Zbrani zapiski za 1. letnik magistrskega študija

Patrik Žnidaršič

11. januar 2025

Opomba k notaciji: Včasih je produkt vektorja s skalarjem napisan kot $\lambda.x$, s piko. Tudi oprator $\vec{\nabla}$ včasih obravnavamo kot vektor, in pišemo gradient s piko spodaj, $\vec{\nabla}.f$. Divergenco označimo z $\vec{\nabla} \cdot f$.

Kazalo

1 Moderna fizika

1.1 Special relativity

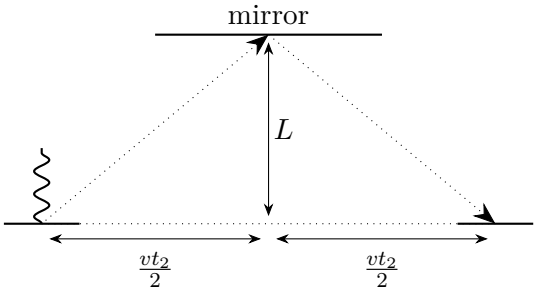
1.1.1 The Michelson-Morley Experiment

Suppose there is some fluid medium at rest in a frame S. Now, we create a plane wave in this medium, traveling in the $\mathbf{e}_x$ direction, with a velocity of c_1 as measured in S. If we switch to a different frame S' with a Galilean transformation $\mathbf{x}' = \mathbf{x} + vt\mathbf{e}_x$, then we expect the plane wave to have a different velocity $c_1 - v$ in S'. This is true for waves in water, or sound waves in air, but, as famously measured by Michelson and Morley, it is not true for light. Their experiment, described below, showed that the speed of light c is the same in any inertial reference frame — a result that contradicts entirely with classical mechanics.

The Michelson-Morley experiment splits a beam of light into two beams traveling orthogonally to each other, which then bounce off a pair of mirrors, before being recombined and the interference pattern of this recombination being measured. Assume that S is the rest frame of the ether, in which light waves are traveling at c. Suppose that we now switch to a frame S', with a relative velocity of v as measured in S. If one of the arms is parallel to the relative velocity vector, then the time of flight of a wave of light should be

$$t_1 = \frac{L}{c+v} + \frac{L}{c-v} = \frac{2cL}{c^2 - v^2},$$

where L is the distance between the beam splitter and the mirror. In S, the other beam of light travels a path that is not parallel to $\mathbf{e}_x$, as shown in figure 1.1.



Slika 1.1: Beam path in S

From Pythagoras' theorem, the length of the beam path is

$$ct_2 = 2\sqrt{L^2 + v^2 t_2^2/4}$$

so

$$t_2 = \frac{2L}{\sqrt{c^2 - v^2}}.$$

As mentioned, Michelson and Morley conducted this experiment, while using the Earth's rotation to perform many measurements in many directions. They found that their

experiment always gave the same result, no matter the relative velocity of the second frame and no matter the orientation of the testbed. This demonstrated beyond any doubt that Max Planck's advisor Philipp von Jolly was very wrong when he stated to Planck in 1874 that there was essentially nothing left to discover in theoretical physics,[1] since, as Feynman stated,

> „It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong."

Classical mechanics and electromagnetism, as complete as they might seem, are wrong.

1.1.2 Lorentz, the new Galileo

Enter Einstein, who formulated special relativity with the following postulates:

1. The principle of relativity: The laws of physics are identical in all inertial frames. This is already a postulate of classical mechanics, but we now have an additional one.

2. The speed of light in vacuum is a law of physics, so it is also the same in all inertial reference frames.

Using these notions, we can find the expressions for the Lorentz transformations, the replacement of Galilean transformations in classical mechanics. Suppose we have two reference frames, one stationary and one moving with velocity $v\mathbf{e}_x$ relative to the first frame. Let $x' = f(x,t)$ be the position of an event in frame S' and $t' = g(x,t)$ be the time of that event, as measured by a clock in S'. Without loss of generality, we can limit our consideration to the case where $x = x' = 0$ at times $t = t' = 0$.

With some inspiration from dimensional analysis, we set $x' = \gamma(v)(x - vt)$ for some function γ. By the principle of relativity, rotating the system should have no effect on the value of γ, but if we rotate frame S' around, the same expression gives $x' = \gamma(-v)(x+vt)$. So γ must be an even function of v.

Now consider a light wave in S, moving with speed c towards the right. Then $x = ct$, and by Einstein's second principle, $x' = ct'$. The above transformation gives us

$$ct' = \gamma(v)(ct - vt).$$

The same transformation formula should also work in reverse, $ct = \gamma(v)(ct'+vt')$. Solving these two equations gives us

$$\gamma(v) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

[1] Source: `https://hsm.stackexchange.com/questions/2129/`

We still need to find $t'(x,t)$. For this, again solve the equations $x' = \gamma(x - vt)$ and $x = \gamma(x' + vt')$, but not for light. We get

$$t' = \gamma(t - \frac{v}{c^2}x),$$

which is the second part of a Lorentz transformation.

Remark. Note that if v is small compared to c, then $\gamma \approx 1$, so we recover Galilean transformations.

Remark. The transformation of time can be tricky to understand. Think about it this way: Time is what is measured by a clock. While in classical mechanics, every pair of clocks which ever agreed will always agree, in special relativity, this only holds for pairs of clocks which are stationary relative to each other (i.e. there exists an inertial reference frame S in which both clocks are stationary). Since the position of a clock doesn't matter, we can think of every reference frame as having a clock at its center, which measures the frame's time.

Definition 1.1.1. An EVENT is a point $(t, \mathbf{x}) \in \mathbb{R}^{1+3}$.

Definition 1.1.2. A WORLDLINE is a function $(t, \mathbf{x}) : \mathbb{R} \to \mathbb{R}^{1+3}$, for which $(t(\lambda), \mathbf{x}(\lambda))$ is an event for any $\lambda \in \mathbb{R}$. We call λ the AFFINE PARAMETER.

Example (time dilation). Let S be a rest frame and S' another frame moving with velocity v relative to S. Consider the time measured by a clock at $x' = 0$. Since $x = vt$, the Lorentz transformation of time gives us

$$t' = \gamma(v)\left(1 - \frac{v^2}{c^2}\right)t = \frac{1}{\gamma}t.$$

Note that $\gamma > 1$, so in the moving system, the clock shows a lower time than in the non-moving system.

Example (length contraction). Consider a stick of length l' in S'. We consider the ends of the stick to be two worldlines with $x'_1 = 0$ and $x'_2 = l'$. We wish to measure the length of the stick in the non-moving frame S. For this, we need to find two simultaneous events in S, one belonging to the first worldline and one belonging to the second.

We'll take the measurement at $t = 0$. From the Lorentz transformation of space, we then get

$$l' = \gamma l,$$

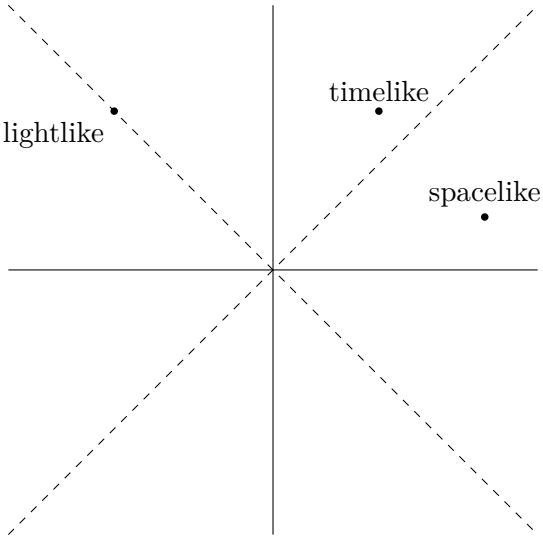
so the stick is shorter in S than in S'.

Note that the expression $s^2 = c^2t^2 - \mathbf{x}^2$ is invariant under Lorentz transformations, as is the difference $\Delta s^2 = c^2\Delta t^2 - \Delta\mathbf{x}^2$ between two events. This allows us to define Δs^2 as a sensible distance-like measure on $\mathbb{R}^{1+3}$, called the MINKOWSKI METRIC, though it is not a metric. Depending on the value of Δs^2, we call a pair of events

- TIMELIKE if $\Delta s^2 > 0$,

- NULL or LIGHTLIKE if $\Delta s^2 = 0$, and
- SPACELIKE if $\Delta s^2 < 0$.

Lightlike events are those which can be connected with a light-speed signal. Two timelike events can then be causally connected, as there was enough time between them to transfer and process information. Spacelike events, on the other hand, cannot be causally connected, since no signal can travel faster than light, so there is no way for the information of the first event to reach the second. We can draw these events on a spacetime diagram, as shown in figure 1.2.



Slika 1.2: Spacetime diagram

1.1.3 4-vector notation

Further on, we will use 4-vectors to denote events and quantities derived from them. We will use the label X^μ to refer to either the μ-th component of the 4-vector X, or to the vector itself; it is generally defined as

$$X^\mu = \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}$$

for an event at time t and coordinate (x, y, z). We also use the convention that indexing 4-vectors with Latin indices only refers to their space coordinates, ignoring the time coordinate. If we write X_μ with the index at the bottom, we mean

$$X_\mu = \eta_{\mu\nu} X^\nu$$

where the double index ν implies a summation, as per the Einstein convention. Above, $\eta_{\mu\nu}$ is the matrix

$$\eta_{\mu\nu} = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix}.$$

This allows us to write the Minkowski dot product simply as $X^\mu X_\mu = X^\mu \eta_{\mu\nu} X^\nu$, so

$$ds^2 = \eta_{\mu\nu} dX^\mu dX^\nu.$$

We can now recognize Lorentz transformations as matrices $\Lambda^\mu{}_\nu$ which allow for transformations

$$X'^\mu = \Lambda^\mu{}_\nu X^\nu.$$

Not all matrices are correct here, as among other things, they need to satisfy

$$X^\rho X_\rho = X'^\mu X'_\mu$$

or

$$X^\rho \eta_{\rho\sigma} X^\sigma = X'^\mu \eta_{\mu\nu} X'^\nu = \eta_{\mu\nu} \Lambda^\mu{}_\rho X^\rho \Lambda^\nu{}_\sigma X^\sigma.$$

This gives us the constraint

$$\eta_{\rho\sigma} = \eta_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma$$

which is what we will use for our definition of a Lorentz transformation, so a Lorentz transformation is any matrix $\Lambda^\mu{}_\nu$ which satisfies the property above.

2 Teorija grafov

2.1 Matchings

Definition 2.1.1. A vertex set $S \subseteq V$ is an INDEPENDENT SET of the graph $G = (V, E)$ if the induced subgraph $G[S]$ is empty. The maximum cardinality of an independent set is the INDEPENDENCE NUMBER $\alpha(G)$.

Definition 2.1.2. A vertex set $T \subseteq V$ is a VERTEX COVER if every edge has at least one of its endings in T. The maximum cardinality of a vertex cover is the VERTEX COVER NUMBER $\beta(G)$.

Definition 2.1.3. An edge set $M \subseteq E$ is a MATCHING if for every distinct $e_1, e_2 \in M$, edges e_1 and e_2 have no common ending. The maximum cardinality of a matching is the MATCHING NUMBER $\alpha'(G)$.

Definition 2.1.4. An edge set $C \subseteq E$ is an EDGE COVER if every vertex of G is covered by at least one edge from C. If the minimum degree $\delta(G)$ is at least 1, we can define the EDGE COVER NUMBER $\beta'(G)$ as the minimum cardinality of an edge cover.

Question 1. Define the independence number, the vertex and edge cover number, and the matching number.

Remark. The complement of an independent set is a vertex cover, so $\alpha(G)+\beta(G) = n(G)$ in every graph G. In a maximum matching, every edge must be covered by different vertices, so $\alpha'(G) \leq \beta(G)$. We can similarly argue that $\alpha'(G) \leq {}^{n(G)}/_{2} \leq \beta'(G)$ and $\alpha(G) \leq \beta'(G)$.

Theorem 2.1.5 (Gallai)**.** *If $\delta(G) \geq 1$, then $\alpha'(G) + \beta'(G) = n(G)$.*

Proof. Take a maximum matching M in G and let $V(M)$ be the vertices covered by M. For every vertex not covered by M, we can take an incident edge and add it to M. This gives an edge cover with

$$|M| + \left|\overline{V(M)}\right| = |M| + (n - 2\,|M|) = n - |M|$$

edges. Since $|M| = \alpha'(G)$, this implies $\beta'(G) + \alpha'(G) \leq n(G)$.

Now take a minimum edge cover C. We claim that for every edge in C, at least one end is covered only once by C. For suppose that $uv \in C$ is an edge and both u and v are covered by other edges in C. If we remove uv, then $C \setminus \{uv\}$ is a smaller cover, which is a contradiction.

The induced subgraph $G[C]$ is then a forest of stars. Suppose it consists of k components. We get $|C| = n - k$, since in a tree, the number of vertices is 1 more than the number of edges. A matching is obtained by choosing one edge from every star, which gives $\alpha'(G) + \beta'(G) \geq n(g)$, thus completing the proof. □

Question 2. State and prove Gallai's theorem.

Definition 2.1.6. Let M be a matching. A path $v_1v_2\dots v_k$ is an M-ALTERNATING PATH if the edges along the path alternate between M and $\overline{M}$. An M-alternating path is M-AUGMENTING if neither end of the path is covered by M.

Remark. Such a path cannot start or end with an edge from M, and the endpoints cannot be part of an edge in M.

Proposition 2.1.7. *If G is a graph, M is a matching and there exists an M-augmenting path P, then M is not a maximum matching.*

Proof. Suppose $P = v_1 \dots v_k$. We know the first and last edge are not in M, so $\left|E(P) \cap \overline{M}\right| = |E(P) \cap M| + 1$. Now let $M' = M \oplus E(P)$ be the symmetric difference of M and $E(P)$. This is clearly a matching. We know that $|M'| = |M| + 1$, so M cannot be maximum. □

Question 3. How can you construct a larger matching from an augmenting path?

Definition 2.1.8. A KÖNIG-EGERVÁRY graph is a graph G with $\alpha'(G) = \beta(G)$.

Theorem 2.1.9 (König)**.** *Let G be a bipartite graph. Then $\alpha'(G) = \beta(G)$. Additionally, if M is a matching in G and there is no M-augmenting path, M is a maximum matching.*

Proof. Let the partite classes of G be A and B. Suppose that M is a matching for which there is no M-augmenting path in G. Define X as the set of all vertices in A that are not covered by the matching, and Y the vertices in B not covered by M. Additionally, let B_1 be the set of vertices in B that can be reached by an M-alternating path from X, and similarly, let A_1 be the set of vertices of A which can be reached from X by an M-alternating path. Finally, define $B_2 = B \setminus (B_1 \cup Y)$ and $A_2 = A \setminus (A_1 \cup X)$.

Observe that on an M-alternating path from X, any edge from A to B is in $\overline{M}$ and any edge from B to A is in M. Out matching provides a one-to-one mapping between A_1 and B_1 and between A_2 and B_2, so $|A_1| = |B_1|$ and $|A_2| = |B_2|$. We also know that $|A_1| + |A_2| = |M|$. Now consider the possible edges between the defined vertex sets.

There is no edge between X and Y, since that would be a (trivial) M-augmenting path. There are also no edges between X and B_2, since an edge xb is an M-alternating path, implying $b \in B_1$. So the only edges from X lead to B_1.

There are also no edges between A_1 and Y, because we can construct an M-augmenting path with such an edge. If $a \in A_1$, then there is an alternating path from X to a, which we could extend with a a-to-Y edge to get an augmenting path. Finally, there are no edges between A_1 and B_2, since that would give an alternating path from X to the B_2-vertex as before.

Then $T = B_1 \cup A_2$ is a vertex cover with $|T| = |B_1| + |A_2| = |A_1| + |A_2| = |M|$. We have thus constructed a vertex cover with $|M|$ vertices, giving

$$\beta(G) \leq |T| = |M| \leq \alpha'(G).$$

The other inequality holds in the general case, so this completes the proof. □

Question 4. State and prove König's theorem.

Corollary 2.1.10. *If G is a bipartite graph, then $\alpha(G) = \beta'(G)$.*

Definition 2.1.11. Let G be a bipartite graph with partite classes A and B. HALL'S CONDITION holds for the set A if for every $S \subseteq A$,

$$|S| \leq |N(S)| = \left| \bigcup_{u \in S} N(u) \right|.$$

Theorem 2.1.12 (Hall)**.** *If G is a bipartite graph with partite classes A and B, then there exists a matching that covers A if and only if Hall's condition holds for A.*

Proof. Let M be a matching covering A and $S \subseteq A$. We can take the pairs matched by M

$$B_S = \{v \in B \mid v \text{ is covered by an edge in } M\}.$$

Clearly, $|S| = |B_S|$ and $B_S \subseteq N(S)$, so $|S| \leq |N(S)|$.

For the other implication, suppose there is no matching covering A. Divide the sets A and B as in the proof of König's theorem, using some matching M. Since the matching doesn't cover A, X is not empty. Now consider $S = A_1 \cup X$. All edges from S lead into B_1, so $N(S) = B_1$, but $|S| = |A_1| + |X| > |B_1| = |N(S)|$. □

Question 5. State and prove Hall's theorem.

Definition 2.1.13. A matching M is a PERFECT MATCHING if it covers all vertices.

Corollary 2.1.14. *In a bipartite graph G, there is a perfect matching if and only if $|A| = |B|$ and A satisfies Hall's condition.*

Definition 2.1.15. Let G be a bipartite graph with partite classes A, B, and $S \subseteq A$. The DEFICIENCY of S is $\operatorname{def}(S) = |S| - |N(S)|$.

Theorem 2.1.16. *Let G be a bipartite graph with partite classes A and B. If M is a maximum matching in G, it covers*

$$\alpha'(G) = |A| - \max_{S \subseteq A} \operatorname{def}(S)$$

vertices of A.

Theorem 2.1.17. *If G is a regular bipartite graph, then G has a perfect matching.*

Proof. The number of edges in the graph is $k \cdot |A| = k \cdot |B|$, so $|A| = |B|$. Let $S \subseteq A$. The number of edges between S and $N(S)$ is exactly $k \cdot |S|$. Every neighbour $u \in N(S)$ has exactly k neighbours, at most k are in S. So at most $k \cdot |N(S)|$ edges are between S and $N(S)$, implying $|S| \leq |N(S)|$. □

Question 6. Show that a regular bipartite graph has a perfect matching.

Theorem 2.1.18. *Let M be a matching in G. Then there is an M-augmenting path in G if and only if M is not a maximum matching in G.*

Proof. We've already proved the right implication. Suppose there is a matching M' with $|M'| > |M|$. Consider the symmetric difference $M \triangle M'$ and denote $G' = G[M \triangle M']$. Clearly the maximum degree $\Delta(G') \leq 2$, from which we know that the components of G' are all paths or cycles.

Any cycle must alternate between edges from M and M', so it is an even cycle, and contains the same number of edges from the two matchings. In any path of even length, there must be the same number of edges in M and in M'. And finally, in a path of odd length, one of the two sets has an extra edge compared to the other. Since $|M'| > |M|$, there must be a path with more edges in M' than in M. Label it G'_1.

This component is an M-augmenting path in G, since if either of its endpoints are covered by M, they must also be covered by the same edge in M', but then M' wouldn't be a matching. □

We can find the maximum matching in polynomial time, with so-called "Blossom algorithms", which find an augmenting path in $O(m\sqrt{n})$. This also means we can determine the edge-cover number $\beta'(G)$ in polynomial time.

2.1.1 Tutte's theorem

Definition 2.1.19. A component of a graph G is ODD if it has an odd number of vertices. We denote the number of odd components in G with $o(G)$.

Theorem 2.1.20 (Tutte)**.** *A graph G has a perfect matching if and only if for any $S \subseteq V(G)$, $|S| \geq o(G - S)$ holds. This is called Tutte's condition.*

Proof. Left to right: Let $S \subseteq V(G)$ and M be a perfect matching in G. Let $H_1, \ldots, H_k$ be the components of $G - S$. If H_i is an odd component, then there exists at least one M-edge between $V(H_i)$ and S. Therefore $|S| \geq o(G - S)$.

Right to left: Suppose that Tutte's condition holds for a graph F but there is no perfect matching in F. Now add edges to F so that this still holds after the addition, and let G be a maximal such graph on $n(F)$ vertices. If we consider Tutte's condition on $S = \varnothing$, we see that there are no odd components in G, which means $n(G)$ is even.

Now take any edge in the complement of G. Since G is maximal, $G + e$ is not a counterexample, so it either has a perfect matching, or Tutte's condition does not hold for it. We know that for any $S \subseteq V(G)$, $|S| \geq o(G - S)$. After having added an edge, at most one pair of components in $G - S$ is joined. If this was a pair of odd components, $o(G - S)$ has reduced, and in all other cases, it has remained the same. This means that

$G + e$ must satisfy Tutte's condition, so it must have a perfect matching as it is not a counterexample.

We will consider two cases. Let U be the set of universal vertices in G, that is, the vertices adjacent to every other vertex, and let $H_1, \ldots, H_k$ be the components of $G - U$. In the first case, suppose every H_i induces a complete graph. We will construct a perfect matching. For the even components, we may create a matching within the component, and for the odd components, we may connect the one remaining vertex with U. We are left over with an even number of vertices in U (since n is even), which we may form a matching with, as $G[U]$ is a complete graph. So in this case, G is not a counterexample, which is a contradiction.

Now suppose there is a non-complete component H_i. Then there exist vertices x, y, z for which $xy \notin E(H_1)$ but both xz and yz are in $E(H_1)$. There is also another vertex $w \in V(G)$ for which $zw \notin E(G)$, since z is not a universal vertex. Let $G_1 = G + xy$ and $G_2 = G + zw$. We know both these graphs have perfect matchings, which must include xy and zw respectively. Denote the perfect matchings with M_1 and M_2 and consider $G[M_1 \triangle M_2]$. Note that every non-isolated vertex in this graph is of degree 2, since it is covered by both perfect matchings. These vertices must of course appear in even cycles.

If xy and zw belong to different components, then we may choose edges in each component, and the edges deleted by the symmetric difference, to form a perfect matching, and we can avoid taking xy or zw. This is a contradiction. Alternatively, if xy and zw belong to the same component, they appear in the same cycle. Without loss of generality they appear in the order $zwxy$ (but not necessarily adjacent). To form a new perfect matching, we will take the edge xz, and one edge set from each side of the cycle. This avoids both new edges xy and zw, so we have a contradiction. □

We also have the Berge-Tutte formula, which states that a maximum matching in G leaves uncovered exactly

$$\max_{S \subseteq V(G)} \{o(G - S) - |S|\}$$

vertices.

2.1.2 Factors

Definition 2.1.21. A FACTOR is a spanning subgraph (it contains all vertices). A k-FACTOR is a k-regular spanning subgraph.

Definition 2.1.22. A CUBIC GRAPH is a 3-regular graph.

Theorem 2.1.23 (Petersen)**.** *Every bridgeless cubic graph has a* 1*-factor.*

Proof. We will prove that Tutte's condition holds. Take $S \subseteq V(G)$. Since the graph is 3-regular, the number of edges between S and $\overline{S}$ is $\left|E(S, \overline{S})\right| \leq 3\,|S|$. If H_i is an odd

component in $G - S$, then $E(V(H_i), S)$ contains at least one edge. Note that

$$2m(H_i) + |E(H_i, S)| = 3n(H_i) = \sum_{v \in H_i} \deg_G(v)$$

where $m(H_i)$ is the number of edges in H_i. Since the RHS is odd, $|E(H_i, S)|$ must be odd. Also, $|E(H_i, S)| \neq 1$, since that would be a cut edge (a bridge) otherwise. Therefore, $|E(H_i, S)| \geq 3$. Then $3|S| \geq \left|E(S, \overline{S})\right| \geq 3\sigma(G - S)$. □

We can also improve this theorem, for example allowing for precisely one bridge in the graph, and the statement still holds. Or we could require that all cut edges lie on the same path, and it would still hold.

Theorem 2.1.24. *If G is a k-regular graph and k is even, then G has a 2-factor.*

Proof. We can limit ourselves to connected graphs. By Euler's theorem, there exists an Eulerian circuit C in G. We can fix a direction for this C and orient the edges according to it. From this, we obtain a directed graph $\vec{G}$. For every vertex v, C enters and exits v exactly $k/2$ times.

Define a bipartite graph F_G with partite classes $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_n\}$ for $n = n(G)$ and the edge set

$$a_i b_j \in E(F_G) \Leftrightarrow v_i v_j \in E(\vec{G}).$$

Clearly, F_G is a $\frac{k}{2}$-regular graph, so it has a perfect matching M. Define a spanning subgraph in G with the following edge set Q:

$$v_i v_j \in Q \Leftrightarrow a_i b_j \in M \vee a_j b_i \in M.$$

Every vertex v_i is covered exactly twice by Q. □

2.2 Connectivity

Definition 2.2.1. The CONNECTIVITY NUMBER $\kappa(G)$ is the minimum number of vertices in $S \subseteq V(G)$ such that $G - S$ is either disconnected or contains only one vertex.

Definition 2.2.2. A graph G is k-CONNECTED if $\kappa(G) \geq k$ or if the removal of $k - 1$ vertices always results in a connected graph with at least two vertices.

Remark. In any graph, $\kappa(G) \leq \delta(G)$. If G is complete, then $\delta(G) = n - 1 = \kappa(G)$.

Remark. If A is an independent set, removing every other vertex gives a disconnected graph, so $\kappa(G) \leq n - \alpha(G) = \beta(G)$.

Theorem 2.2.3. *The minimum number of edges in a k-connected graph of order n is $\lceil \frac{nk}{2} \rceil$, if $n > k \geq 2$.*

Proof. If the graph is k-connected, then $k \leq \kappa(G) \leq \delta(G)$, so

$$m(G) = \frac{1}{2}\sum_{v \in V} \deg_G(v) \geq \frac{nk}{2}.$$

We will show that there exists a k-connected graph of order n with the specified number of edges. For this, we define Harary graphs $H_{n,k}$ as follows.

- If k is even, $H_{n,k} = C_n^{k/2}$ (i.e. the graph we get by connecting all vertices from C_n which are at a distance of at most $k/2$).
- If k is odd and n is even, $H_{n,k}$ is $C_n^{(k-1)/2}$, along with all edges between two opposing vertices of the cycle.
- If both n and k are odd, then we again take $C_n^{(k-1)/2}$ and add the edges between i and $i + \frac{n-1}{2}$ (if we label the vertices $0, 1, \ldots, n-1$).

All these graphs have $m(H_{n,k}) = \lceil \frac{nk}{2} \rceil$.

We will show that all these graphs are k-connected. For the first case, if k is even, let S be a vertex set with $|S| = k - 1$. We define a big gap as $\frac{k}{2}$ consecutive vertices in S, and claim the following:

- If there is no big gap between u and v in a certain direction, then we may find a uv-path in that direction. This is clear from the construction.
- For any $u, v \in V(G) \setminus S$, there is a uv-path in $G - S$. Since there can be only one big gap between them, we can just avoid it by going in the other direction, so this is also clear.

If k is odd and n even, we define a big gap as $\frac{k-1}{2}$ consecutive missing vertices. Similarly as before, if there is no big gap on a path from u to v, we can find a path after removing S. But in this case, both paths may contain a big gap. Let P and Q be the two paths along the cycle. If there are big gaps along both, we know the length of both is at least $\frac{k-1}{2} + 1$. Let u' and v' be the opposite vertices of u and v. Suppose that Q is longer than P, and split it into paths Q_1, Q_2 and Q_3 by v' and u'.

Note that by symmetry, the length of Q_2 (the center region) is also at least $\frac{k+1}{2}$, so the big gap in Q cannot cover both u' and v'. We can then find a u, v-path in $G - S$ using one of these vertices.

We can consider the case of odd k and odd n similarly, it's just more annoying to write down. In all cases, all graphs have precisely $\lceil \frac{kn}{2} \rceil$ edges. □

Definition 2.2.4. A set $F \subseteq E(G)$ is a DISCONNECTING SET if $G - F$ is disconnected.

Definition 2.2.5. An EDGE CUT of A is the set $E(A, \overline{A})$ of edges between A and $\overline{A}$.

