base n expansion, replace n by the smallest ordinal number ω. The "base change" operation in the construction of the Goodstein sequence does not change the ordinal number of the parallel sequences, on the other hand subtracting one is decreasing this sequence. Since the ordinals are well-ordered the parallel sequences terminates with zero and so does the Goodstein sequence. $\mathfrak{Q.E.D.}$

It can be shown that it is not possible to prove the theorem of Goodstein without using cardinal numbers, see [13].

[16] Found by the English mathematician Reuben Goodstein (1912-1985).

3

Discrete Mathematics

3.1 The Pigeonhole principle

The pigeonhole principle seems to be almost trivial, if you put n pigeons to k holes, one of holes contains at least $\lceil n/k \rceil$ pigeons. Here $\lceil x \rceil$ is the smallest number bigger than x. To put the result more formally we have:

Theorem 3.1. *Let $f : A \longmapsto B$ be a map between two finite sets with $|A| > |B|$ than there exists an $b \in B$ with*

$$|f^{-1}(b)| \geq \lceil |A|/|B| \rceil,$$

where $\lceil x \rceil$ is the smallest number bigger than x.

$\mathfrak{Proof}$. If not, we would have $|f^{-1}(b)| < |A|/|B|$ for all $b \in B$ hence

$$|A| = \sum_{b \in B} |f^{-1}(b)| < |A|,$$

a contradiction. $\mathfrak{Q.E.D.}$

Also the principle is very simple, it is a strong tool to prove result in discrete mathematics.

Theorem 3.2. *In any group of n people there are at least two persons having the same number friends (assuming the relations of friendship is symmetric).*

$\mathfrak{Proof}$. If there is a person with $n-1$ friends then everyone is a friend of him, that is, no one has 0 friend. This means that 0 and $n-1$ can not be simultaneously the numbers of friends of some people in the group. The pigeonhole principle tells us that there are at least two people having the same number of friends. □

Theorem 3.3. *In any subset of $M = \{1, 2, \ldots, 2m\}$ with at least $m+1$ elements there are numbers a, b such that a divides b.*

Proof. Let $\{a_1, \ldots, a_{m+1}\} \subseteq M$ and decompose $a_i = 2^{r_i} q_i$, where the q_i are odd numbers. There are only n odd numbers in M hence one q_i appears in the decomposition of two different numbers a_i and a_j. Now a_i divides a_j if $a_i < a_j$. $\mathfrak{Q.E.D.}$

Theorem 3.4. *A sequence of $nm+1$ real numbers contains an ascending sequence of length $m+1$ or a descending sequence of length $n+1$ or both.*

$\mathfrak{Proof}$. For one of the numbers a_i let $f(a_i)$ be the length of the longest ascending subsequence starting with a_i. If $f(a_i) > m$, we have the result. Assume $f(a_i) \leq m$. By the pigeonhole principle there exists an $s \in \{1, \ldots m\}$ and a set A such that

$$f(x) = s \text{ for } x \in A = \{a_{j_i} | i = 1 \ldots n+1\}.$$

Consider two successive numbers a_{j_i} and $a_{j_{i+1}}$ in A. If $a_{j_i} \leq a_{j_{i+1}}$, we would have an ascending sequence of length $s+1$ with starting point a_{j_i}, which implies $f(a_{j_i}) = s+1$. This is a contradiction. Hence the numbers form a descending sequence of length $n+1$. $\mathfrak{Q.E.D.}$

There are many other combinatorial applications of the Pigeonhole principle, see [11]. We like to include here a number theoretical application:

Theorem 3.5. *For any irrational number α the set $\{[n\alpha] | n \in \mathbb{N}\}$ is dense in $[0,1]$, where $[.]$ denotes the fractional part of a number.*

$\mathfrak{Proof}$. Given $\epsilon > 0$ choose $M \in \mathbb{N}$, such that $1/M < \epsilon$. By the pigeonhole principle there must be $n_1, n_2 \in \{1, 2, ..., M+1\}$, such that $n_1\alpha$ and $n_2\alpha$ are in the same integer subdivision of size 1/M (there are only M such subdivisions between consecutive integers). Hence there exits $p, q \in \mathbb{N}$ and $k \in \{0, 1, ..., M-1\}$ such that $n_1\alpha \in (p + k/M, p + (k+1)/M)$ and $n_2\alpha \in (q + k/M, q + (k+1)/M)$.This implies $(n_2 - n_1)\alpha \in (p - q - 1/M, p - q + 1/M)$ and $[n\alpha] < 1/M < \epsilon$. We see that 0 is a limit point of $[n\alpha]$. Now for an arbitrary $p \in (0, 1]$ choose M as above. If $p \in (0, 1/M]$ we are done, if not we have $p \in (j/M, (j+1)/M]$. Setting $k := \sup\{r \in \mathbb{N} | r[na] < j/M\}$, one obtains $|[(k+1)na] - p| < 1/M < \epsilon$. $\square$

3.2 Binomial coefficients

Binomial coefficients play a basic role in combinatorial counting. Here comes the definition:

Definition 3.2.1 *For $n, k \in \mathbb{N}$ with $k \leq n$ the binomial coefficient is defined by*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

where $n! = 1 \cdot 2 \cdot \ldots \cdot n$ is the factorial.

Theorem 3.6. *The Binomial coefficients are given by the Pascal triangle.*[1]

[1] The theorem is attributed to Blaise Pascal (1623-1662).

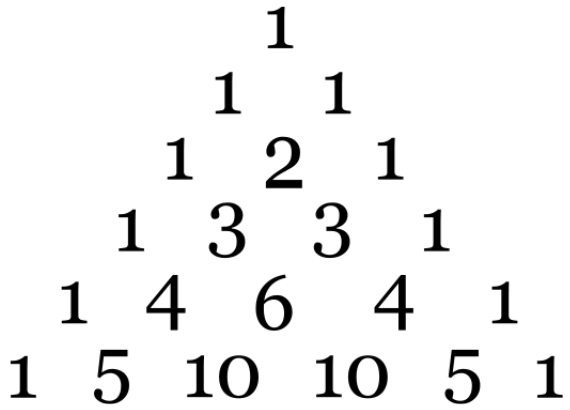

Fig. 3.1. Pascal triangle

𝔓𝔯𝔬𝔬𝔣.

$$\binom{n+1}{k} = \frac{(n+1)!}{k!(n+1-k)!} = \frac{n!(n+1-k)+n!k}{k!(n+1-k)!} = \frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n+1-k)!}$$

$$= \binom{n}{k} + \binom{n}{k-1}$$

𝔔.𝔈.𝔇.

As the first application we count coefficients in the expansion of the binomial.

Theorem 3.7.

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \quad [2]$$

𝔓𝔯𝔬𝔬𝔣. We prove the theorem by induction. $n = 1$ is obvious. Assume the equation for $n \in \mathbb{N}$. We have

$$(a+b)^{n+1} = (a+b)(a+b)^n = (a+b)\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} = \sum_{k=0}^{n+1} \binom{n}{k-1} a^k b^{n+1-k} + \sum_{k=0}^{n+1} \binom{n}{k} a^k b^{n+1-k}$$

$$= \sum_{k=0}^{n+1} \left(\binom{n}{k-1} + \binom{n}{k} \right) a^k b^{n+1-k} = \sum_{k=0}^{n+1} \binom{n}{k} a^k b^{n+1-k} = (a+b)^{n+1}.$$

This is the equation for $n+1$. 𝔔.𝔈.𝔇.

[2] This is due to the English physicist and mathematician Isaac Newton (1643-1727).

The second application of Binomial coefficients can be found in elementary counting tasks.

Theorem 3.8.
(1) There are $\binom{n}{k}$ *possibilities to take* k *Elements from* n *Elements without repetition and ordering.*
(2) There are $k!\binom{n}{k} = \frac{n!}{(n-k)!}$ *possibilities to take* k *Elements from* n *Elements without repetition with ordering.*
(3) There are $\binom{n+1-k}{k}$ *possibilities to take* k *Elements from* n *Elements with repetition but without ordering.*
(4) There are n^k *possibilities to take* k *Elements from* n *Elements with repetition and ordering.*

$\mathfrak{Proof}$. (1) Induction by n. $n = 1$ is obvious. Assume (1) for n and consider a set X with $n+1$ elements. Fix one element $x \in X$. Then by (1) there are $\binom{n}{k}$ possibilities to choose a set with k elements from X without x and there are $\binom{n}{k-1}$ possibilities to choose a set with k elements from X with one element being x. The Pascal triangle gives $\binom{n+1}{k}$ possibilities to choose a set with k elements form a set with $n+1$ elements.
(2) There are $k!$ possibilities to order a set with k elements. Hence (2) follows directly from (1).
In (3) without loss of generality we want to choose k-elements $1 \le a_1 \le \ldots \le a_k \le n$ from $\{1, \ldots, n\}$. With $b_j := a_j + j - 1$ we have $1 \le b_1 < \ldots < b_k \le n + k - 1$. By (1) there are $\binom{n+1-k}{k}$ possibilities to choose these b_j, but the map between the elements a_j and elements b_j is a bijection, so the result follows.
(4) is obvious by induction. $\mathfrak{Q.E.D.}$

3.3 The inclusion exclusion principle and derangements

As the pigeonhole principle above the inclusion exclusion principle is a strong tool in combinatorics.

Theorem 3.9. *Let* $A_1, A_2, \ldots, A_n$, *be finite sets than*

$$|\bigcup_{k=1}^{n} A_k| = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \le i_1 < i_2 < \ldots < i_k < n} |A_{i_1} \cap A_{i_2} \cap \ldots A_{i_k}|.$$

$\mathfrak{Proof}$. Let $A = \bigcup_{k=1}^{n} A_k$ and $\mathbf{1}_B$ be the indicator function of $B \subseteq A$. We show that

$$\mathbf{1}_A = \sum_{k=1}^{n} (-1)^{k+1} \sum_{|I|=k,\ I \subset \{1,\ldots,n\}} \mathbf{1}_{A_I},$$

where $A_I = \bigcap_{i \in I} A_i$. Summing the equation over all $x \in A$ gives the result. By expanding the following product using $\mathbf{1}_{A_i} \mathbf{1}_{A_j} = \mathbf{1}_{A_{\{i,j\}}}$, we have

$$\prod_{k=1}^{n} (\mathbf{1}_A - \mathbf{1}_{A_k}) = \mathbf{1}_A - \sum_{k=1}^{n} (-1)^{k+1} \sum_{|I|=k,\ I \subset \{1,\ldots,n\}} \mathbf{1}_{A_I}.$$

If $x \in A$, one factor in the product is zero, if $x \notin A$ all are zero, hence the function here is identical zero, which proves the claim. $\mathfrak{Q.E.D.}$

We apply the inclusion-exclusion principle to permutations.

Definition 3.3.1 *A permutation of $\{1, \ldots, n\}$ which has no fix point is a derangement.*

Theorem 3.10. *The number of derangements d_n is given by*

$$d_n = n!(\sum_{k=0}^{n} (-1)^k \frac{1}{k!}) \approx n!/e.$$

$\mathfrak{Proof}$. Let A_k be the set of permutations that fix k. We have $|A_{i_1} \cap A_{i_2} \cap \ldots A_{i_k}| = (n-k)!$. Hence by the inclusion exclusion principle

$$c_n = n! - |\bigcup_{k=1}^{n} A_k| = n! - \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \le i_1 < i_2 < \ldots < i_k < n} (n-k)!$$

$$= n! - \sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} (n-k)! = n! - \sum_{k=1}^{n} (-1)^{k+1} \frac{n!}{k!} = n!(\sum_{k=0}^{n} (-1)^k \frac{1}{k!}).$$

The asymptotic formula follows from definition 5.7.1 of the Euler number e. $\mathfrak{Q.E.D.}$

3.4 Trees and Catalan numbers

In this section we have a look at graphs without loops. Therefore a definition.

Definition 3.4.1 *A graph is a forest if it has no loops. A tree is a connected forest. A binary tree is a rooted tree in which each vertex has at most two children.*

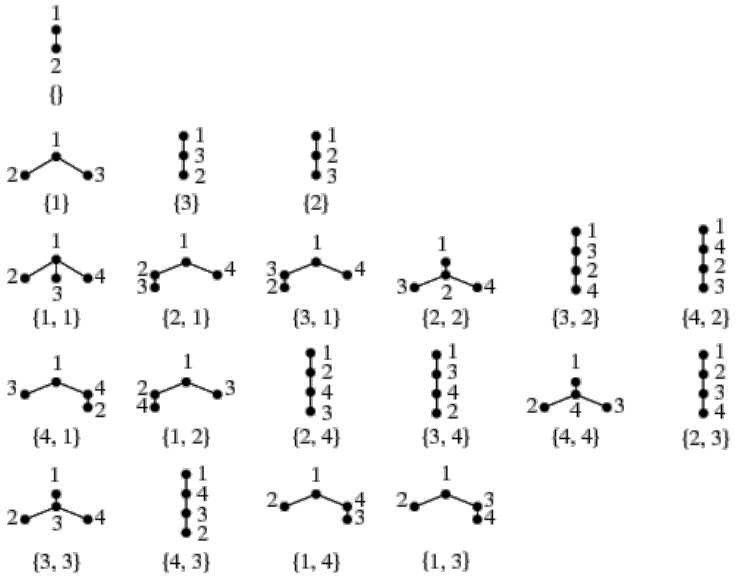

Fig. 3.2. Labeled trees

Theorem 3.11. *There are n^{n-1} different binary trees with n labeled vertices.*[3].

$\mathfrak{Proof}$. We proof a more general result for forests. Let $A = \{1, \dots k\}$ be a set of vertices, and let $T_{n,k}$ be the number of forest on $1, \dots, n$ that are rooted in A. Let F be one of this forests. Assume the $1 \in A$ has i neighbors. If we delete 1 we get a forest $F^{'}$ with roots in $\{2, \dots, k-1\} \cup \{$ neighbors of $i\}$ and there are $T_{n-1,k-1+n}$ such forests. The other way we construct a forest F by fixing i and than choose i neighbors of 1 in $\{k+1, \dots, n\}$ and thus get the forest $F^{'}$. Therefore we have the recursion

$$T_{n,k} = \sum_{i=0}^{n-k} \binom{n-k}{i} T_{n-1,k-1+i}$$

with $T_{0,0} := 1$ and $T_{n,0} := 0$. By induction we will prove that

$$T_{n,k} = kn^{n-k-1}.$$

This gives the result since $T_{n,1} = n^{n-1}$ is the number of tress with one root. Using the formula for $T_{n,k}$ in the first step and the assumption of the induction in the second step we get:

$$T_{n,k} = \sum_{i=0}^{n-k} \binom{n-k}{i} T_{n-1,k-1+i} = \sum_{i=0}^{n-k} \binom{n-k}{i} (k-1+i)(n-1)^{n-1-k-i}$$

$$= \sum_{i=0}^{n-k} \binom{n-k}{i} (n-1)^{i} - \sum_{i=1}^{n-k} \binom{n-k}{i} i(n-1)^{i-1}$$

[3] This is a result of the English mathematician Arthur Cayley (1821-1894).

$$= n^{n-k} - (n-k)\sum_{i=0}^{n-1-k}\binom{n-1-k}{i}(n-1)^i$$
$$= n^{n-k} - (n-k)n^{n-1-k} = kn^{n-k-1}$$

This completes the induction. $\mathfrak{Q.E.D.}$

To determine the number of full binary trees we need a special sequence of numbers.

Definition 3.4.2 *The Catalan numbers C_n are given by the recurrence relation*

$$C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i}$$

with $C_0 = 1$.[4]

The Catalan numbers may by calculated using Binomial coefficients.

Theorem 3.12.

$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

$\mathfrak{Proof}$. Consider the generating function

$$c(x) = \sum_{n=0}^{\infty} C_n x^n,$$

than

$$c(x)^2 = \sum_{k=0}^{\infty}(\sum_{m=0}^{k} C_m C_{m-k})x^k = \sum_{k=0}^{\infty} C_{k+1}x^k = xc(x) + 1$$

with the solution

$$c(x) = \frac{1-\sqrt{1-4x}}{2x}.$$

The other solution has a pole at $x = 0$ which does not give the generating function. By the (generalized) Binomial theorem we get by some simplification,

$$\sqrt{1-4x} = 1 - 2\sum_{n=1}^{\infty}\frac{1}{n}\binom{2n-2}{n-1}x^n = 1 - \sum_{n=1}^{\infty} C_{n-1}x^n,$$

hence

$$c(x) = \sum_{n=0}^{\infty}\frac{1}{n+1}\binom{2n}{n}x^n.$$

$\mathfrak{Q.E.D.}$

[4] This numbers where introduced by the Belgian mathematician Eugene Charles Catalan (1814 - 1894).

Theorem 3.13. *C_n is the number of full binary trees with n vertices.*

𝔓𝔯𝔬𝔬𝔣. Let T_n be the number of binary trees with n vertices. The numbers binary trees, with n vertices and j left subtrees and $n-1-j$ right subtrees with respect to the root, is $T_j T_{n-1-j}$. Summing up over all possible numbers of right and left subtrees gives

$$T_n = \sum_{j=0}^{n-1} T_j T_{n-1-j},$$

but this is the recursion for the Catalan numbers. 𝔔.𝔈.𝔇.

There are more than fifty other combinatorial application of the Catalan numbers, [23].

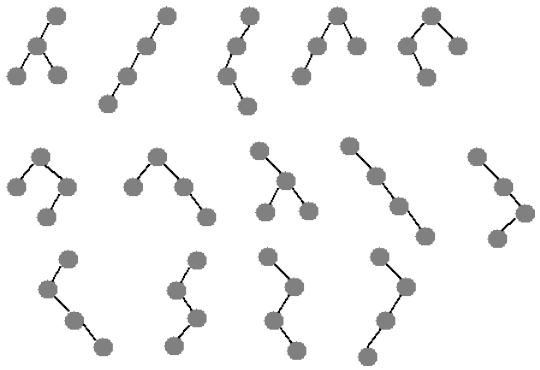

Fig. 3.3. Binary trees with four vertices

3.5 Ramsey theory

Now we come to graphs with colored edges. The following famous theorem is quite simple to prove:

Theorem 3.14. *For any $(r,s) \in \mathbb{N}^2$ there exists at least positive integer $R(r,s)$ such that for any complete graph on $R(r,s)$ vertices, whose edges are colored red or blue, there exists either a complete subgraph on r vertices which is entirely blue, or a complete subgraph on s vertices which is entirely red. Moreover*

$$R(r,s) \leq \binom{r+s-2}{r-1}.$$

[5]

[5] This is a lemma of the British mathematician Frank Plumpton Ramsey (1903-1930).

𝔓𝔯𝔬𝔬𝔣. We prove the result by induction on $r + s$. Obviously $R(r, 2) = r$ and $R(2, s) = s$ which starts the induction. We show

$$R(r, s) \leq R(r - 1, s) + R(r, s - 1).$$

By this $R(r, s)$ exists under the induction hypothesis that $R(r-1, s)$ and $R(r, s-1)$. Consider a complete graph on $R(r - 1, s) + R(r, s - 1)$ vertices. Pick a vertex v from the graph, and partition the remaining vertices into two sets M and N, such that for every vertex w, $w \in M$ if (v, w) is blue, and $w \in N$ if (v, w) is red. Because the graph has $R(r - 1, s) + R(r, s - 1) = |M| + |N| + 1$ vertices, it follows that either $|M| > R(r - 1, s)$ or $|N| > R(r, s - 1)$. In the former case, if M has a red complete graph on s vertices, then so does the original graph and we are finished. Otherwise M has a blue complete graph with $r - 1$ vertices and so $M \cup \{v\}$ has blue complete graph on k vertices by definition of M. The latter case is analog. With the starting values we get the upper bound on $R(r, s)$ by the recursion 3.1 for binomial coefficients. 𝔔.𝔈.𝔇.

The next result is a nice application of the Ramsey's theorem.

Theorem 3.15. *In any party of at least six people either at least three of them are (pairwise) mutual strangers or at least three of them are (pairwise) mutual acquaintances.*

𝔓𝔯𝔬𝔬𝔣. Describe the party as a complete graph on 6 vertices where the edges are colored red if the related people are mutual strangers and blue if not. With this we just have to show $R(3, 3) \leq 6$. Pick a vertex v. There are 5 edges incident to v at least 3 of them must be the same color. Assume (without loss of generality) that these vertices r, s, t are blue. If any of the edges $(r, s), (r, t), (s, t)$ are also blue, we have an entirely blue triangle. If not, then those three edges are all red and we have an entirely red triangle. 𝔔.𝔈.𝔇.

The following figure shows a 2-colored graph on 5 vertices without any complete monochromatic graph on 3 vertices. Hence $R(3, 3) = 6$.

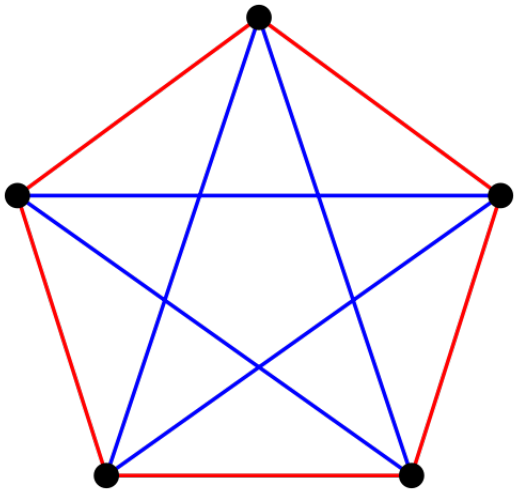

Fig. 3.4. A 2-coloring of complete graph

3.6 Sperner's lemma and Brouwers's fix point theorem

In this section we color graphs with three colors.

Definition 3.6.1 *A Sperner coloring of a triangulation T of triangle ABC is a 3-coloring, such that A, B, C have different colors and the vertices of T on each edge of triangle use only the two colors of the endpoints.*

Theorem 3.16. *Every Sperner colored triangulation contains a triangle whose vertices have all different colors.*[6]

Proof. Let q denote the number of triangles colored AAB or BAA and r be the number of rainbow triangles, colored ABC. Consider edges in the subdivision whose endpoints receive colors A and B. Let x denote the number of boundary edges colored AB and y bd the number of interior edges colored AB (inside the triangle T). We now count in two different ways. First we count over the triangles of the subdivision: For each triangle of the first type, we get two edges colored AB, while for each triangle of second type, we get exactly one such edge. Note that this way we count internal edges of type AB twice, whereas boundary edges only once. We conclude that $2q + r = x + 2y$. Now we count over the boundary of T: Edges colored AB can be only inside the edge between two vertices of T colored A and B. Between vertices colored A and B there must be an odd number of edges colored AB. Hence, x is odd. This implies that r is also odd and especially not zero. Q.E.D.

These combinatorial result can be used to give a simple prove of Brouwers's fix point theorem:

Theorem 3.17. *Every continues function of a disk into itself has a fixed point.*[7]

[6] This is a lemma of the German mathematician Emanuel Sperner (1905-1980).
[7] This theorem was proved by the Dutch mathematician L.E.J. Brouwer (1881-1966)

𝔓𝔯𝔬𝔬𝔣. We prove the result here for a triangles instead of disks. (The original theorem follows from the fact that any disc is homömorphic to any triangle). Let Λ be the triangle in $\mathbb{R}^3$ with vertices $e_1 = (1,0,0)$, $e_2 = (0,1,0)$ and $e_3 = (0,0,1)$, that is

$$\Lambda = \{(a_1, a_2, a_3) \in [0,1]^3 | a_1 + a_2 + a_3 = 1\}.$$

Define a sequence of triangulations $T_1, T_2, T_3, \ldots$ of Λ, such that $\delta(T_k) \longmapsto 0$, where δ is the maximal length of an edge in a triangulation.
Now suppose that a continuous map $f : \Lambda \longmapsto \Lambda$ has no fixed point. Color vertices in the following with the colors $1, 2, 3$. For each $v \in T$, define the coloring of v to be the minimum color i such that $f(v)_i < v_i$, that is, the minimum index i such that the i-th coordinate of $f(v) - v$ is negative. This is well-defined assuming that f has no fixed point because we know that the sum of the coordinates of v and $f(v)$ are the same. We claim that the vertices e_1, e_2 and e_3 are colored differently. Indeed, they maximize a different coordinate and hence are first negative at this coordinate. Furthermore, a vertex on the e_1e_2 edge has $a_3 = 0$, so that $f(v) - v$ has nonnegative third coordinate and is hence colored 1 or 2. The same is true for the other coordinates.
Now applying Sperner's lemma, for any triangulation T_k there is a triangle Λ_k whose vertices have all different colors. Furthermore every sequence in a compact space has a convergent subsequence, see section 6.1. So there is some subsequence of the triangles Λ_k that converges to some point x. By the construction we obtain $f(x)_i = x_i$ for each coordinate, hence x is a fixed point. 𝔔.𝔈.𝔇.

Using induction on the dimension it is possible to prove Sperner's lemma and with this the fix point theorem in higher dimensions [22].

3.7 Euler walks

Let us now take a walk on a graph.

Definition 3.7.1 *An Euler walk is a closed path on a graph that uses each edge exactly once.*

The question under what conditions such a walk exists, was answered by Euler.

Theorem 3.18. *An Euler walk on an finite connected graph exists if and only if the degree of every vertex is even.*[8]

𝔓𝔯𝔬𝔬𝔣. We first proof the "only if" part. Consider a closed Euler walk. We leave every vertex on a different edge. So the degree of each vertex must be even, since it is twice the number of times it has been visited. For the "if" part we use induction on the number of edges k. For $k = 1$ the result is obvious. Assume that we have the result for all graphs with fewer than k edges. Now suppose G is a connected graph with k edges such that the degree of each vertex is even. The degree of each vertex

[8] Proved by the Swiss mathematician Leonard Euler (1707-1783).

is at least two since the graph is connected. So the number of edges is at least the number of vertices, so G is not a tree and hence has a cycle C. If we remove the edges of C from G we obtain a graph G' with fewer than k edges and all vertices of even degree. Now G_2 does not have to be connected. So let $H_1, \ldots, H_i$ be connected components of G_2. By induction hypotheses we have an Euler walk W_j on each H_j. Moreover H_j has a vertex v_j in common with C. We may assume W_j to start and finish with v_j. Now we construct an Euler walk on G in the following way. We walk around C, every time we come to a vertex v_j we follows the walk W_j and than continue around C. Q.E.D.