Remark. An edge cut is a disconnecting set. A minimal disconnected set is an edge cut.

Definition 2.2.6. The EDGE-CONNECTIVITY number of G is the minimum number of edges in a disconnecting set. We denote it by $\kappa'(G)$. A graph is k-EDGE-CONNECTED if the removal of less than k edges always leaves a connected graph.

Theorem 2.2.7. *Suppose G is a simple graph with $n(G) \geq 2$. Then $\kappa(G) \leq \kappa'(G) \leq \delta(G)$.*

Proof. The second inequality is clear. Consider a minimum edge cut $E(A, \overline{A})$ in G. By definition $\left|E(A, \overline{A})\right| = \kappa'(G)$. We will show that there is a vertex cut with at most $\kappa'(G)$ vertices. For that, consider two cases. If $E(A, \overline{A})$ forms a complete bipartite graph, then

$$\left|E(A, \overline{A})\right| = |A| \left|\overline{A}\right| = |A| (n - |A|).$$

Since $0 < |A| < n$, we have $\left|E(A, \overline{A})\right| \geq n - 1$, so $\kappa'(G) \geq n - 1$, but $\kappa(G) \leq n - 1$ and $\kappa(G) \leq \kappa'(G)$.

In the second case, if there are vertices $x \in A$ and $y \in \overline{A}$ which are nonadjacent, then we may choose an endpoint different from x and y in each edge in the edge cut. This gives us a vertex cut with at most $\left|E(A, \overline{A})\right|$ vertices, in which x and y are not cut, and are disconnected. □

Corollary 2.2.8. *If G is k-connected, then G is k-edge-connected.*

Corollary 2.2.9. *The minimum number of edges in a k-edge-connected graph on n vertices is $\lceil \frac{kn}{2} \rceil$.*

Proof. We know that $k \leq \kappa'(G) \leq \delta(G)$, so

$$m(G) = \frac{1}{2} \sum_{v \in V} \deg_G(v) \geq \frac{1}{2} n\delta(G)$$

which means $m(G) \geq \lceil \frac{nk}{2} \rceil$. For the other direction, note that the Harary graphs are k-edge-connected. □

Theorem 2.2.10 (Whitney)**.** *If G is a 2-connected graph, then for every $u, v \in V(G)$, there are two internally disjoint u, v-paths. The converse also holds.*

Proof. For the left implication, suppose $u, v \in V(G - x)$ for some vertex x. There are two disjoint paths from u to v in G, and at most one of them includes x, so there is a path from u to v in $G - x$. This means our graph has no cut vertex, so it is 2-connected.

Now consider the right implication. Let u, v be vertices in G. Induction on $d = d(u, v)$. For $d = 1$, we know that G is 2-edge-connected (since it is 2-connected), so there is no bridge in G. If we remove uv, then the graph must still be connected. Therefore, there is another u, v-path in G.

For a general d, let w be the neighbour of v on the shortest u, v-path. Then $d(u, w) = d - 1$. By the induction hypothesis, there are two internally disjoint paths P, Q from u to

w. Consider two cases. If $v \in V(P)$ (or Q, symmetrically), then we have two internally disjoint u, v-paths in

$$u \xrightarrow{P} v, \qquad u \xrightarrow{Q} w \to v.$$

Otherwise, if v is on neither path, then consider the graph $G - w$. It is still connected, so there is at least one u, v-path R in this graph. If R shares no internal vertex with P or Q, then we may choose $u \xrightarrow{P} w \to v$ and R as the two paths. Finally, if R intersects P and/or Q, identify the last intersection with either, and label it z. Without loss of generality, $z \in Q \cap R$. Then we have two paths

$$u \xrightarrow{P} w \to v, \qquad u \xrightarrow{Q} z \xrightarrow{R} v.$$

□

Theorem 2.2.11 (Expansion lemma)**.** *If G is k-connected and we add a new vertex v and k incident edges to the graph, then we obtain a k-connected graph.*

Proof. Call the new vertex y, and let G' be the new graph. We will prove that every vertex cut in G' contains at least k vertices. For that, let S be a vertex cut in G'. Consider three cases.

- If $y \in S$, then let V_1 and V_2 be components of $G \setminus S$. Every path from V_1 to V_2 passes S, which is also true in G, since $y \in S$. In G, every vertex cut contains at least k vertices, so $S \setminus \{y\}$ contains at least k vertices, and S contains at least $k+1$ vertices.
- If $y \notin S$ and $N(y) \subseteq S$, then y is its own component in $G' \setminus S$. This means $|S| \geq k$.
- Otherwise, there is a vertex y' which is a neighbour of y and belongs to the same component in $G' \setminus S$. If we remove y from G', S remains a vertex cut in G, as no path exists between the components which avoids S. Note that we don't remove the component of y if we delete the vertex, as the component has at least one other vertex (y'). Since G is k-connected, every vertex cut contains at least k vertices, so $|S| \geq k$.

□

Theorem 2.2.12. *If G is a graph with $n(G) \geq 3$, then the following statements are equivalent:*

- *G is 2-connected,*
- *G is connected and there is no cut vertex,*
- *for every $u, v \in V(G)$, there are at least two internally disjoint u, v-paths,*
- *for every $u, v \in V(G)$, there is a cycle through u and v,*

- $\delta(G) \geq 1$ *and, for every two edges* e_1, e_2 *there is a cycle containing* e_1 *and* e_2.

Proof. We already know the first four statements are equivalent. Suppose G is 2-connected. Take edges $e, f \in E(G)$ and label $e = uv, f = u'v'$. Now add two vertices w, w' to G, w connected to u and v, and w' connected to u' and v'. Let G' be the resulting graph. By the expansion lemma, G' is 2-connected, and satisfies the fourth condition, so there is a cycle though w and w'. This cycle must contain all the new edges and the vertices $u, v, u'v'$. We can now replace the added edges with e and f.

Conversely, take vertices $u, v \in V(G)$. Now let $e = uu'$ and $f = vv'$ be two different edges. There is a cycle though e and f by assumption, so there is a cycle though u and v. □

Proposition 2.2.13 (subdivision lemma)**.** *Suppose* G' *is obtained from* G *by subdividing an edge* $uv \in E(G)$ *with a vertex* w*. Then* G *is* 2*-connected if and only if* G' *is* 2*-connected.*

Proof. Left to right: Consider two vertices $x, y \in V(G')$. If neither is equal to w, then we can take the cycle though xy in G. If it contains uv, we can replace it with the path uwv. If one of the vertices x, y is w, then we find a cycle through u and v, and subdivide the edge.

Right to left: Any cycle in G' that contains w must also contain u and v. □

2.2.1 Ear decomposition of a graph

Definition 2.2.14. In a graph G, a path P is an (OPEN) EAR if all internal vertices of P are of degree 2 in G, and for the end vertices of the path, the degree is at least 3.

Definition 2.2.15. An (OPEN) EAR DECOMPOSITION of G is a sequence $P_0, P_1, \ldots, P_k$, where P_0 is a cycle in G, and every other P_i is an ear in the graph $G_i := P_0 \cup P_1 \cup \ldots \cup P_i$. We also require $G_k = G$.

Theorem 2.2.16. *A graph* G *is* 2*-connected if and only if it has an ear decomposition.*

Proof. Right to left: We will prove that G_i is 2-connected for every i by induction. For $i = 0$, we know that a cycle is 2-connected. For the induction step, let u, v be the endpoints of P_{i+1}. Add an edge uv to G_i. It is still 2-connected, as adding an edge cannot decrease connectivity. Now we can repeatedly subdivide this edge, and the resulting graph is still 2-connected by the subdivision lemma.

Left to right: Since G is 2-connected, there exists a cycle C, which we can take as P_0. We will construct the decomposition inductively. If G_i is obtained from P_0 by adding the ears $P_1, \ldots, P_i$, then:

- If $G = G_i$, we're done.

- If G_i is not an induced subgraph of G, we may add an edge from $E(G) \setminus E(G_i)$ with both endpoints in $V(G_i)$. This is a valid ear (as since G_i is 2-connected, $\delta(G_i) \geq 2$), so we can continue the decomposition.
- If G_i is an induced subgraph of G, then there exists a vertex $v \in V(G) \setminus V(G_i)$ which is connected to a vertex $u \in V(G_i)$. Since G is 2-connected, there is a cycle C though the edge uv and some other edge $f \in E(G_i)$. Let u' be the first vertex on this cycle which is in G_i on a $u \to v \to \cdots$ path. We can add this path between u and u' as P_{i+1}. □

Proposition 2.2.17. *A graph G is 2-edge-connected if and only if G is connected and there is no cut edge in G.*

Definition 2.2.18. A CLOSED EAR in a graph G is a cycle for which all but one vertex has degree 2, and the exception has degree at least 4.

Definition 2.2.19. A CLOSED EAR DECOMPOSITION of G is a sequence $P_0, P_1, \ldots, P_k$ such that P_0 is a cycle and P_i is either an open or closed ear in $G_i = P_0 \cup P_1 \cup \ldots \cup P_i$ for all $i \geq 1$.

Theorem 2.2.20. *A graph G is 2-edge-connected if any only if it has a closed ear decomposition.*

Proof. Right to left: Since P_0 is a cycle, it is 2-edge-connected. If G_i is 2-edge-connected, so is G_{i+1}: Take any edge $uv \in E(G_{i+1})$. If $uv \in E(G_i)$, then it is part of a cycle in G_i. If $uv \in E(P_{i+1})$, then either P_{i+1} is a closed ear, so a cycle containing uv, or it is an open ear between some vertices x, y, in which case there is an x, y-path in G_i, which forms a cycle with P_{i+1}.

Left to right: Start the closed ear decomposition with a cycle P_0, and build inductively. If G_i is not an induced subgraph of G, then we can add edges with open ears. If it is, then let v be a vertex in $V(G) \setminus V(G_i)$ connected to a vertex $u \in V(G_i)$. We know uv is in a cycle in G. If u is the only vertex of the cycle in G_i, then we can add it as a closed ear. Otherwise, we can add it as an open ear as in the previous theorem. □

Definition 2.2.21. A STRONG ORIENTATION of an undirected graph G is a digraph $\vec{G}$ which is strongly connected, and which you get from choosing an orientation for each edge in G.

Theorem 2.2.22 (Robbins)**.** *An undirected graph G has a strong orientation if and only if it is 2-edge-connected.*

Proof. Left to right: Suppose that G is not 2-edge-connected, and consider two cases.

- If G is not connected, then there can be no strong orientation.
- If G has a cut edge $e = xy$, then for whatever orientation of e we choose, there is no path between those vertices in the opposite direction.

Right to left: If G is 2-edge-connected, then we have a closed ear decomposition $P_0, \ldots, P_k$. We can orient the edges in P_0 consistently to get a strongly connected graph. Whenever you add an ear (open or closed) to $\vec{G}_i$, orient the new edges consistently. You can show via simple casework that the new digraph is still strongly connected. □

2.2.2 x, y-cuts

Definition 2.2.23. If x and y are nonadjacent vertices in the graph G, then $S \subseteq V(G)$ is an x, y-CUT if x and y belong to different components in $G - S$. We label the minimum size of an x, y-cut in G with $\kappa_G(x, y)$.

Definition 2.2.24. The maximum number of internally vertex-disjoint x, y-paths in a graph G is labeled with $\lambda_G(x, y)$.

Theorem 2.2.25 (Menger's theorem for vertex cuts)**.** *If x and y are nonadjacent vertices in G, then $\kappa_G(x, y) = \lambda_G(x, y)$.*

Proof. Denote $\kappa = \kappa_G(x, y)$ and $\lambda = \lambda_G(x, y)$. We have $\lambda \leq \kappa$ since, if there are λ internally vertex-disjoint x, y-paths in G, we have to remove at least λ vertices to separate x and y, one on each path.

For the other direction, use induction on $n(G)$. If $n(G) = 2$, since x and y are not connected, they must be isolated, so $\kappa = \lambda$. For the induction step, consider two cases. In the first case, if there is a minimum x, y-cut in S such that $S \neq N(x)$ and $S \neq N(y)$, then at least one neighbour of x is not in S (as $N(x)$ is an x, y-cut and S is minimal). The same holds for y.

Consider x, S-paths (paths from x to a vertex in S which only hit S in one vertex). Let V_1 be the union of the vertex sets of all these paths, and let V_2 be the union of the vertex sets of all y, S-paths. We will prove that $V_1 \cap V_2 = S$. Let X_1 be the component of $G - S$ which includes x, and let X_2 be the component of $G - S$ which includes y. Note that every vertex in S must be adjacent to a vertex in X_1 and to a vertex in X_2, as otherwise we could remove the offending vertex from S and get a smaller vertex cut.

Suppose that $w \in V_1 \cap V_2 \setminus S$. Then we have an x, w-path which does not intersect S, and an y, w-path which does not intersect S. This can't happen as S is a cut set, so $V_1 \cap V_2 \subseteq S$. The other inclusion clearly holds. As we've noted before, there is a neighbour of x not in S, so $|V_1 \setminus S| \geq 2$ and similarly $|V_2 \setminus S| \geq 2$.

Let

$$G_1 = G[V_1] \cup (\text{a vertex } y' \text{ adjacent to every vertex of } S).$$

We know $n(G_1) < n(G)$. It is easy to prove that S is a minimum x, y'-cut in G_1, so $\kappa_{G_1}(x, y') = \kappa_G(x, y) = \kappa$, but by the induction hypothesis, $\lambda_{G_1}(x, y') = \kappa_{G_1}(x, y')$. So we have κ internally vertex-disjoint x, y'-paths in G_1. We can similarly define G_2 as $G[V_2]$ with an added vertex x', adjacent to every vertex of S, and find κ internally

vertex-disjoint x', y-paths. This allows us to construct κ internally vertex-disjoint x, y-paths in G by just connecting the G_1 and G_2 paths which share a common vertex in S.

This concludes the first case, now consider the case where all minimum x, y-cuts are either $N(x)$ or $N(y)$. Consider three subcases.

- If there is a vertex $v \in N(x) \cap N(y)$, then $\kappa_{G-v}(x, y) = \kappa - 1$ and we can use the induction hypothesis for $G - v$.
- If there is a vertex $v \notin N[x] \cup N[y]$, then v does not belong to any minimum x, y-cuts, so $\kappa_{G-v}(x, y) = \kappa_G(x, y) = \kappa$. We can again use the induction hypothesis on $G - v$.
- If neither of the above hold, then $V(G) = N[x] \cup N[y]$ and $N[x] \cap N[y] = \varnothing$. Let F be the bipartite graph obtained by taking $V(F) = N(x) \cup N(y)$ and $E(F) = E(N(x), N(y))$. We may find an x, y-path by taking an edge in F and connecting x and y to the endpoints. If we have a matching in F, we can use it to construct that many internally vertex-disjoint paths, so $\lambda \geq \alpha'(F)$. To find an x, y-cut in G, we must remove all edges in F, so we need a vertex cover T of F. By König's theorem, $\beta(F) = \alpha'(F)$, so $\kappa \leq \beta(F) = \alpha'(F) \leq \lambda$, which is what we were trying to prove. □

Definition 2.2.26. For two vertices x, y of G, a set $R \subseteq E(G)$ is an x, y-EDGE CUT if $G - R$ is disconnected and x, y belong to different components of $G - R$. We denote the minimum size of an x, y-edge cut in G by $\kappa'_G(x, y)$. We also define $\lambda'_G(x, y)$ to be the maximum number of edge-disjoint x, y-paths in G.

Theorem 2.2.27 (Menger's theorem for edge cuts)**.** *Let $x, y \in V(G)$. Then $\kappa'_G(x, y) = \lambda'_G(x, y)$.*

Proof. Let G' be obtained by adding two vertices to G, u and v, and edges xu and yv. It is easy to see that a ux, yv-path in $L(G')$ corresponds to an x, y-path in G. By Menger's theorem for vertices,

$$\lambda'_G(x, y) = \lambda_{L(G')}(ux, yv) = \kappa_{L(G')}(ux, yv).$$

Clearly, a vertex cut in $L(G')$ that separates ux and yv corresponds to an edge cut in G that separates x and y, which finishes the proof. □

Lemma 2.2.28. *Let $e \in E(G)$. Then $\kappa(G) - 1 \leq \kappa(G - e) \leq \kappa(G)$.*

Proof. Clearly $\kappa(G - e) \leq \kappa(G)$. Suppose that the strong inequality holds. Then let S be a minimum vertex cut in $G - e$ for $e = xy$, so $|S| = \kappa(G - e)$ and $G - e - S$ is disconnected. Since $|S| < \kappa(G)$, it is not a vertex cut in G, so x and y belong to different components of $G - e - S$. Let x be in the component X and y in the component Y. Consider three cases:

- if $|X| \geq 2$, then $S \cup \{x\}$ is a vertex cut in G,
- if $|Y| \geq 2$, then $S \cup \{y\}$ is a vertex cut in G,
- If $|X| = |Y| = 1$, then $n(G) - 2 = \kappa(G - e) = |S|$. We know $\kappa(G) \leq n(G) - 1$, so $\kappa(G - e) + 1 \geq \kappa(G)$. □

Theorem 2.2.29 (Menger)**.** *In every graph G,*

$$\kappa'(G) = \min_{x \neq y} \lambda'_G(x, y), \qquad \kappa(G) = \min_{x \neq y} \lambda_G(x, y).$$

Proof. For any $x, y \in V(G)$, if S is an x, y-edge-cut, then it is an edge-cut in G, so

$$\min_{x \neq y} \kappa'_G(x, y) \geq \kappa'(G).$$

If S is an edge cut, then it separates two vertices, so

$$\kappa'(G) \geq \min_{x \neq y} \kappa'_G(x, y).$$

By Menger's theorem for edges, $\lambda'_G(x, y) = \kappa'_G(x, y)$.

For the second claim, we analogously show (if G is not complete)

$$\kappa(G) = \min_{x \neq y, xy \notin E(G)} \lambda_G(x, y).$$

It suffices to prove that for any two adjacent vertices x, y that $\lambda_G(x, y) \geq \kappa(G)$. For that, define $G' = G - xy$. Then $\lambda_{G'}(x, y) = \lambda_G(x, y) - 1$ since the single edge was a path. By Menger's theorem for G', we have

$$\kappa_{G'}(x, y) = \lambda_{G'}(x, y) = \lambda_G(x, y) - 1,$$

and by the preceding lemma, $\kappa(G') \geq \kappa(G) - 1$, so

$$\lambda_G(x, y) = \kappa_{G'}(x, y) + 1 \geq \kappa(G).$$ □

2.3 Coloring

Remark. In any graph, $\omega(G) \leq \chi(G) \leq \Delta(G) + 1$.

Remark. As the color classes are independent sets, the number of vertices in a color class is at most $\alpha(G)$, so

$$\chi(G) \geq \frac{n(G)}{\alpha(G)}.$$

Theorem 2.3.1 (Welsh-Powel)**.** *If $d_1 \geq d_2 \geq \cdots \geq d_n$ is the degree sequence of the vertices of G, then*

$$\chi(G) \leq 1 + \max_{i=1,\ldots,n} \{\min\{d_i, i - 1\}\}.$$

Theorem 2.3.2 (Brooks)**.** *If G is connected and not a complete graph or odd cycle, then $\chi(G) \leq \Delta(G)$.*

Proof. Let $k = \Delta(G)$. Consider two cases. First, if G is not k-regular, then there is a vertex v_n with degree at most $k-1$. Define the following vertex order for a greedy coloring. Consider a breadth-first search tree, rooted in v_n. Order the vertices such that every child of the tree precedes its parent. This way, each vertex is preceded by at most $k-1$ of its neighbours, so we can color the entire graph with at most k colors.

In the second case, if G is k-regular, we can quickly write off $k = 1$ as then $G = P_2$, which is a complete graph, and the case $k = 2$, where G must be a cycle. We have excluded odd cycles, and for even cycles, $\chi(G) = \Delta(G) = 2$. So let $k \geq 3$. We will consider three subcases.

If $\kappa(G) = 1$, then there exists a cut vertex $x \in V(G)$. Let V_1, V_2 be disconnected vertex sets in $G - x$ such that $V_1 \cup V_2 = V(G - x)$, and let $G_i = G[V_i \cup \{x\}]$. Since x has neighbours in both V_1 and V_2, the degree of x in either G_i is at most $k-1$ and G_1, G_2 are not regular. So we can color them with at most k colors. We can permute the colors in one of the colorings so that x has the same color in both.

In the second subcase, suppose $\kappa(G) = 2$. Then we have $x, y \in V(G)$ such that $G - \{x, y\}$ is disconnected. Define V_1 and V_2 as before and $G_i = G[V_i \cup \{x, y\}]$. We have $\deg_{G_i}(x) < \deg_G(x) = k$, so neither G_i is regular. Therefore, they are k-colorable, so we have colorings φ_1, φ_2. If $\varphi_1(x) = \varphi_1(y)$ and $\varphi_2(x) = \varphi_2(y)$, or if both of these pairs are nonequal, then we can define a combined k-coloring.

If we cannot find such φ_1, φ_2 however, then without loss of generality every k-coloring of G_1 assigns the same color to x and y, and every k-coloring of G_2 assigns different colors to those vertices. Equivalently, $G_1 + xy$ is not k-colorable, and $G_2 + xy$ is. So $G_1 + xy$ must be k-regular, in which case, we will prove that G_2 can be k-colored with $\varphi_2(x) = \varphi_2(y)$. Let $G_2' = G_2 - \{x, y\}$. It is k-colorable. Since $G_1 + xy$ is k-regular, x and y have precisely one neighbour each in G_2, meaning we can choose a color different from the colors of those neighbours (as $k \geq 3$) and find a k-coloring for G_2 with $\varphi_2(x) = \varphi_2(y)$.

Finally, consider the case $\kappa(G) \geq 3$. We have vertices $u, v \in V(G)$ for which $d(u, v) = 2$, since G is not complete. Let z be a common neighbour for them. We want a vertex order such that in a greedy coloring, $\varphi(u) = \varphi(v)$ and the last vertex in the ordering is z. Let $G' = G - \{u, v\}$. Consider a breadth-first search in G', rooted in z. Since $\kappa(G) \geq 3$, it will not stop immediately. Again, take the order in which every vertex is preceded by at most $k-1$ children, and z is at the end. Now a greedy coloring $u, v, v_1, v_2, \ldots, v_{n-2} = z$ of G will assign $\varphi(u) = \varphi(v) = 1$, and for $v_1, \ldots, v_{n-3}$, at most $k-1$ neighbours are colored before v_i, so it is assigned a color $\leq k$. We know that z has two neighbours with the same color, so we can find a k-coloring for G. □

2.3.1 Mycielski's construction

The Mycielskian $M(G)$ of a graph G is a graph defined as follows:

- label the vertices of G as $v_1, \ldots, v_n$,
- create $n+1$ new vertices $u_1, \ldots, u_n, z$,
- add connections $u_i v_j$ for all pairs $v_i v_j \in E(G)$,
- add connections $u_i z$ for all i.

Label $V = \{v_1, \ldots, v_n\}$ and $U = \{u_1, \ldots, u_n\}$.

Theorem 2.3.3. *If G is a graph with at least one edge, then $\chi(M(G)) = \chi(G) + 1$ and $\omega(M(G)) = \omega(G)$.*

Proof. Since G is a subgraph of $M(G)$, we have $\omega(G) \leq \omega(M(G))$. If z is in a clique of $M(G)$, then this is a clique of order at most 2, which appears in G as well. If u_i is in a clique of $M(G)$, then v_i can't be in the same clique, so we can replace u_i with v_i, and find a clique in G of the same size. Therefore $\omega(M(G)) \leq \omega(G)$.

If we have a coloring of G, then we can paint u_i with the same color as v_i, and use a new color for z. So $\chi(M(G)) \leq \chi(G) + 1$. Now suppose that there exists a $\chi(G)$-coloring of $M(G)$ and label it φ. Without loss of generality, $\varphi(z) = k := \chi(G)$. Then this color does not appear in U, so U is colored with $k-1$ colors. But then since $\chi(G) = k$, the color k must appear in V. We can replace the colors for those v_i which have $\varphi(v_i) = k$ with the color of u_i, and get a proper $(k-1)$-coloring of G, which is impossible as $\chi(G) = k$. $\rightarrow\!\!\!\!\!\!\times\!\!\!\!\!\!\leftarrow$ □

Theorem 2.3.4. *If G is a graph on n vertices and $\chi(G) = k$, then $m(G) \geq \binom{k}{2}$. This is sharp for any $n \geq k$.*

Proof. There is a partition of G into k color classes. Since $\chi(G) = k$, there must be at least one edge between any pair of color classes. This bound is sharp, as we can take K_k and add isolated vertices as required. □

2.3.2 Turàn's theorem

Definition 2.3.5. A graph G is k-PARTITE if the vertex set can be partitioned into k classes $V_1, V_2, \ldots, V_k$ such that every edge is between different classes.

Remark. A graph G is k-partite if and only if G is k-colorable.

Definition 2.3.6. G is a COMPLETE k-PARTITE GRAPH if every edge between the partite classes is present in G.

Definition 2.3.7. THE TURÀN GRAPH $T_{n,k}$ is the complete k-partite graph on n vertices such that the partite classes are of nearly equal size, i.e. if $||V_i| - |V_j|| \leq 1$ for all i, j.

Theorem 2.3.8 (Turàn)**.** *If G is a graph of order n and $\omega(G) \leq r$, then the number of edges in G is at most the number of edges in $T_{n,r}$.*

Proof. Induction on r. If $r = 1$, there are no edges, so $G = T_{n,1}$. For $r > 1$, let $k = \Delta(G)$ and let v be a vertex with degree k. Let $G' = G[N(v)]$. We know $\omega(G') \leq r - 1$, so by the induction hypothesis, $m(G') \leq m(T_{k,r-1})$.

Consider the following construction. Let H be the graph obtained by a complete join of a graph with $n - k$ independent vertices and the graph $T_{k,r-1}$. Clearly, $n(H) = n$ and $\omega(H) = r$, since $\omega(T_{k,r-1}) = r - 1$. Compute

$$m(H) = m(T_{k,r-1}) + (n-k)k \geq m(G') + (n-k)k$$

and $m(G) \leq m(G') + k(n-k)$, noting that in the remainder of the graph, every vertex has degree at most k, and there are $n - k$ vertices there.

We see that H is an r-partite graph. We claim that among the r-partite graphs on n vertices, $T_{n,r}$ has the most edges. Suppose that F is an r-partite graph with $n(F) = n$. If it is not a complete r-partite graph, then we may add edges until it is. If the sizes of the partite classes in F differ by at least 2, then take a vertex u from the larger class V_i and put it into the smaller class V_j. For this, we break $|V_j|$ edges and add $|V_i| - 1$ edges, so we have increased their total number. □

Remark. As $T_{n,r}$ satisfies the theorem's conditions, the bound is sharp. It is actually the unique graph with the maximum number of edges.

Remark. A similar result holds for graphs with $\chi(G) = r$.

2.3.3 Chordal graphs

Definition 2.3.9. A graph G is CHORDAL if there is no induced subgraph that is isomorphic to a cycle C_n for $n \geq 4$.

Definition 2.3.10. A vertex v is a SIMPLICIAL vertex in G if $N_G[v]$ induces a clique.

Lemma 2.3.11 (Voloshin)**.** *If G is a chordal graph, then for every $x \in V(G)$ there exists a simplicial vertex among the vertices farthest from x.*

Proof. Induction on $n(G)$. If $n(G) = 1$, the statement is obvious. Let $n(G) > 1$. If x is a universal vertex in G, then remove x and apply the induction hypothesis to $G - x$. Otherwise, let T be the set of vertices farthest from x, and let H be a component in T. Take S to be the neighbours of vertices in H which are not themselves in H, so $S = N(H) \setminus H$. Finally, let Q be the vertices in the component of $G - S$ which contains x.