Corollary 3.19. *An Euler walk with different starting and ending vertex exists if and only if the degree of these vertices is odd.*

Proof. Just divide resp. identify a starting point. Q.E.D.

We may apply the result to the bridges of Königsberg.

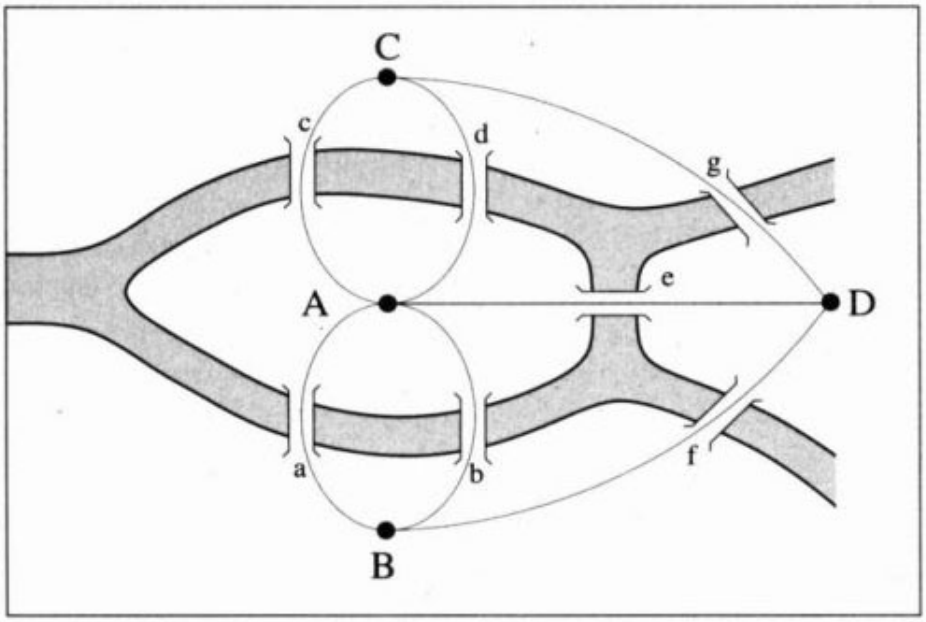

Fig. 3.5. The Bridges of Königsberg

4

Geometry

4.1 Triangles

The most basic geometrical theorem on all triangles is:

Theorem 4.1. *The sum of interior angles in a triangle is* 180°.

𝔓𝔯𝔬𝔬𝔣. Consider a triangle ABC. The line AB is extended and the line BD is constructed so that it is parallel to line segment AC. If two parallel lines are cut by a transversal, alternate interior angles are congruent and corresponding angles are congruent, hence $d = c$ and $a = e$. Now angles on a straight line add up to 180. Hence $a + b + c = 180°$ given the result. 𝔔.𝔈.𝔇.

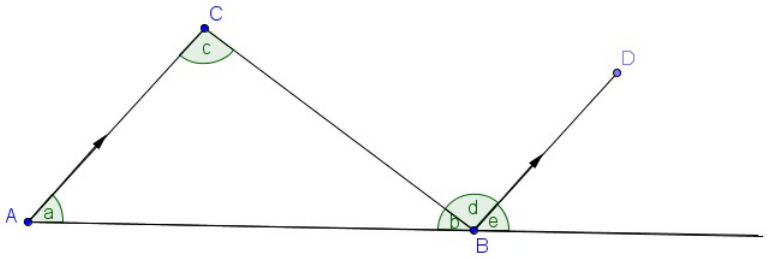

Fig. 4.1. Sum of angles in a triangle

Now we describe some nice ancient results on right triangles.

Theorem 4.2. *A triangle inscribed in a a circle such that one side is a diameter has a right angle opposite to this side.*[1]

𝔓𝔯𝔬𝔬𝔣. Split the triangle into two isosceles triangles by a line from center of the circle to point at the angle in question. The angle is spliced by the line into angles α and

[1] This theorem is attributed to the Greek philosopher Thales of Milet (624-546 BC).

β. Since base angles of equilateral triangles are equal we get $\alpha + (\alpha + \beta) + \beta = 180°$ implying $\alpha + \beta = 90°$. Q.E.D.

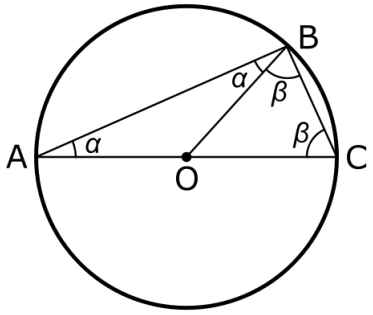

Fig. 4.2. Theorem of Thales

Theorem 4.3. *In a right triangle the altitude is the geometric mean of the two segments of the hypotenuse.* [2]

Proof. By considering angles we see that the triangles ACD and BCA are congurent hence

$$\frac{|AC|}{|CD|} = \frac{|CB|}{|AC|}$$

given

$$|AC| = \sqrt{|CD||CB|}$$

Q.E.D.

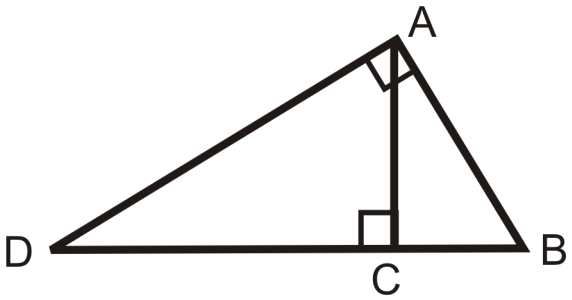

Fig. 4.3. The altitude of a right triangle

Theorem 4.4. *In any right triangle be have*

$$a^2 + b^2 = c^2$$

where c is the length of the hypotenuse and a and b are the length of the legs. [3]

Proof. Construct a square with side length $a + b$ by four not overlapping copies of the triangle with middle square of side length c. By calculating the area of this square directly and as sum of the area of parts we have

[2] The theorem can be found in the "Elements" of Euklid (360-280 B.C.)
[3] This theorem the attributed to ancient Greek mathematician Pythagoras (570-495 BC

$$(a+b)^2 = 4(ab/2) + c^2.$$

given the result. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

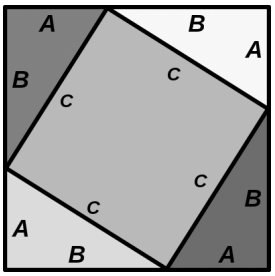

Fig. 4.4. Proof of the theorem of Pythagoras

To study arbitrary triangles we introduce the geometric cosine, compare with 5.7 for the analytic definition

Definition 4.1.1 *Given a right triangle the cosine of an angle* $\cos(\alpha)$ *is the ratio of the length of the adjacent side to the length of the hypotenuse.*

Theorem 4.5. *In any triangle with side length* a, b, c *we have*

$$c^2 = a^2 + b^2 - 2ab\cos\gamma$$

where the angle γ *is opposite to side of length* c. [4]

$\mathfrak{Proof}$. Extend the line of length b by length d to get a right triangle with hight h. Applying Pythagoras gives

$$(b+d)^2 + h^2 = c^2$$

and

$$d^2 + h^2 = a^2$$

which yield

$$c^2 = a^2 + b^2 + 2bd$$

On the other hand using the geometric definition of cos we have

$$\cos\gamma = -\cos(\pi - \gamma) = -d/a.$$

$\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

[4] This theorem also goes back to Euklid (360-280 B.C.)

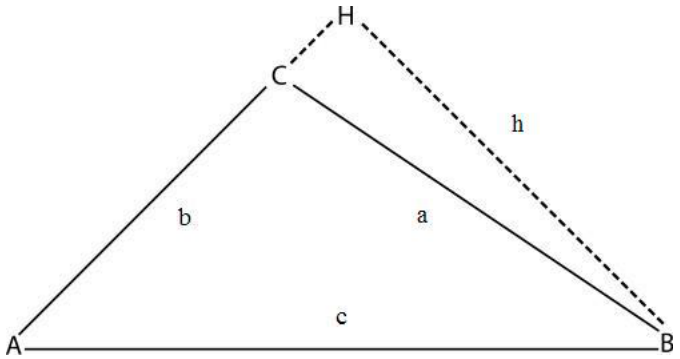

Fig. 4.5. Proof of the low of cosine

Theorem 4.6. *The area of a triangle with side length a, b, c is*

$$A = \sqrt{s(s-a)(s-b)(s-b)}$$

with $s = (a+b+c)/2$.[5].

𝔓𝔯𝔬𝔬𝔣. By the low of cosine we have

$$\cos\gamma = \frac{a^2+b^2-c^2}{2ab}$$

Hence

$$\sin\gamma = \sqrt{1-\cos^2\gamma} = \frac{\sqrt{4a^2b^2-(a^2+b^2-c^2)^2}}{2ab}$$

The altitude h_a of the triangle of the side of length a is given by $b\sin\gamma$. Hence

$$A = \frac{1}{/}2ah_a = \frac{1}{2}ab\sin\gamma = \frac{1}{4}\sqrt{4a^2b^2-(a^2+b^2-c^2)^2}$$

leading to the result. 𝔔.𝔈.𝔇.

4.2 Quadrilaterals

Theorem 4.7. *The sum of the angles in a quadrilateral is* 360°

𝔓𝔯𝔬𝔬𝔣. Divide the quadrilateral into two triangles. The sum of interior angles in each rectangle is 180° summing up to 360°. 𝔔.𝔈.𝔇.

We present here two beautiful results on cyclic quadrilateral (inscribed into a circle) with are easier to handel than arbitrary quadrilaterals.

Theorem 4.8. *In a cyclic quadrilateral the product of the length of its diagonals is equal to the sum of the products of the length of the pairs of opposite sides.*[6]

[5] This is attributed to the ancien mathematician Heron of Alexandria (10-70)
[6] This result is attributed to the Greek astronomer and mathematician Ptolemy (90-168).

𝔓𝔯𝔬𝔬𝔣. On the diagonal BD locate a point M such that angles ACB and MCD be equal. Considering angles we see that the triangles ABC and DMC are similar. Thus we get

$$\frac{|CD|}{|MD|} = \frac{|AC|}{|AB|}$$

or $|AB||CD| = |AC||MD|$. Now, angles BCM and ACD are also equal; so triangles BCM and ACD are similar which leads to

$$\frac{|BC|}{|BM|} = \frac{|AC|}{|AD|}$$

or $|BC||AD| = |AC||BM|$. Summing up the two identities we obtain

$$|AB||CD| + |BC||AD| = |AC||MD| + |AC||BM| = |AC||BD|$$

𝔔.𝔈.𝔇.

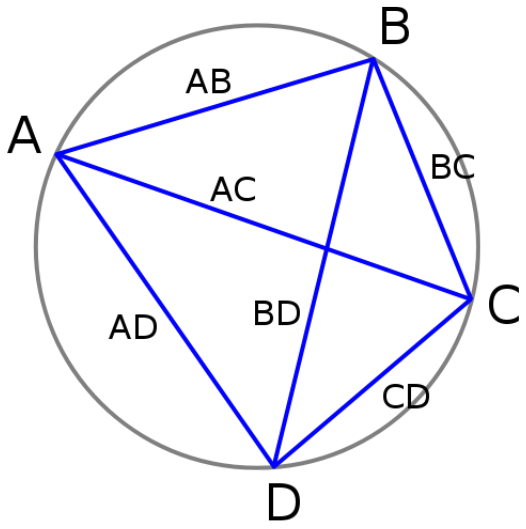

Fig. 4.6. A cyclic quadrilateral

Theorem 4.9. *The area of a cyclic quadrilateral with side length a, b, c, d is given by*

$$A = \sqrt{(s-a)(s-b)(s-c)(s-d)}$$

with $s = (a+b+c+d)/2$. [7]

[7] This result is attributed to the Indian mathematician Brahmagupta (598-668)

𝔓𝔯𝔬𝔬𝔣. Dividing the quadrilateral into two triangles and using trigeometry for the altitude of the triangles we get

$$A = \frac{1}{2}\sin(\alpha)(ab + cd)$$

Multiplying by 4, sqauring and using $\sin^2(\alpha) = 1 - \cos^2(\alpha)$ gives

$$4A^2 = (1 - \cos^2(\alpha))(ab + cd)^2$$

By the low of cosine

$$2\cos(\alpha)(ab + cd)^2 = a^2 + b^2 - c^2 - d^2$$

hence

$$4A^2 = (ab + cd)^2 - \frac{1}{4}(a^2 + b^2 - c^2 - d^2)^2$$

leading to the result. 𝔔.𝔈.𝔇.

Both theorems may be generalized to arbitrary quadrilateral, see [**?**]. The formulation of the theorems and its proves is getting more involved with this. For the shake of simplicity we decided to skip this.

4.3 The golden ratio

The golden ratio has fascinated mathematician, philosophers and artist for at least 2,500 years, here is the definition:

Definition 4.3.1
Two quantities $a, b \in \mathbb{R}$ with $a > b$ are in the golden ration if

$$\frac{a+b}{a} = \frac{a}{b} =: \phi$$

The golden ration ϕ is given by $(\sqrt{5}+1)/2 \approx 1.618$ a solution of $x^2 - x - 1 = 0$.[8]

[8] Ancient mathematicians like Pythagoras (570-495 BC) and Euclid (360-280 B.C.) have know the golden ratio.

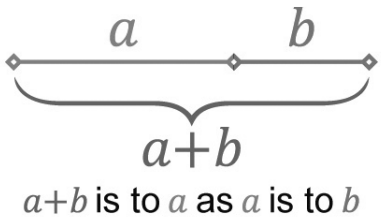

Fig. 4.7. The golden ratio

The golden ratio is related to the regular pentagon, a beautiful figure:

Theorem 4.10. *In a regular pentagon the diagonals intersect in the golden ratio and the ratio of diagonal length to side length is as well the golden ratio.*

$\mathfrak{Proof}$. Denote the edges of the pentagon by A, B, C, D, E and let M be the the point of intersection of the diagonals AC and BD. The triangles ACB and BCM are congruent hence BC/MC=AC/BC. (This can be seen considering the angel sum of the pentagon 3π. The angle BAE is $3/5\pi$ and the angle ABE is $\pi/5$. Hence $AMB = 3/5\pi$ as well). On the other hand $AMDE$ is a parallelogram hence MA=DE=BC. (This can be again seen by considering angels) This gives $MA/MC = AC/MA = AC/BC = \phi$ since $MA + MC = AC$, proving the result. $\mathfrak{Q.E.D.}$

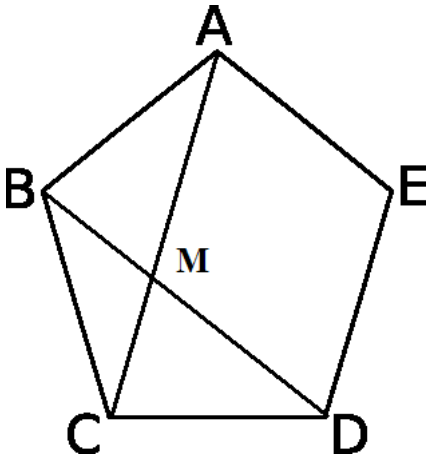

Fig. 4.8. The pentagon

Moreover the golden ratio is intimately related to the fibonacci sequences defined now.

Definition 4.3.2 *The Fibonacci sequence is given by* $f_{n+1} = f_n + f_{n-1}$ *with* $f_0 = 0$ *and* $f_1 = 1$. [9]

[9] Named after the Italian mathematician Fibonacci (1180-1241)

Theorem 4.11.

$$f_n = \frac{1}{\sqrt{5}}(\phi^n - (-\phi)^{-n})$$

where ϕ is the golden ratio.[10]

$\mathfrak{Proof}$. Since $x^{n+1} = x^n + x^{n-1}$ for $x = \phi$ and $x = -\phi^{-1}$ the sequences

$$s_n = c_1\phi^n + c_2(-\phi)^{-n}$$

fulfill the fibonacci recursion. With $s_0 = 0$ and $s_1 = 1$ we get $c_1 = -c_2 = \frac{1}{\sqrt{5}}$. $\mathfrak{Q.E.D.}$

4.4 Lines in the plan

The study of configurations of lines in the plan is one basic subject of combinatorial geometry. The following two lovely theorems form a basis of this field.

Theorem 4.12. *Given a finite number of points in the plane, either all the points lie in the same line; or there is a line which contains exactly two of the points.* [11]

$\mathfrak{Proof}$. We prove the theorem by contradiction. Suppose that we have a finite set of points S not all collinear but with at least three points on each line.
Define a connecting line to be a line which contains at least two points in the collection. Let (P, l) be the point and connecting line that are the smallest positive distance apart among all point-line pairs. By the assumption, the connecting line l goes through at least three points of S, so dropping a perpendicular from P to l there must be at least two points on one side of the perpendicular. Call the point closer to the perpendicular B, and the farther point C. Draw the line m connecting P to C. Then the distance from B to m is smaller than the distance from P to l, since the right triangle with hypotenuse BC is similar to and contained in the right triangle with hypotenuse PC. Thus there cannot be a smallest positive distance between point-line pairsevery point must be distance 0 from every line. In other words, every point must lie on the same line if each connecting line has at least three points. $\mathfrak{Q.E.D.}$

Theorem 4.13. *Any set of $n \geq 3$ noncollinear points in the plane determines at least n different connecting lines. Equality is attained if and only if all but one of the points are collinear.* [12]

[10] This formula is attributed to the French mathematician Jacque Binet (1786-1856), also it was know to Euler and Bernoulli.
[11] This theorem was conjectured by the English mathematician James Joseph Sylvester (1814-1897) and first proved by the Hungarian mathematician Tibor Gallai (1912-1992) .
[12] This was observed by the Hungarian mathematician Paul Erdoš (1913-1996)

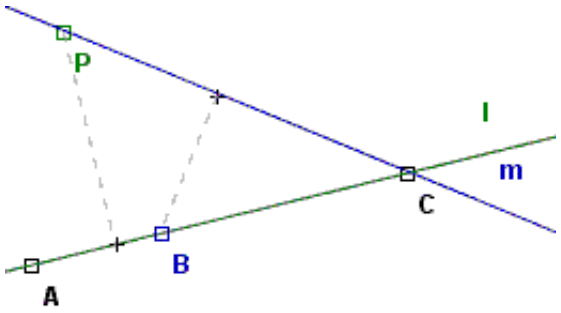

Fig. 4.9. Proof of the theorem of SylvesterGallai

𝔓𝔯𝔬𝔬𝔣. We prove the first part of the theorem by induction. The result is obvious for $n = 3$. If we delete one of the points that lie on an ordinary line, then the number of connecting lines induced by the remaining point set will decrease by at least one. Unless the remaining point set is collinear the result follows. However, in the latter case, our set determines precisely n connecting lines. For the second part of the theorem just consider a near-pencil (a set of $n - 1$ collinear points together with one additional point that is not on the same line as the other points). 𝔔.𝔈.𝔇.

4.5 Construction with compass and ruler

Constructions with compass and ruler are a classical topic in geometry. A nice characterization of constructible numbers is given by:

Theorem 4.14. *We can construct a real number with compass and ruler from* $\{0, 1\}$ *if and only if it is in a field given by quadratic extensions of the rational numbers.*

𝔓𝔯𝔬𝔬𝔣. Given two points a, b on the line we construct $a + b$, $a - b$, ab, a/b by elementary geometry. Hence we can construct the field of all rational numbers from $\{0, 1\}$. Using the theorem of Thales and Pythagoras we see that we also can constructed $\sqrt{a}$ given a hence we can construct all fields given by quadratic extensions of the rationales. On the other hand the only way to construct points in the plane is given by the intersection of two lines, of a line and a circle, or of two circles. Using the equations for lines and circles, it is easy to show that coordinates of points at which they intersect fulfill linear or quadratic equations. Hence all points on the line that are constructible lie a filed given by quadratic extension of the rationales. 𝔔.𝔈.𝔇.

The question which basic constructions are not possible with square and ruler remaind open for a long time. Now we have:

Theorem 4.15. *Squaring the circle, doubling the cube and trisecting all angles is impossible using compass and ruler.*

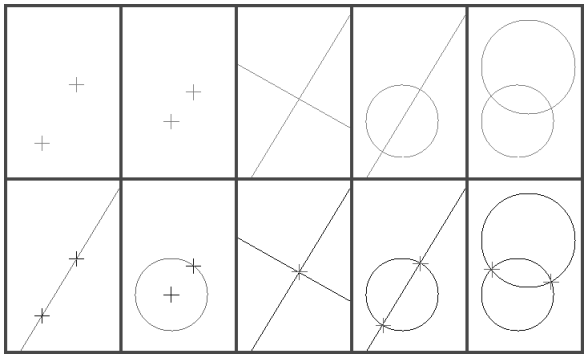

Fig. 4.10. Basic constructions with compass and ruler

𝔓𝔯𝔬𝔬𝔣. If we square the circle we constructed $\sqrt{\pi}$. This is impossible since $\sqrt{\pi}$ is transcendental and hence not in a field given by quadratic extensions. Here we presuppose the transcendence of π, see [15]. If we double the cube we construct $\sqrt[3]{2}$. This is impossible since $\sqrt[3]{2}$ is an extension of degree three and hence not in a field given by quadratic extensions. If we trisect the angle $60°$ we construct $\cos(20°)$. The minimal polynomial of this number is $p(x) = x^3 - 3x - 1$ by trigonometry. Since the polynomial is irreducible $\cos(20°)$ is not in a field given by quadratic extensions. 𝔔.𝔈.𝔇.

4.6 Polyhedra formula

We will formulate and proof Euler's polyhedra formula in graph theoretic language since this prove is nice and short.

Theorem 4.16. *For a connected planar graph G with vertices V, edges E and faces F we have*

$$|V| - |E| + |F| = 2.$$ [13]

𝔓𝔯𝔬𝔬𝔣. Let T be a spanning tree of G, this is a minimal subgraph containing all edges. This graph has no loops.
Consider the dual graph $G^\star$. This is constructed by putting vertices $V^\star$ in each face of F and connecting them by edges $E^\star$ crossing the edges in E. Now consider the edges $T^\star \subseteq E^\star$ in $G^\star$ corresponding to edges $E\backslash T$. The edges of $T^\star$ connect all faces F, since T has no loops. Moreover $T^\star$ itself has no loops. If it had loops vertices inside the loop would be separated from vertices outside the loop which contradicts T to be a spanning tree. Hence $T^\star$ itself is a spanning tree of $G^\star$.
Now let $E(T)$ by the edges of T and $E(T^\star)$ be the edges of $T^\star$. The number of vertices of each tree is the number of edges of the tree plus one (the root!). Hence $E(T) + 1 = |V|$ and $E(T^\star) + 1 = |F|$ and

$$|V| + |F| = |E(T)| + |E(T^\star)| + 2 = |E(T)| + |E(E\backslash T)| + 2 = |E| + 2.$$

Q.E.D.

Using the Polyhedra formula we can characterize the platonic solids (regular convex polyhedron with congruent faces).

Theorem 4.17. *There are fife platonic solids.*[14]

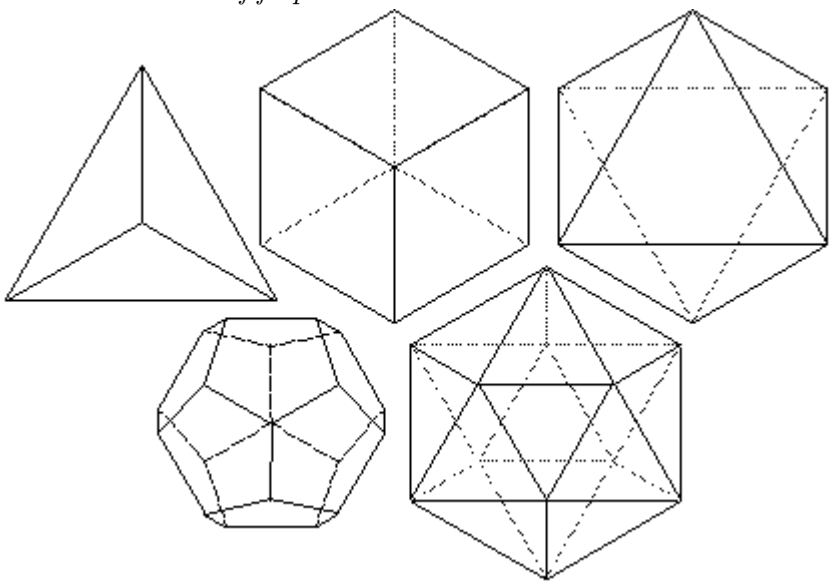

Fig. 4.11. The Platonic solids

Proof. Let n be the number of vertices and edges of a face and m the number of edges adjacent to one vertex. In a platonic solid n and m are constant. Summing over all faces gives $n|F|$ edges where each edge is counted twice, hence $n|F| = 2|E|$. Summing over all vertices gives $m|V|$ edges where each again each edge is counted twice. Hence $m|V| = 2|E|$. Using the Euler formula this gives

$$\frac{1}{m} + \frac{1}{n} = \frac{1}{2} + \frac{1}{|E|}$$

To construct a solid we necessary have $m, n \geq 3$ leading to fife possible platonic solids with $(m, n) = (3, 3), (3, 4), (3, 5), (4, 3), (5, 3)$. All this solids are constructible, see figure 4.1. Foies $(m, n) = (3, 3)$ we get the tetrahedra with $(V, E, F) = (4, 6, 4)$, for $(m, n) = (3, 4)$ the cube with $(V, E, F) = (8, 12, 6)$, for $(m, n) = (4, 3)$ the Octahedra with $(V, E, F) = (6, 12, 8)$, for $(m, n) = (3, 5)$ icosahedra with $(V, E, F) = (12, 30, 20)$ and for $(m, n) = (5, 3)$ the dodecahedral with $(V, E, F) = (20, 30, 12)$. Q.E.D.