We will prove that S induces a clique. Take $u, v \in S$. Then both vertices have neighbours in H and in Q, so there exist two u, v-paths through H and Q, respectively. Consider

the shortest such paths P_H, P_Q. Combined, they form a cycle of order at least 4. There cannot be a chord in $\{u/v\} \cup H$ or $\{u/v\} \cup Q$, since that would give a shorter path, and there is also no edge between Q and H, so there must be a cord $uv \in E(G)$. So any two vertices in S are connected, meaning S is a clique. By the induction hypothesis on $G[S \cup H]$ with a vertex from S, there is a simplicial vertex in H. □

3 Teorija izračunljivosti

3.1 Introduction

A TURING MACHINE is defined to consist of the following components. There is an infinite tape divided into cells, each of which contains a symbol from the chosen alphabet Γ. This alphabet must include a `blank` symbol. At the start, only a finite number of cells in the tape have a character different than `blank`. The machine also possesses a read-write head positioned at some cell, and an internal control state, which determines the instruction to be followed.

Instructions are given as a TRANSITION (partial) function f, which maps

$$(\text{state}, \text{character}) \mapsto (\text{new state}, \text{new character}, \text{motion}).$$

To perform an action, a TM will look for rules matching its current control state and the character currently written at the position of the read-write head. When a matching rule is found, the machine switches to the defined new state, writes the specified character on the tape and moves according to the instruction. We limit its motion to three possibilities: one cell to the left or right, or no motion at all. More formally, we may restate the definition as follows:

Definition 3.1.1. A TURING MACHINE is specified by the following:

- a finite TAPE ALPHABET Γ with $\boxdot \in \Gamma$,
- a finite set Q of STATES with $\texttt{start} \in Q$,
- a transition partial function

$$\delta : Q \times \Gamma \rightharpoonup Q \times \Gamma \times \{-1, 0, +1\}.$$

For a given input alphabet $\Sigma_1 \subseteq \Gamma$ and output alphabet $\Sigma_2 \subseteq \Gamma$, a TM specifies a partial function $f : \Sigma_1^* \rightharpoonup \Sigma_2^*$ if for any $w \in \Sigma_1^*$, running the TM on the input w results in the machine halting in the `halt` state and it has $f(w)$ as the word on the tape, with the head at the leftmost character. We also require $\boxdot \notin \Sigma_1$ or Σ_2. If the machine does not halt in the `halt` state, we say $f(w)$ is not defined.

Definition 3.1.2. A partial function $f : \Sigma_1^* \rightharpoonup \Sigma_2^*$ is COMPUTABLE if there exists a TM that computes it.

Question 1. Describe the basic working of a Turing machine. What does it mean for a function to be computable?

Definition 3.1.3. A LANGUAGE over an alphabet Σ is a subset $L \subseteq \Sigma^*$.

For language recognition we require a subset $\Sigma \subseteq \Gamma$ (the INPUT ALPHABET) and two distinguished states, `accept` and `reject`. A language-recognizing Turing machine accepts a word w if, when the machine is run on the input w, the computation halts in the `accept` state. Similarly, w is rejected if the machine halts in the `reject` state. We say

that a Turing machine M DECIDES or COMPUTES L if for any $w \in \Sigma^*$, $w \in L$ implies that M accepts w and $w \notin L$ implies that M rejects w.

If M correctly accepts words of the language, and does not accept words that are not part of the language, we say that M SEMIDECIDES, SEMICOMPUTES or RECOGNIZES L. A language is DECIDABLE or COMPUTABLE if there is a TM which decides it, and SEMIDECIDABLE, SEMICOMPUTABLE or COMPUTABLY INNUMERABLE if there is a TM which recognizes it. Clearly, every decidable language is also semidecidable.

Question 2. What is a decidable and what is a semidecidable language?

Given a k-tape machine (Γ, Q, δ), we can simulate it by a single-tape machine. We will encode the different tapes by separating them with a special symbol $| \in \tilde{\Gamma}$ and encoding the positions of the read-write heads with another special symbol $\Delta \in \tilde{\Gamma}$. When simulating a computational step of the multi-tape machine, we scan for transitions and implement them manually.

Question 3. How can you simulate a multi-tape Turing machine with a single-tape machine?

Proposition 3.1.4. *There are languages that are not semidecidable.*

Proof. There are only countably many non-equivalent Turing machines, but an uncountable number of languages on any alphabet. □

We can also give an example of a language that is semidecidable but not decidable. To construct it, we will encode a Turing machine $M = (\Gamma, Q, \delta)$ into a string. We encode it in $\langle M \rangle \in \Sigma_u^*$ for the alphabet

$$\Sigma_u = \{0, 1, -1, [,], \|, \cdot\}.$$

We encode every state $q \in Q$ as a word $\langle q \rangle \in \{0, 1, -1\}^l$, where $l \geq \log_3 |Q|$. We require that the encoding of the start state is 0^l, the encoding of the accepting state is 1^l, and the encoding of the rejecting state is $(-1)^l$. Every symbol $a \in \Gamma$ is encoded as $\langle a \rangle \in \{-1, 0, 1\}^m$ for $m \geq \log_3 |\Gamma|$. We require that the encoding of the blank symbol is 0^m.

Finally, to encode δ, consider an instruction $(q, a) \mapsto (q', b, d)$. This will be encoded as a word

$$[\langle q \rangle \cdot \langle a \rangle \parallel \langle q' \rangle \cdot \langle b \rangle \cdot d]$$

with $d \in \{0, 1, -1\}$. This encoding has length $2l + 2m + 7$, and allows us to encode the full Turing machine as the encoding of the start state, followed by a dot $\cdot$, followed by the encoding of the blank symbol, and then followed by the encodings of all transitions one after another. We can encode any word $w \in \Gamma^*$ as a sequence of characters, delimited by $\cdot$, so that we may define a language L_{accept} as follows: the language includes words of the form $\langle M \rangle \cdot \langle w \rangle$, where M is a single-tape Turing machine with tape alphabet $\Gamma \supseteq \Sigma_u$, and $w \in \Sigma_u^*$ is a word which M accepts.

Question 4. How is L_{accept} defined? Describe the universal encoding of a Turing machine.

Theorem 3.1.5. *The language L_{accept} is undecidable.*

Proof. Suppose that it is decidable, so there exists a Turing machine D which decides it. We define a new Turing machine N with input alphabet Σ_u. This machine reads its input string v and converts it to the string $v \cdot \langle v \rangle$. It then proceeds as D on this input, except it switches the accept and reject states.

Consider what N does when given an input of the form $v = \langle M \rangle$ for some Turing machine M. If D rejects $\langle M \rangle \cdot \langle \langle M \rangle \rangle$, then N terminates on the accepting state, and vice versa. Because D decides L_{accept}, we get from N:

- `accept` iff M does not accept $\langle M \rangle$, and
- `reject` iff M accepts $\langle M \rangle$.

Now run N on $\langle N \rangle$. We get `accept` iff N does not accept $\langle N \rangle$, and `reject` iff N accepts $\langle N \rangle$. This is a contradiction. □

Question 5. Show that L_{accept} is undecidable.

Using Σ_u, we can also construct the UNIVERSAL TURING MACHINE.

Theorem 3.1.6. *There exists a Turing machine U over the tape alphabet $\Sigma_u \cup \{\boxdot\}$ that exhibits the following behavior: If we run U on the string $\langle M \rangle \cdot \langle w \rangle$, then the resulting execution satisfies the following.*

- *It terminates if and only if M terminates on input w.*
- *It accepts if and only if M accepts w.*
- *It rejects if and only if M rejects w.*

The idea of the proof is to use a three-tape machine, putting the encoding of M on the first tape, the encoding of the state on the second, and the input on the third. Then simulate the execution of M.

Theorem 3.1.7. *The language L_{accept} is semidecidable.*

Proof. We construct a machine S that does the following. It first reads its input word $v \in \Sigma_u^*$ and checks whether v is of the form $\langle M \rangle \cdot \langle w \rangle$. If v is not of this form, then we reject immediately. Otherwise, we run the universal machine U on the input v and end in the end state of U. □

Question 6. Show that L_{accept} is semidecidable.

Proposition 3.1.8. *If $f : \Sigma_1^* \rightharpoonup \Sigma_2^*$ and $g : \Sigma_2^* \rightharpoonup \Sigma_3^*$ are computable, then so if $g \circ f$.*

Note that the composite of two partial functions is defined as

$$(g \circ f)(w) \simeq \begin{cases} g(f(w)) & f(w)\downarrow \\ \uparrow & f(w)\uparrow \end{cases}$$

Definition 3.1.9. A k-tape Turing machine COMPUTES $f : \Sigma_1^* \times \cdots \times \Sigma_k^* \rightharpoonup \Sigma^*$ if when we run the Turing machine on the configuration with a word on each tape, the machine terminates in the halt state if and only if for all $(w_1, \ldots, w_k)$ in the domain of f, it halts in the configuration with $f(w_1, \ldots, w_k)$ on the first tape and all other tapes blank.

Example. To define the computability for partial functions $f : \mathbb{N} \rightharpoonup \mathbb{N}$, we can represent numbers using words over $\Sigma_b = \{0, 1\}$ and the representation

$$\gamma_{\mathbb{N}}(w) = \sum_{i=0}^{|w|-1} 2^{|w|-i-1} w_i.$$

If we allow for leading zeros, every number is represented by an infinite number of words. Additionally, zero is also represented by the empty word ε. A Turing machine M computes $f : \mathbb{N} \rightharpoonup \mathbb{N}$ if it computes $g : \Sigma_b^* \rightharpoonup \Sigma_b^*$ such that for all words w for which $\gamma_{\mathbb{N}}(g(w)) \simeq f(\gamma_{\mathbb{N}}(w))$. We say that f is COMPUTABLE if there is a Turing machine which computes it.

We could also restrict our representation, for example requiring words to begin with a 1 (we also allow the empty word). Alternatively, we could use e.g. a unary representation.

Definition 3.1.10. A REPRESENTATION of a set X by words over an alphabet Σ is a surjective partial function $\gamma : \Sigma^* \rightharpoonup X$.

Remark. Only countable sets can be represented.

Definition 3.1.11. Given representations $\gamma_1 : \Sigma_1^* \rightharpoonup X_1$ and $\gamma_2 : \Sigma_2^* \rightharpoonup X_2$, a partial function $f : X_1 \rightharpoonup X_2$ is $(\gamma_1 \to \gamma_2)$-COMPUTABLE if there exists a computable partial function $g : \Sigma_1^* \rightharpoonup \Sigma_2^*$ such that for all words w in the domain of γ_1, if $g(w)$ is defined, then $\gamma_2(g(w)) \simeq f(\gamma_1(w))$.

Definition 3.1.12. Two representations γ_1 and γ_2 of the same set X are EQUIVALENT if the identity function id_X is $(\gamma_1 \to \gamma_2)$-computable and $(\gamma_2 \to \gamma_1)$-computable.

Given representations $(\gamma_i : \Sigma_i^* \rightharpoonup X_i)_{i=1,\ldots,k}$, we construct a product representation $\gamma : \Sigma^* \rightharpoonup X_1 \times \cdots \times X_k$, where

$$\Sigma = (\Sigma_1 \cup \Sigma_2 \cup \ldots \cup \Sigma_k) \amalg \{,\}$$

and

$$\gamma(w_1, \ldots, w_k) = (\gamma_1(w_1), \ldots, \gamma_k(w_k)).$$

Definition 3.1.13. A partial function $f : \mathbb{N} \rightharpoonup \mathbb{N}$ is COMPUTABLE if it is $(\gamma_{\mathbb{N}} \times \cdots \times \gamma_{\mathbb{N}} \to \gamma_{\mathbb{N}})$-computable.

Definition 3.1.14. Given a representation $\gamma : \Sigma^* \rightharpoonup X$, a subset $A \subseteq X$ is γ-DECIDABLE if there exists a Turing machine M such that for all w in the domain of γ, if $\gamma(w) \in A$, then M accepts w, and if $\gamma(w) \notin A$, then M rejects w. The same subset is γ-SEMIDECIDABLE if there exists a Turing machine M such that for all w in the domain of γ, M accepts w if and only if $\gamma(w) \in A$.

Proposition 3.1.15. *Given a representation γ of X and $A \subseteq X$, A is γ-decidable if and only if its characteristic function is $(\gamma \to \gamma_b)$-computable, and A is γ-semidecidable if and only if its partial characteristic function is $(\gamma \to \gamma_b)$-computable.*

Remark. Above, γ_b is the representation of $\{0, 1\}$ which maps $0 \mapsto 0$ and $1 \mapsto 1$.

3.1.1 Models of computation

We have many models of computation:

- partial recursive functions,
- λ-calculus,
- Turing machines,
- string rewriting systems,
- unrestricted grammars,
- cellular automata,
- nondeterministic Turing machines,
- quantum Turing machines,
- hypercomputation,
- finite/pushdown automata

Everything on this list, up until hypercomputation, is computationally equivalent. We say that a model is TURING COMPLETE if it can simulate any Turing machine.

There is also the Church-Turing thesis, which states that any Turing complete physically realizable computational model is equivalent to a Turing machine. Informally, the thesis states that the notion of what an algorithm is is equivalent to a Turing machine.

3.2 Computability of natural numbers

Definition 3.2.1. The COMPUTABLE PARTIAL FUNCTIONS is the set

$$\{f : \mathbb{N}^k \rightharpoonup \mathbb{N} \mid k \geq 0, f \text{ is computable}\}.$$

Definition 3.2.2. The PRIMITIVE RECURSIVE FUNCTIONS are the smallest collection $\mathcal{F} \subseteq \{\mathbb{N}^k \rightharpoonup \mathbb{N} \mid k \geq 0\}$ of partial functions that satisfies the following properties:

- $\mathcal{F}$ contains the zero function $Z : \{()\} \to \mathbb{N}$, which maps $Z() = 0$, the successor function $S : \mathbb{N} \to \mathbb{N}$, which maps $x \mapsto x + 1$ and the projection functions: for any $k \geq 1$ and $1 \leq i \leq k$, $U_i^k : \mathbb{N}^k \to \mathbb{N}$ is defined as

$$U_i^k(x_1, \ldots, x_k) = x_i.$$

- $\mathcal{F}$ is closed under composition: if $f : \mathbb{N}^k \rightharpoonup \mathbb{N}$ and $g_1, \ldots, g_k : \mathbb{N}^l \rightharpoonup \mathbb{N}$ are in $\mathcal{F}$, then $f \circ (g_1, \ldots, g_k)$ is in $\mathcal{F}$.
- Primitive recursion: If $f : \mathbb{N}^k \rightharpoonup \mathbb{N}$ and $g : \mathbb{N}^{k+2} \rightharpoonup \mathbb{N}$ are both in $\mathcal{F}$, then so is $R_{fg} : \mathbb{N}^{k+1} \rightharpoonup \mathbb{N}$, defined with $R_{rg}(x_1, \ldots, x_k, 0) \simeq f(x_1, \ldots, x_k)$ and $R_{fg}(x_1, \ldots, x_k, x + 1) \simeq g(x_1, \ldots, x_k, x, R_{fg}(x_1, \ldots, x_k, x))$.

Remark. Since all basic functions are total, every function in $\mathcal{F}$ is total.

Remark. Not every total computable function is primitive recursive. We can show for example that the Ackermann function grows faster than any primitive recursive function.

Definition 3.2.3. The PARTIAL RECURSIVE FUNCTIONS are the smallest collection of partial functions $\mathcal{F} \subseteq \{\mathbb{N}^k \rightharpoonup \mathbb{N}\}_{k \geq 0}$, which satisfies the axioms of partial recursive functions, with an additional one:

- Minimization: If $f : \mathbb{N}^{k+1} \rightharpoonup \mathbb{N}$ is in $\mathcal{F}$, then so is $\mu f : \mathbb{N}^k \rightharpoonup \mathbb{N}$, with $\mu f(x_1, \ldots, x_k)$ equal to the smallest number $n \in \mathbb{N}$ such that $f(x_1, \ldots, x_k, n) = 0$ if it exists and $f(x_1, \ldots, x_k, m)$ is defined for all $m < n$, and $\mu f(x_1, \ldots, x_k)$ undefined otherwise.

Proposition 3.2.4. *A partial function $f : \mathbb{N}^k \rightharpoonup \mathbb{N}$ is computable if and only if there exists a $(k+1)$-tape Turing machine such that for all $x_1, \ldots, x_k \in \mathbb{N}$ and binary words $w_1, \ldots, w_k$ of $x_1, \ldots, x_k$, if we run the Turing machine with $w_1, \ldots, w_k$ on the first k tapes, then it halts if and only if $f(x_1, \ldots, x_k)$ is defined and it halts with the representation of $f(x_1, \ldots, x_k)$ on the last tape, and $w_1, \ldots, w_k$ on the first k tapes.*

Theorem 3.2.5. *The partial recursive functions coincide with the computable partial functions.*

Proof. A partial recursive function is computable: We will show that the family of computable functions satisfies the properties of partial computable functions. Since the partial recursive functions are the smallest such family, every partial recursive function is computable.

Clearly, computable partial functions satisfy composition and include the basic functions. Let's show they are closed under primitive recursion. Suppose we have a $(k+1)$-tape Turing machine M_f computing f and a $(k+3)$-tape Turing machine M_g computing g. We will find a $(k+4)$-tape Turing machine computing R_{fg}, which can then be compressed. On the first $k+1$ tapes, put the arguments to R_{fg}. On the $(k+2)$-th tape, put a counter.

On the next tape, put the result of R_{fg} on the current iteration, and finally, on the last tape, put the result being computed. The Turing machine then proceeds as follows:

1. initialize tape $k+2$ to 0
2. use M_f to compute $f(x_1, \ldots, x_k)$ and write the result on tape $k+4$
3. if the numbers on tapes $k+1$ and $k+2$ are equal, halt
4. copy tape $k+4$ to tape $k+3$
5. apply M_g on tapes $1, \ldots, k, k+2, k+3$ and write the result on tape $k+4$
6. increment tape $k+2$
7. go to step 3

We may similarly construct a machine which performs minimization, which finishes this inclusion.

Before starting the proof of the other inclusion, consider the following. We can encode $\mathbb{N} \times \mathbb{N}$ using $\mathbb{N}$ with the bijection $p(x, y) = \frac{1}{2}(x+y)(x+y+1)+x$, which is also primitive recursive. The functions $q_1, q_2 : \mathbb{N} \to \mathbb{N}$ which reverse p (such that $q_1(p(x, y)) = x$ and $q_2(p(x, y)) = y$) are also primitive recursive.

We can also encode finite sequences with $\lceil \cdot \rceil : \mathbb{N}^* \to \mathbb{N}$, defined as $\lceil n_0, \ldots, n_{k-1} \rceil = 2^{n_0} + 2^{n_0+n_1+1} + \ldots + 2^{n_0+\ldots+n_{k-1}+k-1}$, so that numbers are encoded in a binary sequence with the number of zeros between a pair of ones denoting the sequence. This is clearly a bijection. Additionally, the functions $\sigma : \mathbb{N}^2 \to \mathbb{N}$ mapping

$$\sigma(\lceil w \rceil, i) = \begin{cases} 0 & i \geq |w| \\ w_i + 1 & i < |w| \end{cases}$$

and $l : \mathbb{N} \to \mathbb{N}$ mapping

$$l(\lceil w \rceil) = |w|$$

are primitive recursive.

Now we can prove the other inclusion. Suppose $f : \mathbb{N}^k \rightharpoonup \mathbb{N}$ is computable and that it is computed by a Turing machine M. We assume M is a single-tape machine computing f via representations. Suppose M has a tape alphabet $\Gamma \supseteq \{\boxdot, 0, 1, ,\}$ and state set $Q \supseteq \{\text{start}, \text{halt}\}$. Choose injective functions $r : \Gamma \to \{\text{odd numbers}\}$ and $s : Q \to \{\text{even numbers}\}$. Now suppose we are in a configuration C with a finitely many not necessarily blank symbols $a_0, \ldots, a_{k-1}$, the tape head at a_i, and the current state equal to q. Define an encoding

$$\lceil C \rceil = \lceil r(a_0) \ldots r(a_{i-1}) s(q) r(a_i) \ldots r(a_{k-1}) \rceil.$$

The following functions are primitive recursive:

- step : $\mathbb{N} \to \mathbb{N}$, mapping $\lceil C \rceil$ to $\lceil C' \rceil$, which is the configuration we obtain by taking one step of M on configuration C. If the input of the function is not a valid configuration, it returns 0.
- run : $\mathbb{N}^2 \to \mathbb{N}$, mapping (n, x) to $\text{step}^n(x)$.
- extract : $\mathbb{N} \to \mathbb{N}$, mapping $\lceil s(\text{halt})r(w_0)\dots r(w_k) \rceil$ to n if w is a binary representation of n, and any other input to 0.
- halt? : $\mathbb{N} \to \mathbb{N}$ mapping $\lceil s(\text{halt})r(w_0)\dots r(w_{k+1}) \rceil$ to 0 and all other inputs to 1.
- init : $\mathbb{N}^k \to \mathbb{N}$, mapping $(x_1, \dots, x_k)$ to $\lceil s(\text{start})y_1 \dots y_m \rceil$, where y_i is equal to the result of r on the i-th character of the sequence $(\text{bin}(x_1), \dots, \text{bin}(x_k))$.

Then, f is partial recursive because

$$f(x_1, \dots, x_n) \simeq \text{extract}(\text{run}(\mu(n \mapsto \text{halt?}(\text{run}(n, \text{init}(x_1, \dots, x_k)))), \text{init}(x_1, \dots, x_k)))$$

□

We will consider number representations $\gamma : \mathbb{N} \rightharpoonup X$. There is a canonical representation of $\mathbb{N}$, $n \mapsto n$. Given two representations γ_X of X and γ_Y of Y, we can define the product representation

$$\gamma_X \times \gamma_Y(n) = (\gamma_X(n_1), \gamma_Y(n_2)),$$

where $n = p(n_1, n_2)$ for the projection function p.

Given representations ρ_X of X and ρ_Y of Y, we say that a function $f : X \to Y$ is $(\rho_X \to \rho_Y)$-computable if there is a computable partial function which computes between the representations.

As there are only a countable many partial recursive functions of any arity n, we can enumerate them. We will use the label ϕ_i^n to mean the i-th partial recursive function $\mathbb{N}^n \rightharpoonup \mathbb{N}$ in this enumeration. We require that all functions on the list are computable, and that every partial recursive function occurs on the list, but not necessarily with a unique index. The following definition satisfies these properties. We say that $\phi_e^n(x_1, \dots, x_n)$ is equal to y if there is a Turing machine M with $e = \lceil M \rceil$, which, when run on the input $\text{bin}(x_1), \dots, \text{bin}(x_n)$, halts with $\text{bin}(y)$ on the tape. Otherwise, we say that the expression is undefined.

Note that the function $\mathbb{N} \to \text{Comp}(\mathbb{N}^n \to \mathbb{N})$, defined with the expression

$$e \mapsto \phi_e^n$$

is a number representation of the set of all computable functions.

Proposition 3.2.6. *The function $h : \mathbb{N} \to \mathbb{N}$ defined below is not computable:*

$$h(e) = \begin{cases} \phi_e(e) + 1, & \phi_e(e) \downarrow, \\ 0, & \textit{otherwise.} \end{cases}$$

Proof. Suppose that h is computable, so there is an index e such that $h = \phi_e$. Then $\phi_e(e) = h(e) = \phi_e(e) + 1$, which is a contradiction. □

Theorem 3.2.7 (the universal function)**.** *For any $n \geq 1$, the $(n+1)$-arity function*

$$\phi_u^n(e, x_1, \ldots, x_n) \simeq \phi_e^n(x_1, \ldots, x_n)$$

exists and is computable.

Proposition 3.2.8. *The unary total function below is not computable.*

$$g(e) = \begin{cases} 1, & \text{if } \phi_e \text{ is a total function,} \\ 0, & \text{otherwise.} \end{cases}$$

Proof. Suppose that g is computable and consider

$$h(x) = \begin{cases} \phi_x(x) + 1, & \text{if } \phi_x \text{ is total,} \\ 0, & \text{otherwise.} \end{cases}$$

Since g is computable, so is h (using the universal function). Then $h = \phi_e$ for some e, but $\phi_e(e) + 1 = h(e) = \phi_e(e)$, which is a contradiction. □

Theorem 3.2.9 (S-M-N)**.** *For every $n > m > 0$ there exists an $(m+1)$-ary primitive recursive function $S_n^m : \mathbb{N}^{m+1} \to \mathbb{N}$ such that for all $e, x_1, \ldots, x_n$:*

$$\phi^{n-m}_{S_n^m(e,x_1,\ldots,x_n)}(x_{m+1}, \ldots, x_n) \simeq \phi_e^n(x_1, \ldots, x_n).$$

Theorem 3.2.10 (Kleene's normal form)**.** *There exists a primitive recursive function $U : \mathbb{N} \to \mathbb{N}$ and for each $n \geq 1$ an $(n+2)$-ary primitive recursive function $T^n : \mathbb{N}^{n+2} \to \mathbb{N}$ such that for all $e, x_1, \ldots, x_n$:*

$$\phi_e^n(x_1, \ldots, x_n) \simeq U((\mu T^n)(e, x_1, \ldots, x_n)).$$

3.3 Computable and computably enunmerable sets

Definition 3.3.1. A subset $A \subseteq \mathbb{N}$ is COMPUTABLE if the characteristic function $\chi_A : \mathbb{N} \to \mathbb{N}$ is computable. It is COMPUTABLY ENUMERABLE if the partial characteristic function $\chi_A^p : \mathbb{N} \to \mathbb{N}$ is computable as a partial function.

Definition 3.3.2. KLEENE'S HALTING SET is the set

$$K = \{e \in \mathbb{N} \mid \phi_e(e) \downarrow\}.$$

Proposition 3.3.3. *Kleene's halting set K is not computable.*

Proof. Suppose it is. Then so is the partial function

$$h(e) \simeq \begin{cases} \uparrow, & \chi_K(e) = 1, \\ 0, & \chi_K(e) = 0. \end{cases}$$

Therefore there exists an index e such that $h = \phi_e$. Then $\phi_e(e) \downarrow \Leftrightarrow e \in K \Leftrightarrow h(e) \uparrow \Leftrightarrow \phi_e(e) \uparrow$. This is a contradiction. □

Proposition 3.3.4. *A set $A \subseteq \mathbb{N}$ is computably enumerable if and only if it is the domain of some computable partial function.*

Proof. If a set A is computable enumerable, it is the domain of χ_A^p. Suppose that $A = \mathrm{dom}(f)$ for some computable $f : \mathbb{N} \to \mathbb{N}$. We can compute χ_A^p by

$$\chi_A^p(n) \simeq \begin{cases} 1, & f(n) \downarrow, \\ \uparrow & \text{otherwise.} \end{cases} \quad \square$$

Proposition 3.3.5. *The halting set K is computably enumerable.*

Proof. It is the domain of the computable partial function $e \mapsto \phi_e(e) = \psi_u(e, e)$. □

We can enumerate the computably enumerable sets by

$$W_e := \mathrm{dom}(\phi_e).$$

Lemma 3.3.6. *Suppose that $A \subseteq \mathbb{N}$ is a set for which there exists a total computable function $t : \mathbb{N}^2 \to \mathbb{N}$ such that $x \in A$ if and only if there exists $z \in \mathbb{N}$ with $t(x, z) = 0$. Then A is computably enumerable. The reverse also holds.*

Proof. Suppose that A is computably enumerable, so by Kleene's normal form theorem, $x \in A$ if and only if $T(e, x, z) = 0$ for some z, and for $A = W_e$. We can take $t(x, z) = T(e, x, z)$.

As for the reverse, given a computable t, the partial function $\mu t : \mathbb{N} \to \mathbb{N}$ has A as its domain, so A is computably enumerable. □

Theorem 3.3.7. *A set $A \subseteq \mathbb{N}$ is computable if and only if both A and $\overline{A}$ are computably enumerable.*

Proof. The left-to-right implication is trivial. Suppose both A and $\overline{A}$ are computably enumerable. By the lemma, there exist total computable $s, s' : \mathbb{N} \to \mathbb{N}$ such that $x \in A$ if and only if there exists $z \in \mathbb{N}$ for which $s(x, z) = 0$, and similarly $x \notin A$ if and only if there is a number z such that $s'(x, z) = 0$.