[14] Named after the ancient Greek philosopher Plato (428-348 BC).

4.7 Non-euclidian geometry

Any axiomatic geometry based on an incidence axiom, an order axiom, a congruence axiom a continuity axiom and a parallel axiom, [**?**]. The last axiom is independent of the other and gives three different types of geometries.

Definition 4.7.1 *A geometry is euclidian if for any line L and any point x not on the line there is exactly one line parallel to P that meets x. A geometry is elliptic if such a parallel does not exits, it is hyperbolic if there infinity many such parallels.*

Theorem 4.18. *There are models of all three types of geometries.*

$\mathfrak{Proof}$. A model of the Euclidian geometry obviously is the plane $\mathbb{R}^2$.
Now consider the unit sphere $\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 = 1\}$ and identify antipodal points. This defines the set of all points of a geometry. A line line in this geometry is given by a great circle in $\mathbb{S}^2$ where we again identify antipodal points. It is obvious that this geometry is elliptic since any two great circles intersect in two antipodal points.
To construct a hyperbolic geometry consider the upper half plan $\mathbb{H} = \{(x, y) \in \mathbb{R}^2 | y > 0\}$, this describes the set of all points of the geometry. The lines of the geometry are given by vertical lines and semicircles in $\mathbb{H}$ which meet the axis orthogonally. Given such a semicircle or a vertical line L and a point not on it there obviously exists infinitely many semicircles that meet the point without intersecting L. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}$.

These models may be generalized to n-dimensional geometries based on $\mathbb{R}^n$, $\mathbb{S}^{n+1}$ and $\mathbb{H}^n$, see [**?**]. We include here figures describing a ecliptic and hyperbolic geometry.

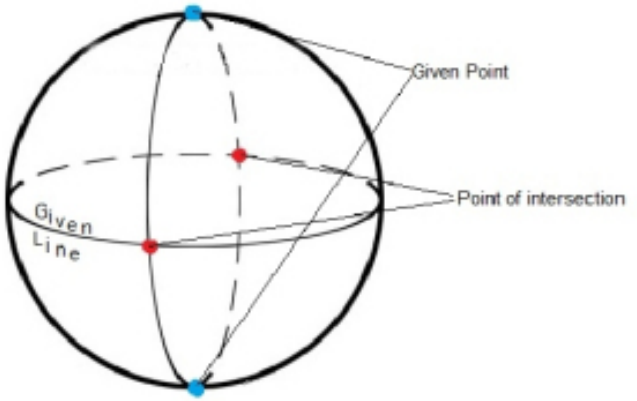

Fig. 4.12. The spherical Model of elliptic geometry

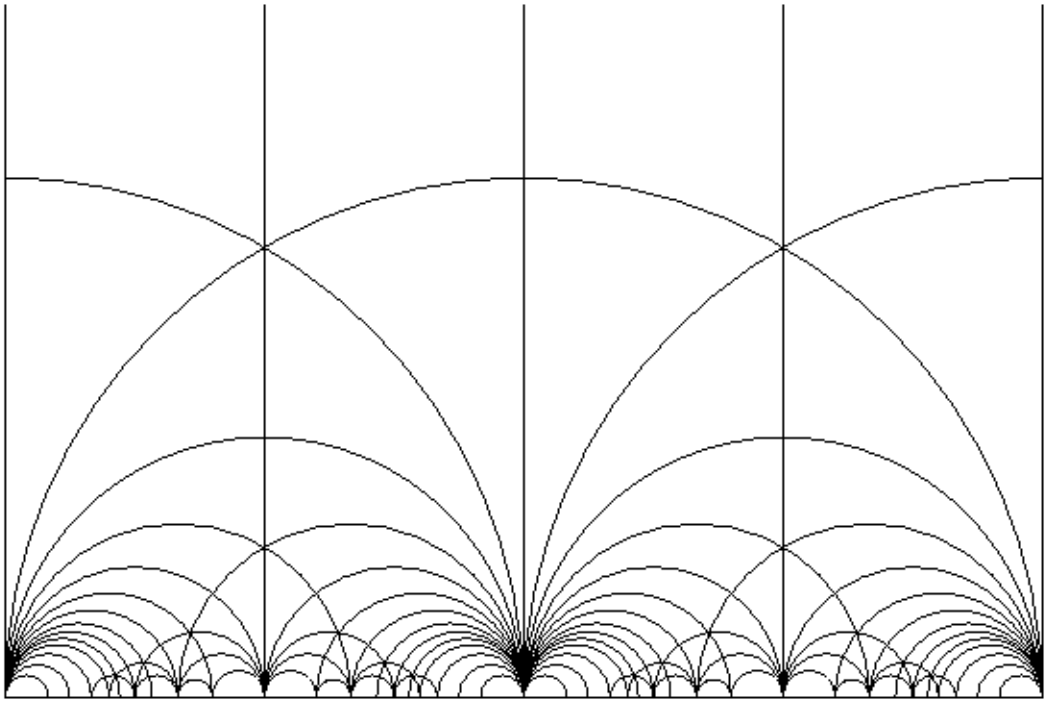

Fig. 4.13. The upper half plane model of hyperbolic geometry

5

Analysis

5.1 Series

As a warm up to this chapter we present beautiful results on infinite series presupposing the notion of convergence and very basic results on limits, see for instance [20].

Theorem 5.1. *For $q \in (0,1)$ we have*

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}. \quad [1]$$

$\mathfrak{Proof}$. By induction we see that

$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}.$$

Taking the limit gives the result. $\mathfrak{Q.E.D.}$

Theorem 5.2. *For the triangle numbers $T_n = n(n+1)/2$ we have*

$$\sum_{n=1}^{\infty} \frac{1}{T_n} = 2. \quad [2]$$

$\mathfrak{Proof}$.

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

[1] This formula was know to the acient Greek mathematician, physicist, and inventer Archimeds (287-212 BC)

[2] Attributed to the German philosopher and mathematician Gottfried Wilhelm von Leibniz (1646-1716)

Theorem 5.3. *Let*

$$H_n := \sum_{k=1}^{n} \frac{1}{k}$$

be the n-th harmonic number. The harmonic series H_n diverges[3]*. On the other hand*

$$\lim_{n \longmapsto \infty} H_n - \log(n) = \gamma$$

where $\gamma \approx 0.57721$ is Euler-Mascheroni constant. [4]*.*

$\mathfrak{Proof}$. First note that

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \sum_{k=0}^{\infty} \sum_{n=1}^{2^k} \frac{1}{2^k + n} \geq \sum_{k=0}^{\infty} 2^k \frac{1}{2^{k+1}} = \sum_{k=0}^{\infty} 1/2 = \infty.$$

Furthermore using elementary properties of integration, see [20] we have

$$H_n - \log(n) = H_n - \int_1^n \frac{1}{x} dx \geq H_n - \sum_{k=1}^{n-1} \frac{1}{x} = \frac{1}{n} > 0$$

and

$$H_{n+1} - \log(n+1) = H_n - \log(n+1) + \frac{1}{n+1} = H_n - \log(n) - \int_n^{n+1} \frac{1}{x} dx + \frac{1}{n+1} \leq H_n - \log(n).$$

This means that sequence $H_n - \log(n)$ is positive and monotone decreasing and hence converges. $\mathfrak{Q.E.D.}$

Theorem 5.4. *We have*

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} = \log(2). \quad [5]$$

$\mathfrak{Proof}$. Considering formal power series we have

$$\frac{\partial \log(x+1)}{\partial x} = \frac{1}{x+1} = \sum_{n=0}^{\infty} (-1)^n x^n.$$

Integrating gives

$$\log(x+1) = \sum_{n=0}^{\infty} (-1)^n \frac{1}{n+1} x^{n+1}.$$

The converges of the series for $x \in (0, 1]$ is guaranteed considering partial sums

[3] This result is due to the french philosopher Nicole Oreseme (1323-1382)
[4] Due to Leonard Euler (1707-1783) and Lorenzo Mascheroni (1750-1800)
[5] Proved by the mathematician Nicholas Mercator (1620 -1687)

$$P_{2k} = \sum_{n=0}^{2k} (-1)^{2k} \frac{1}{n+1} x^{n+1} \quad P_{2k+1} = \sum_{n=0}^{2k+1} (-1)^{2k} \frac{1}{n+1} x^{n+1}$$

which are monotone and bounded and hence convergent. Since the difference $P_{2k} - P_{2k+1}$ goes to zero P_k converges as well. Now setting $x = 1$ gives the result. $\mathfrak{Q.E.D.}$

5.2 Inequalities

The key tool in may analytic proofs are inequalities. We present three important inequalities starting with the arithmetic and geometric mean.

Theorem 5.5. *For positive real numbers a_i for $i = 1, \ldots, n$ we have*

$$\sqrt{a_1 a_2 a_3 \ldots a_n} \leq \frac{a_1 + a_2 + \ldots a_n}{n}. \quad ^6$$

$\mathfrak{Proof}$. Denote the inequality by $I(n)$. Obviously $I(2)$ is correct. We proof $I(n) \Rightarrow I(n-1)$ and $I(n) \wedge I(2) \Rightarrow I(2n)$. The theorem follows by induction. For the first implication let

$$A = \sum_{k=1}^{n-1} a_k/(n-1).$$

By the assumption we have,

$$A \prod_{k=1}^{n-1} a_k \leq (\frac{\sum_{k=1}^{n-1} a_k + A}{n})^n = A^n.$$

Hence

$$\prod_{k=1}^{n-1} a_k \leq A^{n-1}.$$

To see the second implication consider

$$\prod_{k=1}^{2n} a_k = \prod_{k=1}^{n} a_k \prod_{k=n+1}^{2n} a_k \leq^{I(n)} (\sum_{k=1}^{n} a_k/n)^n (\sum_{k=n+1}^{2n} a_k/n)^n$$

$$\leq^{I(2)} (\frac{\sum_{k=1}^{2n} a_k/n}{2})^{2n} = (\frac{\sum_{k=1}^{2n} a_k/n}{2n})^n.$$

$\mathfrak{Q.E.D.}$

The nexte inequalities concern the Euclidian structor of $\mathbb{R}^n$ we now introduce;

[6] Also due to Cauchy (1789 - 1857).

Definition 5.2.1 *Define an inner product* $\langle .,. \rangle$ *on* $\mathbb{R}^n$ *by*

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

This obviously is a positive definite bilinear and symmetric form. The number

$$||x|| = \sqrt{\langle x, x \rangle}$$

is called the Euclidean norm on $x \in \mathbb{R}^n$.

Theorem 5.6.

$$|\langle x, y \rangle| \leq ||x|| \cdot ||y||.^7$$

$\mathfrak{Proof}$. The inequality is trivial true if $y = 0$ thus we may assume $\langle y, y \rangle > 0$.For a number $\lambda \in \mathbb{R}$ we have

$$||x - \lambda y|| = \langle x, x \rangle - \lambda \langle x, y \rangle - \lambda \langle y, x \rangle + \lambda^2 \langle y, y \rangle.$$

Setting $\lambda = \langle x, y \rangle / \langle y, y \rangle$ we get

$$0 \leq \langle x, x \rangle - |\langle x, y \rangle|^2 / \langle y, y \rangle$$

which proves the inequality. $\mathfrak{Q.E.D.}$

Now we prove the triangle inequality, which shows that $||.||$ is in fact a well-defined norm.

Theorem 5.7.

$$||x + y|| \leq ||x|| + ||y||.$$

$\mathfrak{Proof}$.

$$||x+y||^2 = \langle x+y, x+y \rangle = ||x||^2 + ||y||^2 + 2|\langle x, y \rangle| \leq ||x||^2 + ||y||^2 + 2||x|| \cdot ||y|| = (||x|| + ||y||)^2$$

Taking the square root gives the result. $\mathfrak{Q.E.D.}$

5.3 Intermediate value theorem

We assume that the reader is familiar with the notation of continuity, compare [20] and prove:

[7] This is the inequality of the French mathematician August Lois Cauchy (1889-1857) and in the more general setting due to the German mathematician Hermann Amandus Schwarz (1843-1921).

Theorem 5.8. *A continues function $f : [a,b] \longmapsto \mathbb{R}$ takes all values between $f(a)$ and $f(b)$.*[8]

$\mathfrak{Proof}$. Assume without loss of generality $f(a) < f(b)$ and choose a value $\eta \in (f(a), f(b))$. Consider $g(x) = f(x) - \gamma$. Then $g(a) < 0$ and $g(b) > 0$. The set $A = \{x | g(x) < 0\}$ is not empty and bounded so let $\xi = \sup(A)$. Choose a sequence $x_n \in A$ with $x_n \longmapsto \xi$. By continuity $g(x_n) \longmapsto g(\xi)$ hence $g(\xi) \leq 0$. Assume $g(\xi) < 0$ then there exist $x \in (\xi, b]$ with $g(x) < 0$. A contradiction to the definition of ξ. Hence $g(\xi) = f(\xi) - \eta = 0$ and $f(\xi) = \eta$.

Alternative proof using topology, compare with the next section: The closed intervals in $\mathbb{R}$ are the compact and connected sets. The image of a compact and connected set under a continuous map is compact and connected. Hence $f([a,b])$ is an interval containing all values between $f(a)$ and $f(b)$. $\mathfrak{Q.E.D.}$

Corollary 5.9. *A continues function $f : [a,b] \longmapsto [a,b]$ has a fixed point $x \in [a,b]$ with $f(x) = x$.*

The fix point theorem of Dutch mathematician Luitzen Brouwer (1981-1966) generalizes this to continues function of n-dimensional balls or simplices, compare with section 3.6 where we gave a simple combinatorial proof.

5.4 Mean value theorems

First we state here the mean value theorem of differential and than the mean value theorem of integral calculus where we presuppose the notion of a differentiable and integrable real function.

Theorem 5.10. *Let $f : [a,b] \longmapsto \mathbb{R}$ be continues and differentiable in (a,b) than there exits $x \in (a,b)$ with*

$$f^{'}(x) = \frac{f(a) - f(b)}{a - b},$$

for some $x \in (a,b)$.[9]

$\mathfrak{Proof}$. Let

$$g(x) = f(x) - \frac{f(a) - f(b)}{a - b}(x - a)$$

Note that $g(a) = g(b) = f(a)$. If f is not constant g is not constant and hence a maximum or minimum x of g occurs in (a,b). Now we proof that $g^{'}(x) = 0$ leading

[8] This was proofed by the Bohemian mathematician Bernardo Bolzano (1781-1848)
[9] Proved by Cauchy (1789 - 1857)

directly to the result. Assume x is a maximum then $f(x) \leq f(y)$ for all y in a small neighborhood $B_\epsilon(x)$ and hence

$$\frac{g(y)-g(x)}{y-x} \geq 0 \text{ if } y < x \quad \text{and} \quad \frac{g(y)-g(x)}{y-x} \geq 0 \text{ if } y > x$$

Now consider the limit y goes to x. If x is a minimum the argument is analog. $\mathfrak{Q.E.D.}$

Theorem 5.11. *If $f : [a,b] \longmapsto \mathbb{R}$ be continues than there exists $x \in [a,b]$ such that*

$$\int_a^b f(t)\,dt = f(x)(b-a)$$

$\mathfrak{Proof}$. Since f is continues there are numbers m and M such that $f(x) \in [m,M]$ for all $x \in [a,b]$, hence

$$\frac{1}{b-a}\int_a^b f(t)dt \in [m,M].$$

Furthermore the function f on $[a,b]$ takes all values in $[m,M]$ by the intermediate value theorem above. This implies the result. $\mathfrak{Q.E.D.}$

.

5.5 Fundamental theorem of calculus

In the heart of one dimensional analysis we find the fundamental theorem which every high school pupil knows:

Theorem 5.12. *Let f be integrable on $[a,b] \subseteq \mathbb{R}$ and let*

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

than F is continues on $[a,b]$ and differentiable on (a,b) with $F'(x) = f(x)$.[10]

For all $x \in (a,b)$ we have

$$F(x+\epsilon) - F(x) = \int_a^{x+\epsilon} f(t)dt - \int_a^x f(t)dt = \int_x^{x+\epsilon} f(t)dt$$

By the means value theorem of integral calculus there exists $c(\epsilon) \in [x, x+\epsilon]$ such that

$$\int_x^\epsilon f(t)dt = f(c(\epsilon))\epsilon$$

[10] This goes back to Isaac Newton (1643-1727).

Hence we have

$$F^{'}(x) = \lim_{\epsilon \longmapsto 0} \frac{F(x) - F(x+\epsilon)}{\epsilon} = \lim_{\epsilon \longmapsto 0} f(c(\epsilon)) = f(x).$$

$\mathfrak{Q.E.D.}$

Corollary 5.13. *If f is continues and as a antiderivative g with $g' = f$ on an interval $[a,b] \subseteq \mathbb{R}$ than*

$$\int_a^b f(x)dx = g(b) - g(a)$$

$\mathfrak{Proof}$. By our theorem $F(x) = g(x) + c$ on $[a,b]$, with $x = a$ we have $c = -f(a)$ hence $F(b) = g(b) - g(a)$. $\mathfrak{Q.E.D.}$

In fact this result can also be proved under the weaker assumption that f is integrable and has an antiderivative, see [20].

5.6 The Archimedes' constant π

We first define Sinus and Cosine using power series to given an analytic definition of π.

Definition 5.14. *The Sinus and Cosine function are given by*

$$\sin(z) = \sum_{k=0}^{\infty} (-1)^n \frac{1}{(2n+1)!} z^{2n+1} \quad \cos(z) = \sum_{k=0}^{\infty} (-1)^n \frac{1}{(2n)!} z^{2n}$$

for $z \in \mathbb{C}$. The Archimedes' constant is given by

$$\pi = \min\{x > 0 | \sin(x) = 0\}.$$

To prove the converges of the powers series above is an exercise in complex analysis, see [20]. Now we rediscover the geometric meaning of π.

Theorem 5.15. *The area and the half of the circumference of unite circle is given by π.*[11]

$\mathfrak{Proof}$. The upper half of units circle line is given by the function $f(x) = \sqrt{1-x^2}$ for $x \in [-1,1]$. The area is

$$A = \int_{-1}^{1} \sqrt{1-x^2} dx = \int_{\arcsin(-1)}^{\arcsin(1)} \cos^2(t) dt = [\frac{\sin(2t) + 2t}{4}]_{-\pi/2}^{\pi/2} = \pi/2.$$

[11] Know to Archimedes (287-212 BC).

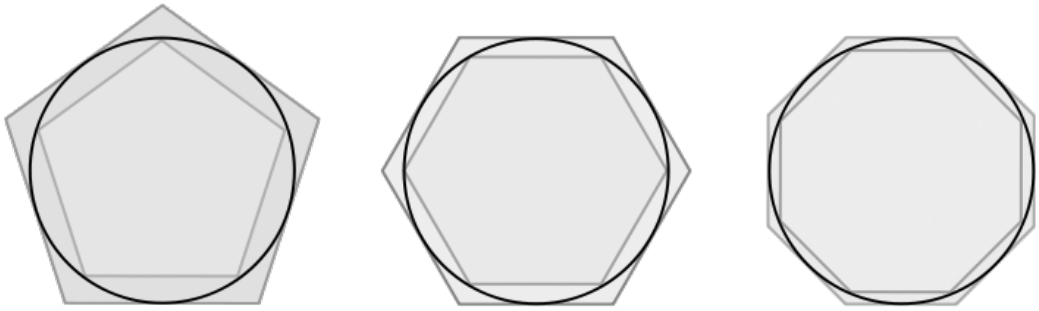

Fig. 5.1. Archimedes approximation of π

The length is

$$l = \int_{-1}^{1} (\sqrt{1 + (\frac{\partial\sqrt{1-x^2}}{\partial x})^2} dx = \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} dx = \int_{\arcsin(-1)}^{\arcsin(1)} 1 dt = \pi.$$

$\mathfrak{Q.E.D.}$

In the following the reader will find four really beautiful presentations of π.

Theorem 5.16.

$$\frac{2}{\pi} = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2+\sqrt{2}}}{2} \cdot \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2} \cdots .$$ [12]

$\mathfrak{Proof}$. We first proof the Euler formula[13]

$$\frac{\sin(x)}{x} = \cos\left(\frac{x}{2}\right) \cdot \cos\left(\frac{x}{4}\right) \cdot \cos\left(\frac{x}{8}\right) \cdots$$

By the doubling formula of sinus $\sin(2x) = 2\sin(x)\cos(x)$ (which you get from the definition by a simple calculation) we have

$$\sin(x) = 2^n \sin(x/2^n) \prod_{i=1}^{n} \cos(x/2^n).$$

Furthermore it is easy to infer from the definition of sinus that:

$$\lim_{n \longmapsto \infty} 2^n \sin(x/2^n) = x$$

given the Euler formula. With $x = \pi/2$ we obtain

$$\frac{2}{\pi} = \cos\left(\frac{\pi}{4}\right) \cdot \cos\left(\frac{\pi}{8}\right) \cdot \cos\left(\frac{\pi}{16}\right) \cdots .$$

[12] The formula is due to the French mathematican Franciscus Vieta (1540-1603).
[13] Due to Leonard Euler (1707-1783).

Now the result follows by the half-angle formula of cosine

$$2\cos(x/2) = \sqrt{2 + 2\cos(x)}$$

and $\cos(4/\pi) = \sqrt{2}/2$. $\mathfrak{Q.E.D.}$

Theorem 5.17.

$$\frac{\pi}{4} = \sum_{k=1}^{\infty} \frac{(-1)^n}{2n+1} \quad [14]$$

$\mathfrak{Proof.}$

$$\frac{\partial \arctan(x)}{\partial x} = \frac{1}{x^2+1} = \sum_{n=0}^{\infty} (-1)^n x^{2n}$$

for $x < 1$. Integrating gives

$$\arctan(x) = \sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1} x^{2n+1}$$

for $x < 1$. Now note that both sides of the equation are continues in $x = 1$. Since $\sin(\pi/4) = \cos(\pi/4)$ tacking the limit gives the result. $\mathfrak{Q.E.D.}$

Theorem 5.18.

$$\frac{\pi^2}{6} = \sum_{k=1}^{\infty} \frac{1}{n^2} \quad [15]$$

$\mathfrak{Proof.}$

$$I = \int_0^1 \int_0^1 \frac{1}{1-xy} dxdy = \sum_{n=0}^{\infty} \int_0^1 \int_0^1 x^n y^n dxsy = \sum_{n=0}^{\infty} \int_0^1 x^n dx \int_0^1 y^n dy = \sum_{k=1}^{\infty} \frac{1}{n^2}$$

On the other hand we ge by substituting $u = (x+y)/2$ and $v = (x-y)/2$

$$I = 4\int_0^{1/2} \int_0^u \frac{1}{1-u^2+v^2} dudv + 4\int_0^{1/2} \int_0^u \frac{1}{1-u^2+v^2} dudv$$

and with

$$\int \frac{1}{a^2+x^2} dx = \frac{1}{a} \arctan x/a$$

[14] First proof by the German philospher and mathematics Gottfried Wilhelm Leibniz. (1646-1716)
[15] This series was calculated by the Swiss mathematican Leonard Euler (1707-1783)

$$I = 4\int_0^{1/2} \frac{1}{\sqrt{1-u^2}} \arctan(\frac{u}{\sqrt{1-u^2}})du + 4\int_{1/2}^{1} \frac{1}{\sqrt{1-u^2}} \arctan(\frac{1-u}{\sqrt{1-u^2}})du$$

Now by substituting $u = \sin\phi$ in the first integral and $u = \cos\phi$ in the second integral this simplifies to

$$I = 4\int_0^{\pi/6} \frac{\phi}{\cos\phi} d\sin\phi + 4\int_0^{\pi/3} \frac{1/2\phi}{\sin\phi} d\cos\phi = 4\int_0^{\pi/6} \phi d\phi + 2\int_0^{\pi/3} \phi d\phi = \frac{\pi^2}{6}.$$

$\mathfrak{Q.E.D.}$

This theorem gives the the value $\zeta(2)$ of the Riemanian Zeta function, see chapter six. With more effort using the Bernoulli numbers it also possible to calculate $\zeta(2n)$ explicitly.

Theorem 5.19.

$$\frac{\pi}{2} = \prod_{k=1}^{\infty} \frac{4k^2}{4k^2-1} \quad ^{16}$$

$\mathfrak{Proof}$. By partitial integration we get the recursion

$$I_n := \int_0^{\pi/2} \sin^n(x)dx = \frac{n-1}{n}\int_0^{\pi/2} \sin^{n-2}(x)$$

with staring values $I_0 = \pi/2$ and $I_1 = 1$. Hence

$$I_{2j} = \frac{\pi}{2}\prod_{k=1}^{j} \frac{2k-1}{2k} \text{ and } I_{2j+1} = \prod_{k=1}^{j} \frac{2k}{2k+1}$$

Since

$$\sin^{2j+1}(x) \leq \sin^{2j}(x) \leq \sin^{2j-1}(x)$$

we get by integrating

$$\prod_{k=1}^{j} \frac{2k}{2k+1} \leq \frac{\pi}{2}\prod_{k=1}^{j} \frac{2k-1}{2k} \leq \prod_{k=1}^{j-1} \frac{2k}{2k+1}.$$

Dividing by the right hand side and taking $j \longmapsto \infty$ gives the result. $\mathfrak{Q.E.D.}$

5.7 The Euler number e

Definition 5.7.1 *The Euler function is given by*

$$e^z := \exp(z) = \sum_{k=0}^{\infty} \frac{1}{n!} z^n$$

for $z \in \mathbb{C}$. The Euler number is

$$e = \exp(1) = \sum_{k=0}^{\infty} \frac{1}{n!}.^{17}$$

Using complex numbers in the definition we obtain Euler's relation between the exponential function and tintometric functions.

Theorem 5.20.

$$e^{iz} = \cos(z) + \sin(z)i \ ^{18}$$

$\mathfrak{Proof}$.

$$e^{iz} = \sum_{k=0}^{\infty} \frac{1}{n!} i^n z^n = \sum_{k=0}^{\infty} \frac{1}{(2n)!} i^{2n} z^{2n} + \sum_{k=0}^{\infty} \frac{1}{(2n+1)!} i^{2n+1} z^{2n+1}$$

$$= \sum_{k=0}^{\infty} \frac{1}{(2n)!} (-1)^n z^{2n} + (\sum_{k=0}^{\infty} \frac{1}{(2n+1)!} (-1)^n z^{2n+1})i = \cos(z) + \sin(z)i$$

since $i^2 = -1$. $\mathfrak{Q.E.D.}$

As a corollary we obtain the fairest formula of mathematics.

Corollary 5.21.

$$e^{2\pi i} = 1,$$

and the following representation of tintometric functions

Corollary 5.22.

$$\cos(z) = \frac{e^{iz} + e^{-iz}}{2} \quad \sin(z) = \frac{e^{iz} - e^{-iz}}{2i}$$

$\mathfrak{Proof}$. Just past in the expression for e^{iz} and e^{-iz} on the rind hand side. $\mathfrak{Q.E.D.}$

Furthermore we obtain:

Corollary 5.23.

$$(\cos(z) + \sin(z)i)^n = \cos(nz) + \sin(nz)i \ ^{19}$$

$\mathfrak{Proof}$. Use the properties of the exponential function. $\mathfrak{Q.E.D.}$

There is another beautiful presentation of the exponential function, which may also be used as its definition.

[17] Introduced by by Euler (1707-1783)

[19] This is the theorem of the French mathematician Abraham de Moivre (1667-1754)

Theorem 5.24.

$$e^z = \lim_{n\longmapsto\infty}(1+z/n)^n$$

$\mathfrak{Proof}$. Using the Binomial theorem

$$\sum_{k=0}^{n}\frac{z^k}{k!}-(1+z/n)^n=\sum_{k=0}^{n}\frac{z^k}{k!}(1-\frac{n!}{(n-k)!n^k})\leq\frac{z^2}{2n}+\sum_{k=3}^{n}\frac{z^k}{k!}(1-\prod_{l=0}^{k-1}(1-l/n))$$

$$\leq\frac{z^2}{2n}+\sum_{k=3}^{n}\frac{z^k}{k!}(1-(1-\frac{k-1}{n})^k)\leq\frac{z^2}{2n}+\sum_{k=3}^{n}\frac{z^k}{k!}\frac{k(k-1)}{n}=\frac{1}{n}(\frac{z^2}{2}+\sum_{k=1}^{n-2}\frac{z^k}{k!})$$

Now this upper estimate tends to zero with $n\longmapsto\infty$. This implies the result. $\mathfrak{Q.E.D.}$

At the end of the section we include the sterling formula which relates the factorials to the Euler number e.

Theorem 5.25.

$$\lim_{n\to\infty}\frac{n!}{\sqrt{2\pi n}\ n^ne^{-n}}=1.^{20}$$

$\mathfrak{Proof}$. Let

$$a_n=\frac{n!}{\sqrt{n}\ n^ne^{-n}}$$

By a simple calculation we have

$$\log(\frac{a_n}{a_{n+1}})=(n+\frac{1}{2})\log(1+\frac{1}{n})-1$$

$$=(n+\frac{1}{2})\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{kn^k}-1=\frac{1}{12n^2}+\text{terms of higher order.}$$

Hence there is $N>0$ such that for all $n>N$

$$0<\log(\frac{a_n}{a_{n+1}})<\frac{1}{6n^2}.$$

Now for $M\geq N$ we have

$$\log(a_N)-\log(a_M)\leq\sum_{n=N}^{M-1}\log(\frac{a_n}{a_{n+1}})\leq\sum_{n=N}^{M-1}\frac{1}{6n^2}\leq\frac{\pi^2}{36}$$

using the Euler series above. This gives

$$a_M\geq\exp(\log a_N-\frac{\pi^2}{36}):=C>0$$

[20] The formula is due to the Scottish mathematician James Sterling (1692-1770).

Thus we see that the sequence a_n is decreasing and bounded from below for $n > N$. Thus a_n converge to $\alpha > 0$. Let $h_n = a_n - \alpha$. We have

$$\frac{(\alpha + h_n)^2}{\alpha + h_{2n}} = \frac{a_n^2}{a_{2n}} = \sqrt{4 \prod_{k=1}^{\infty} \frac{4k^2}{4k^2 - 1}}$$

Now the formula of Wallis, see theorem 5.19, gives

$$\alpha = \lim_{n \longmapsto \infty} \frac{(\alpha + h_n)^2}{\alpha + h_{2n}} = \sqrt{2\pi},$$

which completes the proof. $\mathfrak{Q.E.D.}$

5.8 The Gamma function

Related to the Sterling formula in the end of the last section we like to interpolate the factorials by a analytic. Therefore a definition:

Definition 5.8.1 *The Gamma function* $\Gamma : (0, \infty) \longmapsto \mathbb{R}$ *is given by*

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$$

[21]

The integral in this definition improper but the reader may easily check that limits involved exists. Using complex analysis we might in addition show that Γ is merophorphic with simple poles at negative integers [20]. Here we are interested in the following nice property of Γ:

Theorem 5.26. *The Gamma function fulfills the functional equation*

$$\Gamma(x + 1) = x\Gamma(x)$$

with $\Gamma(1) = 1$*, especially* $\Gamma(n) = (n - 1)!$ *for all* $n \in \mathbb{N}$

$\mathfrak{Proof}$. Obviously

$$\Gamma(1) = \int_0^{\infty} e^{-t} dt = [-e^{-t}]_0^{\infty} = 1$$

and integrating by parts yields

[21] The gamma function was introduced by Euler (1707-1783)

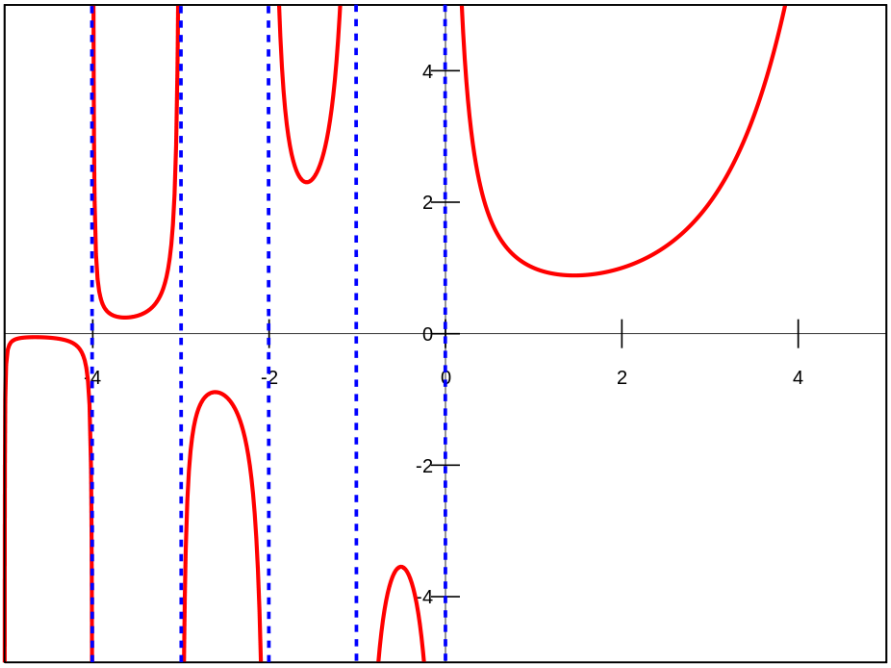

Fig. 5.2. The Gamma function

$$\Gamma(x+1) = \int_0^\infty t^x e^{-t} dt = [-t^x e^{-t}]_0^\infty + x \int_0^\infty t^{x-1} e^{-t} dt = x\Gamma(x).$$

Q.E.D.

Theorem 5.27. *The Gamma function has the representation*

$$\Gamma(x) = \lim_{n \longmapsto \infty} \frac{n^x n!}{x(x+1)(x+2)\dots(x+n)}.$$[22]

Proof. Using theorem 5.24, interchanging limits and substituting we have

$$\Gamma(x) = \lim_{n \longmapsto \infty} \int_0^\infty t^{x-1}(1-t/n)^n dt = \lim_{n \longmapsto \infty} n^x \int_0^1 t^{x-1}(1-t)^n dt$$

$$=: \lim_{n \longmapsto \infty} \Gamma(x,n)$$

Integrating $\Gamma(x,n)$ by parts we find

$$\Gamma(x,1) = \frac{1}{x(x+1)}$$

[22] This represtation is due to the German mathematician and physical scientist Carl Friedrich Gauß(1777-1855).

and

$$\Gamma(x, n+1) = \frac{1}{x}(\frac{n}{n-1})^{x+1}\Gamma(x+1, n-1),$$

yielding

$$\Gamma(x, n) = \frac{n^x n!}{x(x+1)(x+2)\dots(x+n)}$$

$\mathfrak{Q.E.D.}$

There is another beautiful product presentation of γ using the EulerMascheroni constant γ from the beginning of this chapter.

Theorem 5.28.

$$\Gamma(x) = \frac{e^{-\gamma x}}{x}\prod_{n=k}^{\infty}\frac{e^{x/k}}{(1+\frac{x}{k})}$$ [23]

$\mathfrak{Proof}$. Using the last theorem and theorem 5.3 we have

$$\begin{aligned}
\Gamma(x) &= \lim_{n\longmapsto\infty}\frac{n^x}{x(1+x)(1+x/2)\dots(1+x/n)} \\
&= \lim_{n\longmapsto\infty}\frac{n^x}{x}\prod_{k=1}^{n}(1+\frac{x}{k})^{-1} \\
&= \lim_{n\longmapsto\infty}\frac{e^{x(\log(n)-H_n+H_n)}}{x}\prod_{k=1}^{n}(1+\frac{x}{k})^{-1} \\
&= \lim_{n\longmapsto\infty}\frac{e^{x(\log(n)-H_n)}}{x}\prod_{k=1}^{n}(1+\frac{x}{k})^{-1}e^{x/k} \\
&= \frac{e^{-\gamma x}}{x}\prod_{k=1}^{\infty}(1+\frac{x}{k})^{-1}e^{x/k}.
\end{aligned}$$

$\mathfrak{Q.E.D.}$

[23] This represtation is due to the German mathematician Karl Weierstrass (1815-1897)

6

Topology

6.1 Compactness and Completeness

In this chapter we assume that the reader is familiar with very basic topology, namely open and closed sets. We develop the notion of compactness and completeness here.

Definition 6.1.1 *A metric space is compact if any open covers has a finite subcover; it is totally bounded if it can be covered by finitely many open balls of a given radius and it is complete if every Cauchy sequence has a limit in the space. Furthermore a metric space is separable if there a countable dense set.*

Theorem 6.1. *A metric space X is compact if and only if it is complete and totally bounded.*

$\mathfrak{Proof}$. Assume that X is compact. The cover of X by all open balls of radius ϵ has a finite subcover. Hence X is totally bounded. Now assume that X is not complete. Than there is a Cauchy sequence (x_n) that does not converge. The union of the elements of the sequence is closed and the sets

$$O_n = X \setminus \bigcup_{i=n}^{\infty} \{x_i\}$$

form an open cover of X. This cover does not have a finite subcover since it is increasing. A contradiction.
Now assume X is complete and totally bounded. We first proof that X is sequential compact, that is every sequence (x_n) has a convergent subsequence. Since X is totally bounded there is a ball of radius 1 that contains a subsequence of (x_n). By induction we constructed nested subsequences lying in balls of radius $1/2^n$. Using the diagonal process we get a Cauchy subsequence. Since X is complete this sequence converges.
Now we proof that a sequential compact space is separable. If not there exists an $\epsilon > 0$ such that to every countable set there is a point with distance greater than ϵ to this set. By induction we get a sequence of point with distance grater ϵ. This sequence does not contain a convergent subsequence.
Since X is separable every open cover contains a countable subcover.
Now we finish the proof. If X has a countable cover $\{O_i\}$ without a finite subcover,

take the union of the first n sets and choose a point x_n out of this union. This sequence has a convergent subsequence by completeness of X with limit in some O_i. But for n large x_n is not in O_i. Hence the sequence does not converge. A contradiction.$\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

Now we have a look at the spaces $\mathbb{R}^n$ and obtain two nice corollaries.

Corollary 6.2.
A subset K of $\mathbb{R}^n$ is compact if and only if K is closed and bonded.[1]

$\mathfrak{Proof}$. A closed subset of a complete metric space is complete and a bounded subset of $\mathbb{R}^n$ is totally bounded. On the other hand a totaly bounded subset of $\mathbb{R}^n$ is bounded and any complete set is closed. Thus this result is a corollary to the last theorem. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

Corollary 6.3. *A bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.*[2]

$\mathfrak{Proof}$. The sequence is contained in compact and hence bounded and complete ball. This ball is sequential compact as show in the prove of 6.1. □

6.2 Homömorphic Spaces

Topology is concerned with spaces of to bicontinous changes. Therefore the following definition.

Definition 6.2.1 *Two metric spaces are homömorphic if there exists a continues bijection with continues inverse between them.*

Theorem 6.4. $\mathbb{R}$ *is not homömorphic to* $\mathbb{R}^n$ *for* $n \geq 2$.

$\mathfrak{Proof}$. First not that $A = \mathbb{R}^n \backslash \{0\}$ is path connected for $n \geq 2$; there are continuous maps $f : [0,1] \longmapsto A$ with $f(0) = x$ and $f(1) = y$ for all $x, y \in A$. Now assume that $h : \mathbb{R}^n \longmapsto \mathbb{R}$ is a homömorphism. Then $\mathbb{R}\backslash\{h(0)\}$ would be past connected via the continuous maps $h \circ f : [0,1] \mapsto \mathbb{R}\backslash\{h(0)\}$ with $h \circ f(0) = h(x) = a$ and $h \circ f(1) = h(y) = b$ for all $a, b \in \mathbb{R}\backslash\{h(0)\}$. Obviously this is a contradiction.$\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

We are sorry not to know a simple prove of the fact $\mathbb{R}^n$ is in general not homömorphic to $\mathbb{R}^m$ for $m \neq n$. This can be shown using the machinery of algebraic topology, see [**?**] for instance. The topological difference between the slat spaces and the spere is simple to prove.

Theorem 6.5. *There are no $n, m \in \mathbb{N}$ such that the sphere $\mathbb{S}^n$ is homömorphic to $\mathbb{R}^m$.*

[1] Found by the German mathematician Eduard Heine (1821-1881) and the French mathematician Emile Borel (1871-1956) .
[2] This is the theorem of Bolzano (1781-1848) and Weierstrass (1815-1897).

$\mathfrak{Proof}$. $\mathbb{S}^n$ is compact, since it is closed and bounded in $\mathbb{R}^{n+1}$. Now the image of a compact space under a homömorphism is obviously compact. On the other hand $\mathbb{R}^m$ is not compact, because it is not bunded. $\mathfrak{Q.E.D.}$

Now we consider the Hilbert cube $[0,1]^{\infty}$[3] with the product metric

$$d((x_i, y_i) = \sum_{i=1}^{\infty} |x_i - y_i| 2^{-i}$$

and prove the following beautiful theorem:

Theorem 6.6. *A compact separable metric spaces X is homömorphic to a subset of the Hilbert cube $[0,1]^{\infty}$.*

$\mathfrak{Proof}$. With out less of generality we may assume that the diameter of X is less than one by multiplying the metric with a constant factor if necessary. By separability of X there is a dense sequence x_n. Now define

$$F : X \longmapsto [0,1]^{\infty} \text{ by } F(x) = (d(x, x_i))$$

Obviously this map is injective and continuous, hence a homeṁorphism to a subset of the Hilbert cube. $\mathfrak{Q.E.D.}$

6.3 A Peano curve

It follows from the prove of theorem 6.4 in the last section that the unit interval $[0,1]$ is not homömorphic to the unite square $[0,1]^2$. On the other hand we have the following surprising result that shows that there are exits space filling curves.

Theorem 6.7. *There is a continues surjection from $[0,1]$ to $[0,1]^2$*[4].

$\mathfrak{Proof}$. Define a metric on $[0,1]^2$ by

$$d((x_1, y_1), (x_2, y_2) = \max\{|x_1 - x_2|, |y_1 - y_2|\}$$

and a metric on the set of continues functions $C([0,1], [0,1]^2)$ by

$$d_{\infty}(f, g) = \sup\{d(f(t), g(t)) | t \in [0,1]\}.$$

Using uniform convergence the reader will easily show that a Cauchy Sequence (f_n) in $C([0,1], [0,1]^2)$ converges to a continuous function. Hence the space is complete. Now construct a sequence (f_n) in $C([0,1], [0,1]^2)$ as show in figure 6.1. We will show

[3] Introduced by the German mathematician David Hilbert (1826-1943)
[4] Constructed by the Italian mathematician Giuseppe Peano (1858-1932)

that this is a Cauchy sequence. Note that $d_\infty(f_n, f_{n+1}) \leq 2^{-n}$ since $f_n(t)$ and $f_{n+1}(t)$ are in the same subsquare of diameter 2^{-n} for $t \in [i/2^n, (i+1)/2^n]$. Thus for $m > n$

$$d_\infty(f_n, f_m) \leq \sum_{i=n}^{m-1} 2^{-i} \leq \sum_{i=n}^{\infty} 2^{-i} = 2^{-n+1}.$$

By completeness there is a continuous function $f \in C([0,1],[0,1]^2)$ such that $f_n \to f$. It remains to show that f is surjective. Fix $x \in [0,1]^2$. Given $\epsilon > 0$ choose N sufficient large such that $2^{-N} < \epsilon$ and $d_\infty(f_N, f) < \epsilon/2$. Then there exists $t \in [0,1]$ with $d(f_N(t); x) \leq 2^{-N}$. So

$$d(f(t), x) \leq d(f(t), f_N(t)) + d(f_N(t), x) \leq \epsilon.$$

We have shown thus shown that x is in the closure of $f([0,1])$, but this set is compact hence $x \in f([0,1])$. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

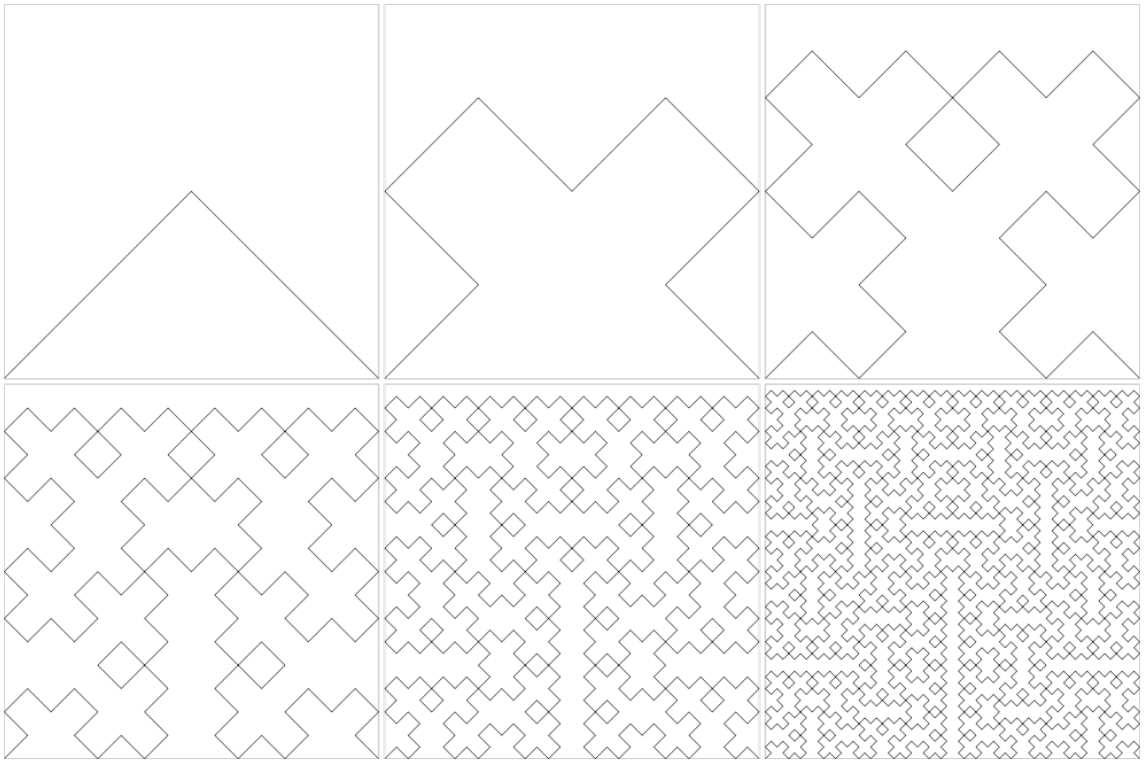

Fig. 6.1. Construction of the Peano curve

6.4 Tychonoff's theorem

In section 6.2 we have shown that any compactum can be embedded into the Hilbert cube. The question if this space is itself compact is answered by the following theorem.

Theorem 6.8. *If X_i for $i \in I$ is a family of compact metric spaces than*

$$\prod_{i \in I} X_i$$

is compact. [5]

We use the converges of filters (which were defined in section 2.5) to characterize compactness.

Definition 6.4.1 *A filter F on a metric space X converges to $x \in X$ ($F \to x$), if any neighborhood of x is contained in the filter F.*

Theorem 6.9. *A metric space X is compact if and only if every ultrafilter on X is convergent.*

$\mathfrak{Proof}$. Let F be an ultrafilter. Consider $\{\overline{A} | A \in F\}$. The intersection over finite subcollections of this family is not empty. Hence by compactness there is an $x \in X$ with

$$x \in \bigcap_{A \in F} \overline{A}.$$

Let N be the by the system of all neighborhoods of x and $B \in N$ than $A \cap B \neq \emptyset$ for all $A \in F$. This implies $N \subseteq F$ since F is an ultrafilter. Hence $F \to x$.
On the other hand suppose $\mathbf{O}$ is an open cover of X with no finite subcover. Consider $B = \{(X \backslash O_1) \cap \ldots \cap (X \backslash O_n) | O_i \in \mathbf{O}\}$. By theorem 2.15 there exists an ultrafilter F that contains B. By the assumption $F \to x \in X$. So there exists $O \in \mathbf{O}$ with $x \in O$ and $O \in F$ but $X \backslash O \in F$ as well. Hence F is not an ultrafilter. A contradiction. $\mathfrak{Q.E.D.}$

$\mathfrak{Proof}$. of Theorem 6.7. Let F be a ultrafilter on $X = \prod_{i \in I} X_i$ so $F_i := pr_i F$ is an ultrafilter on X_i. Since X_i is compact $F_i \to x_i$ for some $x_i \in X$. From this it is easy to see that $F \to (x_i)_{i \in I} \in X$. Hence every ultrafilter in X converges and thus X is compact. $\mathfrak{Q.E.D.}$

[5] This theorem is due to the Russian mathematician Andrei Tychonoff (1906-1993)

6.5 The Cantor set

We discus here uncountable of (in some topological sense) small sets.

Definition 6.5.1 *The (middle-third) Cantor set is given by*

$$C = \{\sum_{n=1}^{\infty} s_n 3^{-n} | s_n \in \{0, 2\}\}$$

6

Theorem 6.10. *C is uncountable, compact, totally disconnected and perfect.*

$\mathfrak{Proof}$. Obviously C is uncountable since $\{0, 2\}^{\infty}$ is. We may write C in the following way

$$C = \bigcap_{k=1}^{\infty} \bigcup_{s_1, s_2, \ldots, s_k \in \{0,2\}} [\sum_{n=1}^{k} s_n 3^{-n}, \sum_{n=1}^{k} s_n 3^{-n} + 3^{-n}]$$

We see that C is closed as the intersection of closed sets. Moreover C is obviously bounded since $C \subseteq [0, 1]$. Hence C is compact by corollary 6.2. Furthermore the length of C is given by

$$\ell(C) = \lim_{k \longmapsto \infty} (2/3)^k = 0$$

hence C contains no interval and thus is totally disconnected. Now consider the ϵ interval

$$[\sum_{n=1}^{\infty} s_n 3^{-n} - \epsilon, \sum_{n=1}^{\infty} s_n 3^{-n} + \epsilon]$$

around a point in $x \in C$. If m is large enough the interval contains

$$\{\sum_{n=1}^{m} s_n 3^{-n} + \sum_{n=m+1}^{\infty} t_n 3^{-n} | t_n \in \{0, 2\}\}$$

hence x is an accumulation point of C and C is perfect. $\mathfrak{Q.E.D.}$

The following theorem shows the Cantor set is universal.