Then the following function is both computable and total:

$$g := \mu((x, z) \mapsto \min(s(x, z), s'(x, z))).$$

So

$$\chi_A(x) = \begin{cases} 1, & s(x, g(x)) = 0, \\ 0, & \text{otherwise}, \end{cases}$$

is computable. □

Corollary 3.3.8. *The set $\overline{K}$ is not computably enumerable.*

Theorem 3.3.9. *The following are equivalent for a set $A \subseteq \mathbb{N}$:*

- *A is computably enumerable,*
- *$A = \varnothing$ or A is the range of a total computable function,*
- *A is the range of a computable partial function.*

Proof. One to two: Suppose $A \neq \varnothing$, and select any $a_0 \in A$. By the above lemma, there exists a computable $s : \mathbb{N}^2 \to \mathbb{N}$ satisfying

$$y \in A \Leftrightarrow \exists z.s(y, z) = 0.$$

Using the pairing function $p : \mathbb{N}^2 \to \mathbb{N}$, we have that A is the range of the total computable

$$x \mapsto \begin{cases} y, & s(y, z) = 0, \\ a_0 & \text{otherwise}. \end{cases}$$

We used $(y, z) = p^{-1}(x)$ above.

Two to three is trivial. Three to one: Suppose A is the range of ϕ_e. Let T and U be as in Kleene's normal form theorem. Then the total function $s : \mathbb{N}^2 \to \mathbb{N}$ is computable:

$$s(w, y) = \begin{cases} 0 & T(e, x, z) = 0 \wedge w = U(z), \\ 1 & \text{otherwise}, \end{cases}$$

where $y = p(x, z)$. Then $w \in A$ if and only if there exists a y such that $s(w, y) = 0$. □

Theorem 3.3.10. *The following are equivalent for a set $A \subseteq \mathbb{N}$:*

- *A is computable,*
- *$A = \varnothing$ or A is the range of an increasing total computable function,*
- *A is finite or A is the range of a strictly increasing computable total function.*

Proof. We will only prove the equivalence of the first and third statement. One to three: Suppose A is computable and infinite. Then A is the image of the function which maps n to the n-th smallest element of A, by searching through $m = 0, 1, 2, \ldots$ using χ_A.

Three to one: Suppose A is the range of $f : \mathbb{N} \to \mathbb{N}$, which is strictly increasing. Then we can compute $\chi_A(n)$ by checking whether n is in the image of $f(x)$ for $x = 0, 1, 2, \ldots$, until we find either n or a number larger than it. □

Corollary 3.3.11. *Every infinite computably enumerable set has an infinite computable subset.*

Proof. Let A be infinite and computably enumerable. Then A is the image of a computable total function f. Define $g(0) = f(0)$ and $g(n+1) = f(m)$ where m is the smallest number such that $f(m) > g(n)$. □

We define E_e as the domain of ϕ_e, and

$$\mathcal{C} = \{f : \mathbb{N} \rightharpoonup \mathbb{N} \mid f \text{ is computable}\}.$$

Definition 3.3.12. A subset $B \subseteq \mathcal{C}$ is DECIDABLE if

$$I_B = \{e \in \mathbb{N} \mid \phi_e \in B\}$$

is a computable set. Also, B is SEMIDECIDABLE if I_B is computably enumerable.

Lemma 3.3.13 (Reduction lemma)**.** *If A is reducible to B and B is computable, then so is A. If A is reducible to B and B is computably enumerable, then so it A.*

Theorem 3.3.14 (Rice)**.** *A set $B \subseteq \mathcal{C}$ is decidable if and only if $B = \varnothing$ or $B = \mathcal{C}$.*

Proof. The right-to-left implication is trivial. Left-to-right: Assume B is neither $\varnothing$ nor $\mathcal{C}$. We'll show that I_B is not computable. Without loss of generality, we assume the everywhere undefined partial function $f_\varnothing$ is in $\mathcal{C} \setminus B$.

Choose some function $g \in B$. Define

$$f(x, y) \simeq \begin{cases} g(y), & x \in K \\ \uparrow, & \text{otherwise} \end{cases}$$

Clearly f is computable. By the s-m-n theorem, there is a total computable $S : \mathbb{N} \to \mathbb{N}$ such that $\phi_{s(x)}(y) \simeq f(x, y)$. Then for $x \in K$, $\phi_{s(x)}(y) \simeq g(y)$ and for $x \notin K$, $\phi_{s(x)}(y) \downarrow$. Since $g \in B$, we have $s(x) \in I_B$ for any $x \in K$. As $f_\varnothing$ is in $\mathcal{C} \setminus B$, we have $\phi_{s(x)} \notin I_B$ for any $x \notin K$. So s is a reduction of K to I_B. But K is not computable, so I_B isn't either. □

Definition 3.3.15. For partial functions $\mathbb{N} \rightharpoonup \mathbb{N}$, $f' \subseteq f$ if the graph of f' is a subset of the graph of f. We say that f is FINITE if its domain is finite.

Theorem 3.3.16 (Rice-Shapiro)**.** *If $B \subseteq \mathcal{C}$ is semidecidable, then for all $f \in \mathcal{C}$, then $f \in B$ if and only if there exists $f' \in B$ such that f' is finite and $f' \subseteq f$.*

Proof. We prove that if either implication of the equivalence fails, then B is not semidecidable.

Suppose that the left-to-right implication fails, so there exists a function $f \in B$ such that all finite subfunctions of f aren't in B. Note that f is necessarily infinite. Because K is computably enumerable, there is some total computable $t : \mathbb{N}^2 \to \mathbb{N}$ such that $x \in K$ if and only if there exists a number z for which $t(x, z) = 0$. Define

$$g(x, z) \simeq \begin{cases} f(z), & \text{if } t(x, y) \neq 0 \text{ for all } 0 \leq y \leq z \\ \uparrow, & \text{otherwise} \end{cases}$$

By the s-m-n theorem, there is a total computable function $s : \mathbb{N} \to \mathbb{N}$ such that $\phi_{s(x)}(z) \simeq g(x, z)$. If $x \in K$, then $\phi_{s(x)}$ is finite and a subfunction of f, so by assumption $s(x) \notin I_B$ (since $\phi_{s(x)} \notin B$). If $x \notin K$, then $\phi_{s(x)} = f$, so $s(x) \in I_B$ because $f \in B$. Then s is a reduction of $\overline{K}$ to I_B. By the reduction lemma, I_B is not computably enumerable (as $\overline{K}$ isn't).

Now suppose that the right-to-left implication fails. Let f' be a finite subfunction of f such that $f' \in B$ and $f \notin B$. Define a computable partial function

$$g(x, z) \simeq \begin{cases} f(z), & \text{if } z \in \operatorname{dom}(f') \text{ or } x \in K \\ \uparrow, & \text{otherwise} \end{cases}$$

By the s-m-n theorem, there is a total computable $s : \mathbb{N} \to \mathbb{N}$ such that $\phi_{s(x)}(z) = g(x, z)$. Since $f' \subseteq f$, we have:

- if $x \in K$, then $\phi_{s(x)} = f$, but $f \notin B$, so $s(x) \notin I_B$,
- if $x \notin K$, then $\phi_{s(x)} = f'$, but since $f' \in B$, so $s(x) \in I_B$.

So s is a reduction of $\overline{K}$ to I_B, and I_B can't be computably enumerable. □

3.3.1 Varieties of non-computable sets

Observe that the following are equivalent for a set $A \subseteq \mathbb{N}$:

- A is not computably enumerable,
- for any $W_e \subseteq A$ there exists some $n \in A \setminus W_e$.

A set $A \subseteq \mathbb{N}$ is PRODUCTIVE if there exists a computable total function $g : \mathbb{N} \to \mathbb{N}$ such that for all $e \in \mathbb{N}$, if $W_e \subseteq A$, then $g(e) \in A \setminus W_e$. We call g an OUTSIDER FINDER.

Proposition 3.3.17. *Every productive set is not computably enumerable.*

Proposition 3.3.18. *The set $\overline{K}$ is productive.*

Proof. We show that $g = \mathrm{id}_{\mathbb{N}}$ is an outsider finder for $\overline{K}$. Suppose that $W_e \subseteq \overline{K}$. Also suppose that $g(e) = e \notin \overline{K}$. Then $e \in K$, so $\phi_e(e) \downarrow$, which means $e \in W_e \subseteq \overline{K}$, which is a contradiction, meaning that $e \in \overline{K}$. Now if $e \in \overline{K}$, then $e \in \{f \mid f \notin W_f\}$, so $e \notin W_e$. All this implies $g(e) \in \overline{K} \setminus W_e$. □

Lemma 3.3.19. *If A reduces to B and A is productive, then so is B.*

Proof. Let g be an outsider finder of A. We claim that $f \circ g \circ h$ is an outsider finder for B, where h is defined in the following way. Consider the map $(e, x) \mapsto \phi_e(f(x))$. By the s-m-n theorem, there is a computable $h : \mathbb{N} \to \mathbb{N}$ for which $\phi_{h(e)}(x) \simeq \phi_e(f(x))$. To prove $f \circ g \circ h$ is an outsider finder for B, suppose $W_e \subseteq B$. Then $f^{-1}(W_e) \subseteq A$, but

$$f^{-1}(W_e) = f^{-1}(\operatorname{dom} \phi_e) = \operatorname{dom} \phi_e \circ f = \operatorname{dom} \phi_{h(e)} = W_{h(e)}.$$

So $f(g(h(e))) \subseteq B \setminus W_e$ because $g(h(e)) \in A \setminus W_{h(e)}$. □

Definition 3.3.20. A set A is CREATIVE if it is computably enumerable and its complement is productive.

Theorem 3.3.21. *If $\varnothing \subsetneq B \subsetneq \mathcal{C}$ and the everywhere undefined function $f_\varnothing \in B$, then I_B is productive.*

Proof. In the proof of Rice's theorem, we reduced K to $\overline{I_B}$, so $\overline{K}$ to I_B. Because $\overline{K}$ is productive, so is I_B. □

Theorem 3.3.22. *If $\varnothing \subsetneq B \subsetneq \mathcal{C}$ is such that I_B is computably enumerable, then is it creative.*

Theorem 3.3.23. *Every productive set has an infinite computably enumerable subset.*

Proof. Suppose that A is productive with outsider finder g. We will define a function $f : \mathbb{N} \to \mathbb{N}$ which satisfies $f(0) = e_0$ with $W_{e_0} = \varnothing$, and $f(n+1) = e_{n+1}$ with $W_{e_{n+1}} = W_{e_n} \cup \{g(e_n)\}$. By the s-m-n theorem, there exists a computable total function $h : \mathbb{N} \to \mathbb{N}$ such that

$$\phi_{h(x)}(y) = \simeq \begin{cases} 0 & y \in W_x \vee y = g(x), \\ \uparrow & \text{otherwise.} \end{cases}$$

So $W_{h(e)} = W_e \cup \{g(e)\}$, and we can find an f such that $W_{f(0)} = \varnothing$ and $f(n+1) = h(f(n))$.

We then have $g \circ f$, a total computable function with an image that is an infinite subset of A. Then that image is a computably enumerable subset of A. □

Definition 3.3.24. A set $A \subseteq \mathbb{N}$ is IMMUNE if it is infinite and it has no infinite computably enumerable subset.

Remark. If A is immune, it is not computably enumerable, and not productive.

Definition 3.3.25. A set $A \subseteq \mathbb{N}$ is SIMPLE if it is computably enumerable and its complement is immune.

Theorem 3.3.26 (Post)**.** *There exists a simple set.*

Proof. Consider a partial function $f : \mathbb{N} \to \mathbb{N}$, defined as follows: $f(e) = \phi_e(z)$ if z is the smallest number such that $\phi_e(x) \downarrow$ for any $x \leq z$ and $\phi_e(z) \geq 2e$, and $f(e)$ is undefined if no such z exists. This is clearly computable, so its image is computably enumerable. Define $A = \operatorname{im} f$. We will prove that A is simple.

When $f(e) = n$, we have $n \geq 2e$, so the numbers $\{0, 1, \ldots, 2m-1\}$ can only appear as values $f(e)$ when $e < m$. So for every $m \geq 0$, at least m distinct numbers from that set belong to $\overline{A}$, meaning that $\overline{A}$ is infinite.

Let B be an infinite computably enumerable set. We know that B must be the image of some total ϕ_e. Then $f(e) \downarrow$, since because B is infinite, it must contain some number larger than $2e$. So indeed $B \nsubseteq A$. □

3.4 Computation with continuous data

A type 2 Turing machine (T2M for short) with k input tapes, n working tapes and one output tape is a Turing machine with $k + n + 1$ tapes where:

- Input tapes start with no blank symbols, are only infinite in ore direction, and their heads are read-only and only move to the right.
- Working tapes start blank on all but a finite number of squares. They are infinite in both directions, and have regular heads.
- The output tape is initially blank and has a write-only head that can only move right.

Formally, a T2M is specified by

- the tape alphabet Γ with $\boxdot \in \Gamma$,
- an input/output alphabet $\Sigma \subseteq \Gamma \setminus \{\boxdot\}$,
- a finite set Q of control states,
- the transition function

$$\delta : Q \times \Sigma^k \times \Gamma^n \rightharpoonup Q \times \{0,1\}^k \times \Gamma^n \times \{-1, 0, 1\}^n \times \{\Sigma \cup \boxdot\}.$$

We say that a T2M COMPUTES an infinite word $p \in \Sigma^\omega$ given input $(p^1, \ldots, p^k) \in (\Sigma^\omega)^k$ if when we run the machine on input tapes containing $p^1, \ldots, p^k$, it outputs p on the output tape. An ω-word is COMPUTABLE if it is computed by some T2M with not input tapes.

A T2M M COMPUTES a partial function $f : (\Sigma^\omega)^k \rightharpoonup \Sigma^\omega$ if:

- for any $(p^1, \ldots, p^k) \in \operatorname{dom} f$, M computes $f(p^1, \ldots, p^k)$ given input $(p^1, \ldots, p^k)$,
- if M is given $(p^1, \ldots, p^k)$ as input, it will compute some $p \in \Sigma^\omega$ only if $(p^1, \ldots, p^k) \in \operatorname{dom} f$.

For any recognition machines for an ω-language, we assume distinguished halting states `accept` and `reject`. A single input tape T2M ACCEPTS $p \in \Sigma^\omega$ if, when run on input p, it halts in the accepting state. Similarly, it rejects p, if it halts in the rejecting state.

Given a word $p \in \Sigma^\omega$, we can define $\operatorname{Prefix}(p) \subseteq \Sigma^*$ as the set of all prefixes of p.

Theorem 3.4.1. *The following are equivalent:*

- *p is computable via a T2M,*
- $\operatorname{Prefix}(p)$ *is decidable,*
- $\operatorname{Prefix}(p)$ *is semidecidable.*

Proof. 1 to 2: Suppose that there is a T2M M computing p. To decide whether w is in $\operatorname{Prefix}(p)$ for some word w, we may run M until it produces $|w|$ characters, then compare that result to w.

2 to 3: Trivial.

3 to 1: Suppose S is an ordinary TM that semidecides $\operatorname{Prefix}(p)$. We build a T2M that computes p as follows. Suppose we have already output n symbols of p. To find the next symbol, for each of the m symbols $b_1, \ldots, b_m \in \Sigma$, we run S (in parallel) on the input $p_0 \ldots p_{n-1} b_i$. Exactly one of these will halt in the accepting state, so when it does, output that symbol. □

3.4.1 Topological aspects of computing with ω-words

Theorem 3.4.2. *If $L \subseteq \Sigma^\omega$ is semidecidable, then for any $p \in L$, there exists $n \geq 0$ such that for any infinite word $q \in \Sigma^\omega$, if $q|_n = p|_n$.*

Proof. Let M be a T2M that semidecides L. Consider any $p \in L$, and let n be 1 plus the position of the read head when M enters the accept state if run on p. Note that the read head can only move right, so M could only access the first n characters during its execution. If we give it another ω-word with the same n-prefix, it will take the same actions and accept. □

Theorem 3.4.3. *If a partial function $f : \Sigma^\omega \rightharpoonup \Sigma^\omega$ is computable, then for every $p \in \operatorname{dom} f$ and for every $n \geq 0$ there exists an $m \geq 0$ such that for all $q \in \operatorname{dom} f$, if $q|_m = p|_m$, then $f(q)|_n = f(p)|_n$.*

Proof. Let M be a T2M that computes f. Consider any $p \in \operatorname{dom} f$ and $n \geq 0$. Let m be 1 plus the position of the input head at the time M writes the n-th symbol to the output tape.

Now consider any $q \in \operatorname{dom} f$ which agrees with p on the first m characters. Then the execution of M on q follows the same steps as on p, so it produces the same first n characters of output. □

We introduce a topology on Σ^ω, which is just the infinite product topology of Σ (which is discrete). This topology is metrizable for the metric

$$d(p, q) = 2^{-i},$$

where i is the smallest number such that $p_i \neq q_i$. We of course take $d(p, p) = 0$. Also, for a word $p \in \Sigma^\omega$ and number $n \geq 0$, define the CYLINDER SET $\langle p|_n \rangle$, where for a finite word w,

$$\langle w \rangle = \{q \in \Sigma^\omega \mid q|_{|w|} = w\}.$$

Note that the collection of cylinder sets is a countable basis for Σ^ω.

Remark. Theorem 3.4.2 states: A semidecidable language is an open set.

Remark. Theorem 3.4.3 states: Computable functions are continuous with respect to the subspace topology on $\operatorname{dom} f$.

Proposition 3.4.4. *If $L \subseteq \Sigma^\omega$ is decidable, it is clopen.*

Proof. The complement is semidecidable. □

Theorem 3.4.5. *An ω-language L is decidable if and only if it is clopen.*

Definition 3.4.6. A subset $Z \subseteq \Sigma^\omega$ is G_δ if it is a countable intersection of open sets.

Theorem 3.4.7. *If a partial function $f : \Sigma^\omega \rightharpoonup \Sigma^\omega$ is computable, then its domain of definition is a G_δ-subset of Σ^ω.*

Proof. Suppose that M is a T2M which computes f. For every $n \geq 0$, define

$$D_n = \{p \in \Sigma^\omega \mid M \text{ produces} \geq n \text{ output characters when run on } p\}.$$

Note that $\operatorname{dom} f \subseteq D_n$ and that D_n is semidecidable (and hence open). Clearly $\operatorname{dom} f$ is the intersection of all D_n. □

Theorem 3.4.8. *The topological space Σ^ω is compact.*

Proof. Tychonoff. □

Corollary 3.4.9. *The compact subsets of Σ^ω are exactly the closed sets.*

Lemma 3.4.10. *Every clopen set is decidable.*

Proof. Let L be a clopen set. Since the cylinder sets form a basis and L is open, we have

$$L = \bigcup \{\langle w \rangle \mid w \in \Sigma^*, \langle w \rangle \subseteq L\}.$$

So this family of cylinders is an open cover of L. Because L is closed, it is compact, so there is a finite subcover

$$L = \langle w_1 \rangle \cup \ldots \cup \langle w_k \rangle .$$

We can decide L by checking whether the prefix of a word is equal to any of $w_1, \ldots, w_k$. □

3.4.2 Computing with real numbers

We can represent real numbers as ω-words via infinite decimal expansions (or representations in other bases), so with the alphabet $\{0, 1, \ldots, 9, -, .\}$, with at most one $-$ at the start, and exactly one decimal point. The problem is that this is a poor representation, as many useful algorithms cannot be written with it.

Definition 3.4.11. A TYPE 2 REPRESENTATION of a set X is a surjective partial function $\gamma : \Sigma^\omega \rightharpoonup X$. We say that p is a NAME for an element $x \in X$ if $\gamma(p) = x$, and that x is COMPUTABLE if it has a computable name.

We now introduce the Cauchy representation of $\mathbb{R}$. First, give a type 1 representation of the dyadic rationals $\mathbb{Q}_d$, which we represent by a finite word $\pm d_{m-1} \ldots d_0 . d_{-1} \ldots d_{-n}$ with every $d_i \in \{0, 1\}$ and $m, n \geq 0$. For a word u representing a dyadic rational, define $q_d(u)$ as its rational value, interpreted as we usually interpret binary numbers.

We can now define the Cauchy representation $\gamma_c : \Sigma_c^\omega \rightharpoonup \mathbb{R}$ for $\Sigma_c = \{-, ., 0, 1, ;\}$. The domain of γ_c consists of ω-words of the form

$$p = u_0; u_1; u_2; \ldots,$$

where $u_i \in \operatorname{dom} q_d$ and the sequence $(q_d(u_i))_i$ is a fast Cauchy sequence, i.e. it satisfies

$$|q_d(u_m) - q_d(u_n)| \leq \frac{1}{2^n}$$

for all $m \geq n \geq 0$. For such a name p, we define $\gamma_c(p) = \lim q_d(u_n)$.

Proposition 3.4.12. *The representation γ_c is surjective, and if $p \in \operatorname{dom} \gamma_c$ is as above, then for any $n \geq 0$, $|q_d(u_n) - \gamma_c(p)| \leq 2^{-n}$.*

Proposition 3.4.13. *A real number is γ_c-computable if and only if it is computable with respect to the decimal or binary representation.*

Definition 3.4.14. A name p is CLOSE if for every $n \geq 0$, we have $|q_d(u_n) - \gamma_c(p)| \leq 2^{-(n+1)}$.

Lemma 3.4.15. *Every $x \in \mathbb{R}$ has a close name. If p is a close name for x, then for every $n \geq 0$, every x' with $|x' - x| < 2^{-(n+1)}$ has a name of the form*

$$u_0; u_1; u_2; \ldots; u_n; u'_{n+1}; u'_{n+2}; \ldots$$

Theorem 3.4.16 (continuity theorem)**.** *If $f : \mathbb{R} \rightharpoonup \mathbb{R}$ is computable with respect to γ_c, then f is continuous on its domain.*

Proof. Suppose f is computable, so it has a realiser $g : \Sigma_c^\omega \rightharpoonup \Sigma_c^\omega$. Let $x \in \operatorname{dom} f$ and $\varepsilon > 0$. We are searching for a suitable δ. Let $p = u_0; u_1; \ldots$ be a close name for x, and let $r = g(p)$. Since g is a realiser for f, r is a name for $f(x)$. Therefore r has the form $r = v_0; v_1; \ldots$.

Let N be such that $2^{-N} < \varepsilon/2$. We have $|q_d(v_N) - f(x)| \leq 2^{-N} < \varepsilon/2$ by one of the preceding propositions. Let n be the length of the string $v_0; v_1; \ldots; v_N;$. By the topological continuity theorem, there exists an $m \geq 0$ such that for all $p' \in \operatorname{dom} g$, if $p'|_m = p|_m$, then $g(p')|_n = v_0; v_1; \ldots, v_N;$. Now take $M \geq 0$ such that the prefix $u_0; u_1; \ldots; u_M;$ of p has length $m' \geq n$, and define $\delta = 2^{-(M+1)}$.

Then for $x' \in \operatorname{dom} f$ with $|x' - x| < \delta$, we have a name for x' of the form $p' = u_0; \ldots; u_M; u'_{M+1}; \ldots$, since p is a close name for x. Since $x' \in \operatorname{dom} f$, $g(p')$ is a name for $f(x')$, and $g(p')|_n = v_0; v_1; \ldots; v_N;$, so $|q_d(v_N) - f(x')| \leq 2^{-N} < \varepsilon/2$. Then $|f(x) - f(x')| < \varepsilon$. □

4 Uvod v funkcionalno analizo

4.1 Normirani in Banachovi prostori

Definicija 4.1.1. Naj bo X vektorski prostor nad poljem $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Preslikava $\|\cdot\| : X \to \mathbb{R}$ je NORMA, če velja:

- $\|x\| \geq 0$,
- $\|x\| = 0 \Leftrightarrow x = 0$,
- $\|\lambda x\| = |\lambda| \, \|x\|$,
- $\|x + y\| \leq \|x\| + \|y\|$.

Opomba. Velja $|\|x\| - \|y\|| \leq \|x - y\|$, iz česar sledi, da je norma zvezna (celo Lipschitzova za $L = 1$).

Norma porodi metriko $d(x, y) = \|x - y\|$ na prostoru X, ki je invariantna na translacije, in za katero velja

$$d(\lambda x, \lambda y) = |\lambda| \, d(x, y).$$

Zaprto kroglo radija r s središčem v točki x označimo z $B(x, r)$, odprto kroglo pa z $\mathring{B}(x, r)$. Zaradi zveznosti norme je zaprtje odprte krogle natanko pripadajoča zaprta krogla.

Definicija 4.1.2. Normiran prostor je BANACHOV, če je poln za inducirano metriko.

Trditev 4.1.3. *Seštevanje in množenje vektorjev s skalarjem sta zvezni operaciji.*

Definicija 4.1.4. Algebra A je NORMIRANA ALGEBRA, če je normiran vektorski prostor in če velja $\|xy\| \leq \|x\| \, \|y\|$. Če ima normirana algebra enoto, zahtevamo še $\|e\| = 1$.

Primer. Naj bo X Hausdorffov topološki prostor in $\mathcal{C}_b(X)$ množica zveznih omejenih funkcij $X \to \mathbb{F}$. Če jo opremimo s supremum normo, postane normirana algebra za seštevanje in množenje po točkah. Preverimo lahko, da je celo Banachov prostor.

Posledica 4.1.5. *Če je X kompakten Hausdorffov prostor, je $\mathcal{C}(X)$ Banachova algebra.*

Trditev 4.1.6. *Naj bo X normiran prostor in Y (vektorski) podprostor v X. Veljata naslednji točki.*

- *Če je Y poln, potem je Y zaprt v X.*
- *Če je X Banachov prostor, potem je Y Banachov natanko tedaj, ko je Y zaprt v X.*

Dokaz. Prva točka: Naj bo $y \in \overline{Y}$. Obstaja zaporedje $(y_n)_n$ v Y, ki konvergira k y. To zaporedje je Cauchyjevo, torej ima limito, ki je enaka y. Sledi $y \in Y$, zato $Y = \overline{Y}$.

Druga točka: V desno smo ravno dokazali. V levo naj bo $(y_n)_n$ Cauchyjevo zaporedje v Y. Potem je Cauchyjevo tudi v X, kjer ima limito, saj je X poln. Ker je Y zaprt, je $y \in Y$, torej $y_n \to y$ v Y. □

Primer. Naj bo X lokalno kompakten Hausdorffov prostor ter $\mathcal{C}_0(X)$ množica vseh funkcij v $\mathcal{C}(X)$, za katere za vsak $\varepsilon > 0$ obstaja kompakt $K_\varepsilon \subseteq X$, da je $|f|$ zunaj K_ε strogo manjša od ε.

Pri $X = \mathbb{R}$ so to natanko vse funkcije, katerih limita v obeh neskončnostih je enaka 0, pri $X = \mathbb{N}$ pa je to natanko prostor c_0 zaporedij, ki konvergirajo k 0.

Dokažemo lahko, da je $\mathcal{C}_0(X)$ zaprt dvostranski ideal v $\mathcal{C}_b(X)$ in zato Banachova algebra.

Primer. Množica c vseh konvergentnih zaporedij je Banachov prostor za supremum normo.

4.1.1 Napolnitve normiranih prostorov

Naj bo X normiran prostor, ki ni poln, in naj bo $\tilde{X}$ množica vseh Cauchyjevih zaporedij v X. To je vektorski prostor za operacije po komponentah. Definiramo

$$\|(x_n)_n\| = \lim_{n\to\infty} \|x_n\| .$$

Ta izraz je dobro definiran, saj velja $|\|x_n\| - \|y_n\|| \leq \|x_n - y_n\|$, torej je zaporedje norm Cauchyjevo in konvergira v $\mathbb{R}$. To pa ni norma, ker obstaja veliko zaporedij z limito norm enako 0. Na $\tilde{X}$ zato vpeljemo ekvivalenčno relacijo

$$(x_n)_n \sim (y_n)_n \Leftrightarrow x_n - y_n \xrightarrow[n\to\infty]{} 0.$$

Sedaj definiramo $\hat{X} = \tilde{X}/\sim$. V $\hat{X}$ potem vpeljemo operaciji seštevanja in množenja s skalarjem, ki delujeta na predstavnikih, ter podobno vpeljemo normo.

Izrek 4.1.7. *Prostor $(\hat{X}, \|\cdot\|)$ je Banachov in vsebuje X kot gost podprostor.*

Dokaz. Dokaz o polnosti izpustimo. Definiramo vložitev $j : X \to \hat{X}$ s predpisom $x \mapsto [(x)_n]$. To je očitno linearna preslikava, za katero velja $\|j(x)\| = \|x\|$, torej je tudi izometrija. Pokazali bomo, da je $j(X)$ gost podprostor v $\hat{X}$.