Theorem 6.11. *Any compact, totally disconnected and perfect metric space is homeomorphic to the Cantor set C.*

$\mathfrak{Proof}$. In a totally disconnected metric space any point is contained inside a set of arbitrarily small diameter which is both closed and open. So consider a cover of the space by such sets of diameter smaller than one. Take a finite subcover. Since any finite intersection of such sets is still both closed an open by taking all possible

[6] The Cantor set Irish mathematician by Henry Smith (1826-1883) and introduced by Georg Cantor

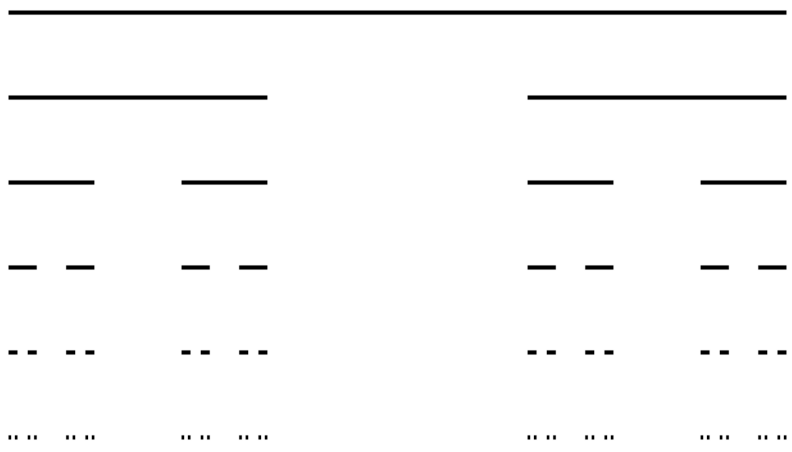

Fig. 6.2. Construction of the Cantor set

intersection we obtain a partition of the space into finitely many closed and open sets of diameter smaller one. Since the space is perfect no element of this partition is a point so a further division is possible. Proceeding this procedure we obtain a nested sequence of finite partitions into closed and open sets of positive diameter less than $1/2^n$ for $n \in \mathbb{N}$. Mapping elements of each partition inside the nested sequence of contracting intervals

$$[\sum_{n=1}^{k} s_n 3^{-n}, \sum_{n=1}^{k} s_n 3^{-n} + 3^{-n}]$$

we construct a homeomorphism of the space onto C. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

6.6 Baires' category theorem

We introduce her a notion of the size of a set in the topological sense.

Definition 6.6.1 *A subset of a metric space is nonwhere dense if there is no neighborhood on which the set is dense. A subset of a metric space is called of first category or meager if it is the union of countable many nonwhere dense closed set. Otherwise it is called of second category or fat.*

Theorem 6.12. *In a complete metric space a countable intersection of open and dense sets is dense.*[7]

[7] This is due to French mathematician Rene Loise Baire (1974-1932)

𝔓𝔯𝔬𝔬𝔣. Let $\{O_i\}_{i\in\mathbb{N}}$ be a family of open and dense subsets of a complete metric space X. For an arbitrary open set B_0 in X we choose inductively open balls B_{i+1} for radius at most $1/i$ such that $\bar{B}_{i+1} \subseteq B_i \cap O_{i+1}$. By this construction we have

$$\bigcap_{i=1}^{\infty} B_i \subseteq B_0 \cap \bigcup_{i=1}^{\infty} O_i.$$

The first intersection is not empty since the centers of the balls converges by completeness. 𝔔.𝔈.𝔇.

Corollary 6.13. *A complete metric space is of second category.*

Now we prove that differential functions are exceptional in the space of continuous functions.

Theorem 6.14. *The set of somewhere differentiable functions $D(I)$ on an interval I is meager in the fat space $C(I)$ of all continuous function on I with the supremum norm.*[8]

A beautiful example of a continuous and nowhere differentiable function is

$$f(x) = \sum_{k=0}^{\infty} (1/2)^k \cos(2^k \pi x)$$

see, [25].[9]

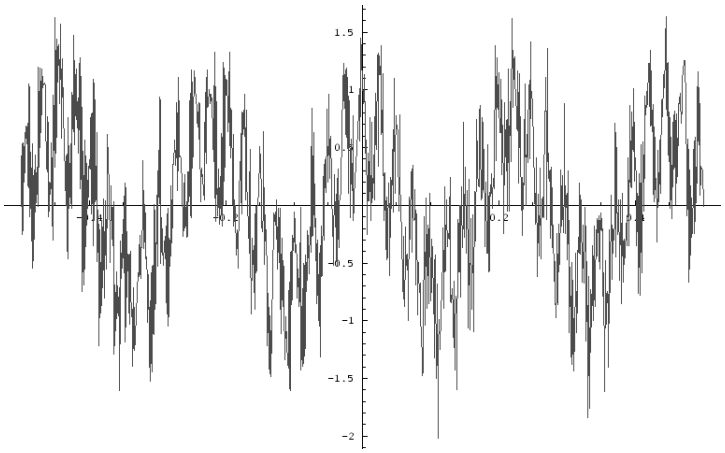

Fig. 6.3. The Weierstrass function

𝔓𝔯𝔬𝔬𝔣. First note that $C(I)$ is fat since it is a complete metric space. Now let

[8] Proofed by the Polish mathematician Stefan Banach (18921945)
[9] This type of functions were introduced by the German mathematician Karl Weierstrass (1815-1897).

$$D(I) = \{f \in C(I) | f \text{ is differentable for some } x \in I\}$$

and

$$A_{n,m} = \{f \in C(I) | \exists x \in I \ : \ \frac{|f(t) - f(x)|}{|t - x|} \leq n \text{ if } |x - t| \leq 1/m\}$$

Note that if $f \in D$ then $f \in A_{n,m}$ for some n, m hence $D(I)$ is the union of all $A_{n,m}$. We now proof that all $A_{n,m}$ are closed and non where dense. Consider a cauchy sequence with $f_n \longmapsto f$ in $A_{n,m}$. For each i there is an x_i with

$$\frac{|f_i(t) - f(x_i)|}{|t - x|} \leq n \text{ if } |x - t| \leq 1/m$$

By the theorem of Bolzano Weierstrass there is a convergent subsequence of x_i with limit x. Now

$$\lim_{i \longmapsto \infty} \frac{|f_i(t) - f(x_i)|}{|t - x_i|} = \frac{|f(t) - f(x)|}{|t - x|} \leq n \text{ if } |x - t| \leq 1/m$$

Hence $f \in A_{n,m}$. So $A_{n,m}$ is closed. Now we proof that $A_{n,m}$ does not contain an open ball $B_\epsilon(f)$ and is thus nowhere dense. Consider a piecewise linear function p with $|p - f| < \epsilon/2$ and choose $k > 2(M + n)/\epsilon$ where M is a bound on the modulus of the derivatives of p. There is a continues piecewise linear function $\phi(x)$ with $|\phi(x)| < 1$ and $\phi'(x) = \pm k$ where the function is differentiable. Let

$$g(x) = p(x) + \epsilon/2\phi(x)$$

On the one hand $g \in B_\epsilon(x)$ by construction. On the other hand we show $g \notin A_{n,m}$ proving $B_\epsilon(f) \not\subseteq A_{n,m}$. If g is differentiable in x then $g'(x) = |p'(x) \pm k\epsilon/2|$. Since $p'(x) \leq M$ and $g'(x) > n$ there is $l > m$ such that g is linear with slope greater l on $[x - 1/l, x + 1/l]$. Hence

$$\frac{|g(t) - g(x)|}{|t - x|} > n \text{ if } |x - t| \leq 1/l < 1/m$$

giving $g \notin A_{n,m}$. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

6.7 Banachs' fix point theorem and fractals

Theorem 6.15. *Let X be a complete metric spaces and $f : X \longmapsto X$ be a contraction than there is a unique fix point $\bar{x} \in X$ with $f(\bar{x}) = \bar{x}$. Moreover for all $x \in X$ we have*

$$\lim_{n \longmapsto \infty} f^n(x) = \bar{x}. \quad [10]$$

[10] A result of the Polish mathematician Stefan Banach (1892- 1945).

𝔓𝔯𝔬𝔬𝔣. Let $\lambda \in (0,1)$ be the contraction constant. By induction we see that

$$d(f^n(x), f^n(y)) \leq \lambda^n d(x,y).$$

Now $(f^n(x))$ is a Cauchy sequence for all $x \in X$ since

$$d(f^m(x), f^n(x)) \leq \sum_{k=0}^{m-n-1} d(f^{n+k+1}(x), f^{n+k}(x))$$

$$\leq \sum_{k=0}^{m-n-1} \lambda^{n+k} d(f(x), x) \leq \frac{\lambda^n}{1-\lambda} d(f(x), x) \to 0$$

and does thus converge by completeness. Let $\bar{x}$ be the limit. Since

$$d(\bar{x}, f(\bar{x})) \leq d(\bar{x}, f^n(x)) + d(f^n(x), f^{n+1}(x)) + d(f^{n+1}(x), f(\bar{x}))$$
$$\leq (1+\lambda) d(\bar{x}, f^n(x)) + \lambda^n d(x, f(x)) \to 0$$

$\bar{x}$ is a fix point. Since

$$d(f^n(x), \bar{x}) \leq \frac{\lambda^n}{(1-\lambda)} d(f(x), x))$$

for all x the fix point is unique. 𝔔.𝔈.𝔇.

The following application of the fix point theorem provides a basic construction in fractal geometry, see [7].

Theorem 6.16. *Let (X,d) be a complete metric spaces and T_i for $i = 1, \ldots n$ be contractions on X than there is a unique compact attractor K set with*

$$K = \bigcup_{i=1}^{n} T_i(K).$$

𝔓𝔯𝔬𝔬𝔣. Consider the space of all $\mathfrak{K}$ of all compact subset of X. With the Hausdorff metric

$$d_H(A,B) = \max\{\sup_{a \in A} \inf_{b \in B} d(a,b), \sup_{b \in B} \inf_{a \in A} d(a,b)\}$$

$\mathfrak{K}$ becomes a complete metric space, since X is complete metric. Now consider the operator

$$\mathfrak{O}(K) = \bigcup_{i=1}^{n} T_i(K)$$

on $\mathfrak{K}$. This a contraction since T_i are contractions. Now the fixpoint theorem gives the result. 𝔔.𝔈.𝔇.

Example 6.17. On $\mathbb{R}^2$ consider the contractions

$$T_1 x = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} x \quad T_2 x = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} x + \begin{pmatrix} 1/2 \\ 0 \end{pmatrix} \quad T_3 x = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} x + \begin{pmatrix} 1/4 \\ \sqrt{3}/4 \end{pmatrix}$$

than the attractor for these maps is the Sierpinski triangle.[11]

[11] Named afte Polish mathematician Waclaw Franciszek Sierpin'ski (1882-1969).

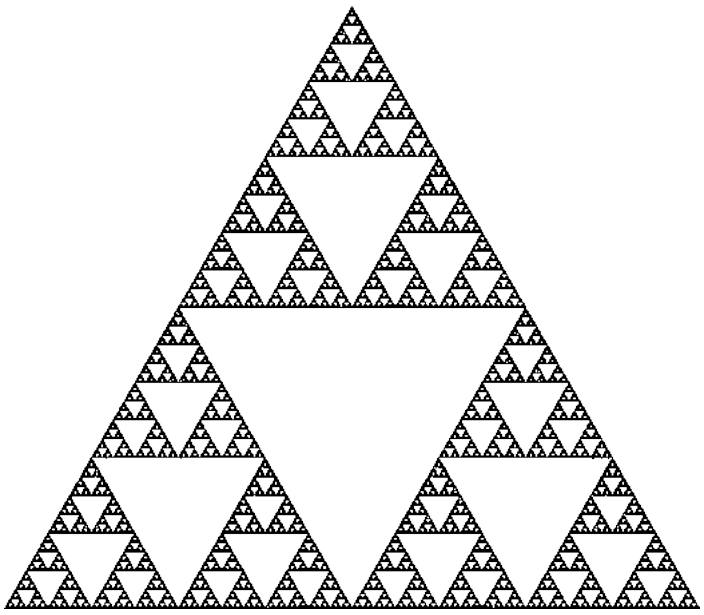

Fig. 6.4. The Sierpin'ski triangle given by three contractions

At the end of this section we like to remark that the fix point theorem may also be used to prove the existence of solution of differential equations, see [2] for further reading.

7

Algebra

7.1 The Determinant and eigenvalues

Let us begin this chapter with a few beautiful results from linear algebra. We assume the reader to be familiar with basic concepts, like real vector spaces, in this field. We first introduce the determinant of a matrix.

Definition 7.1.1 *The determinant of a matrix $A \in \mathbb{R}^{(n,n)}$ is given by*

$$\det(A) := \sum_{\pi} sgn(\pi) \prod_{i=1}^{n} a_{i\pi(i)}$$

where the sum is taken over all permutations π of $\{1, \ldots, n\}$. $sgn(\pi)$ is the sign given by

$$sgn(\pi) = \prod_{1 \leq k,l \leq n} \frac{\pi(k) - \pi(l)}{k - l} \in \{1, -1\}.$$

The following theorem is rather technical to prove nevertheless it is simple and fundamental.

Theorem 7.1.

$$\det(AB) = \det(A) \det(B)$$

$\mathfrak{Proof}$. Let $A = (a_{ij}), B = (b_{ij})$ and $C = AB = (c_{ij})$. Then $c_{ij} = \sum_k a_{ik} b_{kj}$. Therefor

$$det(AB) = \sum_{\pi} \sum_{k_1, \ldots, k_n} \operatorname{sgn}(\pi) a_{1k_1} b_{k_1\pi(1)} \cdots a_{nk_n} b_{k_n\pi(n)}$$

If $k_i = k_j$ for $i \neq j$ corresponding terms in the sum cancel due to the sign. Hence the sum is taken over all permutations σ with $\sigma(i) = k_i$. Hence

$$\det(AB) = \sum_{\pi} \sum_{\sigma} \operatorname{sgn}(\pi) a_{1\sigma(1)} b_{\sigma(1)\pi(1)} \cdots a_{n\sigma n} b_{\sigma(n)\pi(n)}$$

$$=^{\pi=\tau\rho} \sum_{\tau, \rho \in S_n} \operatorname{sgn}(\tau\rho) a_{1\sigma(1)} b_{\sigma(1)\tau\rho(1)} \cdots a_{n\sigma n} b_{\sigma(n)\tau\rho(n)}$$

$$= \sum_{\tau,\rho\in S_n} \operatorname{sgn}(\rho) a_{1\rho(1)} \cdots a_{n\rho(n)} \operatorname{sgn}(\tau) b_{1\tau(1)} \cdots b_{n\tau(n)} = \det(A)\det(B)$$

$\mathfrak{Q.E.D.}$

Theorem 7.2. *The determinant is multilinear and alternating with $det(E_n) = 1$ for the unit matrix E_n. Moreover the determinant is unique with these properties.*

We have to show that $\det(A)$ is linear in each collum of the matrix A. The maps $(a_1, \ldots, a_n) \mapsto \prod_{i=1}^n a_{i\pi(i)}$ are obviously linear and the sum of linear maps is linear hence det is linear. Now let $a_i = a_j$ for $i \neq j$. Let τ_{ij} be the transposition of i and j. By the definition of the sign we have $\operatorname{sgn}(\tau\pi)) = -\operatorname{sgn}(\pi)$. Pasting this into the definition the determinant is alternating. $\det(E_n) = 1$ is obvious. Now let D be another multilinear, alternating form with $D(E_n) = 1$ and consider the alternating multlinaere form $D - \det$. By assumption $(D - \det)(E_n) = 0$, which implies $D - \det = 0$ using linearity. Hence $D = \det$. $\mathfrak{Q.E.D.}$

Corollary 7.3. *The (signed) volume of the parallelotop spanned by the vectors $a_1, \ldots, a_n \in \mathbb{R}^n$ is given by $\det(a_1 \ldots a_n)$.*

$\mathfrak{Proof}$. We naturally expect a volume to be multilinear, alternating with $det(E_n) = 1$ and the determinant is unique with these properties. $\mathfrak{Q.E.D.}$

Corollary 7.4. *A matrix $A \in \mathbb{R}^{n\times n}$ is invertible if and only if $\det(A) \neq 0$.*

$\mathfrak{Proof}$. If a matrix A is invertible we have $\det(A)\det(A^{-1}) = \det(E_n) = 1$ which implies $\det(A) \neq 0$. On the other hand if two columns in A are identical $\det(A) = 0$, since det the alternating and linear. Now, if A is not invertible one column is a linear combination of the others. Using linearity we obtain $\det(A) = 0$. $\mathfrak{Q.E.D.}$

Now we have a look at Eigenvalues

Definition 7.1.2 *$\lambda \in \mathbb{R}$ is an eigenvalue of a matrix $A \in \mathbb{R}^{(n,n)}$ if there is a eigenvector $v \neq 0$ such that*

$$Av = \lambda v.$$

We may find eigenvalues using the determinant.

Theorem 7.5. *λ is an eigenvalue of A if and only if $\det(A - \lambda E_n) = 0$.*

$\mathfrak{Proof}$. $\det(A - \lambda E_n) = 0$ if an only if the matrix $A - \lambda E_n$ is not invertible. This is the case if and only if there is a vector $v \neq 0$ such that $(A - \lambda E_n)v = 0$ which means $Av = \lambda v$.

Corollary 7.6. *If $\Xi(x) = \det(A - \lambda E_n x)$ has n different real roots the matrix A is diagonalliseable.*

$\mathfrak{Proof}$. Under the assumption we have n different eigenvalues with n linear independent eigenvectors. With respect to this basis A obviously has diagonal form. $\mathfrak{Q.E.D.}$

7.2 Group theory

On of the basic algebraic structures is given by a group:

Definition 7.2.1 *A group is a set with an inner operation that is associative and has an unique neutral element e and inverse elements. The order of a group is its cardinality. The symmetric group S_n contains all permutation of the set $\{1, \ldots, n\}$. The inner operation in this groups is the composition.*

The first beautiful theorem in group theory concerns the order of finite groups.

Theorem 7.7. *For a finite group G the order of each subgroup H divides the order the G.*[1]

$\mathfrak{Proof}$. If two coset sets $aH := \{ah|h \in H\}$ and $bH := \{bh|h \in H\}$ are not disjoint we have $ah_1 = bh_2$ for some $h_1, h_2 \in H$. This implies $b^{-1}a \in H$ and $b^{-1}aH = H$ hence $aH = bH$. We have thus shown that the set of coset sets $G/H = \{gH|g \in G\}$ forms a partition of G. Since $|gH| = |H|$ this implies $|G| = |G/H| \cdot |H|$. $\mathfrak{Q.E.D.}$

Now we have a look at maps which pressers the group structure.

Definition 7.2.2 *A group homomorphism $f : G_1 \longmapsto G_2$ fulfills $f(g\bar{g}) = f(g)f(\bar{g})$ for all $g, \bar{g} \in G$. $K(f) = \{g \in G_1|f(g) = e\}$ is the kernel and $I(f)$ the image of the homomorphism. A bijective homomorphic is called isomorphism. If two groups are isomorphic we write $G_1 \cong G_2$.*

The following nice result shows that the symmetric groups is universal.

Theorem 7.8. *A group G of order n is isomorphic to a subgroup of the symmetric group S_n.*[2]

$\mathfrak{Proof}$. For $g \in G$ consider the map $p_g : G \longmapsto G$ given by $p_g(x) = gx$. This is obvious a permutation of G. Now consider the set $H := \{p_g|g \in G\}$. Since $p_{g_1} \circ p_{g_2}^{-1} = p_{g_1 g_2^{-1}} \in H$ this is a subgroup of the group of all permutation $S(G)$ of G. On the other hand by $p_{g_1 g_2} = p_{g_1} p_{g_2}$ the group H is homomorphic to G.$\mathfrak{Q.E.D.}$

To state the main result about the structure of homomorphisms we need the following definition

Definition 7.2.3 *A subgroup N of a group G is normal if $aN = Na$ for all $a \in G$*

With this we have:

[1] Found by the Italian mathematician Joseph-Louis de Lagrange (1736-1813).
[2] This is a result of the English mathematician Arthur Cayley (1821-1894) .

Theorem 7.9. *1. Isomorphism theorem: If $f : G_1 \longmapsto G_2$ is a group homomorphism than*

$$G_1/K(f) \cong I(f).$$

2. Isomorphism theorem: Let S be a subgroup and N be a normal subgroup of a group G than

$$S/(S \cap N) \cong SN/N.$$

3. Isomorphism theorem: Let $H \subseteq K \subseteq G$ be normal subgroups of G than

$$(G/K)/(H/K) \cong G/H.$$

$\mathfrak{Proof}$. 1. The map $\bar{f}$ given by $\bar{f}(g\mathrm{K}(f)) = f(g)$ for all $g \in G_1$ is well defined cause $\mathrm{K}(f)$ is a normal subgroup. It is obvious that $\bar{f}$ is an homomorphism and surjective. Moreover $\mathrm{K}(\bar{f}) = \{gK | \bar{f}(gK) = 1\} = K$ which is the neutral element in $G_1/\mathrm{K}(f)$, implying that $\bar{f}$ is surjective.
2. Define a homomorphism ϕ from S to SN/N by $\phi(s) = sN$. We have $\phi(S) = SN/N$ and Kernal$(\phi) = S \cap N$. Thus the result follows form the first isomorphism theorem.
3. Let p and q be the natural homomorphism form G to G/H and G to G/K. Let ϕ be the homomorphic from G/K to G/H with $\phi \circ q = p$. Since p is surjective ϕ is surjective. The kernel of ϕ is $\mathrm{K}(p)/K = H/K$. Thus the result follows form the first isomorphism theorem. $\mathfrak{Q.E.D.}$

The last theorem may be generalized to other algebraic structures like vector spaces or rings, which we introduce in the next section.

7.3 Rings

Definition 7.3.1 *A commutative group $(G, +)$ with an associative and distributive multiplication is called a ring.*

The integers $\mathbb{Z}$ with addition and multiplication is a well known ring. With the Euclidean algorithm described now the integers form a so called Euclidian ring.

Definition 7.3.2 *Given two integers $a, b \in \mathbb{Z}$ with $|a| > |b|$ the Euclidean algorithm computes recursively quotients $q_k \in \mathbb{Z}$ and reminders $r_k \in \mathbb{Z}$ with $|r_{k+1}| < |r_k|$ such that*

$$r_{k-2} = q_k r_{k-1} + r_k$$

where $r_0 = a$ and $r_k = b$.

Theorem 7.10. *The Euclidian algorithm terminates with $r_N = 0$ for some N and r_{N-1} is the greatest common divisor of a and b.*[3]

[3] Due to Eulcid.

𝔓𝔯𝔬𝔬𝔣. The algorithm terminates, since $|r_{k+1}| < |r_k|$. Obviously r_{N-1} divides r_{N-2}. Since it divides both it also divides r_{N-3}. By backward recursion r_{N-1} divides a and b and is hence a common divisor. Now let c be an arbitrary common divisor of a and b; $a = mc$ and $b = nc$. Now $r_0 = a - q_0 b = (m - q_0 n)c$ which means c divides r_0 and by recursion c divides r_{N-1} hence r_{N-1} is the greatest common divisor. 𝔔.𝔈.𝔇.

Now we introduces the ring of concurrence classes.

Definition 7.3.3 *Two integers a, b are called congruent modulo m (or a=b mod m for short) if there is an integer r such that $a - b = rm$. A concurrence class a modulo m (or a mod m for short) is the set $a + m\mathbb{Z}$ containig all integeres congruent to a. The set of all concurrence classes $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z} = \{0, 1, \ldots, m-1\}$ forms a ring with respect to addition and multiplication of integers.*

We prove here the following beautiful reminder theorem:

Theorem 7.11. *Let m_i $i = 1 \ldots n$ be coprime integers. Than for integers a_i for $i = 1 \ldots n$ there is a unique solution $x \in \mathbb{Z}_m$ with $m = m_1 m_2 \ldots m_n$ of the equations*

$$x = a_i \ mod \ m_i$$

for $i = 1 \ldots n$.[4]

𝔓𝔯𝔬𝔬𝔣. Note that m_i and m/m_i are coprime. Using the euclidian algorithm we find numbers t_i and s_i such that

$$t_i m_i + s_i(m/m_i) = 1$$

for $i = 1, \ldots n$. Define

$$x = \sum_{i=1}^{n} a_i s_i (m/m_i)$$

Since $s_i(m/m_i) = 1$ modulo m_i and $s_i(m/m_i) = 0$ modulo m_j for $j \neq i$, the number x fulfils the congruences. 𝔔.𝔈.𝔇.

The prove we provided her shows that the remainder theorem in fact holds in all Euclidian rings.

7.4 Units, fields and the Euler function

Definition 7.4.1 *In a ring R we multiplicative identity element 1 an element a is invertible (or a unit) if there exits $b \in R$ such that $ba = ab = 1$. If all $a \in R$ with $a \neq 0$ are a units R is a field.*

Of course there are no units beside 1 in $\mathbb{Z}$. $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are well known to be number fields. In the ring of concurrences classes $\mathbb{Z}_m$, introduced in the last section, the question for units is more exciting. In fact we have:

[4] Attributed to the Chinese mathematician Sun Tzu ($\sim$ 300).

Theorem 7.12. *$a \in \mathbb{Z}_m$ is unit if and only if a and m are coprime. $\mathbb{Z}_m$ is a field if an only if p is a prime number.*

$\mathfrak{Proof}$. Assume $ab = 1$ modulo m where a and m are not coprime. We have

$$ab = 1 + mr$$

for some $r \in \mathbb{Z}$ and there is exists $t \in \mathbb{N}$ with $t \neq 1$ which divides m and a. But than t divedes 1. A contradiction.
On the other hand if a and m are coprime us the Euclidean algorithm (backwards) to obtain $b, r \in \mathbb{Z}$ such that

$$ab + tr = 1.$$

This implies $ab = 1$ modulo m. The fist statement implies the second since p is prime if and only if all $a \neq p$ are coprime to p. $\mathfrak{Q.E.D.}$

Now we introduce the Euler function.

Definition 7.4.2 *For $m \in \mathbb{N}$ the Euler function $\phi(m)$ counts the number of coprimes of m. By the last theorem this is the number of units in $\mathbb{Z}_m$.*

The Euler function is given by the following beautiful formula.