Naj bo $[(x_n)_n] \in \hat{X}$. Za poljuben $\varepsilon > 0$ obstaja n_ε, da za vse $n, m \geq n_\varepsilon$ velja $\|x_n - x_m\| < \varepsilon$. Za $m = n_\varepsilon$ dobimo $j(x_{n_\varepsilon}) = [(x_{n_\varepsilon})_n]$, in velja

$$\|j(x_{n_\varepsilon}) - [(x_n)_n]\| = \lim_{n\to\infty} \|x_{n_\varepsilon} - x_n\| \leq \varepsilon. \qquad \square$$

Posledica 4.1.8. *Prostor X je Banachov natanko tedaj, ko je $j(X) = \hat{X}$.*

Dokaz. Ideja. Izometrije ohranjajo polnost, polni podprostori so zaprti. Če so gosti, so enaki celoti. $\square$

4.1.2 Osnovne konstrukcije

Naj bo X vektorski prostor nad $\mathbb{F}$. Normi $\|\cdot\|_1$ in $\|\cdot\|_2$ sta EKVIVALENTNI, če obstajata $\alpha, \beta > 0$, da za vse $x \in X$ velja

$$\alpha \|x\|_1 \leq \|x\|_2 \leq \beta \|x\|_1 .$$

Topologiji, ki jih normi porodita, sta enaki, zato je identiteta $(X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ linearni homeomorfizem.

Definicija 4.1.9. Normirana prostora X in Y sta IZOMORFNA, če obstaja linearni homeomorfizem med njima.

Če sta normi ekvivalentni, je $(X, \|\cdot\|_1)$ Banachov natanko tedaj, ko je $(X, \|\cdot\|_2)$ Banachov. Če je $Y \subseteq X$ podprostor in X normiran, je tudi Y normiran, če normo zožimo na Y. Vložitev Y v X je izometrija.

Lema 4.1.10. *Če je $Y \subseteq X$ podprostor in X normiran, je $\overline{Y}$ podprostor.*

Dokaz. Naj bosta $x, y \in \overline{Y}$ in $\alpha, \beta \in \mathbb{F}$. Potem obstajata zaporedji $x_n \to x$ in $y_n \to y$. Velja $\alpha x_n + \beta y_n \to \alpha x + \beta y$. □

Če je X normiran in $Y \leq X$, lahko na X vpeljemo relacijo $x_1 \sim x_2 \Leftrightarrow x_1 - x_2 \in Y$. V kvocientni prostor vpeljemo

$$\|x + Y\| = \inf\{\|x + y\| \mid y \in Y\}.$$

Trditev 4.1.11. *Naj bo X normiran prostor in $Y \leq X$.*

- *$\|\cdot\|$ je polnorma na X/Y.*
- *$\|\cdot\|$ je norma na X/Y natanko tedaj, ko je Y zaprt v X.*
- *Če je X Banachov, je kvocient Banachov.*

Dokaz.

- Točka je le vprašanje preverjanja; dokaz izpustimo.
- S sklicem na prvo točko moramo dokazati le, da je $\|x + Y\| = 0 \Leftrightarrow x \in Y$. Če je $\|x + Y\| = 0$, je $d(x, Y) = 0$, ker pa je Y zaprta, to pomeni $x \in Y$. Podobno v obratno smer.
- Naj bo $(x_n + Y)_n$ Cauchyjevo zaporedje v X/Y. Poiskali bomo podzaporedje $(x_{n_k} + Y)_k$, ki bo konvergiralo. Ker je prvotno zaporedje Cauchyjevo, bo tudi konvergiralo k isti limiti.

  Vemo, da za poljuben $\varepsilon > 0$ obstaja $n_\varepsilon \in \mathbb{N}$, za katerega za vse $n, m \geq n_\varepsilon$ velja $\|x_n - x_m + Y\| < \varepsilon$. Sedaj induktivno konstruiramo podzaporedje $(x_{n_k} + Y)_k$, da za vsak k velja $\left\|x_{n_{k+1}} - x_{n_k} + Y\right\| < 2^{-k}$. Ko to zaporedje imamo, po definiciji

infimuma obstaja tak $y_k \in Y$, da je $\left\|x_{n_{k+1}} - x_{n_k} + y_k\right\| < 2^{-k}$. Definiramo $z_1 = 0$ in $z_{k+1} = z_k + y_k \in Y$. Tedaj

$$\left\|(x_{n_{k+1}} + z_{k+1}) - (x_{n_k} + z_k)\right\| = \left\|x_{n_{k+1}} - x_{n_k} + y_k\right\| < 2^{-k}.$$

Za $w_k = x_{n_k} + z_k$ velja $\|w_{k+1} - w_k\| < 2^{-k}$, hkrati pa je

$$\|w_{m+k} - w_m\| = \left\|\sum_{i=0}^{k-1} w_{m+i+1} - w_{m'i}\right\| \leq \sum_{i=0}^{k-1} \|w_{m+i+1} - w_{m+i}\| < \sum_{i=0}^{k-1} 2^{-m-i}$$

kar je manjše od 2^{1-m}. Sedaj je $(w_m)_m$ Cauchyjevo v X, torej obstaja limita $x \in X$. Zato je

$$\|(x_{n_i} + Y) - (x + Y)\| \leq \|x_{n_i} - x + z_i\| = \|w_i - x\| \to 0. \qquad \square$$

Trditev 4.1.12. *Naj bo Y zaprt podprostor normiranega prostora X. Tedaj je X Banachov natanko tedaj, ko sta tako Y kot X/Y Banachova.*

Dokaz. V desno smo ravno dokazali. V levo: Naj bo $(x_n)_n$ Cauchyjevo zaporedje v X. Vemo, da je $\|x + Y\| \leq \|x\|$, torej je tudi $(x_n + Y)_n$ Cauchyjevo zaporedje in ima limito $x + Y$.

Za vsak $n \in \mathbb{N}$ obstaja tak y_n, da je

$$\|x_n - x + y_n\| < \|x_n - x + Y\| + \frac{1}{n}.$$

Ker velja

$$\|y_n - y_m\| \leq \|y_n + x_n - x\| + \|y_m + x_m - x\| + \|x_m - x_n\|,$$

je $(y_n)_n$ Cauchyjevo, torej ima limito $y \in Y$. Sledi $\lim x_n = x - y$. $\square$

Posledica 4.1.13. *Vsak končnorazsežen normiran prostor je Banachov.*

Dokaz. Indukcija na $d = \dim X$. Za $d = 1$ izberimo $x \in X$ z $\|x\| = 1$. Tedaj za vsak $y \in X$ velja $y = \lambda x$ in $\|y\| = |\lambda|$. Naj bo $(y_n)_n$ Cauchyjevo v X. Potem je $y_n = \lambda_n x$ in $\|y_n - y_m\| = |\lambda_n - \lambda_m|$, zato je $(\lambda_n)_n$ Cauchyjevo v $\mathbb{F}$ in ima limito λ. Seveda $\lambda_n x \to \lambda x$.

Recimo, da so vsi končnorazsežni normirani prostori dimenzije $d-1$ ali manj polni. Naj bo $Y \leq X$ poljuben enorazsežen prostor. Po predpostavki sta Y in X/Y Banachova. $\square$

Produkt $X \times Y$ lahko opremimo z eno od spodnjih norm:

- $\|(x,y)\|_\infty = \max\{\|x\|_X, \|y\|_Y\}$
- $\|(x,y)\|_p = \left(\|x\|_X^p + \|y\|_Y^p\right)^{1/p}$

Produkt je Banachov natanko tedaj, ko sta X in Y Banachova.

4.2 Linearni funkcionali

Izrek 4.2.1. *Naj bo $T : X \to Y$ linearna preslikava med normiranima prostoroma. Naslednje trditve so ekvivalentne:*

- *T je zvezna na X*
- *T je zvezna v $x_0 \in X$*
- *T je zvezna v 0*
- *obstaja $C > 0$, da za vse $x \in X$ velja $\|Tx\| \leq C \|x\|$*
- *T je Lipschitzova*
- *T je enakomerno zvezna*

Za omejen operator $T : X \to Y$ definiramo

$$\|T\| = \inf\{C > 0 \mid \forall x. \|Tx\| \leq C \|x\|\}.$$

Ta infimum obstaja in je dejansko minimum. Potem velja

$$\|Tx\| \leq \|T\| \|x\|$$

za vsak $x \in X$. Velja

$$\|T\| = \sup_{\|x\|=1} \|Tx\| = \sup_{\|x\|\leq 1} \|Tx\| = \sup_{\|x\|<1} \|Tx\|.$$

Množico vseh omejenih linearnih operatorjev $X \to Y$ označimo z $B(X, Y)$, in jo opremimo z zgornjo normo.

Trditev 4.2.2. *Naj bodo X, Y, Z normirani prostori.*

- *$B(X, Y)$ je normiran prostor.*
- *Če $T \in B(X, Y)$ in $S \in B(Y, Z)$, potem je $ST \in B(X, Z)$ in $\|ST\| \leq \|S\| \|T\|$.*

Dokaz. Samo druga točka. Ker je omejenost ekvivalentna zveznosti, je kompozitum v $B(X, Z)$. Za vsak $x \in X$ velja

$$\|STx\| \leq \|S\| \|Tx\| \leq \|S\| \|T\| \|x\|. \qquad \square$$

Definicija 4.2.3. Dualni prostor prostora X je $X^* = B(X, \mathbb{F})$.

Izrek 4.2.4. *Naj bo X normiran in Y Banachov prostor. Tedaj je $B(X, Y)$ Banachov prostor.*

Dokaz. Naj bo $(T_n)_n$ Cauchyjevo v $B(X,Y)$. Izberimo $\varepsilon > 0$. Obstaja n_ε, da za $m, n \geq n_\varepsilon$ velja

$$\|(T_n - T_m)x\| \leq \|T_n - T_m\| \, \|x\| < \varepsilon \, \|x\|$$

za poljuben $x \in X$. Zaporedje $(T_n x)_n$ je Cauchyjevo, zato obstaja $Tx = \lim T_n x \in Y$. S tem dobimo po točkah definiran operator $T : X \to Y$.

Enostavno se prepričamo, da je T linearen. Ker je $|\|T_n\| - \|T_m\|| \leq \|T_n - T_m\|$, je tudi zaporedje norm Cauchyjevo, in zato omejeno. Torej je $\|T_n x\| \leq \|T_n\| \, \|x\| \leq M \, \|x\|$, in je T omejen.

Za konec za poljuben $x \in X$ in $n, m \geq n_\varepsilon$ dobimo $\|T_n x - T_m x\| < \varepsilon \, \|x\|$, in če vzamemo limito $n \to \infty$,

$$\|Tx - T_m x\| \leq \varepsilon \, \|x\| \, .$$

Torej je $\|T - T_m\| \leq \varepsilon$ za $m \geq n_\varepsilon$ in zato $T_m \to T$. □

Posledica 4.2.5. *Dualni prostor je vedno Banachov.*

Izrek 4.2.6. *Naj bo X normiran prostor in $Y \leq X$. Naj bo $T : Y \to Z$ omejen operator in Z poln. Tedaj obstaja natanko en omejen linearen operator $S : \overline{Y} \to Z$, da je $S|_Y = T$. Velja še $\|S\| = \|T\|$.*

Dokaz. Naj bo $x \in \overline{Y}$. Radi bi definirali Sx. Obstaja zaporedje $(x_n)_n$ v Y, da bo $x_n \to x$. Definiramo $Sx = \lim Tx_n$.

Če je tudi $(x'_n)_n$ zaporedje, ki konvergira k x, je

$$\lim(Tx_n - Tx'_n) = \lim T(x_n - x'_n) = 0,$$

ker je T zvezen. Torej je S dobro definiran. Očitno je tudi linearen, in velja $S|_Y = T$. Za enoličnost predpostavimo, da je tudi S' tak operator. Potem za $x \in \overline{Y}$ velja $Sx_n = S'x_n$, in sta limiti Sx in $S'x$ posledično tudi enaki.

Velja

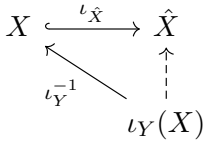
$$\|Sx\| = \|\lim Tx_n\| = \lim \|Tx_n\| = \lim \|T\| \, \|x_n\| = \|T\| \, \|x\| \, ,$$

torej je S omejen in $\|S\| \leq \|T\|$. Obrat je očiten. □

Posledica 4.2.7. *Naj bosta $S, T : X \to Y$ omejena operatorja, ki se ujemata na gostem podprostoru. Potem je $S = T$.*

Posledica 4.2.8. *Naj bo X normiran prostor, ki je gost podprostor v Banachovem prostoru Y. Tedaj sta $\hat{X}$ in Y izometrično izomorfna.*

Dokaz. Naj bosta ι_Y in $\iota_{\hat{X}}$ vložitvi.



Preslikavi ι_Y^{-1} in $\iota_{\hat{X}}$ sta izometriji, torej je tak tudi njun kompozitum. To preslikavo lahko enolično razširimo do izometrije prostorov Y in $\hat{X}$. □

Definicija 4.2.9. Naj bo $T : X \to X$ omejen operator med normiranima prostoroma. Tedaj je T OBRNLJIV, če je bijektiven in T^{-1} omejen.

Definicija 4.2.10. Operator $T : X \to Y$ je NAVZDOL OMEJEN, če obstaja $c > 0$, da je $\|Tx\| \geq c\,\|x\|$ za vsak $x \in X$.

Opomba. Vsak navzdol omejen operator je injektiven.

Trditev 4.2.11. *Naj bo $T \in B(X, Y)$. Naslednji trditvi sta ekvivalentni.*

- *Obstaja operator $T^{-1} : TX \to X$.*
- *T je navzdol omejen.*

4.2.1 Banachov izrek

Definicija 4.2.12. Naj bo X vektorski prostor. Preslikava $p : X \to \mathbb{R}$ je SUBLINEARNI FUNKCIONAL, če je $p(x + y) \leq p(x) + p(y)$ in $p(\lambda x) = \lambda p(x)$ za poljubne $x, y \in X$ ter $\lambda \geq 0$.

Primer. Vsaka polnorma je sublinearni funkcional.

Izrek 4.2.13 (realni Hahn-Banachov izrek)**.** *Naj bo $Y \leq X$ vektorski prostor in $p : X \to \mathbb{R}$ sublinearni funkcional. Naj bo $f : Y \to \mathbb{R}$ tak linearni funkcional, da za vsak $y \in Y$ velja $f(y) \leq p(y)$. Tedaj obstaja linearni funkcional $F : X \to \mathbb{R}$, da je $F|_Y = f$ in $F(x) \leq p(x)$ za vsak $x \in X$.*

Dokaz. Prvo obravnavajmo primer, kjer je $\dim X/Y = 1$. Tedaj je $X = Y \oplus \mathbb{R}x_0$ za neki $x_0 \in X \setminus Y$. Vse možne linearne razširitve f do F so oblike

$$F(x) = F(y + \lambda x_0) = F(y) + \lambda F(x_0),$$

torej je razširitev enolično določena z $F(x_0) =: \alpha$. Za $y_1, y_2 \in Y$ velja

$$f(y_1 + y_2) \leq p(y_1 + y_2) = p(y_1 + y_2 + x_0 - x_0) \leq p(y_1 + x_0) + p(y_2 - x_0)$$

oziroma

$$f(y_2) - p(y_2 - x_0) \leq p(y_1 + x_0) - f(y_1),$$

torej je

$$\sup_{y \in Y} \left(f(y) - p(y - x_0)\right) \leq \inf_{y \in Y} \left(p(y + x_0) - f(y)\right).$$

Za α lahko izberemo katerokoli vrednost med tema številoma. Potem definiramo $F(y + tx_0) = f(y) + t\alpha$ in ločimo primere.

- Če je $t = 0$, je $F(y) \leq p(y)$ po predpostavki, saj je $F(y) = f(y)$.

- Če je $t > 0$, je $\alpha \leq p(y/t + x_0) - f(y/t)$, kar množimo s t, in dobimo $f(y) + t\alpha \leq p(y + tx_0)$.
- Če je $t < 0$, je $\alpha \geq f(-y/t) - p(-y/t - x_0)$, kar množimo z $(-t)$ in zaključimo kot zgoraj.

Za splošen primer tvorimo množico

$$\mathcal{A} = \{(Y_1, f_1) \,|\, Y \leq Y_1 \leq X,\, f_1|_Y = f, \forall y_1 \in Y_1 . f_1(y_1) \leq p(y_1)\}$$

in definiramo relacijo

$$(Y_1, f_1) \preccurlyeq (Y_2, f_2) \Leftrightarrow Y_1 \leq Y_2 \wedge f_2|_{Y_1} = f_1.$$

To je očitno delna urejenost. Naj bo $\{(Y_i, f_i)\}_{i \in I}$ neka veriga v $\mathcal{A}$. Vzemimo $Z = \bigcup_{i \in I} Y_i$. To je podprostor, ker je veriga urejena. Definiramo še preslikavo $g : Z \to \mathbb{R}$ z $g(z) = f_i(z)$ za nek $i \in I$, za katerega je $z \in Y_i$. To je dobro definiran funkcional, za katerega velja $g(z) \leq p(z)$ za vse $z \in Z$.

Očitno je (Z, g) zgornja meja za verigo. Po Zornovi lemi ima $\mathcal{A}$ maksimalen element $(\tilde{Y}, \tilde{f})$. Če je $\tilde{Y} \neq X$, potem obstaja $x \in X \setminus \tilde{Y}$. Po prvem koraku dokaza lahko $\tilde{f}$ razširimo na linearno ogrinjačo množice $\{\tilde{Y}, x\}$, kar je protislovje z maksimalnostjo. Torej $X = \tilde{Y}$. □

Lema 4.2.14. *Naj bo X kompleksen vektorski prostor. Potem veljajo naslednje točke.*

- *Če je $f : X \to \mathbb{R}$ $\mathbb{R}$-linearen funkcional, je $\tilde{f} : X \to \mathbb{C}$, definiran z $\tilde{f}(x) = f(x) - if(ix)$ $\mathbb{C}$-linearen funkcional, in velja $\operatorname{Re} \tilde{f} = f$.*
- *Če je $g : X \to \mathbb{C}$ $\mathbb{C}$-linearen funkcional in $\operatorname{Re} g = f$, potem je $g = \tilde{f}$.*
- *Če je p polnorma, potem je $|f(x)| \leq p(x)$ za vse $x \in X$ natanko tedaj, ko za vse $x \in X$ velja $\left|\tilde{f}(x)\right| \leq p(x)$.*
- *Če je X normiran in $f : X \to \mathbb{R}$ omejen, je $\tilde{f}$ omejen in $\left\|\tilde{f}\right\| = \|f\|$.*

Dokaz. Prva točka je enostavna. Za drugo točko le pogledamo $g(x) = f(x) + if_1(x)$ ter $ig(x) = g(ix)$, s čimer pokažemo $f_1(x) = -f(ix)$.

Za tretjo točko opazimo

$$|f(x)| = \left|\operatorname{Re} \tilde{f}(x)\right| \leq \left|\tilde{f}(x)\right| \leq p(x),$$

v drugo smer pa

$$\left|\tilde{f}(x)\right| = e^{i\varphi}\tilde{f}(x) = \tilde{f}(e^{i\varphi}x) = \operatorname{Re} \tilde{f}(e^{i\varphi}x)$$

za nek kot φ. Potem

$$\left|\tilde{f}(x)\right| = f(e^{i\varphi}x) \leq p(e^{i\varphi}x) = \left|e^{i\varphi}\right| p(x) = p(x).$$

Za zadnjo točko poglejmo $|f(x)| \leq \|f\| \, \|x\|$. Potem je $p(x) = \|f\| \, \|x\|$ polnorma na X, za katero velja $|f(x)| \leq p(x)$. Po prejšnji točki velja $\left|\tilde{f}(x)\right| \leq p(x)$, torej je $\tilde{f}$ omejen z $\|f\|$. Po drugi strani je $\left\|\tilde{f}(x)\right\| \leq q(x)$ za $q(x) = \left\|\tilde{f}\right\| \|x\|$, in zato po prejšnji točki $|f(x)| \leq q(x)$ in $\|f\| \leq \left\|\tilde{f}\right\|$. □

Izrek 4.2.15 (kompleksni Hahn-Banach)**.** *Naj bo X vektorski prostor nad $\mathbb{F}$, $Y \leq X$ in p polnorma na X. Če je $f : Y \to \mathbb{F}$ linearni funkcional, da za vse $y \in Y$ velja $|f(y)| \leq p(y)$, potem obstaja linearni funkcional $F : X \to \mathbb{F}$, za katerega je $F|_Y = f$ in $|F(X)| \leq p(x)$ za vsak $x \in X$.*

Dokaz. Za $\mathbb{F} = \mathbb{R}$ po realni verziji izreka obstaja funkcional $F : X \to \mathbb{R}$, ki razširja f in za katerega je $F(x) \leq p(x)$ za $x \in X$. Velja $-F(x) = F(-x) \leq p(-x) = p(x)$, torej $|F(x)| \leq p(x)$.

Če pa je $\mathbb{F} = \mathbb{C}$, vzemimo $\mathbb{R}$-linearen funkcional $f_1 = \operatorname{Re} f$, za katerega po lemi velja $|f_1(y)| \leq p(y)$ za $y \in Y$. Tega lahko razširimo do $F_1 : X \to \mathbb{R}$, nato pa vzamemo funkcional $\tilde{F}$, za katerega velja $\operatorname{Re} \tilde{F} = F$. S pomočjo leme hitro vidimo, da $\tilde{F}$ razširja f in $\left|\tilde{F}(x)\right| \leq p(x)$. □

Izrek 4.2.16 (Hahn-Banachov izrek za normirane prostore)**.** *Naj bo $Y \leq X$ podprostor normiranega prostora X in $f : Y \to \mathbb{F}$ omejen. Tedaj obstaja $f : X \to \mathbb{F}$, da je $F|_Y = f$ ter $\|F\| = \|f\|$.*

Posledica 4.2.17. *Naj bo X normiran in $x \in X$ neničeln vektor. Tedaj obstaja $F \in X^*$, da je $\|f\| = 1$ in $f(x) = \|x\|$.*

Dokaz. Naj bo $Y = \mathbb{F} \cdot x$. Funkcional $g(\lambda x) = \lambda \|x\|$ je zvezen, linearen in definiran na Y z $\|g\| = 1$. Po Hahn-Banachu ga lahko razširimo na funkcional $X \to \mathbb{F}$. □

Posledica 4.2.18. *Naj bo X normiran in $x \in X$. Tedaj je*

$$\|x\| = \max\{|f(x)| \mid f \in X^*, \|f\| = 1\}.$$

Posledica 4.2.19. *Naj bo $Y \leq X$ zaprt podprostor normiranega prostora X. Naj bo $x_0 \in X \setminus Y$ ter $d = d(x, Y)$. Tedaj obstaja $f \in X^*$, da je $f(x_0) = 1$, $f|_Y = 0$ in $\|f\| = 1/d$.*

Dokaz. Naj bo $g \in (X/Y)^*$ tak, da je $\|g\| = 1$ in $g(x_0 + Y) = \|x_0 + Y\| = d$. Definirajmo $f = g \circ \pi$, kjer je π kvocientna projekcija. Za $y \in Y$ velja $f(y) = 0$, ker je $y + Y = 0$ in g linearna. Funkcional f je omejen, ker je kompozitum omejenih funkcionalov, velja $f(x_0) = d$. Izračunamo

$$\|f\| = \sup_{\|x\|<1} |f(x)| = \sup_{\|x\|<1} |g(\pi(x))| = \sup_{\|x+Y\|<1} g(x + Y) = \|g\| = 1,$$

torej je iskani funkcional $\frac{1}{d}f$. □

Izrek 4.2.20. *Naj bo $Y \leq X$ podprostor normiranega prostora X. Tedaj je*

$$\overline{Y} = \bigcap \{\ker f \mid f \in X^*, Y \subseteq \ker f\}.$$

Dokaz. Naj bo Z presek iz izreka. Če je $Y \subseteq \ker f$, je $\overline{Y} \subseteq \ker f$, saj je jedro zaprto. Torej je $\overline{Y} \subseteq Z$. Recimo, da $\overline{Y}$ ni enako Z. Tedaj obstaja $z \in Z \setminus \overline{Y}$, torej po prejšnji posledici obstaja $f \in X^*$, da je $f(z) = 1$ in $f|_{\overline{Y}} = 0$. Velja $Y \subseteq \ker f$, torej $z \in \ker f$, kar je protislovje z $f(z) = 1$. □

4.2.2 Adjungirani operator in drugi dual

Definicija 4.2.21. Za normiran prostor X in $f \in X^*$ definiramo $\hat{x}(f) = f(x)$.

Trditev 4.2.22. *Velja $\hat{x} \in X^{**}$ in $\|\hat{x}\| = \|x\|$.*

Dokaz. Velja $|\hat{x}(f)| = |f(x)| \leq \|f\| \, \|x\|$, zato $\|\hat{x}\| \leq \|x\|$. Po Hahn-Banachu velja f z $\|f\| = 1$ in $f(x) = \|x\|$. Tedaj $\hat{x}(f) = \|x\|$, torej $\|\hat{x}\| \geq \|x\|$. □

Prostor X vložimo v X^{**} s preslikavo $i(x) = \hat{x}$.

Trditev 4.2.23. *Preslikava i je linearna izometrična vložitev.*

Naj bo $A : X \to Y$ omejen linearen operator. Za $f \in Y^*$ definiramo adjungirani operator operatorja A v smislu Banachovih prostorov, A^*, z

$$A^* f = f \circ A.$$

Trditev 4.2.24. *Naj bo $A : X \to Y$ omejen linearen operator med normiranima prostoroma. Tedaj je $A^* : Y^* \to X^*$ tudi omejen in $\|A^*\| = \|A\|$.*

Dokaz. Izračunamo

$$|A^* f(x)| = |f(Ax)| \leq \|f\| \, \|Ax\| \leq \|f\| \, \|A\| \, \|x\| \,,$$

torej $\|A^* f\| \leq \|A\| \, \|f\|$. Iz tega sledi, da je tudi A^{**} omejen z $\|A^{**}\| \leq \|A^*\| \leq \|A\|$. Vemo pa

$$\|Ax\| = \left\| \widehat{Ax} \right\| = \left\| A^{*+} \hat{x} \right\| \leq \|A^{**}\| \, \|\hat{x}\| = \|A^{**}\| \, \|x\|$$

torej je $\|A\| \leq \|A^{**}\|$ in so norme enake. □

Definicija 4.2.25. Normiran prostor je REFLEKSIVEN, če je i_X surjekcija.

Opomba. Refleksivni prostori so Banachovi, saj je X^{**} Banachov.

4.3 Temeljni izreki funkcionalne analize

Izrek 4.3.1 (Baire)**.** *Naj bo (X, d) poln metrični prostor in $(U_n)_n$ števna družina odprtih gostih množic v X. Tedaj je presek $\bigcap_n U_n$ gost v X.*

Dokaz. Dokazujemo, da za vsak $x \in X$ obstaja $r > 0$, da je presek

$$\mathring{B}(x, r) \cap \bigcap_{n\in\mathbb{N}} U_n \neq \varnothing.$$

Naj bosta $x \in X$ ter $r > 0$ poljubna. Induktivno bomo konstruirali zaporedji $(x_n)_n$ in $(r_n)_n$ z naslednjimi lastnostmi:

- $B(x_{n+1}, r_{n+1}) \subseteq U_n \cap \mathring{B}(x_n, r_n)$,
- $r_n \leq 1/n$.

Postavimo $x_1 = x$ in $r_1 = \min\{1, r\}$. Recimo, da smo že konstruirali zaporedji do n-tega člena. Ker je U_n gosta odprta množica, je $U_n \cap \mathring{B}(x_n, r_n)$ neprazna odprta množica, zato obstajata $r_{n+1} \leq 1/n+1$ in x_{n+1}, da je $\mathring{B}(x_{n+1}, 2r_{n+1}) \subseteq U_n \cap \mathring{B}(x_n, r_n)$. Prva lastnost sedaj velja, ker je $B(x_{n+1}, r_{n+1}) \subseteq \mathring{B}(x_{n+1}, 2r_{n+1})$.