Theorem 7.13.

$$\phi(m) = m \prod_{p|m} (1 - \frac{1}{p}) \quad [5]$$

$\mathfrak{Proof}$. Obviously if p is prime $\phi(p) = p - 1$. Moreover there are p^{n-1} numbers between 1 and p^n dividable by p namely the numbers $1p, 2p, 3p, 4p, \ldots, p^{n-1}p$. Hence

$$\phi(p^n) = p^n - p^{n-1} = p^n(1 - \frac{1}{p}).$$

We now show that ϕ is multiplicative for coprime numbers p and q. Let c be a coprime to pq with reminders of division $\bar{p}$ and $\bar{q}$:

$$c = \bar{p} \bmod p \qquad c = \bar{q} \bmod q \quad (\star)$$

Obviously $\bar{p}$ is coprime to p and $\bar{q}$ is coprime q. Assume now that $\bar{p}$ and $\bar{q}$ are given. Than by the Chinese reminder theorem the there is a unique c such that the equations $(\star)$ holds. Hence $\phi(pq) = \phi(p)\phi(q)$. The theorem now follows from prime factorization of m. $\mathfrak{Q.E.D.}$

Another remarkable property of the Euler function we like to prove is:

[5] A formula know to Euler.

Theorem 7.14. *If a and m are coprime than*

$$a^{\phi(m)} = 1 \ mod \ m. \ ^{6}$$

Proof. Consider the group $\mathbb{Z}_m^\times$ of all multiplicative invertible elements in $\mathbb{Z}_m$. An element $b \in \mathbb{Z}_m$ is in $\mathbb{Z}_m^\times$ if b and m are coprime. Hence $|\mathbb{Z}_m^\times| = \phi(m)$. Now consider the subgroup $A = \{a, a^2, a^3, \ldots, a^{|A|}\}$ of $\mathbb{Z}_m^\times$ generated by a. Note that $a^{|A|} = 1$ in $\mathbb{Z}_m$. By the theorem of Lagrange $|A|$ divides $\phi(m)$. Hence we have in $\mathbb{Z}_m$

$$a^{\phi(m)} = a^{k|A|} = (a^{|A|})^k = 1.$$

Q.E.D.

Corollary 7.15. *If p is prime*

$$a^{p-1} = 1 \ mod \ p \ ^{7}$$

7.5 The fundamental theorem of algebra

Theorem 7.16. *Any non-constant polynomial p with complex coefficient has a complex root.*[8]

Proof. We show that there is a $c \in \mathbb{C}$ with $p(c) = 0$. The set $K = \{z||z| < R\}$ is compact hence $f(z) = |p(z)|$ has a minimum $c \in K$. Is R large enough we have $|p(z)| \geq |p(c)|$ for all $z \notin K$ hence $|p(c)| \leq |p(z)|$ forall $z \in \mathbb{Z}$. Assume $p(c) \neq 0$ Consider $h(z) = p(z+c)/p(c)$. We show that there exists an u with $|h(u)| < 1$. Than $|p(c+u)| < |p(u)|$ which is a contradiction. h is of the form

$$h(z) = 1 + bz^k + z^k g(z)$$

where g is a polynomial with $g(0) = 0$. Choose d with $d^k = -1/b$ than

$$|h(td)| \leq 1 - t^k + t^k |d^k g(dt)|.$$

Since g is continuous we have

$$h(td) \leq 1 - 1/2t^k$$

if t is small enough. This completes the proof. Q.E.D.

[6] This was proved by Euler
[7] Known to Pierre de Fermat (1608-1665).
[8] First proved by the German mathematician Johann Carl Friedrich Gauß(1646-1716)

Theorem 7.17. *Any complex polynomial p of degree n may by written in the form*

$$p(z) = C\prod_{i=1}^{n}(z - z_i)$$

with $C, z_i \in \mathbb{C}$.

$\mathfrak{Proof}$. If $p(z_0) = 0$ we have $\mathfrak{Q.E.D.}$

$$p(z) = p(z) - p(z_0) = \sum_{k=1}^{n} a_i(x^k - x_0^k) = (x - x_0)\sum_{k=1}^{n} a_i\frac{x^k - x_0^k}{x - x_0}$$

Furthermore

$$\frac{x^k - x_0^k}{x - x_0} = \sum_{i=0}^{k-1} x^i x_0^{k-1-i}$$

hence $f(x) = (x - x_0)g(x)$ where g has degree $n - 1$. Now the result follows by induction using the existence theorem. $\mathfrak{Q.E.D.}$

7.6 Roots of polynomial

Solving quadratic equation is taught in school. Nevertheless we think that the $p-q$-formula is a beautiful result with a simple prove:

Theorem 7.18.

$$z^2 + pz + q = 0$$

has the roots

$$z_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}^{9}$$

$\mathfrak{Proof}$.

$$(z + \frac{p}{2})^2 = z^2 + pz + \frac{p^2}{4} = \frac{p^2}{4} - q$$

Take the roots and subtract $p/2$. $\mathfrak{Q.E.D.}$

Any cubic equation $ax^3 + bx^2 + cx + d = 0$ may be reduced to an equation of the type $z^3 + px + q = 0$ diving by a and substituting $x = z - b/3a$. The solution of these equations are given by a formula which is as beautiful as the $p - q$-formula for quadratic equations.

[9] The formula was first intrdouced by the Indian mathematician Brahmagupta(598-668).

Theorem 7.19.

$$z^3 + px + q = 0$$

has the roots

$$z_{1,2,3} = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} - \frac{q^3}{27}}} + \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} - \frac{q^3}{27}}} \quad [10]$$

where the two square roots are chosen to be the same, and the two cube roots are chosen such that their product is $-p/3$.

$\mathfrak{Proof}$. Choose variables u, v such that $u + v = z$. Substituting gives

$$u^3 + v^3 + (3uv + p)(u + v) + q = 0$$

Choose u, v such that $3uv + p = 0$. Note that by this choice $uv = -p/3$. We obtain

$$u^6 + qu^3 - \frac{p^3}{27} = 0$$

$$v^6 + qv^3 - \frac{p^3}{27} = 0.$$

These are quadratic equation in u^3 and v^3 with solutions

$$u^3 = -\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{q^3}{27}}$$

$$v^3 = -\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{q^3}{27}},$$

given the formula stated in the theorem. $\mathfrak{Q.E.D.}$

We do not understand why the $p - q$-formula for the cubic equation is not taught is school. Now we have a look at the quartic equation $ax^3 + bx^2 + cx + dx + e = 0$. Diving the equation by a and substituting. $x = z - b/4a$ we obtain a equation of the form $z^4 + pz^2 + qz + r = 0$ which may by solved using the following theorem.

Theorem 7.20.

$$z^4 + pz^2 + q^z + r = 0$$

has the roots

$$z_1 = \frac{1}{2}(\sqrt{-\theta_1} + \sqrt{-\theta_2} + \sqrt{-\theta_3})$$

$$z_2 = \frac{1}{2}(\sqrt{-\theta_1} - \sqrt{-\theta_2} - \sqrt{-\theta_3})$$

[10] This is the formula of the Italian Gerolamo Cardano (1501-1576) .

$$z_3 = \frac{1}{2}(-\sqrt{-\theta_1} + \sqrt{-\theta_2} - \sqrt{-\theta_3})$$
$$z_4 = \frac{1}{2}(-\sqrt{-\theta_1} - \sqrt{-\theta_2} + \sqrt{-\theta_3})$$

where $\theta_1, \theta_2, \theta_3$ *are the roots of the resolvent*

$$z^3 - 2pz + (p^2 - 4r)z + q^2 = 0.$$[11]

Proof. First note that $z_1 + z_2 + z_3 + z_4 = 0$ since

$$\prod_{i=1}^{4}(z - z_i) = z^4 + pz^2 + qz + r. \quad (\star)$$

Let $\theta_1 = (z_1 + z_2)^2 = (z_3 + z_4)^2$, $\theta_2 = (z_1 + z_3)^2 = (z_2 + z_4)^2$ $\theta_3 = (z_1 + z_4)^2 = (z_2 + z_3)^2$. Solving this system for $z_{1,2,3,4}$ we get the formulas stated in the theorem. It remains to show that $\theta_1, \theta_2, \theta_3$ are the roots of the resolvent. Using $(\star)$ we get by some computation

$$-2p = \theta_1 + \theta_2 + \theta_3$$
$$p^2 - 4r = \theta_1\theta_2 + \theta_1\theta_3 + \theta_2\theta_3$$
$$q^2 = \theta_1\theta_2\theta_3.$$

Hence

$$\prod_{i=1}^{3}(z - \theta_i) = z^3 - 2pz + (p^2 - 4r)z + q^2.$$

Q.E.D.

It was proved by Able[12]that equations of degree fife or higher can in general not be solved by radicals. The prove of this result need to much perpetration or is to long (with some messy calculations) to present it here. Using the Galois theory[13] it is even possible to decide whether equations are solvable by radicals, see [14].

[11] The first solution of quartic equations is attributed the the Italian mathematician Lodovico Ferrari (1522-1565). The approach used here is due to the Italian mathematician Joseph-Louis de Lagrange (1736-1813)
[12] The Norwegian mathematician Niels Henrik Abel (1802-1829)
[13] Which goes back to the French mathematician Evariste Galois (1811-1832)

8

Number theory

8.1 Prime numbers

Theorem 8.1. *There are infinitely many prime numbers.*[1]

$\mathfrak{Proof}$. Assume that there are finitely many primes $P = \{p_1, \ldots, p_n\}$. The number

$$N = \prod_{i=1}^{N} p_i + 1 > \max\{p_i | i = 1, \ldots n\}$$

has a prime divisor p. $p \in P$ implies $p|1$. A contradiction. $\mathfrak{Q.E.D.}$

Theorem 8.2. *Every natural number $n \geq 2$ can be uniquely, up to ordering of factors, represented as the product of primes.*[2]

$\mathfrak{Proof}$. We first prove existences than uniqueness by contradiction. Assume that n is the smallest number that is not a product of primes. Than $n = ab$ with a and b smaller than n. a and b are products of primes so n is, a contradiction. Now let s be the smallest number that can be written in two ways $s = p_1 \cdot \ldots \cdot p_n = q_1 \cdot \ldots \cdot q_m$. Than p_1 divides $q_1 \cdot \ldots \cdot q_m$ hence $p_1 = q_j$ for some j by the property of primes. Now s/p_1 has two factorization and is smaller than s, a contradiction. $\mathfrak{Q.E.D.}$

Theorem 8.3.

$$\sum_{p \ prime} \frac{1}{p} = \infty \quad [3]$$

$\mathfrak{Proof}$. Assume the contrary. Then there is an k such that

[1] This was proved by ancient Greek mathematician Euclid (360-280 B.C.).
[2] The fundamental theorem of arithmetics is also due to Euclid .
[3] A result of Euler (1707-1883).

$$\sum_{i=k+1}^{\infty} \frac{1}{p_i} < 1/2$$

Given $N \in \mathbb{N}$ let N_s be the number of $n \leq N$ that only have prim factors in $\{p_1, \ldots, p_k\}$ and let $N_b = N - N_a$ the number of $n \leq N$ that have prime factors bigger that p_k. First we estimate

$$N_b \leq \sum_{i=k+1}^{\infty} \lfloor \frac{N}{p_i} \rfloor < N/2$$

Furthermore we write every number $n \leq N$ that only has small prime factors in the form $n = a_n b_n^n$ where a_n is square free. There are 2^k different possibilities for a_n and less then $\sqrt{N}$ possibilities for b_n. Hence and $N_s \leq 2^k\sqrt{N}$ and $N = N_b + N_a < 2^k\sqrt{N} + N/2$. This is a contradiction for N large enough, for instance $N = 2^{2k+2}$. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

Theorem 8.4. *If p is prime if and only if $(p-1)! + 1$ is a multiple of p.*[4]

$\mathfrak{Proof}$. If p is prime than $2, \ldots, p-2$ are relatively prime to p. All of these integers have unique distinct multiplicative inverse modulo p since $\mathbb{Z}_p$ is a field with respect to multiplication. If we group these numbers with their inverses into pairs we get

$$1 \cdot 2 \cdot 3 \cdot \ldots \cdot (p-2) = 1 \bmod p$$

. Multiplying by $(p-1)$ gives

$$1 \cdot 2 \cdot 3 \cdot \ldots \cdot (p-1) = -1 \bmod p,$$

which proves the if part of the result. On the other hand if p is not a prime the greatest common divisor of $(p-1)!$ and p is bigger than one. Hence $(p-1)! + 1$ can not be a multiple of p. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

Theorem 8.5. *Every prime number of the form $n = 4m + 1$ is the sum of two squares.*[5]

$\mathfrak{Proof}$. Consider

$$S := \{(x, y, z) \in \mathbb{Z}^3 | 4xy + z^2 = p \quad x > 0 \quad y > 0\}$$

and the map

$$f : S \longmapsto S \quad f(x, y, z) = (y, x, -z).$$

[4] This is a result of the Italian mathematician Joseph-Louis Lagrange (1736-1813).
[5] Another result of the of Lagrange (1736-1813)

First note that S is finite. f is an involution and it has no fix points since $4xy \neq p$. If T is the subset of S with $z > 0$ we have $f(T) = S\backslash T$ hence as well $f(U) = S\backslash U$ for

$$U := \{(x, y, z) \in S | (x - y) + z > 0\}.$$

This implies that U and T have the same cardinality. Now consider

$$g : U \longmapsto U \quad f(x, y, z) = (x - y + z, y, 2y - z).$$

A simple calculation shows that g is well defined and an involution with unique fixed point $(\frac{p-1}{2}, 1, 1)$. Hence the cardinality of U is odd. At the end consider

$$h : T \longmapsto T \quad f(x, y, z) = (y, x, z).$$

h is an involution of a set with odd cardinality at has thus a fixpoint (x, y, z) with

$$p = 4x^2 + z^2 = (2x)^2 + z^2.$$

$\mathfrak{Q.E.D.}$

Definition 8.1.1 *A prime number is a Mersenne prime if it is of the form $M(n) = 2^n - 1$. A natural number is perfect if it is the sum of its proper divisors.*

Theorem 8.6. *All even perfect numbers are given by $2^{n-1}M(n)$ where $M(n)$ is a Mersenne prime.*[6]

$\mathfrak{Proof}$. Let σ be the sum of all divisors of a number. Assume $p = M(n) = 2^n - 1$ is prime and $N = 2^{n-1}(2^n - 1)$. We have to show that $\sigma(N) = 2N$. σ is obviously multiplicative and $\sigma(2^n - 1) = p + 1$ hence

$$\sigma(N) = \sigma(2^{n-1})\sigma(2^n - 1) = (2^n - 1)(p + 1) = (2^{n-1} - 1)2^n = 2N.$$

Now assume that N is an even perfect number with $N = 2^{n-1}m$ for m odd. We have $\sigma(N) = 2^{n-1}\sigma(m)$ and since n is perfect $\sigma(N) = 2^n m$ hence $2^n m = (2^n - 1)\sigma(m)$ and form this

$$\sigma(m) = m + \frac{m}{2^n - 1}.$$

Since both $\sigma(m)$ and m are integers, $d = m/(2n - 1)$ has to be an integer. Thus d itself divides m. But

$$\sigma(m) = m + \frac{m}{2n - 1} = m + d$$

is the sum of all of the divisors of m. Certainly 1 divides m, so we are forced to conclude that d = 1; if this were not the case, then we would have a contradiction. Therefore, $m = 2n - 1$, and in particular, m has no other positive divisors than one and itself, so $2n - 1$ is a prime. $\mathfrak{Q.E.D.}$

[6] Due to Euler .

Theorem 8.7. *For all $s \in \mathbb{C}$ with real part bigger than one we have:*

$$\zeta(s) := \sum_{n=1}^{\infty} n^{-s} = \prod_{p \ prime} \frac{1}{1-p^{-s}} \quad [7]$$

$\mathfrak{Proof}$. Every integer n be be uniquely written as

$$n = \prod_{\text{p prime}} p^{c_p(n)}.$$

Hence

$$\sum_{n=1}^{\infty} n^s = \sum_{n=1}^{\infty} \prod_{\text{p prime}} p^{-sc_p(n)} = \prod_{\text{p prime}} \sum_{c_p=0}^{\infty} p^{-sc_p} = \prod_{\text{p prime}} \frac{1}{1-p^{-s}}$$

$\mathfrak{Q.E.D.}$

The function $\zeta(s)$ for $s \in \mathbb{C}$ is called the Riemanian zeta function. It is related to the number $\pi(n)$ of primes smaller than n. In fact if the Riemanian hypothesis that all no trivial zeros of ζ have real part $1/2$ is true, we get

$$|\pi(x) - Li(x)| \leq c\sqrt{x}\log(x).$$

for some constant $c > 0$ where

$$Li(x) = \int_2^x \frac{dt}{\log t} \sim \frac{x}{\log(x)}.$$

Without a proof of this conjecture we the have following weaker estimate on the number of primes

$$|\pi(x) - Li(x)| \leq cxe^{-a\sqrt{\log(x)}},$$

for some constant $a, c > 0$, see for instance [16].

8.2 Diophantine equations

Theorem 8.8. *All Pythagorean triples $(a, b, c) \in \mathbb{N}^3$ fulfilling*

$$a^2 + b^2 = c^2$$

are (up to exchanging a and b) given by

$$(n(\frac{u^2-v^2}{2}), nuv, n(\frac{u^2+v^2}{2})$$

where $n \in \mathbb{N}$ and $u > v$ are odd numbers.[8]

[7] The function was introduced by the German mathematican Bernhard Riemann (1826-1866).
[8] A result by the ancient Greek mathematician Euclid (360-280 B.C.)

𝔓𝔯𝔬𝔬𝔣. A simple calculation shows that the triples given in our theorem are pythagorean. We have to show that all triples are given in this way. A pythagorean triple (a, b, c) is called primitive if a, b, c do not have a common factor and b is even. Obviously if we found all primitive triples than all triples are given by (na, nb, nc) or (nb, na, nc) with $n \in \mathbb{N}$. Let (a, b, c) be a primitive triple than $b^2 = (c+a)(c-a)$ and $(c+a)$ and $(c-a)$ do not have a common factor. Hence $c+a$ and $c-a$ are them self squares $c+a = u^2$ and $c-a = v^2$. Solving the system gives $c = \frac{u^2+v^2}{2}$, $a = \frac{u^2-v^2}{2}$ and $b = u + v$. 𝔔.𝔈.𝔇.

Theorem 8.9. *There are no triples $(a, b, c) \in \mathbb{N}^3$ such that*

$$a^4 + b^4 = c^4$$

𝔓𝔯𝔬𝔬𝔣. We show that if there is a triple with $x^4 + y^4 = z^2$ than there is a triple with $u^4 + v^4 = w^2$ with $w < z$. The result than follows from the contradiction coming from the infinite descendent. Without loss of generality we may assume that x, y, z are coprime. By the last theorem there are relative prime numbers m, n such that

$$\begin{aligned} x^2 &= 2mn \\ y^2 &= m^2 - n^2 \\ z &= m^2 + n^2. \end{aligned}$$

Since $n^2 + y^2 = m^2$ the triple (n, y, m) is coprime pythagorean. Again we have relative prime numbers s, t such that

$$\begin{aligned} n &= 2rs \\ y &= r^2 - s^2 \\ m &= r^2 + s^2. \end{aligned}$$

Now we have

$$m\frac{n}{2} = \frac{2nm}{4} = (\frac{y}{2})^2$$

Since m and $n/2$ are relatively prime they hence have to be both squares. Similarly,

$$rs = \frac{2rs}{2} = n/2$$

is a squares, so r and s are squares. So set $r = u^2$, $s = v^2$ and $m = w^2$. Substitution into $m = r^2 + s^2$ gives $u^4 + v^4 = w^2$. We have

$$w^4 < w^4 + n^2 = m^2 + n^2 = z$$

hence $w < z$ which completes the proof. 𝔔.𝔈.𝔇.

Today we now know that the conjecture of Fermat [9] is true, there are no $(a, b, c) \in \mathbb{N}^3$ such that

$$a^n + b^n = c^n$$

for $n > 2$, [26, 24].

8.3 Partition of numbers

Definition 8.3.1 *Let $p(n)$ be the number of ways expressing n as the sum of integers less or equal to n disregarding order.*

Theorem 8.10.

$$\sum_{n=0}^{\infty} p(n)x^n = \prod_{k=1}^{\infty} \frac{1}{1-x^k}$$

$\mathfrak{Proof}$.

$$\prod_{k=1}^{\infty} \frac{1}{1-x^k} = \prod_{k=1}^{\infty} \sum_{i=0}^{\infty} x^{ki}$$

$$= (1 + x^1 + x^{1+1} + x^{1+1+1} + x^{1+1+1+1} + \ldots)$$
$$(1 + x^2 + x^{2+2} + x^{2+2+2} + x^{2+2+2+2} + \ldots)$$
$$(1 + x^3 + x^{3+3} + x^{3+3+3} + x^{3+3+3+3} + \ldots)$$
$$\ldots$$
$$= \sum_{n=0}^{\infty} p(n)x^n$$

since each product leading to the monomial x^n corresponds to a possible sum expression of n. $\mathfrak{Q.E.D.}$

Theorem 8.11.

$$\prod_{n=1}^{\infty}(1-x^n) = \sum_{k=-\infty}^{\infty} (-1)^k x^{w(k)}$$

where $w(k) = k(3k-1)/2$ are the (generalized) pentagonal numbers.[10]

$\mathfrak{Proof}$. In the product we obtain the sum of terms $(-1)^s x^{n_1+\ldots n_s}$. The total coefficient $f(n)$ of x^n will be $f(n) = e(n) - d(n)$ where $e(n)$ is the number of partitions of n with even s and $d(n)$ is the number of partings with odd s. Let $P(n)$ be the set of pairs (s, g) where s is a natural number and g is a decreasing mapping on $\{1, \ldots s\}$ such that

[9] Conjectured by the French mathematician Pierre de Fermat (1608-1665)
[10] Another result by Euler.

$$\sum_{x\in\{1,\dots,s\}} g(x) = n.$$

We obviously we have $|P(n)| = f(n)$. Define $E(n)$ and $D(n)$ in the same way such that such that $|E(n)| = e(n)$ and $|D(n)| = d(n)$. Suppose $(s, g) \in P(n)$ where n is not a (generalized) Pentagonal number. Since g is decreasing, there is a unique integer $k \in \{1, \dots s\}$ such that $g(j) = g(1) - j + 1$ for $j \in \{1, \dots k\}$ and $g(j) < g(1) - j + 1$ for $j \in \{k + 1, \dots, s\}$. If $g(s) \le k$ define $\bar{g}$ on $\{1, \dots, s - 1\}$ by

$$\bar{g}(x) = \left\{ \begin{array}{c} g(x) + 1 \text{ if } x \in \{1, \dots, g(s)\} \\ g(x) \text{ if } x \in \{g(s) + 1, \dots, s - 1\} \end{array} \right\}$$

If $g(s) > k$ define $\bar{g}$ on $\{1, \dots, s + 1\}$ by

$$\bar{g}(x) = \left\{ \begin{array}{c} g(x) - 1 \text{ if } x \in \{1, \dots, k\} \\ g(x) \text{ if } x \in \{k + 1, \dots, s\} \\ k \text{ if } x = s + 1. \end{array} \right\}$$

In both cases $\bar{g}$ is decreasing with

$$\sum_{x\in\{1,\dots,s+1\}} \bar{g}(x) = n.$$

The mapping takes an element having odd s to an element having even s , and vice versa. We have thus constructed a bijection from $E(n)$ to $D(n)$ implying $f(n) = 0$. Now if $n = m(3m + 1)/2$ is a pentagonal number our mapping is a bijection excluding the single element (m, g) with $g(x) = 2m + 1 - x$ resp. $g(x) = 2m - x$ if $n = m(3m - 1)/2$ leading to $f(n) = 1$ if m is even and $f_n = -1$ if n is odd. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

Theorem 8.12. *$p(n)$ is given by the recursion*

$$p(n) = \sum_{k=1}^{\infty} (-1)^{k+1} p(n - w(k)) + p(n - w(-k))$$

with $p(0) = 1$ and $p(n) = 0$ for $n < 0$.[11]

We have

$$(\sum_{k=0}^{\infty} p(k)x^k)(1 + \sum_{n=1}^{\infty} (-1)^n (x^{w(n)} + x^{w(-n)}) = 1$$

Using the Cauchy product this gives

$$\sum_{k=1}^{\infty} p(k)x^k + \sum_{n=1}^{\infty} (-1)^n (x^{w(n)} + x^{w(-n)}) + \sum_{k=1}^{\infty} \sum_{j=1}^{k} (-1)^j p(k-j) x^{k-j} (x^{w(j)}) + x^{w(-j)}) = 0$$

Now collecting coefficients of x^n gives the result. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

[11] This was also know to Euler

8.4 Bernoulli numbers

Definition 8.4.1 *The Bernoulli numbers B_n are defined as the Taylor coefficient of*

$$f(z) = \frac{z}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} z^n \quad [12]$$

Theorem 8.13. *The Bernoulli numbers are given by the recursion*

$$\sum_{n=0}^{m} \binom{m+1}{n} B_n = 0$$

with $B_0 = 1$, especially

$$B_1 = -1/2 \quad B_2 = 1/6 \quad B_4 = -1/30 \text{ and } B_{2n+1} = 0$$

for all $n \geq 1$.

𝔓𝔯𝔬𝔬𝔣.

$$f(z) = (\sum_{k=0}^{\infty} \frac{z^k}{(k+1)!})^{-1}$$

hence

$$\begin{aligned} 1 &= (\sum_{k=0}^{\infty} \frac{z^k}{(k+1)!})(\sum_{n=0}^{\infty} \frac{B_n}{n!} z^n) \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{m} \frac{B_n}{n!(m-n+1)!} z^m \\ &= \sum_{m=0}^{\infty} [\sum_{n=0}^{m} \binom{m+1}{n} B_n)] \frac{z^m}{(m+1)!} \end{aligned}$$

given the recursion by comparing coefficients. A direct calculation gives B_1, B_2, B_4. To see that $B_{2n+1} = 0$ just note that the function

$$\frac{x \cosh(x/2)}{2 \sinh(x/2)} = f(x) + x/2 = 1 + \sum_{n=2}^{\infty} B_n \frac{x^n}{n!}$$

is odd. 𝔔.𝔈.𝔇.

Theorem 8.14. *For all $l, n \in \mathbb{N}$*

$$f_l(n) = \sum_{k=1}^{n} k^l = \frac{1}{l+1} \sum_{m=0}^{l} \binom{l+1}{m} B_m (n+1)^{l+1-m} \quad [13]$$

[12] Inroduced by Jakob Bernoulli (1654-1704)

$\mathfrak{Proof}$. Using the geometric series and the definition of e^x we get

$$\sum_{k=0}^{n} e^{kx} = \frac{e^{(n+1)x}-1}{x}\frac{x}{e^x-1}$$

$$= (\sum_{k=0}^{\infty} \frac{(n+1)^{k+1}}{(k+1)!}x^k)(\sum_{n=0}^{\infty} \frac{B_n}{n!}x^n)$$

$$= \sum_{l=0}^{\infty}(\frac{1}{l+1}\sum_{m=0}^{l}\binom{l+1}{m}B_m(n+1)^{l+1-m})\frac{x^l}{l!}.$$

On the other hand using the definition of e^x directly

$$\sum_{k=0}^{n} e^{kx} = \sum_{k=0}^{n}\sum_{l=0}^{\infty}\frac{(kx)^l}{l!} = \sum_{l=0}^{\infty}\frac{f_l(n)}{l!}x^l$$

giving the result by comparing coefficients. $\mathfrak{Q.E.D.}$

8.5 Irrational numbers

Theorem 8.15. *The roots of unitary integer polynomials*

$$P(x) = \sum_{k=0}^{n-1} a_k x^k + x^n \qquad \text{with } a_i \in \mathbb{Z}$$

are either integers or irrational numbers.

$\mathfrak{Proof}$. If $x = p/q$, with p and q relative prime, is a root of $P(x)$ we have

$$0 = q^n P(p/q) = \sum_{k=0}^{n-1} a_k q^{n-k} p^k + p^n$$

Hence q divides p^n which implies $q = \pm 1$. $\mathfrak{Q.E.D.}$

Corollary 8.16. *If p is a prime $\sqrt[n]{p}$ is irrational.*[14]

$\mathfrak{Proof}$. If $\sqrt[n]{p}$ is an integer p is not prime. $\mathfrak{Q.E.D.}$

[14] This was know to Pythagoras (570-495 BC)

Theorem 8.17. *The Euler constant e is irrational.*

$\mathfrak{Proof}$. If we had $e = a/b$ with $a, b \in \mathbb{N}$ than

$$N = n!(e - \sum_{k=0}^{n} \frac{1}{k!})$$

would we a natural number for all $n \geq b$. On the other hand we have

$$N = n!(\sum_{k=0}^{\infty} \frac{1}{k!} - \sum_{k=0}^{n} \frac{1}{k!}) = \sum_{k=n+1}^{\infty} \frac{n!}{k!} < \sum_{k=0}^{\infty} (\frac{1}{n+1})^k = 1/n$$

A contradiction. $\mathfrak{Q.E.D.}$

Theorem 8.18. *π is irrational.*[15]

$\mathfrak{Proof}$. We show a stronger result: π^2 is irrational. Assume that $\pi^2 = a/b$ and let

$$F(x) = b^n(\sum_{k=0}^{n} (-1)^k \pi^{2n-2k} f^{(2k)}(x))$$

with $f(x) = x^2(1-x)^2/2$. Since $f^{(k)}(0) = (-1)^{(k)} f^k(0)$ are internes by our assumption $F(0)$ and $F(1)$ are integers. Differentiation gives

$$F^{'}(x)\sin(\pi x) - \pi F(x)\cos(\pi x)]^{'} = (F^{''}(x) + \pi^2 F(x))\sin(\pi x)$$

$$= b^n \pi^{2n+2} f(x)\sin(\pi x) = \pi^2 a^n f(x)\sin(\pi x)$$

Therefore

$$N := \pi \int_0^1 a^n f(x)\sin(\pi x) = F(0) + F(1)$$

is an positive integer. On the other hand we have $f(x) < 1/n!$ for $x \in (0, 1)$ hence $0 < N < \pi a^n/n! < 1$ for n big enough. This is a contradiction. $\mathfrak{Q.E.D.}$

It was proved by Hermite[16] that e and by Lindemann[17] that π are transcendental, see [8] and [15]. We can not include these results here, because we do not have a proof which is short enough.

[15] Proved by Swiss mathematian Johann Heinrich Lambert (1728-1777).
[16] The French mathematician Charles Hermite (1822-1901).
[17] The German mathematician Carl Louis Ferdinand von Lindemann (1852-1939).

8.6 Pisot numbers

Definition 8.6.1 *A Pisot number $\alpha = \alpha_1 > 1$ is the root of an unitary integer minimal polynomial P of degree $d \geq 2$ such that all other roots $\alpha_2, \alpha_3, \ldots, \alpha_d$ of P have modulus strictly less than* 1.[18]

Theorem 8.19.

$$||\alpha^n||_{\mathbb{N}} = \min_{k \in \mathbb{N}} |\alpha^n - k| \leq (d-1)\Theta^n$$

with $\Theta = \max\{|\alpha_i||i = 2, \ldots n\}$.[19]

$\mathfrak{Proof}$. We prove that for all $n \geq 1$

$$s_n = \sum_{i=1}^{d} \alpha_i^n \in \mathbb{Z}$$

holds, which directly implies the result. We have

$$P(x) = \prod_{i=1}^{d}(x - \alpha_i) = \sum_{k=0}^{d-1} a_k x^k + x^d \qquad a_k \in \mathbb{Z}.$$

Taking the logarithmic derivative gives

$$\prod_{i=1}^{d} \frac{1}{(x-\alpha_i)} = \frac{\sum_{k=1}^{d-1} a_k k x^{k-1} + d x^{d-1}}{\sum_{k=0}^{d-1} a_k x^k + x^d}.$$

Substituting $1/x$ for x and multiplying by $1/x$ and using geometric series for the first therm gives

$$\sum_{n=0}^{\infty} s_n x^n = \frac{\sum_{k=1}^{d-1} a_{d-k+1}(d-k+1)x^{k-1} + a_0 x^{d-1}}{\sum_{k=0}^{d-1} a_{d-k} x^k + a_0 x^d}$$

Comparing coefficients implies $s_n \in \mathbb{Z}$. $\mathfrak{Q.E.D.}$

Example 8.20. Examples of Pisot numbers are given by the real solutions of

$$x^n - x^{n-1} - \ldots - x - 1 = 0$$

for $n \geq 2$. Especially the golden means ϕ is a Pisot numbers and we have

$$\phi^{20} = 15126.99993.$$

[18] Introduced by the French mathematician Charles Pisot (1910-1984).
[19] Proved by Pisot.

8.7 Dirichlets' theorem

Theorem 8.21. *For any irrational number α there are infinitely many $q, p \in \mathbb{Z}$ such that*

$$|a - \frac{p}{q}| < \frac{1}{q^2}. \quad [20]$$

$\mathfrak{Proof}$. Let {x} denote the fractional part and $[x]$ the integer part of a number x. For $N \geq 1$ consider the numbers $\{i\alpha\} \in [0,1]$ for $i = 1 \ldots N+1$. If we divide $[0,1]$ into N disjoint intervals of length $1/N$ two of these numbers $\{i\alpha\}$ and $\{j\alpha\}$, with $0 \leq i < j \leq N+1$, must lie in the same interval. Hence

$$0 \leq \alpha(i-j) - ([\alpha i] - [\alpha j]) = \{i\alpha\} - \{j\alpha\} \leq \frac{1}{N}.$$

With $p = i - j$ and $q = [\alpha i] - [\alpha j]$ we have

$$|a - \frac{p}{q}| \leq \frac{1}{Nq} \leq \frac{1}{q^2}.$$

By choosing N successively larger we find this way infinity many approximates. $\mathfrak{Q.E.D.}$

8.8 Lioville numbers

Theorem 8.22. *For an irrational algebraic number $\alpha \in \mathbb{R}$ with minimal polynomial of degree n there is a constant $c > 0$ such that*

$$|\alpha - \frac{p}{q}| \geq \frac{c}{q^n}$$

for all $p, q \in \mathbb{Z}$.[22]

$\mathfrak{Proof}$. Let

$$f(x) = \sum_{i=0}^{n} a_i x^i$$

be the minimal polynomial of α. Since f is minimal we have

$$f(x) = g(x)(x - \alpha) \text{ with } g(\alpha) \neq 0$$

Hence by continuity there exits $\rho > 0$ such that $g(x) \neq 0$ for all $x \in [\alpha - \rho, \alpha + \rho]$. Let $c = \min\{\rho, M^{-1}\}$ where $M = \max\{|g(x)||x \in [\alpha - \rho, \alpha + \rho]\}$. Given $p, q \in \mathbb{Z}$ there are two different cases. Either we directly have

[20] This is the Theorem of the German mathematian Johann Peter Gustav Lejeune Dirichlet (1805-1859).[21]
[22] This was proved by the French mathematician Joseph Liouville (1809-1882)

$$|\alpha - \frac{p}{q}| \geq \rho \geq \frac{c}{q^n} \text{ or } |\alpha - \frac{p}{q}| < \rho$$

implying $g(p/q) \neq 0$, but this gives

$$|\alpha - \frac{p}{q}| = |\frac{f(p/q)}{g(p/q)}| = \frac{|\sum_{i=0}^{n} a_i p^i q^{n-i}|}{|q^n g(p/q)|} \geq \frac{1}{q^n M} \geq \frac{c}{q^n}.$$

$\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

Definition 8.8.1 *A number $\alpha \in \mathbb{R}$ is called a Lioville number if for $n > 0$ there are $p, q \in \mathbb{Z}$ with*

$$|\alpha - \frac{p}{q}| < \frac{1}{q^n}.$$

For instance

$$\lambda = \sum_{k=0}^{\infty} \frac{1}{2^{j!}}$$

is a Lioville number.

Theorem 8.23. *All Lioville numbers are transcendental.*

$\mathfrak{Proof}$. First we show that Lioville numbers are irrational. Assume $\alpha = c/d$ is rational and $2^{n-1} > d$ then for all $p, q \in \mathbb{Z}$

$$|\alpha - \frac{p}{q}| \geq |\frac{c}{d} - \frac{p}{q}| \geq \frac{1}{qd} > \frac{1}{2^{n-1}q} \geq \frac{1}{q^n}$$

hence α is not Lioville. Now assume α is an irrational algebraic Lioville number. Choose the constant $c > 0$ from Theorem 5.1 and r with $2^r > 1/c$. Since α is Lioville there are p, q

$$|\alpha - \frac{p}{q}| < \frac{1}{q^{n+r}} \leq \frac{1}{2^r q^n} \leq \frac{c}{q^n}$$

a contradiction to the theorem of Lioville. $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

9

Probability theory

9.1 The birthday "paradox"

Definition 9.1.1 *A finite probability space X with probability P is Laplace if for all events $A \subseteq X$*

$$P(A) = \frac{|A|}{|B|}.$$[1]

Theorem 9.1. *The probability that from n people at least two have the same birthday is*

$$1 - \frac{365!}{(365 - n!)365^n}$$

assuming that the birthdays are equidistributed on 365 days. Especially $p(23) \approx 0.57$ $p(50) \approx 0.97$ and $p(100) \approx 0.99999$.

$\mathfrak{Proof}$. There are 365^n possibilities for n people to have their birthday. There are $365 \cdot 364 \cdot \ldots \cdot (365 - n + 1)$ possibilities for n people to have not the same birthday. Since we assume that the birthdays are equidistributed the experiment is Laplace and the probability for n people not to have the same birthday is $(365!/((365 - n!)365^n)$. Considering the complement gives the result. $\mathfrak{Q.E.D.}$

9.2 Bayes' theorem

Definition 9.2.1 *Let P be a probability and A, B be two events than the condition probability A under the condition B is defined as,*

$$P(A|B) := \frac{P(A \cap B)}{P(B)}.$$

[1] Introduced by the French mathematician Pierre-Simon Laplace (1749-1827).

Theorem 9.2. *If $\{B_i|i = 1 \dots n\}$ is a partition of a probability space we have*

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum P(A|B_j)P(B_j)}.$$ [2]

$\mathfrak{Proof}$. First we have the result on total probability

$$P(A) = P(\bigcup(B_i \cap A)) = \sum P(B_i \cap A) = \sum P(A|B_i)P(B_i)$$

Furthermore

$$P(B_i|A) = \frac{P(A \cap B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{P(A)}$$

given the result pasting in the total probability $P(A)$. $\mathfrak{Q.E.D.}$

9.3 Buffons' needle

Theorem 9.3. *If a needle of length l the thrown on paper lined parallel, where the distance of the lines is $d > l$, the probability P that the needle intersects a line is*

$$P = \frac{2l}{\pi d}.$$ [3]

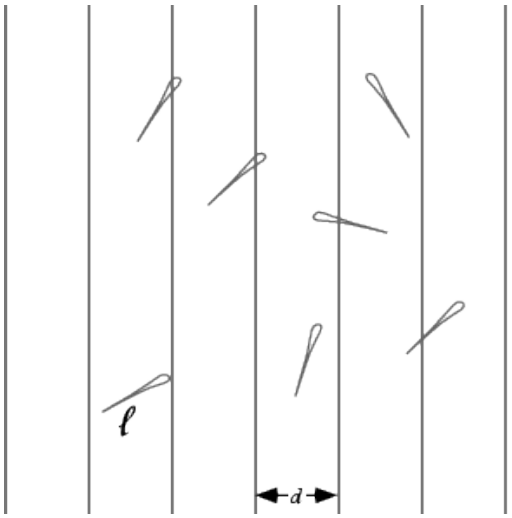

Fig. 9.1. Buffons' needle

$\mathfrak{Proof}$. We assume that the needle has an angle $\alpha \in [0, 2\pi]$ relative to the lines, the case of negative angles gives the same probability by symmetry. The height of this

[2] This is the result of the English mathematician Thomas Bayes (1702-1761).
[3] Prooved by the French mathematician Joseph Emile Barbier (1839-1889)

needle is $l\sin(\alpha)$ hence the probability that it intersects a line is $P(\alpha) = l\sin(\alpha)/d$. Since the angles are equidistributed by the construction of experiment we get p as the mean value over the angles

$$P = \frac{\int_0^{\pi/2} P(\alpha)d\alpha}{\pi/2} = \frac{2l}{\pi d}\int_0^{\pi/2} \sin(\alpha)d\alpha = \frac{2l}{\pi d}.$$

$\mathfrak{Q.E.D.}$

9.4 Expected value, variance and the law of large numbers

Definition 9.4.1 *A discrete random variable X is a map from a probability with a probability P to the natural numbers. The probability that the random variable attends the value $n \in \mathbb{N}$ is*

$$P(X = n) = P(\{w \in Omega|X(w) = n)$$

The expected value is given by

$$E(X) = \sum_{n\in\mathbb{N}} nP(X = n)$$

and the variance is given by

$$V(X) = \sum_{n\in\mathbb{N}} (n - E(X))^2 P(X = n)$$

For a continuous random variable taking values in $\mathbb{R}$ we define he expected value is given by

$$E(X) = \int P(X = x)xdx$$

and

$$V(X) = \int (x - E(X))^2 P(X = x)dx$$

Obviously is expectation value is linear and $V(aX) = a^2V(X)$. Furthermore for independent random variables X and Y the variance is additive $V(X+Y) = V(X)+V(Y)$. Now we are prepared to prove prove an important inequality in probability theory.

Theorem 9.4. *Let X be a random variable with finite expected value $E(X)$ and finite variance $V(X)$. For any $\epsilon > 0$ we have*

$$P(|X - E(X)| \geq \epsilon) \leq \frac{V(X)}{\epsilon^2}.$$[4]

[4] This inequality is due to the Russian mathematicain Lvovich Chebyshev (1821-1894).

𝔓𝔯𝔬𝔬𝔣. We prove the result for discrete random variables. The prove for continues random variables is along the same line using integration instead of summation. Let $m(x) = P(X = x)$ be the distribution function of X. We have

$$P(|X - E(X)| \geq \epsilon) = \sum_{|x-E(X)|\geq\epsilon} m(x).$$

Furthermore using the variance of X we obtain

$$V(X) = \sum_{x\in\mathbb{N}} (x - E(X))^2 m(x) \geq \sum_{|x-E(X)|\geq\epsilon} (x - E(X))^2 m(x)$$

$$\geq \sum_{|x-E(X)|\geq\epsilon} \epsilon^2 m(x) = \epsilon^2 P(|X - E(X)| \geq \epsilon).$$

given the result. 𝔔.𝔈.𝔇.

With this theorem we get a short prove for the law of large numbers.

Theorem 9.5. *Let (X_i) be independent and identically distributed random variables with expected values $E(X_i) = \mu < \infty$ and variance $V(X_i) = \sigma^2 < \infty$. For $S_n = X_1 + \ldots + X_n$ we have*

$$\lim_{n\longmapsto\infty} P(|\frac{S_n}{n} - \mu| > \epsilon) = 0^5$$

𝔓𝔯𝔬𝔬𝔣. By the the linearity of the expectation value we have $E(S_n/n) = E(S_n)/n = n\mu/n = \mu$ and by the properties of the variance we have $V(S_n/n) = V(S_n)/n^2 = n\sigma^2/n^2 = \sigma^2/n$. Hence the last theorem gives

$$P(|\frac{S_n}{n} - \mu| > \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}$$

for all $\epsilon > 0$. For fixed $\epsilon > 0$ we obtain the result. 𝔔.𝔈.𝔇.

9.5 Binomial and Poisson distribution

Definition 9.5.1 *For $n \in \mathbb{N}$ and $p \in (0, 1)$ is Binomial distribution $X_{n,p}$ on $\{0, \ldots n\}$ is given by*

$$P(X_{n,p} = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

For $\lambda > 0$ the Poisson distribution X_λ on $\mathbb{N}$ is given by

$$P(X_\lambda = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

[5] This was proved by the Swiss mathematician Jacob Bernoulli

It is no difficult to see that the expected value and the variance of these random variables is given by $E(X_{n,p}) = np$, $V(X_{n,p}) = np(1-p)$ and $E(X_\lambda) = V(X_\lambda) = \lambda$. We prove here that the Poisson distribution approximates the Binomial distribution in the following sense:

Theorem 9.6. *If* $\lim_{n\longmapsto\infty} np_n = \lambda$ *with* $p_n \in (0,1)$ *the Binomial distribution* X_{n,p_n} *approaches the Poisson distribution* X_λ*, in the following sense*

$$\lim_{n\longmapsto\infty} P(X_{n,p_n} = k) = P(X_\lambda = k)$$

for all $k > 0$.

$\mathfrak{Proof}$. Let $\lambda_n = np_n$. We have

$$P(X_{n,p_n} = k) = \binom{n}{k}(\frac{\lambda_n}{n})^k(1-\frac{\lambda_n}{n})^{n-k}$$

$$= \frac{1}{k!}\frac{n!}{(n-k)!n^k}\lambda_n^k(1+\frac{-\lambda_n}{n})^n)(1-\frac{\lambda_n}{n})^{-k}$$

Now consider the limit $n \longmapsto \infty$. The first factor is constant and the second factor goes to one. By theorem the third factor converges to $e^{-\lambda}$ since $\lim_{n\longmapsto\infty}\lambda_n = \lambda$ by assumption. The last factor goes to one since $\lim_{n\longmapsto\infty}\lambda_n/n = 0$. All in all the limit is given by $\frac{\lambda^k e^{-\lambda}}{k!}$ which is the distribution of X_λ. $\mathfrak{Q.E.D.}$

9.6 Normal distribution

Theorem 9.7.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x-\mu^2}{2\sigma^2}}$$

it the distribution of a random variable N_{μ,σ^2} *with expectation value* $E(X_{\mu,\rho}) = \mu$ *and Variance* $Var(X_{\mu,\rho}) = \sigma^2$.[6]

$\mathfrak{Proof}$.

$$(\int_{-\infty}^{\infty} f(x)dx)^2 = \frac{1}{2\pi\sigma^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2+(y-\mu)^2}{2\sigma^2}}dxdy$$

$$= \frac{1}{2\pi\sigma^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{\bar{x}^2+\bar{y}^2}{2\sigma^2}}d\bar{x}d\bar{y} = \frac{1}{2\pi\sigma^2}\int_0^{\infty}\int_0^{2\pi} re^{-\frac{r^2}{2\sigma^2}}d\theta dr$$

$$= \int_0^{2\pi}\frac{1}{2\pi}d\theta\int_0^{\infty}\frac{re^{-\frac{r^2}{2\sigma^2}}dr}{\sigma^2} = 1,$$

[6] The normal distribution was introduced by Carl Friedrich Gauß(1646-1716).

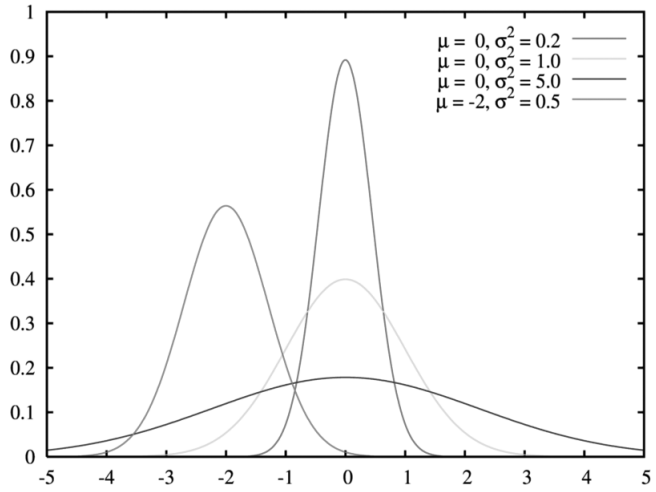

Fig. 9.2. The normal distribution

using the substitution $\bar{x} = x - \mu$ $\bar{y} = y - \mu$ and polar coordinates afterwards. This proves that f is the density of a probability distribution. Moreover we have:

$$E(N_{\mu,\rho}) = \int_{-\infty}^{\infty} xf(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} xe^{-\frac{x-\mu^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (\mu + \bar{x}) e^{-\frac{x^2}{2\sigma^2}} dx = \mu$$

$$Var(N_{\mu,\rho}) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x-\mu)^2 \left(e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) dx = \sigma^2.$$

$\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

Definition 9.6.1 *A sequence (X_i) of random variables converges in distribution to X if*

$$\lim_{i \longmapsto \infty} P(X_i \leq x) = P(X \leq x)$$

for all points of continuity of the commutative distribution function $F(X) = P(X \leq x)$.

Theorem 9.8. *Let (X_i) be a sequence of independent, identical distributed random variables with $E(X_i) = \mu < \infty$ and $0 < V(X_i) = \sigma^2 < \infty$. The mean random variable*

$$\frac{1}{n} \sum_{i=1}^{n} X_i$$

is asymptotically distributed as the normal distribution $N(\mu, \sigma/n)$. [7]

[7] First proved by the French mathematician Pierre-Simon de Laplace (1749-1827).

𝔓𝔯𝔬𝔬𝔣. We prove that the standardized random variable

$$Z_n = \frac{X_1 + \ldots X_n - n\mu}{\sigma\sqrt{n}} = \sum_{i=1}^{n} \frac{Y_i}{\sqrt{n}},$$

where $Y_i = (X_i - \mu)/\sigma$ with mean zero and variant one converges in distribution to the standard normal distribution $N(0,1)$. The result than follows by back transformation of the standardization.
The characteristic function of a standardized random variable Y with mean zero and variance 1 is given by

$$\phi_Y(t) = E(e^{itY}) = 1 - \frac{t^2}{n} + \text{therms of higher order}$$

using power series for the exponential. Hence the characteristic function of Z_n is

$$\phi_{Z_n}(t) = E(e^{itZ_n}) = \prod_{i=1}^{n} E(e^{it\frac{Y_n}{\sqrt{n}}}) = (1 - \frac{t^2}{2n} + \text{therms of higher order})^n$$

using multiplicativity of the expectation value. Taking the limit gives

$$\lim_{n \longmapsto \infty} \phi_{Z_n}(t) = e^{-t^2/2} = \phi_{N(1,0)}(t)$$

It remains to show that convergence of the characteristic function implies convergence in distribution. Suppose that a sequence of random variables X_n has characteristic functions $\pi_n(t) \mapsto \phi(t)$ for each t and ϕ is a continuous function at 0. Than for all $\epsilon > 0$ there exists a $c < \infty$ such that $P[|X_n| > c] \leq \epsilon$ for all n. Hence the sequence of random variables X_n is tight in the sense that any subsequence of it contains a further subsequence which converges in distribution to a proper cumulative distribution function. $\phi(t)$ is the characteristic function of this limit, since converges in distribution obviously implies converges of the characteristic function. Thus, since every subsequence has the same limit, X_n converges in distribution to X. 𝔔.𝔈.𝔇.