S tem smo konstruirali želeni zaporedji. Ker je

$$x_n \in B(x_n, r_n) \subseteq U_{n-1} \cap \mathring{B}(x_{n-1}, r_{n-1}) \subseteq \mathring{B}(x_{n-1}, r_{n-1}) \subseteq \mathring{B}(x_m, r_m)$$

za $m < n$, je $d(x_m, x_n) \leq r_m \leq 1/m$. Torej je zaporedje Cauchyjevo in obstaja limita x_0 v X. Velja

$$x_0 \in \bigcap_{n\in\mathbb{N}} B(x_{n+1}, r_{n+1}) \subseteq \bigcap_{n\in\mathbb{N}} U_n \cap \mathring{B}(x_n, r_n) \subseteq \mathring{B}(x_1, r_1) \cap \bigcap_{n\in\mathbb{N}} U_n = \mathring{B}(x, r) \cap \bigcap_{n\in\mathbb{N}} U_n,$$

s čimer je dokaz zaključen. □

Posledica 4.3.2. *Naj bo X poln metrični prostor in $(A_n)_n$ zaporedje zaprtih množic, da je $X = \bigcup_n A_n$. Tedaj obstaja $m \in \mathbb{N}$, da je $\mathring{A}_m \neq \varnothing$.*

Dokaz. Če je $\mathring{A}_j = \varnothing$ za vse j, potem je A_j^{c} odprta in gosta. Potem je $\bigcap_j A_j^{\mathsf{c}}$ gost, torej $\bigcup_j A_j = \left(\bigcap_j A_j^{\mathsf{c}}\right)^{\mathsf{c}} \neq X$. □

Naj bosta X in Y normirana prostora ter $\mathcal{F} \subseteq B(X, Y)$. Recimo, da obstaja $M \geq 0$, za katerega je $\|T\| \leq M$ za vse $T \in \mathcal{F}$. Tedaj za vsak $x \in X$ velja

$$\|Tx\| \leq \|T\| \, \|x\| \leq M \, \|x\| \,,$$

oziroma $\mathcal{F}x \subseteq B(0, M \, \|x\|)$. Pravimo, da je $\mathcal{F}$ OMEJENA PO TOČKAH.

Izrek 4.3.3 (princip enakomerne omejenosti)**.** *Naj bo X Banachov in Y normiran prostor. Naj bo $\mathcal{F} \subseteq B(X,Y)$ po točkah omejena družina. Potem je kot množica operatorjev enakomerno omejena v $B(X,Y)$.*

Dokaz. Definiramo

$$A_n = \{x \in X \mid \forall T \in \mathcal{F}.\, \|Tx\| \leq n\} = \bigcap_{T \in \mathcal{F}} f_T^{-1}([0,n])$$

za $f_T(x) = \|Tx\|$. Ker je A_n presek zaprtih množic, je zaprt. Ker je $\mathcal{F}$ omejena po točkah, za poljuben $x \in X$ obstaja $m \in \mathbb{N}$, da je $\|Tx\| \leq m$ za poljuben $T \in \mathcal{F}$, oziroma $x \in A_m$. Torej je X enak uniji množic A_n, in po posledici Bairovega izreka obstaja tak $n_0 \in \mathbb{N}$, da ima A_{n_0} notranjo točko, in posledično vsebuje odprto kroglo $\mathring{B}(x_0, r)$.

Izberimo poljuben $x \in B(0,1)$ ter definirajmo $y = x_0 + \frac{r}{2}x$. Velja $y \in \mathring{B}(x_0, r)$, torej je $\|Ty\| \leq n_0$ za poljuben $T \in \mathcal{F}$. Sledi

$$\|Tx\| = \left\| T\frac{2y - x_0}{r} \right\| \leq \frac{2}{r}\|Ty\| + \frac{1}{r}\|Tx_0\| \leq \frac{2}{r}n_0 + \frac{1}{r}\|Tx_0\| =: M.$$

Torej je $\|T\| \leq M$ za vsak $T \in \mathcal{F}$. □

Izrek 4.3.4 (o šibki omejenosti)**.** *Naj bo $A \subseteq X$ podmnožica normiranega prostora X. Potem je A omejena v X natanko tedaj, ko je za vsak $f \in X^*$ množica $\{f(x) \mid x \in A\}$ omejena v $\mathbb{F}$.*

Dokaz. V desno: Ker je A omejena, obstaja M, da je $\|x\| \leq M$ za vsak $x \in A$. Potem je $|f(x)| \leq \|f\|\,\|x\| \leq M\,\|f\|$ za $f \in X^*$.

V levo: Oglejmo si vložitev v drugi dual, $i(x) = \hat{x}$. Množica $\{f(x) \mid x \in A\}$ je omejena natanko tedaj, ko je omejena množica $\{\hat{x}(f) \mid \hat{x} \in i(A)\}$, ker pa je $i(A)$ omejena po točkah, je po principu enakomerne omejenosti $i(A)$ tudi enakomerno omejena, torej obstaja $M \geq 0$, da je $\|x\| = \|\hat{x}\| \leq M$ za vse $x \in A$. □

Posledica 4.3.5. *Naj bo X Banachov prostor in Y normiran prostor. Naj bo $A \subseteq B(X,Y)$ taka, da za vsak $f \in Y^*$ in vsak $x \in X$ obstaja $M_{f,x} \geq 0$, da je $|f(Tx)| \leq M_{f,x}$ za vsak $T \in A$. Tedaj je A omejena.*

Dokaz. Po izreku o šibki omejenosti je množica $\{Ty \mid T \in A\}$ omejena v Y. Torej je A omejena po točkah, ker pa je X Banachov, je A omejena po principu enakomerne omejenosti. □

Lema 4.3.6. *Naj bo X Banachov prostor in $(x_n)_n$ zaporedje vektorjev, za katere velja $\sum \|x_n\| < \infty$. Tedaj vrsta $\sum x_n$ konvergira v X.*

Dokaz. Označimo $s_n = \sum_{i=1}^{n} x_i$. Za $n > m$ velja

$$\|s_n - s_m\| = \|x_n + x_{n-1} + \cdots + x_{m+1}\| \leq \|x_{m+1}\| + \cdots + \|x_n\| \leq \sum_{k \geq m} \|x_k\| \xrightarrow[m\to\infty]{} 0.$$

Torej je zaporedje $(s_n)_n$ Cauchyjevo in zato konvergentno. □

Izrek 4.3.7 (o odprti preslikavi)**.** *Naj bo T omejen surjektiven linearen operator med Banachovima prostoroma X in Y. Tedaj je T odprta preslikava.*

Dokaz. Dokaz poteka v štirih korakih. V prvem koraku dokažimo, da če odprta podmnožica $U \subseteq X$ vsebuje 0, ima $\overline{T(U)}$ notranjo točko. Ker je U odprta, obstaja $\delta > 0$, da je

$$\delta \mathring{B}(0,1) = \mathring{B}(0,\delta) \subseteq U.$$

Poljuben $x \in X$ je v

$$x \in 2\|x\| \mathring{B}(0,1) = \frac{2\|x\|}{\delta} \delta \mathring{B}(0,1) \subseteq mU$$

za $m \geq \frac{2\|x\|}{\delta}$, torej je

$$X \subseteq \bigcup_{n\in\mathbb{N}} nU.$$

Velja

$$Y = TX = \bigcup_{n\in\mathbb{N}} T(nU) = \bigcup_{n\in\mathbb{N}} \overline{T(nU)}$$

in po Bairovem izreku obstaja n_0, da ima $\overline{T(n_0 U)}$ notranjo točko. Množenje s skalarjem je homeomorfizem, zato velja $\overline{T(n_0 U)} = n_0 \overline{TU}$ in ima tudi $\overline{TU}$ notranjo točko.

V drugem koraku pokažimo, da je ob isti predpostavki točka 0 notranja za $\overline{TU}$. Ker je U odprta v X, obstaja $\mathring{B}(0,\varepsilon) \subseteq U$. Vzemimo $V = \mathring{B}(0,\varepsilon/2)$. Potem je množica $V - V \subseteq \mathring{B}(0,\varepsilon)$, kjer smo vzeli definicijo

$$A - B := \{a - b \,|\, a \in A, b \in B\}.$$

Po dokazanem v prvem koraku ima $\overline{TV}$ notranjo točko, torej obstaja odprta množica $W \subseteq \overline{TV}$. Množica

$$\bigcup_{w\in W} (W - \{w\}) = W - W \subseteq \overline{TV} - \overline{TV} \subseteq \overline{TV - TV} = \overline{T(V - V)} \subseteq \overline{TU}$$

je unija odprtih množic in vsebuje 0. Druga vključitev zgoraj velja, ker je funkcija $m : Y \times Y \to Y$, $m(y_1, y_2) = y_1 - y_2$, zvezna.

V tretjem koraku spet ob isti predpostavki pokažimo, da je TU odprta okolica za 0. Najprej pokažimo poseben primer, če je $U = \mathring{B}(0,\varepsilon)$. Definiramo $\varepsilon_0 = \varepsilon/2$ in zapišemo

$$\varepsilon_0 = \sum_{i=1}^{\infty} \varepsilon_i$$

za neko zaporedje $(\varepsilon_i)_i$ pozitivnih števil. Po drugem koraku za vsak i obstaja $\eta_i > 0$, da je $\mathring{B}(0,\eta_i) \subseteq \overline{T(\mathring{B}(0,\varepsilon_i))}$. Sedaj izberimo $y \in \mathring{B}(0,\eta_0)$ in pokažimo $y = Tx$ za neki $x \in \mathring{B}(0,\varepsilon)$.

Če je $\|x\| < \varepsilon_i$, velja $\|Tx\| \leq \|T\|\,\varepsilon_i$, torej je

$$\mathring{B}(0,\eta_i) \subseteq T(\mathring{B}(0,\varepsilon_i)) \subseteq \mathring{B}(0,\varepsilon_i\,\|T\|).$$

Sledi $\eta_i \leq \varepsilon_i\,\|T\|$ in zaporedje η_i konvergira k 0. Po predpostavki je $y \in \mathring{B}(0,\eta_i) \subseteq \overline{T(\mathring{B}(0,\varepsilon_0))}$. Po definiciji zaprtja η_1-okolica za y seka $T(\mathring{B}(0,\varepsilon_0))$, torej obstaja $x_0 \in X$, da je $\|x_0\| < \varepsilon_0$ in $\|y - Tx_0\| < \eta_1$. Sedaj velja $y - Tx_0 \in \mathring{B}(0,\eta_1) \subseteq \overline{T(\mathring{B}(0,\varepsilon_1))}$, in spet po definiciji zaprtja η_2-okolica za $y - Tx_0$ seka $T(\mathring{B}(0,\varepsilon_1))$, torej obstaja $x_1 \in X$ z $\|x_1\| < \varepsilon_1$ ter $\|y - Tx_0 - Tx_1\| < \eta_2$.

Postopek ponavljamo, s čimer dobimo zaporedje $(x_n)_n$, za katerega velja $\|x_n\| < \varepsilon_n$ ter $\|y - T(x_1 + \cdots + x_n)\| < \eta_{n+1}$. Ker je $\sum \|x_n\| < \sum \varepsilon_n < \infty$, vrsta $\sum x_n$ konvergira absolutno in zato v Banachovem prostoru X konvergira proti nekemu vektorju x. Ker je T omejen operator, konvergira tudi vrsta $\sum Tx_n$ in velja $Tx = \sum Tx_n$. To mora biti enako y, saj $\eta_i \to 0$. Dodatno je $\|x\| \leq \sum \|x_n\| < 2\varepsilon_0 = \varepsilon$, torej je $T(\mathring{B}(0,\varepsilon))$ okolica za 0.

Če pa U ni take oblike, obstaja $\varepsilon > 0$, da je $\mathring{B}(0,\varepsilon) \subseteq U$, in po ravno dokazanem T-slike te množice vsebuje odprto kroglo okoli 0. Zato jo vsebuje tudi TU.

V zadnjem koraku dokažimo, da je slika odprte $U \subseteq X$ odprta v Y. Naj bo $y \in TU$. Obstaja $x \in U$, da je $y = Tx$, torej je $V = U - x$ odprta okolica za 0 in je po prejšnjem koraku TV okolica za 0 v Y. Potem je $TU = TV + Tx$ okolica za y. □

Posledica 4.3.8. *Naj bo T omejen linearen bijektiven operator med Banachovima prostoroma. Potem je njegov inverz tudi omejen.*

Posledica 4.3.9. *Naj bo X vektorski prostor in $\|\cdot\|_1, \|\cdot\|_2$ dve normi na X, za kateri je X Banachov prostor. Potem sta normi bodisi ekvivalentni bodisi neprimerljivi.*

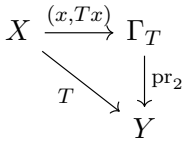
Dokaz. Če je $\|x\|_1 \leq c_1\,\|x_2\|$, je identiteta $I : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$ omejena, in je njen inverz tudi omejen. □

Lema 4.3.10. *Naj bosta X, Y normirana prostora in $f : X \to Y$ zvezna preslikava. Tedaj je $\Gamma_f^{\text{zap}} \subseteq X \times Y$, če ta produkt opremimo z normo $\|(x,y)\|_\infty = \max\{\|x\|, \|y\|\}$.*

Dokaz. Naj gre $(x_n, f(x_n)) \to (x,y)$. Velja $x_n \to x$ in $f(x_n) \to y$, ker pa je f zvezna, tudi $f(x_n) \to f(x)$. □

Izrek 4.3.11 (o zaprtem grafu)**.** *Naj bo $T : X \to Y$ linearna preslikava, X in Y Banachova prostora ter Γ_T zaprt v $X \times Y$. Potem je T omejena.*

Dokaz. Graf je zaprt, torej je Banachov. Oglejmo si spodnji diagram.



Projekcija $\mathrm{pr}_1 : \Gamma_T \to X$ je omejena in bijektivna, torej je tudi njen inverz $x \mapsto (x, Tx)$ omejen po posledici izreka o odprti preslikavi. Velja $T = \mathrm{pr}_2 \circ \mathrm{pr}_1^{-1}$. □

4.4 Hilbertovi prostori

Omejimo se na primer $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

Definicija 4.4.1. Naj bo X vektorski prostor nad $\mathbb{F}$. Preslikava $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{F}$ je SKALARNI PRODUKT, če zadošča

- $\langle x, x \rangle \geq 0$ (realno in nenegativno),
- $\langle x, x \rangle = 0$ natanko tedaj, ko je $x = 0$,
- $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$,
- $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

Trditev 4.4.2 (Paralelogramska enakost)**.** *Naj bo X prostor s polskalarnim produktom. Za $x, y \in X$ velja*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Trditev 4.4.3. *Skalarni produkt je zvezna preslikava.*

Izrek 4.4.4 (Jordan, von Neumann)**.** *Če v normiranem prostoru velja paralelogramska enakost, je norma porojena s skalarnim produktom.*

Definicija 4.4.5. Prostor X s skalarnim produktom je HILBERTOV PROSTOR, če je za porojeno normo Banachov prostor.

Naj bo X prostor s skalarnim produktom in $\hat{X}$ napolnitev X kot normiran prostor. Ker norma na X ustreza paralelogramski enakosti, zaradi zveznosti norme to velja tudi na $\hat{X}$ in je norma na $\hat{X}$ porojena s skalarnim produktom. Torej je Hilbertov prostor. Če $x_n \to x$ in $y_n \to y$, velja

$$\langle x, y \rangle = \lim_{n \to \infty} \langle x_n, y_n \rangle .$$

Definicija 4.4.6. Vektorja x in y sta PRAVOKOTNA, če $\langle x, y \rangle = 0$. Označimo $x \perp y$.

Definicija 4.4.7. Množici A in B sta PRAVOKOTNI, če je $\langle a, b \rangle = 0$ za vsak $a \in A$ ter $b \in B$.

Izrek 4.4.8 (Pitagora)**.** *Naj bo X vektorski prostor s skalarnim produktom. Če sta vektorja x in y pravokotna, je $\|x\|^2 + \|y\|^2 = \|x+y\|^2$.*

Izrek 4.4.9. *Naj bo H Hilbertov prostor in K neprazna zaprta konveksna množica v H. Tedaj za vsak $x \in H$ obstaja natanko en $k \in K$, da je $d(x,K) = \|x,k\|$.*

Dokaz. Brez škode za splošnost lahko privzamemo $x = 0$. Označimo $d = \inf\{\|y\| \mid y \in K\}$. Po definiciji infimuma obstaja zaporedje $(k_n)_n$, da $\|k_n\| \to d$. Izberimo $\varepsilon > 0$. Potem obstaja $N \in \mathbb{N}$, da za $n \geq N$ velja $\|k_n\|^2 \leq d^2 + \varepsilon^2/4$. Po paralelogramski enakosti velja

$$\|k_n - k_m\|^2 = 2\|k_n\|^2 + 2\|k_m\|^2 - 4\left\|\frac{k_n + k_m}{2}\right\|^2 \leq 4d^2 + \varepsilon^2 - 4d^2 = \varepsilon^2,$$

torej je $(k_n)_n$ Cauchyjevo in ima limito $k \in \overline{K} = K$. Velja $\|k\| = d$.

Za enoličnost še enkrat uporabimo paralelogramsko enakost. Če je $k' \in K$ še en vektor s $\|k'\| = d$, potem

$$\left\|k - k'\right\|^2 = 2\|k\|^2 + 2\left\|k'\right\|^2 - 4\left\|\frac{k + k'}{2}\right\|^2 \leq 0,$$

saj je K konveksna in $k, k' \in K$. □

Izrek 4.4.10. *Naj bo M zaprt podprostor Hilbertovega prostora H, $x \in H$ ter $x_0 \in M$. Tedaj velja*

$$x - x_0 \perp M \Leftrightarrow d(x,M) = \|x - x_0\|.$$

Dokaz. Recimo, da za $x_0 \in M$ velja $d(x,M) = \|x - x_0\|$ in da $x - x_0$ ni pravokoten na M. Tedaj obstaja $y \in M$, da je $\langle x - x_0, y\rangle \neq 0$. Brez škode za splošnost lahko privzamemo, da je ta skalarni produkt pozitiven, sicer y zavrtimo za potreben kot. Potem je

$$\|x - (x_0 + \varepsilon y)\|^2 = \|x - x_0\|^2 - 2\varepsilon \operatorname{Re}\langle x - x_0, y\rangle + |\varepsilon|^2 \|y\|^2$$

za poljuben $\varepsilon \in \mathbb{C}$, saj je $(x_0 + \varepsilon y) \in M$. Če izberemo dovolj majhen $\varepsilon \in \mathbb{R}$, dobimo

$$\|x - (x_0 + \varepsilon y)\|^2 < \|x - x_0\|^2,$$

kar je protislovje, saj je x_0 najbližji vektor iz M. Torej je $x - x_0 \perp M$. Tedaj za vsak $y \in M$ velja

$$\|x - y\|^2 = \|x - x_0\|^2 + \|x_0 - y\|^2 \geq \|x - x_0\|.$$ □

Definicija 4.4.11. Naj bo X prostor s skalarnim produktom. Za $x \in X$ definiramo $\{x\}^\perp = \{y \in X \mid x \perp y\}$, za $A \subseteq X$ pa

$$A^\perp = \bigcap_{x \in A} \{x\}^\perp.$$

Lema 4.4.12. *Za $A \subseteq X$ je $A^\perp$ vedno zaprt podprostor v X.*

Dokaz. Dovolj je pokazati, da je $\{x\}^\perp$ zaprt podprostor za katerikoli $x \in X$. Očitno je podprostor. Za zaprtost vzemimo zaporedje $(y_n)_n$ v $\{x\}^\perp$, ki konvergira k $y \in \overline{\{x\}^\perp}$. Potem je

$$\langle y, x\rangle = \langle y - y_n, x\rangle + \langle y_n, x\rangle = \langle y - y_n, x\rangle$$

in zato

$$|\langle y, x\rangle| \leq \|y - y_n\| \, \|x\| \xrightarrow[n\to\infty]{} 0$$

po Cauchy-Schwarzu. Torej je $y \in \{x\}^\perp$. □

Izrek 4.4.13. *Naj bo M zaprt podprostor v Hilbertovem prostoru H. Za $x \in H$ definiramo $Px \in M$ kot tisti vektor, ki je najbližji x med vektorji iz M. Potem velja:*

- *P je linearen operator $H \to M$,*
- *$\|Px\| \leq \|x\|$,*
- *$P^2 = P$,*
- *$\operatorname{im} P = M$ in $\ker P = M^\perp$,*
- *$H = M \oplus M^\perp$ in $M^{\perp\perp} = M$.*

Dokaz. Prva točka: Vzemimo $z \in M$. Potem je

$$\langle (\alpha x + \beta y) - (\alpha Px + \beta Py), z\rangle = \alpha \langle x - Px, z\rangle + \beta \langle y - Py, z\rangle = 0,$$

torej po prejšnjem izreku $P(\alpha x + \beta y) = \alpha Px + \beta Py$.

Za drugo točko izračunamo

$$\|x\|^2 = \|x - Px + Px\|^2 = \|x - Px\|^2 + \|Px\|^2 \geq \|Px\|^2 .$$

Tretja točka je očitna. Za četrto je seveda $\operatorname{im} P = M$, za $x \in \ker P$ velja $x - Px \in M^\perp$, torej ($\ker Px = 0$) tudi $x \in M^\perp$. Če pa je $x \in M^\perp$, je $x = x - 0 \in M^\perp$, torej $Px = 0$ po definiciji P.

Za zadnjo točko razcepimo $x = Px + (x - Px)$. Ker za $A \subseteq H$ vedno velja $A \cap A^\perp \subseteq \{0\}$ in ker je $0 \in M$, je $H = M \oplus M^\perp$. Preslikava $I - P$ je pravokoten projektor na $M^\perp$. Velja $M^{\perp\perp} = \ker(I - P) = \operatorname{im} P = M$. □

Idempotent P iz izreka je pravokotni projektor na M vzdolž $M^\perp$. Množica $M^\perp$ se imenuje ORTOGONALNI KOMPLEMENT M.

Posledica 4.4.14. *Za $A \subseteq H$ je $A^{\perp\perp} = \overline{\operatorname{Lin} A} =: [A]$.*

Dokaz. Seveda je $A \subseteq [A]$. Velja $[A]^\perp \subseteq A^\perp$ in $A^{\perp\perp} \subseteq [A]^{\perp\perp} = [A]$ ter celo $[A]^\perp = A^\perp$, saj za $x \in A^\perp$ velja $x \perp A$ in bo x pravokoten tudi na linearno ogrinjačo A in njeno zaprtje. Torej $A^{\perp\perp} = [A]^{\perp\perp}$. □

Naj bo X prostor s skalarnim produktom in $y \in X$. Definiramo $f_y : X \to \mathbb{F}$ z

$$f_y(x) = \langle x, y \rangle .$$

Lema 4.4.15. *Preslikava f_y je omejen linearen funkcional z $\|f_y\| = \|y\|$.*

Dokaz. Očitno je linearen. Velja

$$\|f_y(x)\| = |\langle x, y \rangle| \leq \|x\| \, \|y\|$$

po Cauchy-Schwarzu. Če je x linearno odvisen od y, velja enakost. □

Izrek 4.4.16 (Riesz)**.** *Naj bo H Hilbertov prostor in $f \in H^*$. Tedaj obstaja natanko en $y \in H$, da je $f(x) = \langle x, y \rangle$ in $\|f\| = \|y\|$.*

Dokaz. Za enoličnost: če je $f_y = f_z$, potem za vsak x

$$\langle x, y - z \rangle = 0,$$

torej $y = z$.

Če je $f = 0$, vzamemo $y = 0$. Sicer $f \neq 0$, in je $\ker f$ zaprt podprostor (praslika zaprte množice), torej $H = \ker f \oplus (\ker f)^\perp$. Obstaja $z \in (\ker f)^\perp$, da je $f(z) = 1$. Za $x \in H$ potem velja

$$x = \underbrace{x - f(x)z}_{\in \ker f} + \underbrace{f(x)z}_{\in (\ker f)^\perp} ,$$

torej $\langle x, z \rangle = \langle f(x)z, z \rangle = f(x) \langle z, z \rangle$. Potem je $f(x) = \left\langle x, z / \|z\|^2 \right\rangle$. □

Posledica 4.4.17. *Za vsak $f \in (l^2)^*$ obstaja natanko en $y \in l^2$, da je*

$$f(x) = \sum_{n=1}^{\infty} x_n \overline{y_n}$$

Trditev 4.4.18. *Preslikava $J : H \to H^*$, podana s predpisom $Jy = f_y$ za $f_y(x) = \langle x, y \rangle$, je poševno linearen izometrični izomorfizem.*

Dokaz. Vemo $\|f_y\| = \|y\|$, torej je J izometrija. Po Rieszovem izreku je surjektivna in injektivna, poševna linearnost pa je preprost račun. □

Izrek 4.4.19. *Naj bo H Hilbertov prostor. Potem je tudi H^* Hilbertov s skalarnim produktom $\langle f, g\rangle_{H^*} = \langle y_g, y_f\rangle_H$.*

Dokaz. Preprosto preverjanje. □

Izrek 4.4.20. *Naj bo H Hilbertov prostor in $K \leq H$ podprostor. Tedaj ima vsak $f \in K^*$ natanko eno Hahn-Banachovo razširitev na H.*

Dokaz. Omejen funkcional $f : K \to \mathbb{F}$ lahko razširimo do $g : \overline{K} \to \mathbb{F}$, pri čemer se ohrani norma. Po Rieszovem izreku obstaja natanko en $y \in \overline{K}$, da je $g(x) = \langle x, y\rangle$ za $x \in \overline{K}$. Definiramo $F(x) = \langle x, y\rangle$ za $x \in H$. Seveda je $F|_K = f$ in $\|F\| = \|y\| = \|f\|$.

Recimo, da je F' še ena Hahn-Banachova razširitev f. Po Rieszu obstaja natanko en $y' \in H$, da je $F'(x) = \langle x, y'\rangle$. To velja tudi za $x \in \overline{K}$, torej $F'(x) = g(x)$, oziroma $\langle x, y'\rangle = \langle x, y\rangle$ za vse $x \in \overline{K}$. Sledi $y - y' \perp \overline{K}$. Potem je

$$\left\|F'\right\|^2 = \|y\|^2 = \left\|y' - y + y\right\|^2 = \left\|y' - y\right\|^2 + \|y\|^2 = \|g\|^2 + \left\|y - y'\right\|^2.$$

Torej $\|y' - y\| = 0$. □

Posledica 4.4.21. *Vsak Hilbertov prostor je refleksiven (vložitev v H^{**} je surjektivna).*

4.4.1 Ortonormirani sistemi

Definicija 4.4.22. Naj bo X prostor s skalarnim produktom. Množica $E \subseteq X$ je ORTONORMIRAN SISTEM, če je $\|e\| = 1$ za vsak $e \in E$ ter $e \perp f$ za vsaka $e, f \in E$.

Opomba. Če velja le druga zahteva, je E ORTOGONALNA MNOŽICA.

Lema 4.4.23. *Vsaka ortogonalna množica je linearno neodvisna.*

Definicija 4.4.24. Naj bo H Hilbertov prostor. Ortonormiran sistem $E \subseteq H$ je KOMPLETEN ali BAZA Hilbertovega prostora H, če je maksimalen v množici vseh ortonormiranih sistemov (glede na $\subseteq$).

Trditev 4.4.25. *Vsak ortonormiran sistem v Hilbertovem prostoru lahko dopolnimo do kompletnega ortonormiranega sistema.*

Dokaz. Če je $(F_\alpha)_\alpha$ veriga v množici vseh ortonormiranih sistemov, ki vsebujejo E, je

$$\bigcup_\alpha F_\alpha$$

zgornja meja za to verigo, ki je očitno ortonormiran sistem. Po Zornovi lemi obstaja kompleten ortonormiran sistem, ki vsebuje E. □

Opomba. Kaj je rumeno in ekvivalentno aksiomu izbire? Zornova limona!