Corollary 9.9. *The Binomial distribution $X_{n,p}$ can be approximated by the normal distribution $N(np, np(1-p))$.*

Corollary 9.10. *The sum of n independent Poisson random variables $X_{\lambda/n}$ can be approximated by a normal distribution $N(\lambda, \lambda)$*

10

Dynamical Systems

10.1 Periodic orbits

Theorem 10.1. *If $f : [a, b] \longmapsto [a, b]$ is continues and has a periodic orbit of period three than its has periodic orbits all periods.*[1]

$\mathfrak{Proof}$. Without loss of generality assume $p = f^3(p) < f(p) < f^2(p)$. Let $k \geq 2$ be an integer. Define inductively $I_n = [f(a), f^2(a)]$ for $n = 0, \ldots k-2$ and $I_n = [p, f(p)]$ for $n = k-1$ and $I_{n+k} = I_n$ (if $k = 1$ chose $I_n = [p, f(p)]$ for all n). Now by induction there are compact intervals Q_n such that $f^n(Q_n) = I_n$ and $Q_{n+1} \subseteq Q_n$. Notice that $Q_k \subseteq Q_0 = I_0$ and $F_k(Q_k) = Q_0 = I_0$. Now by the intermediate value theorem the continues map $G = F^k$ has a fixed point $p_k \in Q_k$. p_k can not have period less than k. Otherwise $F^{k-1}(p_k) = b$ which contradicts $F^{k+1}(p_k) \in [f(a), f^2(a)]$ by construction. Hence p_k is a periodic point of period k for F. $\mathfrak{Q.E.D.}$

In fact Sharkosky [21] proved earlier much more than this: Consider the order

$$3 \preceq 5 \preceq 7 \preceq 9 \preceq 11 \preceq 13 \preceq 15 \preceq ... \preceq 2 \cdot 3 \preceq 2 \cdot 5 \preceq \cdot 7 \preceq 2 \cdot 9 \preceq ...$$

$$\preceq 2^2 \cdot 3 \preceq 2^2 \cdot 5 \preceq 2^2 \cdot 7 \preceq 2^2 \cdot 9 \preceq ... \preceq 2^3 \cdot 3 \preceq ... \preceq 2^5 \preceq 2^4 \preceq 2^3 \preceq 2^2 \preceq 2 \preceq 1$$

If f has a periodic orbit of period t it has periodic orbits of all periods $s \preceq t$. The proof is not short enough to include it.

10.2 Chaos and Shifts

In the following definition we mathematically formalize the meaning of the term "chaos".

Definition 10.2.1 *Let (X, d) be a metric space and let $T : X \longmapsto X$ be a continuous transformation. The system (X, T) is transitive if there is a dense orbit $(f^n(x))$. A system is chaotic if it is transitive and perotic orbits of f are dense in X. The system is sensitive if there is a constant C such that*

[1] This was proved in 1975 by Li and Yorke under the title "Period three implies chaos"

$$\limsup_{n\longmapsto\infty} d(f^n(x), f^n(y)) > C$$

for all $x \in X$ and a least one y in each neighborhood of x.

The reason why we do not need to suppose sensitivity in the definition of chaos is the following result:

Theorem 10.2. *A chaotic system (on an infinite space) is sensitive.*

$\mathfrak{Proof}$. Considers two periodic orbits which have a distance at least $D > 0$ from each other. Note that for for all $x \in X$ there hence is a periodic orbit which distance at least $D/2$ from x. We will prove sensitivity for $C = D/8$.
Fix an arbitrary point $x \in X$ and a neighborhood N of this point. Since periodic points are dense there is such a point in $N \cap B_C(x)$ with period lets say n and there is periodic point q with distance at least $4C$ from x. The set

$$V := \bigcap_{i=0}^{n} f^{-i}(B_C(f^i(q)))$$

is open and not empty since $q \in V$, Since f is transitive there exists a $y \in U$ such that $f^k(y) \in V$. Now choose j such that $1 \leq nj - k \leq n$. By construction we have

$$f^{nj}(y) = f^{nk-k}(f^k(y)) \in f^{nk-k}(V) \subseteq B_C(f^{nj-k}(q)))$$

Since p is periodic of period n we obtain

$$d(f^{nj}(p), f^{nj}(y)) = d(p, f^{nj}(y)) \geq d(x, f^{nj-k}(q)) - d(f^{nj-k}(q), f^{nj}(y)) - d(p, x)$$

$$\geq 4C - C - C = 2C$$

Thus by the triangle inequality either $d(f^{nj}(x), f^{nj}(p)) \geq C$ or $d(f^{nj}(x), f^{nj}(y)) > C$. In either case of found a point in N whose nj^{th} iterate is more than a distance $C > 0$ from this iterate of x. $\mathfrak{Q.E.D.}$

We found this beautiful theorem with a short prove in resent paper [4].

A model of a chaotic dynamical system is given by the shift map on a sequences space, which we now introduce.

Definition 10.2.2 *For $a > 1$ consider the sequence space $\Sigma = \{1, \ldots, a\}^{\mathbb{Z}}$ (or the one sided sequences space $\Sigma = \{1, \ldots, a\}^{\mathbb{N}}$) with the metric*

$$d((s_k), (t_k)) = \sum_{n=0}^{\infty} |s_k - t_k| 2^{-|k|}$$

The shift map σ on the sequence space is given by $\sigma((s_k) = (s_{k+1})$.

It is easy to see that the space Σ with the metric d is a Cantor set (compare with section 6.5) and that σ is a continuous function. The following beautiful theorem is an important result in chaotic dynamic:

Theorem 10.3. *The system (Σ, σ) is chaotic.*

Proof. Consider a sequence $(d_k) \in \Sigma$ which has every block with digits from $\{1, \ldots, a\}$ as its tail; first a block of length 1, than a^2 block of length 2 and so on. Now choose an arbitrary open set O in Σ be the definition of the metric there is a cylinder set $[s_1, \ldots, s_n] \subseteq O$. We will find the block $s_1, \ldots, s_n$ in (d_k). Hence $\sigma^m(d_k) \in O$ for some m. We thus see that $\{\sigma^m(d_k) | m \geq 0\}$ is dense in Σ and the system is transitive. Furthermore a sequence build by the repetition of a block $(s_1, \ldots, s_n)$ is periodic which respect to σ and is contained in the cylinder set $[s_1, \ldots, s_n]$. Hence periodic orbits are dense. □

10.3 Conjugated dynamical systems

Definition 10.3.1 *A dynamical systems (Y, g) is semi-conjugated to (X, f) if there is a continuous surjection $\pi : X \longmapsto Y$ such that $\pi \circ f = g \circ \pi$. The systems are conjugated if if π is an homömorphis.*

Theorem 10.4. *If (Y, g) is semi-conjugates to a chaotic system (X, f), it is chaotic.*

$\mathfrak{Proof}$. Let P be the set of periodic points which is dense in X and $D = (f^n(x))$ be a dense orbit in X. Be continuity of the surjection π the sets $\pi(P)$ and $\pi(D)$ are dense Y. IF $\pi(x) \in \pi(P)$ and x has period n we have $g^n(\pi(x)) = \pi(fn(x)) = \pi(x)$. Hence the points in $\pi(P)$ are periodic the respect to g and dense in Y. Moreover we have $\pi(D) = (\pi(f^n(x))) = g^n(\pi(x))$ which means that the orbit of $\pi(x)$ is dense in Y. □We present examples of systems in dimension one, two and three which are conjugated to a shift and hence chaotic.

Definition 10.3.2 *The continuous map $f : [0, 1] \longmapsto [0, 1]$ given by*

$$t(x) = 1 - 2|x - \frac{1}{2}|$$

is called the tent maps

Definition 10.3.3 *Let f be a continuous (or smooth if you like) map on $\mathbb{R}^2$ which acts on $[0, 1]^2$ as*

$$f(x, y) = \{ \begin{matrix} (1/3x, 3y) & \text{if } y \leq 1/3 \\ (-1/3x + 1, -3y + 3) & \text{if } y \geq 2/3 \end{matrix}$$

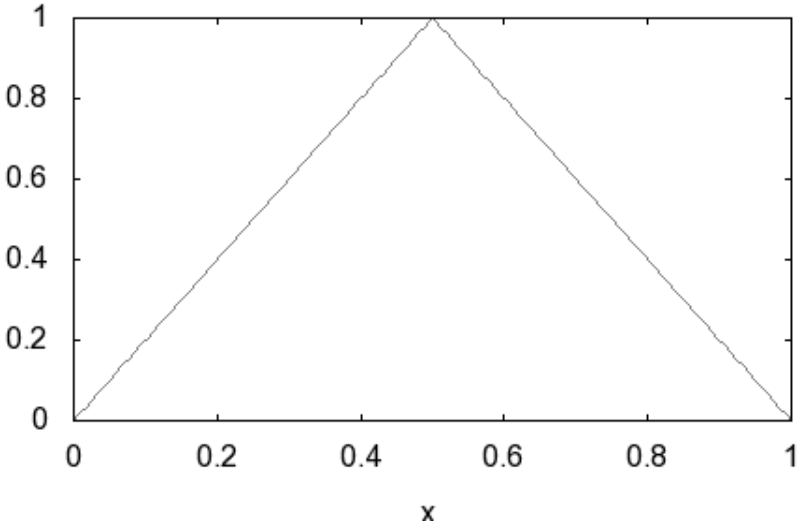

Fig. 10.1. The tent map

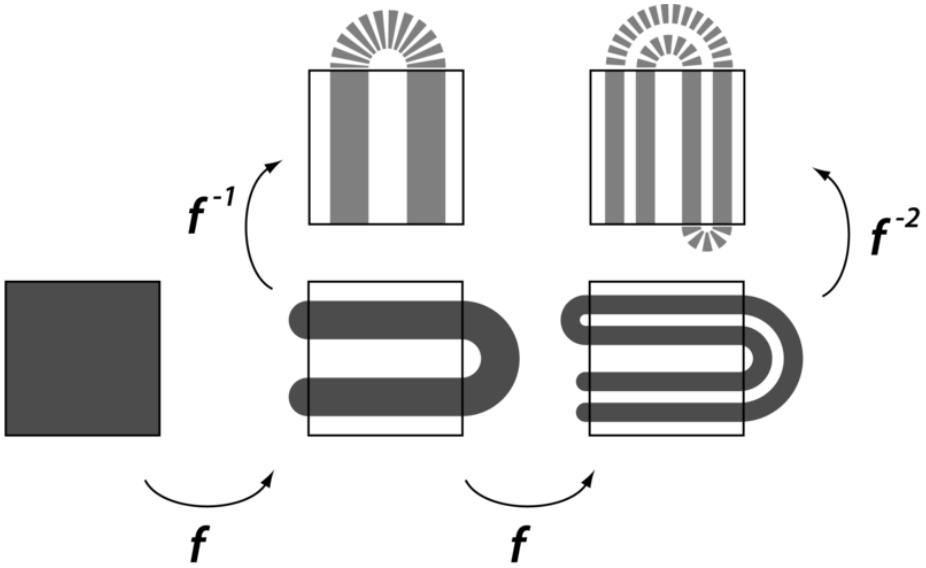

Fig. 10.2. Construction of the horseshoe

A map of this type is called a linear horseshoe with invariant set

$$\Lambda = \bigcap_{n=-\infty}^{\infty} f^n([0,1]^2)$$

Definition 10.3.4 *Let* $\mathbb{T}^2 = \mathbb{D} \times \mathbb{S} = \{(x, y, z) | x^2 + y^2 \leq 1, t \in [0, 1)\}$ *be the full torus and* h *be the map on* $\mathbb{T}^2$ *be given by*

$$h(x, y, z) = (\frac{1}{10}x + \frac{1}{2}\sin(2\pi z), \frac{1}{10}y + \frac{1}{2}\cos(2\pi z), 2z)$$

and consider the set

$$\Psi = \bigcap_{n=-0}^{\infty} h^n(\mathbb{T}^2)$$

The dynamical system (Ψ, h) is called a Solenoid.

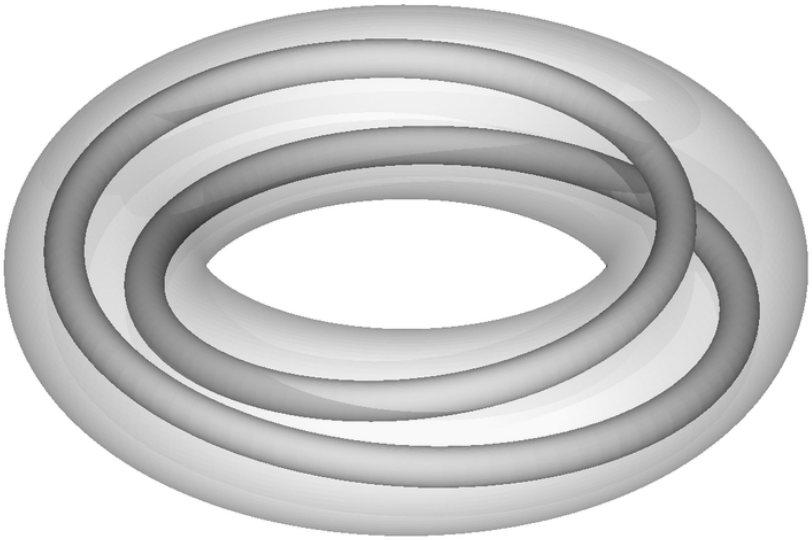

Fig. 10.3. Construction of the Soleonoid

Theorem 10.5. *The dynamical systems $([0,1],t)$, (Λ, f) and (Ψ, h) are (semi-) conjugated to a shift and hence chaotic.*

$\mathfrak{Proof}$. Let $I_0 = [0, 1/2]$ and $I_2 = [1/2, 1]$ and consider the map $\pi_t : \{1,2\}_0^{\mathbb{N}} \longmapsto [0,1]$ given by

$$\pi_{tf}((s_k)) = \bigcap_{k=0}^{\infty} f^{-k}(I_{s_k})$$

Note that $\bigcap_{k=0}^{m} f^{-k}(I_{s_k})$ is a nested sequence of intervals of length $1/2^{m+1}$, hence the map is well defined and continuous. It is onto since these intervals cover $[0,1]$ if we consider all in $(s_k) \in \{1,2\}^{m+1}$. Moreover we have $\pi_t((s_{k+1})) = f(\pi_f(s_k)))$ given the result.

Decompose $[0,1]^2 = Q_1 \cup Q_2$

Decompose $\mathbb{T}^2 = T_1 \cup T_2 = \mathbb{D}^2 \times [0, 1/2) \cup \mathbb{D}^2 \times [1/2, 1)$. Let $\pi_h : \{1, 2\}^{\mathbb{Z}} \longmapsto \mathbb{T}^2$ be given by

$$\pi_h((s_k)) = \bigcap_{k=-\infty}^{\infty} h^{-k}(T_{s_k})$$

Note that $\bigcap_{k=0}^{\infty} h^{-k}(T_{s_k})$ is a disc $\mathbb{D}^2 \times \{z((s_k))\}$ where z as a map of the sequence space is continuous and surjective. We have intersections of discs choosing one disc at each 'level' to correspond to a backward sequence for $\mathfrak{Q}.\mathfrak{E}.\mathfrak{D}.$

10.4 Julia sets and the Mandelbrot set

Definition 10.4.1 *For $c \in \mathbb{C}$ consider the map $f_c(z) = z^2 + c$ on the complex plan. The Julia set J_c is the boundary of the filled Julia set; that is, those points $z \in \mathbb{C}$ whose orbits under iterations of f_c remain bounded.*
The Madelbrot set is given by

$$\mathfrak{M} = \{c \in \mathbb{C} | J_c \text{ is connected}\}$$

It is possible to show that J_c is a Cantor set if it is not connected, compare with section. This the case for all $c \in \mathbb{C}$ that are not in the Madelbrot set. We here on a simple and beautiful cauterization of the Mandelbot set, which allows numerical appropriations of this set.

Theorem 10.6.

$$\mathfrak{M} = \{c \in \mathbb{C} | f_c^n(0) \text{ is bounded}\}$$

Here we need a lemma before we prove the result.

Lemma 10.7.

10.5 Recurrence

Definition 10.5.1 *In a metric space X the Borel σ algebra $\mathfrak{B}$ is the smallest sigma-Algebra which contains all open and closes sets. A set function $\mu : \mathfrak{B} \longmapsto [0, 1]$ which is σ-additive and fulfills $\mu(X) = 1$ and $\mu(X \backslash B) = 1 - \mu(B)$ is a Borel probability measure. A function is $T : X \longmapsto X$ is measurable if $T^{-1}(B) \in \mathfrak{B}$ for all $B \in \mathfrak{B}$ and a measure μ is invariant if $\mu \circ T^{-1} = \mu$.*

Theorem 10.8. *Let X be a metric space $T : X \longmapsto X$ be a measurable transformation and μ a invariant Borel probability measure measure. Let A be a Borel set with $\mu(A) > 0$, then for almost all $x \in A$ the orbit $\{T^n(x) | n \geq 0\}$ returns to A infinitely often.*[2]

[2] Proved by the French mathematician Jules Henri Poincare (1854-1912)

$\mathfrak{Proof}$. Let $F = \{x \in A | T^n(x) \notin A \quad \forall n \geq 1\}$. Obviously it is enough if we show that $\mu(F) = 0$. First we prove that $T^{-n}(F) \cap T^{-m}(F) = \emptyset$ for $n > m$. If there is an $x \in T^{-n}(F) \cap T^{-m}(F)$ then $x \in T^m(F) \subseteq A$ and $x \in T^n(F) = T^{n-m}(T^m(x)) \subseteq A$. This is a contradiction to the definition of A. Hence the sets $T^{-n}(F)$ are disjoint for $n \geq 1$ and by σ-additivity and invariance of μ we get

$$\sum_{n=1}^{\infty} \mu(F) = \sum_{n=1}^{\infty} \mu(T^{-n}(F)) = \mu(\bigcup_{n=1}^{\infty} T^{-n}(F)) \leq \mu(X)$$

Since μ is finite this implies $\mu(F) = 0$. $\mathfrak{Q.E.D.}$

10.6 Birkhoffs' ergodic theorem and normal numbers

The prove Birkhoffs' ergodic theorem we use a lemma which may be consider as an ergodic theorem itself.

Lemma 10.9. *Let X be a metric space $T : X \longmapsto X$ be a measurable transformation and μ be an invariant Borel probability measure measure and let f be integrable with respect to μ. Let*

$$s_n(x) = \sum_{i=0}^{n} f(T^i(x))$$

and $A = \{x | \sup s_n(x) > 0\}$. Than $\int_A f d\mu \geq 0$.

$\mathfrak{Proof}$. Let $S_n(x) = \max\{s_0(x), \ldots, s_n(x)\}$, $S_n^\star(x) = \max\{S_n(x), 0\}$ and $A_n = \{x | S_n(x) > 0\}$. Now we have

$$S_{n+1} = S_n^\star \circ T + f$$

hence

$$\begin{aligned} \int_{A_n} f &= \int_{A_n} S_{n+1} - S_n^\star \circ T \geq \int_{A_n} S_n - S_n^\star \circ T \\ &= \int_{A_n} S_n^\star - S_n^\star \circ T = \int_X S_n^\star - S_n^\star \circ T = 0 \end{aligned}$$

since we integrate with respect to measure T invariant measure. Since $A = \bigcup A_n$ the result follows. $\mathfrak{Q.E.D.}$

Theorem 10.10. *Let X be a metric space $T : X \longmapsto X$ be a measurable transformation and μ an ergodic Borel probability measure measure and let f be integrable with respect to μ than*

$$\lim_{n \longmapsto \infty} \frac{1}{n+1} \sum_{i=0}^{n} f(T^i(x)) = \int f \, d\mu$$

for μ-almost all x. [3]

[3] This is the ergodic theorem of the American mathematician George David Birkhoff (1884-1944)

$\mathfrak{Proof}$. Let $a_n(x) = s_n(x)/(n+1)$. If a_n does not converge almost surely we have

$$\liminf a_n(x) < b < a < \limsup a_n$$

on a set E of positive measure. Since E is invariant we can assume $E = X$ and by rescaling the function f we may assume $a = 1$ and $b = -1$. By the last theorem $\int f \geq 0$. Now replacing f with $-f$ we get $\inf f \leq 0$. The same argument works for all $f + c$ with $0 < c < 1$ implying $\int c = 0$. This is a contradiction. We conclude that $a_n(x)$ converges to a function $F(x)$ for almost all x. Note that F is T invariant and integrable since $||a_n||_1 \leq ||f||_1$. Moreover by ergodicity of μ we get $\int F = \int f$ given the result. $\mathfrak{Q.E.D.}$

Definition 10.6.1 *A real number $x \in (0,1]$ is normal in base $b \geq 2$ if it has a b-expansion*

$$x = \sum_{k=1}^{\infty} x_k b^k$$

with

$$\lim_{n \longmapsto \infty} \frac{|\{k | x_k = i \quad k = 1 \dots n\}|}{n} = \frac{1}{b}$$

for all $i = 0, \dots b-1$.

Corollary 10.11. *All most all numbers in $\mathbb{R}$ are normal to all bases.*[4]

$\mathfrak{Proof}$. Fix a base $b \geq 2$. Consider the transformation $T : [0,1) \longmapsto [0,1)$ given by $Tx = x$ modulo b. The Lebesgue measure ℓ is ergodic with respect to T. By the ergodic theorem we have

$$\lim_{n \longmapsto \infty} \frac{|\{k | T^k x \in [i/b, (i+1)/b) \quad k = 1 \dots n\}|}{n} = \ell([i/b, (i+1)/b)) = \frac{1}{b}$$

but $T^k x \in [i/b, (i+1)/b)$ if and only if $x_k = i$. $\mathfrak{Q.E.D.}$

[4] The result was first proved by Felix Edouard Justin Emile Borel (1871-1956)

11

Conjectures

At the end we state ten conjectures we would like the reader to prove.

Conjecture 1 *There are infinitely many twin primes.*

Conjecture 2 *Every even number is the sum of two primes.*

Conjecture 3 *Between two square numbers there always is a prime number.*

Conjecture 4 *The Riemanian hypothesis is true.*

Conjecture 5 *Natural numbers can not be factorized in polynomial time.*

Conjecture 6 *There are infinity many even and no odd perfect numbers.*

Conjecture 7 *The Euler-Mascheroni constant γ and the Catlan constant C are irrational (and transcendent).*

Conjecture 8 *All irrational algebraic numbers are normal*

Conjecture 9 *The numbers e and π are normal.*

Conjecture 10 *The map $f(n) = n/2$ for n even and $f(n) = 3n+1$ for n odd iterates all natural numbers to one.*

References

1. P. Abad and J. Abad, *The hundred greates theorem of mathematics*, http://personal.stevens.edu/ nkahl/Top100Theorems.html, 1999.
2. V.I. Arnold, *Ordinary differntial equations*, Springer, Berlin, 1980.
3. K. Appel, and W. Haken, Every Planar Map is Four-Colorable, Providence, RI: American Mathematical Society, 1989.
4. J. Banks, J. Brooks, G. Cairns, G. Davis, P. Stacey, The American Mathematical Monthly, Vol. 99, No. 4, 332-334, 1992.
5. L.E.J. Brouwer, "Über eineindeutige, stetige Transformationen von Flächen in sich" Math. Ann. , 69 176180, 1910.
6. P.J. Cohen, Set Theory and the Continuum Hypothesis. W. A. Benjamin, 1966.
7. K. Falconer, *Fractal Geometry - Mathematical Foundations and Applications*, Wiley, New York, 1990.
8. C. Hermite, Ouvres de Charles Hermite, Gauthier-Villars (reissued by Cambridge University Press, 2009).
9. H. Herrlich, Axiom of Choice, Springer Lecture Notes in Mathematics, Springer Verlag Berlin Heidelberg, 2006.
10. K. Gödel, The Consistency of the Continuum-Hypothesis, Princeton University Press, 1940.
11. R. Grimaldi, Discrete and Combinatorial Mathematics: An Applied Introduction, 4th edn., Reading, Mass. : Addison-Wesley Longman, 1998.
12. B. Kleiner and J. Lott, *Notes on Perelman's Papers*, arxiv:math.DG/0605667, 2006.
13. L. Kirby and J. Paris, J., *Accessible independence results for Peano arithmetic*, Bulletin London Mathematical Society 14, 285293, 1982.
14. S. Lang, Algebra, Springer verlag, 2005.
15. F. Lindemann, *Über die Zahl* π, Mathematische Annalen 20, 213-225, 1882.
16. W. Narkiewicz, *The development of prime number theory*, Springer-Verlag Berlin, 2000.
17. G. Perelman, *The entropy formula for the Ricci flow and its geometric applications*, arxiv:math.DG/0211159, 2002.
18. G. Perelman, *Ricci flow with surgery on three-manifolds*, arxiv:math.DG/0303109, 2003.
19. G. Perelman, *Finite extinction time for the solutions to the Ricci flow on certain three-manifolds*, arxiv:math.DG/0307245,2003.
20. W. Rudin, Real and Complex Analysis, McGraw-Hill Science/Engineering/Math; 3 edition, 1986.
21. O.M. Sarkovskii, *Co-Existence of Cycles of a Continuous Mapping of a Line onto Itself.*, Ukranian Math. Z. 16, 61-71, 1964.
22. D. R. Smart, Fixed Point Theorems, Cambridge Tracts in Mathematics, Cambridge University Press, 1980.
23. R. Stanley, Enumerative combinatorics. Vol. 2, Cambridge Studies in Advanced Mathematics, 62, Cambridge University Press, 1999.
24. R. Taylor and A. Wiles, *Ring-theoretic properties of Hecke Algebras*, Annals of Mathematics, 141, 553-572, 1995.
25. Karl Weierstrass, ber continuirliche Functionen eines reellen Arguments, die fr keinen Werth des letzeren einen bestimmten Differentialquotienten besitzen, Collected works; English translation: On continuous functions of a real argument that do not have a well-defined differential quotient, in: G.A. Edgar, Classics on Fractals, Addison-Wesley Publishing Company, 1993.
26. A. Wiles, *Modular Elliptic Curves and Fermat's last theorem*, Annals of Mathematics 141, 443-551, 1995.
27. R. Wilson, The finite simple groups, Berlin, New York: Springer-Verlag, 2009.

Index