Posledica 4.4.26. *Vsak Hilbertov prostor ima bazo.*

Trditev 4.4.27. *Naj bo $\{e_1, \ldots, e_n\}$ ortonormiran sistem v Hilbertovem prostoru H. Naj bo P_n ortogonalna projekcija na $M_n = \mathrm{Lin}\{e_1, \ldots, e_n\}$. Tedaj za $x \in H$ velja*

$$P_n x = \sum_{k=1}^{n} \langle x, e_k \rangle \, e_k.$$

Dokaz. Naj bo x_0 ta vsota. Tedaj za $1 \leq j \leq n$ velja

$$\langle x_0, e_j \rangle = \sum_{k=1}^{n} \langle x, e_k \rangle \, \langle e_k, e_j \rangle = \langle x, e_j \rangle \,,$$

torej je $x - x_0 \perp M_n$. Po definiciji je $x_0 = P_n x$. □

Trditev 4.4.28 (Besselova neenakost)**.** *Naj bo $(e_n)_n$ števen ortonormiran sistem v prostoru s skalarnim produktom X. Tedaj za vsak $x \in X$ velja*

$$\|x\|^2 \geq \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 \,.$$

Dokaz. Definiramo

$$x' = \sum_{k=1}^{n} \langle x, e_k \rangle \, e_k.$$

Enostavno lahko preverimo $x - x' \perp x'$, torej po Pitagori

$$\|x\|^2 = \|x - x'\|^2 + \|x'\|^2 \geq \|x'\|^2 = \sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \,.$$

To velja za vsak n, torej tudi v limiti. □

Posledica 4.4.29. *Naj bo X prostor s skalarnim produktom in $E \subseteq X$ ortonormiran sistem. Naj bo $x \in X$. Tedaj je $\{e \in E \mid \langle x, e \rangle \neq 0\}$ kvečjemu števna.*

Dokaz. Definiramo

$$E_n = \left\{ e \in E \mid |\langle x, e \rangle| \geq \frac{1}{n} \right\}.$$

Potem je $\langle x, e \rangle \neq 0$ natanko tedaj, ko je $e \in E_n$ za vsak $n \in \mathbb{N}$. Trdimo, da so vsi E_n končni. Sicer obstaja $m \in \mathbb{N}$, da je E_m neskončna, torej vsebuje števno neskončno podmnožico $(e_k)_k$. Potem je

$$\|x\|^2 \geq \sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 = \infty,$$

ker je $|\langle x, e_k \rangle| \geq 1/m$ za vse k. □

Posledica 4.4.30. *Če je $E \subseteq X$ kompleten ortonormiran sistem, za vsak $x \in X$ velja*

$$\|x\|^2 \geq \sum_{e \in E} |\langle x, e \rangle|^2 .$$

Dokaz. Po prejšnji posledici je $\langle x, e \rangle \neq 0$ za kvečjemu štetno mnogo $e \in E$. Na njih uporabimo Besselovo neenakost. □

Izrek 4.4.31. *Za ortonormiran sistem $E \subseteq H$ so naslednje trditve ekvivalentne.*

- *E je kompleten,*
- *$E^\perp = \{0\}$,*
- *$[E] = H$ (zaprta linearna ogrinjača),*
- *za vsak $x \in H$ velja*
$$x = \sum_{e \in E} \langle x, e \rangle \, e,$$
- *za poljubna $x, y \in H$ velja*
$$\langle x, y \rangle = \sum_{e \in E} \langle x, e \rangle \, \langle y, e \rangle ,$$
- *(Parsevalova enakost) za vsak $x \in H$ velja*
$$\|x\|^2 = \sum_{e \in E} |\langle x, e \rangle|^2 .$$

Dokaz. 1 v 2: Recimo, da $E^\perp \neq \{0\}$. Vzamemo $x \in E^\perp \setminus \{0\}$ in definiramo $E \cup \{x/\|x\|\}$, kar je ortonormiran sistem, ki vsebuje E. $\rightarrow\!\!\!\!\!\!\times\!\!\!\!\!\!\leftarrow$

2 v 1: Recimo, da E ni KONS. Potem obstaja KONS E', ki vsebuje E, in obstaja $x \in E' \setminus E$. Ampak $x \in E^\perp = \{0\}$. $\rightarrow\!\!\!\!\!\!\times\!\!\!\!\!\!\leftarrow$

Ekvivalentnost 2 in 3: Velja $[E] = H \Leftrightarrow E^\perp = [E]^\perp = H^\perp = \{0\}$.

2 v 4: Vzemimo $x \in H$. Vemo, da obstaja kvečjemu štetno mnogo $e \in E$, da je $\langle x, e \rangle \neq 0$. Te vektorje oštevilčimo v $(e_n)_n$. Dokažimo, da je

$$x = \sum_{n=1}^{\infty} \langle x, e_n \rangle \, e_n .$$

Vrsta res konvergira, saj za delne vsote s_n in $m > n$ velja

$$\|s_m - s_n\|^2 = \left\| \sum_{k=n+1}^{m} \langle e, e_k \rangle \, e_k \right\|^2 = \sum_{k=n+1}^{m} |\langle x, e_k \rangle|^2 \leq \sum_{k=n+1}^{\infty} |\langle x, e_k \rangle|^2 .$$

Ker ta vrsta konvergira po Besselovi neenakosti, je zaporedje $(s_n)_n$ Cauchyjevo v H. Zato $s_n \to x_0 \in H$. Ker je

$$\langle x_0, e_j \rangle = \sum_{k=1}^{\infty} \langle x, e_k \rangle \langle e_k, e_j \rangle = \langle x, e_j \rangle ,$$

velja $x - x_0 \perp e_j$ za vsak j, torej $x = x_0$.

4 v 5: Preprost račun.

5 v 6: Preprost račun.

6 v 2: Če je $E^\perp \neq \{0\}$, obstaja $x \in E^\perp$, različen od 0. Po Parsevalu za x in E velja

$$0 \neq \|x\|^2 = \sum_{e \in E} |\langle x, e \rangle|^2 = 0,$$

kar je protislovno. $\rightarrow\!\!\!\!\times\!\!\!\!\leftarrow$ □

Trditev 4.4.32. *Poljubni ortonormirani bazi Hilbertovega prostora imata isto kardinalnost.*

Dokaz. Naj bosta $E, F \subseteq H$ bazi. Če je $|E| < \infty$, rezultat vemo iz linearne algebre. Sicer za $e \in E$ tvorimo $F_e = \{f \in F \mid \langle e, f \rangle \neq 0\}$. Ta množica je kvečjemu številno neskončna, po drugi strani pa velja

$$F = \bigcup_{e \in E} F_e,$$

saj je E baza. Torej $|F| \leq |E|\,|\mathbb{N}| = |E|$ in podobno v drugo smer. □

Definicija 4.4.33. Dimenzija Hilbertovega prostora je enaka kardinalnosti katerekoli njene baze.

Lema 4.4.34. *V separabilnem metričnem prostoru je vsaka družina paroma disjunktnih odprtih krogel kvečjemu številno neskončna.*

Dokaz. Naj bo $\{\mathring{B}(x_i, \varepsilon_i) \mid i \in I\}$ družina paroma disjunktnih odprtih krogel. Naj bo S številna gosta množica v danem metričnem prostoru. Zaradi gostosti je $\mathring{B}(x_i, \varepsilon_i) \cap S \neq \varnothing$ za vse i. Izberimo $y_i \in \mathring{B}(x_i, \varepsilon_i) \cap S$ in definiramo $\varphi : I \to S$ z $\varphi(i) = y_i$. Ker so krogle med seboj disjunktne, je φ injekcija, torej $|I| \leq |\mathbb{N}|$. □

Trditev 4.4.35. *Neskončnorazsežen Hilbertov prostor je separabilen natanko tedaj, ko je* $\dim H = \aleph_0$.

Dokaz. V desno: Naj bo E KONS v H. Za $e, e' \in E$ velja

$$\left\|e - e'\right\|^2 = \left\langle e - e', e - e'\right\rangle = \|e\|^2 - 2\operatorname{Re}\left\langle e, e'\right\rangle + \left\|e'\right\|^2 = 2.$$

Tvorimo $S = \{\mathring{B}(e, \frac{\sqrt{2}}{2}) \mid e \in E\}$. Te krogle so paroma disjunktne, torej imamo injekcijo $\varphi : E \to S$. Po prejšnji lemi je S največ števna, torej $|E| \leq \aleph_0$.

V levo: Ker je $\dim H = \aleph_0$, obstaja števen KONS $(e_n)_n$. Vsak $x \in X$ lahko razvijemo v Fourierovo vrsto

$$x = \sum_{n=1}^{\infty} \langle x, e_n\rangle\, e_n.$$

Za poljuben $\varepsilon > 0$ potem obstaja n_ε, da za vse $n \geq n_\varepsilon$ velja

$$\left\| x - \sum_{k=1}^{n} \langle x, e_k\rangle\, e_k \right\| < \varepsilon.$$

Skalarne produkte $\langle x, e_k\rangle$ lahko aproksimiramo z $\lambda_k \in \mathbb{Q}$, če je $\mathbb{F} = \mathbb{R}$, oziroma $\lambda_k \in \mathbb{Q} + i\mathbb{Q}$, če je $\mathbb{F} = \mathbb{C}$. V obeh primerih zahtevamo $|\langle x, e_k\rangle - \lambda_k| < \varepsilon/n$. Tako dobimo števno gosto množico

$$\left\{ \sum_{k=1}^{n} \lambda_k e_k \mid n \in \mathbb{N}, \lambda_k \in \mathbb{Q} + i\mathbb{Q} \right\}. \qquad \square$$

Definicija 4.4.36. Linearna preslikava $U : H \to K$ je IZOMORFIZEM Hilbertovih prostorov (tudi UNITARNI OPERATOR), če je surjektivna in če velja

$$\langle Ux, Uy\rangle = \langle x, y\rangle .$$

Trditev 4.4.37. *Naj bo $U : X \to Y$ linearna izometrija med prostoroma s skalarnim produktom. Tedaj U ohranja skalarni produkt.*

Dokaz. Računamo

$$\|U(x+y)\|^2 = \|Ux\|^2 + \|Uy\|^2 + 2\operatorname{Re}\langle Ux, Uy\rangle$$
$$\|U(x+y)\|^2 = \|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\operatorname{Re}\langle Ux, Uy\rangle$$

torej $\operatorname{Re}\langle x, y\rangle = \operatorname{Re}\langle Ux, Uy\rangle$. Če namesto x pišemo ix, dobimo še $-\operatorname{Im}\langle x, y\rangle = -\operatorname{Im}\langle Ux, Uy\rangle$. $\square$

Lema 4.4.38. *Naj bo $U : H \to K$ izomorfizem Hilbertovih prostorov. Potem U slika KONS v KONS.*

Dokaz. Naj bo $\{e_i\}_{i\in I}$ KONS za H. Potem je $\langle Ue_i, Ue_j\rangle = \langle e_i, e_j\rangle = \delta_{ij}$. Če je $y \in K$ tak, da je $y \perp Ue_i$ za vse i, potem je $0 = \langle y, Ue_i\rangle = \left\langle U^{-1}y, e_i\right\rangle$, torej $y = 0$. $\square$

Izrek 4.4.39. *Hilbertova prostora H in K sta izomorfna natanko tedaj, ko je* $\dim H = \dim K$.

Dokaz. V desno sledi iz leme. V levo: Dokazali bomo, da je $H \cong l^2(I)$, kjer je $(e_i)_{i\in I}$ KONS za H. Definiramo $U : H \to l^2(I)$ z $Ux = \hat{x}$, kjer je

$$\hat{x} : i \mapsto \langle x, e_i \rangle .$$

Potem je U linearna izometrija, saj

$$\|Ux\|^2 = \sum_{i\in I} |\hat{x}(i)|^2 = \sum_{i\in I} |\langle x, e_i\rangle|^2 = \|x\|^2$$

po Parsevalovi enakosti, hkrati pa je U tudi surjektivna, saj za $f \in l^2(I)$ lahko definiramo

$$x = \sum_{i\in I} f(i)e_i,$$

kar konvergira, ker je $f \in l^2(I)$. Seveda $\hat{x} = f$. Bijekcijo med indeksnima množicama lahko razširimo do izomorfizma Hilbertovih prostorov. □

Posledica 4.4.40. *Neskončnorazsežen separabilen Hilbertov prostor je izomorfen l^2.*

Naj bosta H in K Hilbertova prostora. Produkt $H\times K$ opremimo s skalarnim produktom

$$\langle (x_1, y_1), (x_2, y_2)\rangle = \langle x_1, x_2\rangle + \langle y_1, y_2\rangle ,$$

s čimer je prostor $H \times K$ poln, torej Hilbertov. Označimo $H \times K = H \oplus K$, konstrukciji pravimo ORTOGONALNA DIREKTNA VSOTA.

Za neskončne direktne vsote množico

$$\left\{ (x_n)_n \,|\, x_n \in H_n, \sum_{n=1}^{\infty} \|x_n\|^2 < \infty \right\},$$

kjer so H_n Hilbertovi prostori, opremimo s skalarnim produktom

$$\langle (x_n)_n, (y_n)_n\rangle = \sum_{n=1}^{\infty} \langle x_n, y_n\rangle .$$

Dobljen prostor označimo z

$$\bigoplus_{n=1}^{\infty} H_n,$$

in ga imenujemo ORTOGONALNA DIREKTNA VSOTA prostorov $(H_n)_n$. Je tudi Hilbertov prostor.

4.4.2 Stone-Weierstrassov izrek

Lema 4.4.41. *Naj bo X neprazna množica in V vektorski prostor funkcij na X, ki loči točke in vsebuje konstante. Tedaj za vsaka različna $x, y \in X$ in $\alpha, \beta \in \mathbb{R}$ obstaja $f \in V$, da je $f(x) = \alpha$ ter $f(y) = \beta$.*

Lema 4.4.42. *Naj bo K kompakten Hausdorffov prostor in $V \subseteq \mathcal{C}(K)$ realen vektorski prostor, ki je podmreža v $\mathcal{C}(K)$. Naj V vsebuje konstante in loči točke na K. Tedaj za vsak $g \in \mathcal{C}(K)$, $a \in K$ ter $\varepsilon > 0$ obstaja $f \in V$, da je $f(a) = g(a)$ in $f(x) > g(x) - \varepsilon$ za vse $x \in K$.*

Dokaz. Po prejšnji lemi za vsak $x \in K$ obstaja $f_x \in V$, da je $f_x(a) = g(a)$ in $f_x(x) = g(x)$. Zaradi zveznosti zato obstaja odprta okolica $U_x \ni x$, da je $f_x(y) > g(y) - \varepsilon$ za vse $y \in U_x$. Dobimo odprto pokritje K, zaradi kompaktnosti obstaja končno podpokritje $U_{x_1} \cup \ldots \cup U_{x_n} = K$. Definiramo $f = \max\{f_{x_1}, \ldots, f_{x_n}\}$. □

Izrek 4.4.43 (mrežni Stone-Weierstrass)**.** *Naj bo K kompakten Hausdorffov prostor ter V realen vektorski podprostor $\mathcal{C}(K)$, ki je podmreža, loči točke K in vsebuje konstante. Tedaj je V gost v $\mathcal{C}(K)$ glede na supremum normo.*

Dokaz. Naj bo $g \in \mathcal{C}(K)$ in $\varepsilon > 0$. Iščemo $f \in V$, da je $\|f - g\|_\infty < \varepsilon$, oziroma $f - g < \varepsilon$ in $g - f < \varepsilon$. Po prejšnji lemi za vsak $x \in K$ obstaja $f_x \in V$, da je $f_x > g - \varepsilon$ in $f_x(x) = g(x)$. Zaradi zveznosti obstaja odprta okolica V_x za x, da je $|f_x(y) - g(y)| < \varepsilon$ za vse $y \in V_x$. Množice V_x tvorijo pokritje za K, torej obstaja končno podpokritje $V_{x_1} \cup \ldots \cup V_{x_n}$. Potem definiramo $f = \min\{f_{x_1}, \ldots, f_{x_n}\}$. Za vsak $m = 1, \ldots, n$ velja $f_{x_m} > g - \varepsilon$, torej $f > g - \varepsilon$. Za poljuben $y \in K$ obstaja m, da je $y \in V_m$, torej $f(y) \leq f_{x_m}(y) < g(y) + \varepsilon$. □

Lema 4.4.44. *Obstaja zaporedje polinomov, ki na $[0, 1]$ konvergira enakomerno proti funkciji $\sqrt{x}$.*

Dokaz. Funkcijo $x \mapsto 1 - \sqrt{1 - x}$ razvijemo v Taylorjevo vrsto na $[0, 1)$. Velja

$$1 - (1 - x)^{1/2} = 1 - \sum_{k=0}^{\infty} \binom{1/2}{k} (-1)^k x^k.$$

S $S_n(x)$ označimo n-to delno vsoto zgornje vrste. Za $x \in (0, 1)$ je zaporedje $(S_n(x))_n$ strogo naraščajoče in velja $\lim S_n(x) \leq 1$, saj je $1 - (1 - x)^{1/2} \leq 1$ za te x. Ker so S_n zvezni, je

$$S_n(x) = \lim_{x \to 1} S_n(x) \leq 1,$$

torej je zaporedje $(S_n(1))_n$ naraščajoče in navzgor omejeno. Definirajmo $f(x) = \lim S_n(x)$. Na $[0, 1)$ se f in $1 - \sqrt{1 - x}$ ujemata, hkrati pa je za $x \in [0, 1]$ tudi

$$0 \leq f(x) - S_n(x) = \sum_{k=n+1}^{\infty} \binom{1/2}{k} (-1)^k x^k \leq \sum_{k=n+1}^{\infty} \binom{1/2}{k} (-1)^k = f(1) - S_n(1) \xrightarrow[n \to \infty]{} 0.$$

Torej S_n konvergira k f enakomerno na $[0,1]$, in je zato f zvezna in se na tem intervalu ujema z $1-\sqrt{1-x}$. Dokazali smo, da obstaja zaporedje polinomov, ki konvergirajo k f enakomerno. Enostavno ga lahko transformiramo v zaporedje, ki enakomerno konvergira k $\sqrt{x}$. □

Izrek 4.4.45 (Stone-Weierstrass, realna verzija)**.** *Naj bo K kompakten Hausdorffov prostor in $A \subseteq \mathcal{C}(K)$ podalgebra, ki loči točke in vsebuje konstante. Tedaj je A gosta v $\mathcal{C}(K)$.*

Dokaz. Pokazali bomo, da je $\overline{A}$ podmreža v $\mathcal{C}(K)$. Potem bo po mrežni različici izreka gosta v $\mathcal{C}(K)$. Ker bo zaprta, bo $\overline{A} = \mathcal{C}(K)$.

Očitno je $\overline{A}$ podalgebra. Ker velja $\max\{f,g\} = \frac{1}{2}(f+g+|f-g|)$ in $\min\{f,g\} = \frac{1}{2}(f+g-|f-g|)$, je dovolj pokazati, da je $\overline{A}$ zaprta za absolutne vrednosti. Naj bo $f \in \overline{A}$. Oglejmo si $g = f/\|f\|_\infty$. Ker je $g^2 \leq 1$, in ker po lemi obstaja zaporedje polinomov $(p_n)_n$, ki konvergirajo enakomerno na $[0,1]$ proti $\sqrt{x}$, velja $(p_n \circ g^2)(x) \to |g(x)|$ enakomerno. Torej $|g| \in \overline{A}$, ampak $|g| = |f|/\|f\|_\infty$, torej tudi $|f| \in \overline{A}$. □

Izrek 4.4.46 (Weierstrass)**.** *Naj bo $K \subseteq \mathbb{R}$ kompakt in $f : K \to \mathbb{R}$ zvezna. Tedaj za vsak $\varepsilon > 0$ obstaja polinom p, da je $|f(x) - p(x)| < \varepsilon$ za vse $x \in K$.*

Izrek 4.4.47 (Stone-Weierstrass, kompleksna verzija)**.** *Naj bo K kompakten Hausdorffov prostor in $A \subseteq \mathcal{C}(K)$ podalgebra v algebri kompleksnih funkcij, ki*

- *loči točke K,*
- *vsebuje konstante,*
- *je sebi-adjungirana: $f \in A$ pomeni $\overline{f} \in A$.*

Tedaj je A gosta v $\mathcal{C}(K)$.

Dokaz. Naj bo A_0 algebra realnih funkcij, ki so vsebovane v A. Očitno A_0 vsebuje konstante. Za $x, y \in K$ obstaja $f \in A$, da je $f(x) \neq f(y)$, torej ena od realnih funkcij $\frac{f+\overline{f}}{2}, \frac{f-\overline{f}}{2i} \in A$ loči točki x, y. Torej A loči točke. Po realni verziji izreka je A_0 gosta v $\mathcal{C}(K, \mathbb{R})$.

Naj bo $f \in \mathcal{C}(K, \mathbb{C})$ poljubna in $\varepsilon > 0$. Pišimo $f = g + ih$ za realni g, h. Po razno dokazanem obstajata $g_1, h_1 \in A_0$, da je $\|g - g_1\| < \varepsilon/2$ in $\|h - h_1\| < \varepsilon/2$. Potem je $\|f - (g_1 + ih_1)\| < \varepsilon$. □

4.5 Omejeni operatorji med Hilbertovimi prostori

Definicija 4.5.1. Naj bosta H, K Hilbertova prostora. Preslikava $u : H \times K \to \mathbb{F}$ se imenuje SESKVILINEARNA FORMA, če velja

- $u(\alpha x + \beta y, z) = \alpha u(x,z) + \beta u(y,z)$,

- $u(x, \alpha y + \beta z) = \overline{\alpha} u(x, y) + \overline{\beta} u(x, z)$.

Definicija 4.5.2. Seskvilinearna forma je ZVEZNA (ali OMEJENA), če obstaja $M \geq 0$, da $|u(x, y)| \leq M \|x\| \|y\|$.

Primer. Za $A \in B(H, K)$ definiramo $u(x, y) = \langle Ax, y \rangle$. To je seskvilinearna forma, velja $|u(x, y)| \leq \|A\| \|x\| \|y\|$.

Izrek 4.5.3. *Naj bo $u : H \times K \to \mathbb{F}$ omejena seskvilinearna forma. Potem obstajata natanko določeni $A \in B(H, K)$ in $B \in B(K, H)$, da*

$$u(x, y) = \langle Ax, y \rangle = \langle x, By \rangle .$$

Dokaz. Fiksiramo $x \in H$ in definiramo $f_x : K \to \mathbb{F}$ z $f_x(y) = \overline{u(x, y)}$. To je očitno omejen linearen funkcional na K. Po Rieszovem izreku obstaja natanko določen $z_x \in K$, da je

$$\overline{u(x, y)} = f_x(y) = \langle y, z_x \rangle$$

za vsak $y \in K$. Definiramo $z_x = Ax$. To nam da preslikavo $A : H \to K$, da je $\overline{u(x, y)} = \langle y, Ax \rangle = \overline{Ax, y}$. Enostavno lahko preverimo, da je A linearna, ker velja

$$\|Ax\| = \|z_x\| = \|f_x\| \leq M \|x\| ,$$

pa je tudi omejena (tu M pride iz definicije omejene seskvilinearne forme). Če je $\langle Ax, y \rangle = \langle A'x, y \rangle$, potem velja $\langle Ax - A'x, y \rangle = 0$ za vsak y, torej $Ax = A'x$ za vse x. Sledi, da je A enolično določena. Podobno za B. □

Definicija 4.5.4. Naj bo A omejen operator $H \to K$. Tedaj operatorju $B : H \to K$, za katerega velja $\langle Ax, y \rangle = \langle x, By \rangle$, pravimo ADJUNGIRANI OPERATOR operatorja A. Označimo ga z A^*.

Trditev 4.5.5. *Preslikava $U \in B(H, K)$ je izomorfizem Hilbertovih prostorov natanko tedaj, ko je U obrnljiv in $U^* = U^{-1}$.*

Dokaz. V levo: Če je U obrnljiv, je surjektiven. Potem je

$$\langle Ux, Uy \rangle = \langle x, U^*Uy \rangle = \left\langle x, U^{-1}Uy \right\rangle = \langle x, y \rangle ,$$

torej U ohranja skalarni produkt in je zato izomorfizem.

V desno: Če je U izomorfizem, je obrnljiv. Velja

$$\langle x, y \rangle = \langle Ux, Uy \rangle = \langle x, U^*Uy \rangle$$

za poljubna x, y, torej je $U^* = U^{-1}$. □

Opomba. Če sta $A, B \in B(H, K)$, potem velja

- $(A + B)^* = A^* + B^*$,

- $(\alpha A)^* = \overline{\alpha} A^*$,
- $A^{**} = A$.

Zadnje je res, ker je

$$\langle Ax, y\rangle = \langle x, A^*y\rangle = \overline{\langle A^*y, x\rangle} = \overline{\langle y, A^{**}x\rangle} = \langle A^{**}x, y\rangle\,.$$

Opomba. Preslikava $i : A \mapsto A^*$ je involucija.

Trditev 4.5.6. *Naj bosta $A \in B(H,K)$ in $B \in B(K,L)$ omejena operatorja. Tedaj $(BA)^* = A^*B^*$.*

Dokaz. Preprost račun. □

Posledica 4.5.7. *Operator $A \in B(H,K)$ je obrnljiv natanko tedaj, ko je $A^* \in B(K,H)$ obrnljiv. Če je A obrnljiv, velja $(A^*)^{-1} = (A^{-1})^*$.*

Dokaz. Operator A je obrnljiv natanko tedaj, ko obstaja $B \in B(K,H)$, da velja $AB = I_K$ in $BA = I_H$. Potem je

$$(AB)^* = B^*A^* = I_K,$$
$$(BA)^* = A^*B^* = I_H,$$

torej je A^* obrnljiv in $(A^*)^{-1} = B^* = (A^{-1})^*$. Podobno v drugo smer, kjer upoštevamo še $A^{**} = A$. □

Trditev 4.5.8. *Če je $A \in B(H,K)$, je $\ker A^* = (\operatorname{im} A)^\perp$.*

Dokaz. Velja

$$x \in \ker A^* \Leftrightarrow A^*x = 0 \Leftrightarrow \forall y.\, \langle A^*x, y\rangle = 0 \Leftrightarrow \forall y.\, \langle x, Ay\rangle = 0 \Leftrightarrow x \in (\operatorname{im} A)^\perp.$$ □

Posledica 4.5.9. *Za $A \in B(H,K)$ velja*

- $\ker A = (\operatorname{im} A^*)^\perp$,
- $(\ker A)^\perp = \overline{\operatorname{im} A^*}$,
- $(\ker A^*)^\perp = \overline{\operatorname{im} A}$.

Dokaz. Prva točka je očitna. Če na njej uporabimo komplement in upoštevamo $X^{\perp\perp} = \overline{\operatorname{Lin} X}$, dobimo drugo točko. Podobno za tretje. □

Definicija 4.5.10. Operator $A \in B(H)$ je SEBI ADJUNGIRAN, če je $A^* = A$. Je NORMALEN, če je $A^*A = AA^*$. Je UNITAREN, če je $A^*A = AA^* = I$.

Opomba. Operator A je unitaren natanko tedaj, ko je A izomorfizem prostora H.

Opomba. Vsak unitaren operator je normalen.

Opomba. Sebi adjungiran operator je vedno normalen.

Če je $A \in B(H)$, lahko definiramo

$$\operatorname{Re} A = \frac{A + A^*}{2},$$
$$\operatorname{Im} A = \frac{A - A^*}{2i},$$

tako dobimo $A = \operatorname{Re} A + i \operatorname{Im} A$. Oba ta operatorja sta sebi adjungirana.

Trditev 4.5.11. *Naj bo H kompleksen Hilbertov prostor in $A \in B(H)$. Tedaj je $A = A^*$ natanko tedaj, ko je $\langle Ax, x\rangle \in \mathbb{R}$ za vsak $x \in H$.*

Dokaz. V desno: $\langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle}$.

V levo: Po predpostavki je $\langle A(x+y), x+y\rangle \in \mathbb{R}$. Velja

$$\langle A(x+y), x+y\rangle = \langle Ax, x\rangle + \langle Ay, y\rangle + \langle Ax, y\rangle + \langle Ay, x\rangle ,$$

torej $\langle Ax, y\rangle + \langle Ay, x\rangle \in \mathbb{R}$ za vsaka x, y. Pišimo $\langle Ax, y\rangle = \alpha + i\beta$ in $\langle Ay, x\rangle = \gamma - i\beta$.

Če menjamo $y \to iy$, dobimo $i \langle Ay, x\rangle - i \langle Ax, y\rangle \in \mathbb{R}$. To je enako

$$i \langle Ay, x\rangle - i \langle Ax, y\rangle = i(\gamma - i\beta - \alpha - i\beta) = i(\gamma - \alpha) + 2\beta,$$

torej $\gamma = \alpha$. □

Izrek 4.5.12. *Naj bo H Hilbertov prostor in $A \in B(H)$ sebi adjungiran operator. Tedaj velja*

$$\|A\| = w(A) := \sup_{\|x\|=1} |\langle Ax, x\rangle| .$$

Dokaz. Velja $|\langle Ax, x\rangle| \le \|Ax\| \, \|x\| \le \|A\| \, \|x\|^2$, torej je $w(A) \le \|A\|$. Računamo

$$\begin{aligned} \langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle &= 2 \langle Ax, y\rangle + 2 \langle Ay, x\rangle \\ &= 2 \langle y, Ax\rangle + 2 \langle Ax, y\rangle \\ &= 4 \operatorname{Re} \langle Ax, y\rangle , \end{aligned}$$

torej

$$\begin{aligned} 4 \, |\operatorname{Re} \langle Ax, y\rangle| &\le |\langle A(x+y), x+y\rangle| - |\langle A(x-y), x-y\rangle| \\ &\le w(A) \left(\|x+y\|^2 + \|x-y\|^2 \right) \\ &= 2w(A) \left(\|x\|^2 + \|y\|^2 \right) \end{aligned}$$

oziroma

$$|\mathrm{Re}\,\langle Ax, y\rangle| \leq \frac{1}{2} w(A) \left(\|x\|^2 + \|y\|^2 \right).$$

Izberemo x z $\|x\| = 1$. Če je $A = 0$, izrek velja, sicer pa lahko vzamemo tak x, da $Ax \neq 0$. Potem nastavimo $y = \frac{Ax}{\|Ax\|}$, in dobimo

$$\left| \mathrm{Re} \left\langle Ax, \frac{Ax}{\|Ax\|} \right\rangle \right| \leq w(A),$$

zato $\|Ax\|^2 \leq w(A)$ za vsak x z $\|x\| = 1$ in $Ax \neq 0$. Sedaj lahko naredimo supremum po enotski sferi. □

Trditev 4.5.13. *Naj bo H kompleksen Hilbertov prostor in $A \in B(H)$ tak, da je $\langle Ax, x\rangle = 0$ za vsak $x \in H$. Tedaj je $A = 0$.*

Dokaz. Ker je H kompleksen, je $A = A^*$. Po izreku je

$$\|A\| = \sup_{\|x\|=1} |\langle Ax, x\rangle| = 0.$$ □

Izrek 4.5.14. *Za $A \in B(H)$ velja $\|A^*A\| = \|A\|^2$.*

Dokaz. Računamo

$$\|A^*A\| = \sup_{\|x\|=1} |\langle A^*Ax, x\rangle| = \sup_{\|x\|=1} |\langle Ax, Ax\rangle| = \sup_{\|x\|=1} \|Ax\|^2 = \|A\|^2.$$ □

Opomba. Tej enakosti pravimo C^*-aksiom.

5 Statistika 2

5.1 Ocenjevanje v linearnih modelih

Splošni linearni model je oblike $X = Z\beta + \varepsilon$, kjer je $Z \in \mathbb{R}^{n\times d}$ znana konstantna matrika, $\beta \in \mathbb{R}^d$ neznani parameter, ε pa je neopazljiv slučajni šum. Privzamemo, da velja $E(\varepsilon) = 0$. V splošnem za varianco ε ne privzamemo ničesar, v standardnih linearnih regresijskih modelih pa privzamemo, da je diagonalna.

Privzemimo splošni linearni model in naj bo B vektorski podprostor v $\mathbb{R}^d$. Naj bo $x \in \mathbb{R}^n$ realizacija slučajnega vektorja X. RESTRINGIRANA OCENA za $\beta \in B$ na podlagi x po metodi najmanjših kvadratov je tak vektor $\hat{\beta}_B$, za katerega je

$$\left\| x - Z\hat{\beta}_B \right\|^2 = \min_{b\in B} \|x - Zb\|^2 .$$

Pišimo $\hat{\beta} = \hat{\beta}_B$. Vemo, da je $Z\hat{\beta}$ ravno pravokotna projekcija vektorja x na podprostor $ZB \subseteq \mathbb{R}^n$. Določena je z zahtevo $x - Z\hat{\beta} \perp ZB$. Za $B = \mathbb{R}^d$ to velja natanko v primeru $Z^T(X - Z\hat{\beta}) = 0$, torej $Z^TZ\hat{\beta} = Z^Tx$. V primeru, ko je Z polnega ranga, je Z^TZ obrnljiva in $\hat{\beta} = (Z^TZ)^{-1}Z^Tx$. Če pa je jedro Z netrivialno, imamo rešitev več.

Če na stolpcih Z izvedemo prirejeno Gram-Schmidtovo ortogonalizacijo, lahko kljub temu poiščemo rešitev. Označimo s S_i rezultat ortogonalizacije na i-tem stolpcu Z. V primeru, ko je ta stolpec v linearni ogrinjači prejšnjih, nastavimo $S_i = 0$. S tem dobimo ortogonalne vektorje $S_1, \ldots, S_d$. Dobimo razcep $Z = SP$, kjer je S matrika iz zloženih stolpcev S_i, P pa zgornje trikotna matrika s pozitivnimi števili na diagonali, in zato obrnljiva. Velja $Z^TZ = P^TJP$, kjer je J diagonalna matrika, na diagonali katere so kvadrati norm stolpcev S_i. S tem lahko definiramo posplošen inverz $(Z^TZ)' = P^{-1}JP^{-T}$. Potem $\hat{\beta} = (Z^TZ)'Z^Tx$ reši enačbo $Z^TZ\hat{\beta} = Z^Tx$.

Opomba. Vsaki matriki M, ki ustreza $Z^TZMZ^T = Z^T$ pravimo POSPLOŠEN INVERZ matrike Z^TZ.

Izračunajmo

$$E(\hat{\beta}(X)) = (Z^TZ)'Z^TE(X) = (Z^TZ)'Z^TZ\beta = P^TJP^{-T}\beta.$$

Če Z nima polnega ranga, potem $\hat{\beta}$ ni nepristranska cenilka za β. V tem primeru pravzaprav ne obstaja nepristranska linearna cenilka za β, namreč, če je $U : \mathbb{R}^n \to \mathbb{R}^d$ taka, mora veljati $E(UX) = \beta$, hkrati pa $E(UX) = UZ\beta$, iz česar sledi, da je Z injektivna, torej matrika polnega ranga.

Trditev 5.1.1. *Naj bo $e(\beta, \varepsilon) = L\beta$ ocenjevana funkcija, kjer je $L : \mathbb{R}^n \to \mathbb{R}^m$ linearna preslikava. Dalje naj bo $U : \mathbb{R}^n \to \mathbb{R}^m$ nepristranska linearna cenilka za $L\beta$. Tedaj je $L = UZ$ in je $UZ\hat{\beta}$ nepristranska linearna cenilka za $L\beta$.*

Dokaz. Kot zgoraj je $L\beta = UZ\beta$ za vse β. Izračunajmo

$$Z\hat{\beta}(X) = Z(Z^TZ)'Z^TX = SJS^TX = SS^TX.$$

Sledi

$$E(UZ\hat{\beta}(X)) = USS^TZ\beta = USS^TSP\beta = USP\beta = UZ\beta. \qquad \square$$

Izrek 5.1.2 (Gauss-Markov)**.** *Privzemimo linearni regresijski model $X = Z\beta + \varepsilon$, kjer je* $\operatorname{var}(\varepsilon) = \sigma^2 I$*. Naj bo $U : \mathbb{R}^n \to \mathbb{R}^m$ linearna preslikava. Tedaj ima $UZ\hat{\beta}$ med vsemi nepristranskimi linearnimi cenilkami za $UZ\beta$ enakomerno najmanjšo disperzijo.*

Dokaz. Naj bo $W : \mathbb{R}^n \to \mathbb{R}^m$ druga nepristranska linearna cenilka za $UZ\beta$. Po prejšnji trditvi je $WZ = UZ$. Primerjati želimo $\operatorname{var}(WX)$ in $\operatorname{var}(UZ\hat{\beta}) = \operatorname{var}(WZ\hat{\beta})$ upoštevaje $UZ = WZ$.

Velja $\operatorname{var}(WX) = W \operatorname{var}(X) W^T = \sigma^2 WW^T$ in podobno kot v prejšnjem dokazu

$$\operatorname{var}(WZ\hat{\beta}) = \operatorname{var}(WSS^T X) = \sigma^2 WSS^T W^T.$$

Trdimo, da za poljuben $\xi \in \mathbb{R}^m$ velja

$$\left\langle WW^T\xi, \xi \right\rangle \geq \left\langle WSS^T W^T \xi, \xi \right\rangle.$$

Upoštevaje $\left\langle WW^T\xi, \xi \right\rangle = \left\langle W^T\xi, W^T\xi \right\rangle$ in $\left\langle WSS^T W^T\xi, \xi \right\rangle = \left\langle S^T W^T\xi, S^T W^T\xi \right\rangle$ je dovolj za poljuben w pokazati

$$\langle w, w \rangle \geq \left\langle S^T w, S^T w \right\rangle.$$

To velja, ker je $\|S_i\| \in \{0, 1\}$. □

5.1.1 Ocenjevanje v normalnem linearnem regresijskem modelu

Privzamemo $X = Z\beta + \varepsilon$ za $\varepsilon \sim N(0, \sigma^2 I)$. To je parametričen model, vendar β in σ^2 porazdelitve ne določata enolično, če Z nima polnega ranga.

Če dodatno zahtevamo $x - Z\hat{\beta}(x) \perp \operatorname{im} Z$, pa je vsaj $Z\hat{\beta}(x)$ enolično določen. Izračunajmo

$$\|x - Z\beta\|^2 = \|x\|^2 - 2\langle x, Z\beta \rangle + \|Z\beta\|^2 = \left\|x - Z\hat{\beta} + Z\hat{\beta}\right\|^2 - 2\langle x, Z\beta \rangle + \|Z\beta\|^2.$$

Upoštevaje pravokotnost je to enako

$$\|x - Z\beta\|^2 = \left\|x - Z\hat{\beta}\right\|^2 + \left\|Z\hat{\beta}\right\|^2 - 2\langle x, Z\beta \rangle + \|Z\beta\|^2$$

oziroma

$$\|x - Z\beta\|^2 = \left\|x - Z\hat{\beta}\right\|^2 + \left\|Z(Z^T Z)' Z^T x\right\|^2 - 2\left\langle Z^T x, \beta \right\rangle + \|Z\beta\|^2.$$

Ker je gostota porazdelitve enaka

$$f_X(x) = \left(2\pi\sigma^2\right)^{-n/2} \exp\left(\frac{-1}{2\sigma^2} \|x - Z\beta\|^2\right),$$

torej $\left\|x - Z\hat{\beta}\right\|^2$ in $Z^T x$ tvorita zadostno statistiko za dano porazdelitev. Označimo $T(x) = \left(Z^T x, \left\|x - Z\hat{\beta}\right\|^2\right)$. Drugemu členu dvojice označimo z $\operatorname{SSR}(x)$.

Posledica 5.1.3. *Statistike, ki so od vzorca odvisne le preko T, so avtomatično nepristranske cenilke z enakomerno najmanjšo disperzijo za svojo pričakovano vrednost.*

Izrek 5.1.4. *Statistika $U(X) = (Z\hat{\beta}(X), \frac{1}{n-r}\,\mathrm{SSR}(X))$, kjer je $r = \operatorname{rang} Z$, je nepristranska cenilka za $(Z\beta, \sigma^2)$, ki ima med vsemi nepristranskimi cenilkami enakomerno najmanjšo disperzijo. Dalje sta $Z\hat{\beta}(X)$ in $\mathrm{SSR}(X)$ neodvisni, ter $\mathrm{SSR}(X)/\sigma^2 \sim \chi^2_{n-r}$.*

Dokaz. Vemo že, da je $Z\hat{\beta}$ nepristranska cenilka z enakomerno najmanjšo disperzijo. Če je $Y = (Y_1, \ldots, Y_n) \sim N(\nu, \sigma^2 I)$, potem je vsaka funkcija $(Y_1, \ldots, Y_m)$ neodvisna od vsake funkcije $(Y_{m+1}, \ldots, Y_n)$. Če je še $\nu_1 = \cdots = \nu_m = 0$, je

$$\frac{1}{\sigma^2} \sum_{i=1}^{m} Y_i^2 \sim \chi^2_m.$$

Konstruirajmo ortogonalno matriko $\tilde{S} \in O(n)$ na sledeči način. Postavimo $\tilde{S}_i = S_i$ za tiste i, za katere je $\|S_i\| = 1$, za ostale stolpce pa te vrednosti dopolnimo do ortonormirane baze. Velja

$$\mathrm{SSR}(X) = \left\| X - Z\hat{\beta}(X) \right\|^2 = \left\| \tilde{S}^T X - \tilde{S}^T S S^T X \right\|^2 = \left\| \tilde{S}^T X - \tilde{S}^T S S^T \tilde{S} \tilde{S}^T X \right\|$$

ker je $\tilde{S}$ ortogonalna. Izračunamo

$$\mathrm{SSR}(X) = \left\| (I - \begin{bmatrix} J \\ 0 \end{bmatrix} \begin{bmatrix} J & 0 \end{bmatrix}) \tilde{S}^T X \right\|^2 = \left\| \begin{bmatrix} I - J & \\ & I \end{bmatrix} \tilde{S}^T X \right\|^2,$$

torej je $\mathrm{SSR}(X)/\sigma^2 \sim \chi^2_{n-r}$. Takoj se prepričamo, da je $Z\hat{\beta}(X) = SS^T X$ odvisen od preostalih r komponent vektorja $\tilde{S}^T X$, torej sta $Z\hat{\beta}(X)$ in $\mathrm{SSR}(X)$ neodvisna slučajna vektorja. □

5.2 Ocenjevanje za velike vzorce

Naj bodo $X, X_1, X_2, \ldots$ slučajni vektorji, definirani na nekem skupnem verjetnostnem prostoru.

- $(X_n)_n$ konvergira k X SKORAJ GOTOVO, če je $P(\lim X_n = X) = 1$.
- $(X_n)_n$ konvergira k X V VERJETNOSTI, če za vsak $\varepsilon > 0$ velja $\lim P(\|X_n - X\| > \varepsilon) = 0$.
- $(X_n)_n$ konvergira k X V L^2, če je $\lim E(\|X_n - X\|_2^2) = 0$ ($\|\cdot\|_2$ je funkcijska norma).
- $(X_n)_n$ konvergira k X V PORAZDELITVI, če velja $\lim F_{X_n}(x) = F_X(x)$ za vsako točko x, v kateri je F_X zvezna.

Lema 5.2.1. *Naj bodo $X, X_1, X_2, \ldots$ slučajni r-vektorji. Potem velja naslednje.*

- *X_n konvergira k X v verjetnosti natanko tedaj, ko ima vsako podzaporedje zaporedja $(X_n)_n$ nadaljnje podzaporedje, ki konvergira skoraj gotovo.*
- *X_n konvergira k X v porazdelitvi natanko tedaj, ko za vsak $\xi \in \mathbb{R}^r$ velja $\langle X_n, \xi\rangle \xrightarrow{d} \langle X, \xi\rangle$.*

Trditev 5.2.2. *Z enakimi oznakami, če X_n konvergira k X skoraj gotovo ali v L^2 smislu, potem konvergira tudi v verjetnosti. Če konvergira v verjetnosti, potem konvergira tudi v porazdelitvi. Če je X konstanten vektor, potem sta konvergenca v porazdelitvi in v verjetnosti ekvivalentni.*

Trditev 5.2.3. *Naj bodo $X, X_1, X_2, \ldots$ slučajni r-vektorji in naj bo $g : \mathbb{R}^r \to \mathbb{R}^s$ Borelova funkcija, ki je zvezna skoraj povsod glede na P_X. Potem, če X_n konvergira k X skoraj gotovo, ali v verjetnosti, ali v porazdelitvi, potem tudi $g(X_n)$ konvergira k $g(X)$ v enakem smislu.*

Trditev 5.2.4 (Slucki)**.** *Naj bodo $X, X_1, X_2, \ldots$ in $Y_1, Y_2, \ldots$ slučajne spremenljivke. Dalje naj bo $c \in \mathbb{R}$. Privzemimo $X_n \xrightarrow{d} c$ in $Y_n \xrightarrow{d} c$. Tedaj $X_n + Y_n \xrightarrow{d} X + c$, $X_nY_n \xrightarrow{d} cX$ in, če $c \neq 0$, $X_n/Y_n \xrightarrow{d} X/c$.*

Remark. Trditev lahko posplošimo na slučajne vektorje.

Primer. Naj bodo $X_1, X_2, \ldots$ neodvisne enako porazdeljene slučajne spremenljivke z disperzijo σ^2 in pričakovano vrednostjo μ. Po centralnem limitnem izreku velja

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow[n\to\infty]{d} N(0,1).$$

Potem po krepkem zakonu velikih števil

$$\left(\frac{1}{n}\sum_i X_i, \frac{1}{n}\sum_i X_i^2\right) \xrightarrow[n\to\infty]{s.g.} (\mu, \sigma^2 + \mu^2).$$

Upoštevaje zveznost potem velja

$$S^2 = \frac{n}{n-1}\left(\frac{1}{n}\sum_i X_i^2 - \overline{X}^2\right) \xrightarrow[n\to\infty]{s.g.} \sigma^2$$

za

$$S := \sqrt{\frac{1}{n-1}\sum_i (X_i - \overline{X})^2}.$$

Po izreku Sluckega je potem

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \xrightarrow[n\to\infty]{d} N(0,1).$$

Trditev 5.2.5. *Naj bodo $Y, X_1, X_2, \ldots$ slučajni r-vektorji, $c \in \mathbb{R}^r$ ter $(a_n)_n$ zaporedje pozitivnih realnih števil z $a_n \to \infty$. Naj bo $g : \mathbb{R}^r \to \mathbb{R}$ Borelova funkcija, diferenciabilna pri c. Če velja*

$$a_n(X_n - c) \xrightarrow[n\to\infty]{d} Y$$

in $dg(c) \neq 0$, potem velja tudi

$$a_n(g(X_n) - g(c)) \xrightarrow[n\to\infty]{d} \left\langle \vec{\nabla}.g(c), Y \right\rangle Y.$$

Dokaz. Spomnimo se: Za vsak $\varepsilon > 0$ obstaja $\delta > 0$, da iz $\|x - c\| \leq \delta$ sledi

$$|g(x) - g(c) - dg(c)(x - c)| \leq \varepsilon \|x - c\|.$$

Izberimo neki $\varepsilon > 0$ in mu pridružimo δ. Pišimo $Z_n = a_n(g(X_n) - g(X)) - a_n dg(c)(X_n - c)$. Pokazali bomo $Z_n \xrightarrow{p} 0$. Ko to vemo, lahko Z_n prištejemo $a_n dg(c)(X_n - c)$ in bo rezultat sledil.

Pokažimo, da za poljuben $\eta > 0$ velja $\lim P(|Z_n| > \eta) = 0$. Velja

$$\begin{aligned} P(|Z_n| > \eta) &= P(|Z_n| > \eta \wedge \|X_n - c\| > \delta) + P(|Z_n| > \eta \wedge \|X_n - c\| \leq \delta) \\ &\leq P(\|X_n - c\| > \delta) + P(|Z_n| > \eta \wedge |Z_n| \leq a_n \varepsilon \|X_n - c\|) \\ &\leq P(\|X_n - c\| > \delta) + P(a_n \varepsilon \|X_n - c\| > \eta). \end{aligned}$$

Ker $a_n^{-1} \to 0$ in $a_n(X_n - c) \xrightarrow{d} Y$, po Sluckem velja $X_n - c \xrightarrow{d} 0$. Sledi $P(\|X_n - c\| > \delta) \xrightarrow{p} 0$. Ker je norma zvezna, velja tudi $\|a_n(X_n - c)\| \xrightarrow{d} \|Y\|$, zato

$$\limsup P(|Z_n| > \eta) \leq 0 + \limsup P(\|a_n(X_n - c)\| > \eta/\varepsilon).$$

To je enako $1 - F_{\|Y\|}(\eta/\varepsilon)$, če je $F_{\|Y\|}$ zvezna v točki η/ε.

Točk nezveznosti je lahko največ števno mnogo. Izberemo lahko torej tako zaporedje $(\varepsilon_i)_i$, ki pada proti 0, za katero je η/ε_i točka zveznosti $F_{\|Y\|}$. Za vsak ε_i potem velja

$$\limsup P(|Z_n| > \eta) \leq P(\|Y\| > \eta/\varepsilon_i) \xrightarrow[i\to\infty]{} 0.$$

Torej limita $\lim P(|Z_n| > \eta)$ obstaja in je enaka 0. □

Posledica 5.2.6 (metoda δ)**.** *Z enakimi oznakami, če $a_n(X_n - c) \xrightarrow[n\to\infty]{d} N(0, \Sigma)$, potem $a_n(g(X_n) - g(c)) \xrightarrow[n\to\infty]{d} N(0, dg(c)\Sigma dg(c)^T)$.*

Izrek 5.2.7 (krepki zakon velikih števil)**.** *Naj bodo $X_1, X_2, \ldots$ neodvisni enako porazdeljeni slučajni vektorji s pričakovano vrednostjo $\mu \in \mathbb{R}^r$. Potem*

$$\frac{1}{n} \sum_i X_i \xrightarrow[n\to\infty]{s.g.} \mu.$$

Izrek 5.2.8 (centralni limitni izrek)**.** *Naj bodo $X_1, X_2, \dots$ neodvisni enako porazdeljeni slučajni vektorji z variančno matriko $\Sigma > 0$ ter pričakovano vrednostjo $\mu \in \mathbb{R}^r$. Tedaj*

$$\sqrt{n}\left(\frac{1}{n}\sum_i X_i - \mu\right) \xrightarrow[n\to\infty]{d} N(0, \Sigma).$$

5.2.1 Doslednost

Naj bodo $X_1, X_2, \dots$ neodvisne replikacije slučajne spremenljivke $\Omega \to \mathbb{R}$. Spomnimo se, da tako zaporedje X_i lahko modeliramo na $S = \Omega^{\mathbb{N}}$ s predpisom $X_i(s) = X(\omega_i)$ za $s = (\omega_n)_n$. Seveda je P na S definirana s predpisom

$$P(A_1 \times A_2 \times \cdots \times A_k \times \Omega \times \Omega \times \cdots) = P(A_1) \cdots P(A_k).$$

Porazdelitev X_i pripada privzetemu modelu dopustnih (enorazsežnih) porazdelitev $\mathcal{P}$. Naj bo $e : \mathcal{P} \to \mathbb{R}^r$ ocenjevana funkcija. Zaporedje cenilk $T_n : \mathbb{R}^n \to \mathbb{R}^r$ je za e

- KREPKO DOSLEDNO, če $T_n(X_1, \dots, X_n) \xrightarrow{s.g.} e(P_X)$,
- ŠIBKO DOSLEDNO, če $T_n(X_1, \dots, X_n) \xrightarrow{p} e(P_X)$,
- L^2-DOSLEDNO ali SKN-DOSLEDNO, če $E(\|T_n(X_1, \dots, X_n) - e(P_x)\|^2) \to 0$

za vsako dopustno porazdelitev $P_X \in \mathcal{P}$ in vsako zaporedje $X_i \sim P_X$ neodvisnih cenilk.

Izrek 5.2.9 (šibki zakon velikih števil Markova)**.** *Naj bo $X_1, X_2, \dots$ zaporedje nekoreliranih slučajnih spremenljivk z enako pričakovano vrednostjo μ in disperzijami σ_i^2, za katere je $\sup \sigma_i^2 < \infty$. Tedaj za vsak $\varepsilon > 0$ velja $\lim P(\left|\overline{X} - \mu\right| > \varepsilon) = 0$.*

Dokaz. Velja

$$P(\left|\overline{X} - \mu\right|^2 > \varepsilon^2) \le \frac{E(\left|\overline{X} - \mu\right|^2)}{\varepsilon^2} = \frac{D(\overline{X})}{\varepsilon^2} = \frac{\sigma_1^2 + \cdots + \sigma_n^2}{n^2\varepsilon^2} \le \frac{n}{n^2\varepsilon^2} \sup_i \sigma_i^2$$

za disperzijo D. To konvergira k 0. □

Naj bodo $\mathcal{P}_1, \mathcal{P}_2, \dots$ družine dopustnih porazdelitvenih zakonov na $B(\mathbb{R}), B(\mathbb{R}^2), \dots$ za vektorje $X_1, (X_1, X_2), \dots$, ki so med seboj usklajeni; če je $P_{(X_1,\dots,X_{n+1})} \in \mathcal{P}_{n+1}$, je tudi $P_{(X_1,\dots,X_n)} \in \mathcal{P}_n$. Model je parametričen, če obstajajo usklajene bijektivne koresponden-ce $P_n \to \Theta$.

Definicija 5.2.10. Naj bodo $e_n : \mathcal{P}_n \to \mathbb{R}^r$ ocenjevane funkcije in $T_n : \mathbb{R}^n \to \mathbb{R}^r$ Borelove preslikave. Zaporedje $(T_n)_n$ je za $(e_n)_n$

- ŠIBKO DOSLEDNO, če za vsak η

$$\lim_{n\to\infty} P(\left\|T_n(X^{(n)}) - e_n(P_n)\right\| > \eta) = 0,$$

- L^2 DOSLEDNO ali SKN DOSLEDNO, če

$$\lim_{n\to\infty} E\left(\left\|T_n(X^{(n)}) - e_n(P_n)\right\|^2\right) = 0$$

za vsako zaporedje $(P_n \in \mathcal{P}_n)_n$ in vsak vektor $X = (X_1, \ldots)$ z lastnostjo $P_{X^{(n)}} = P_n$ za vse n.

Izrek 5.2.11. *Naj λ_{min} označuje najmanjšo lastno vrednost matrike. Če $\lambda_{min}(Z^TZ) \xrightarrow[n\to\infty]{d} \infty$, je zaporedje cenilk za regresijski parameter β SKN dosledno.*

Trditev 5.2.12. *Naj bo $T : \mathbb{R}^n \to \mathbb{R}^m$ statistika in X slučajni vektor z vrednostmi v $\mathbb{R}^n$. Tedaj je $E(\|TX - E(TX)\|^2) = \mathrm{sl}(\operatorname{var} TX)$.*

Dokaz. Račun. Uporabimo dejstvo $\|t\|^2 = \mathrm{sl}(tt^T)$. □

Izrek 5.2.13. *Privzemimo naslednji linearni model:*

$$\mathcal{P}_n = \{P_{X^{(n)}} \mid X^{(n)} = Z^{(n)}\beta + \varepsilon^{(n)}, \operatorname{var}\varepsilon^{(n)} < \infty\}.$$

Naj bodo $L_n \in \mathbb{R}^{m\times d}$ fiksne realne matrike. Če velja

- $\sup_n \lambda_{max}(\operatorname{var}\varepsilon^{(n)}) < \infty$,
- $\lim \lambda_{max}((Z^TZ)') = 0$,
- $\sup_n \max_i \|(L_n)_i\| < \infty$,
- $L_n = U_nZ^{(n)}$

za vse n, je zaporedje cenilk $T_nX^{(n)} = L_n(Z^TZ)'Z^TX$ SKN-dosledno za $e_n(\beta, \varepsilon^{(n)}) = L_n(Z^TZ)'Z^TZ\beta$.

Dokaz. Račun. Nujno naredi. □

5.2.2 Pristranske cenilke

Privzemimo parametrični model s prostorom parametrov $\Theta \subseteq \mathbb{R}^d$ in naj bo $e : \Theta \to \mathbb{R}^r$ ocenjevana funkcija. PRISTRANSKOST cenilke $T : \mathbb{R}^n \to \mathbb{R}^r$ za e je $b_\theta(T) = E_\theta(T(X)) - e(\theta)$. Kvaliteto take cenilke merimo s srednjo kvadratno napako $\mathrm{SKN}(\theta) = E_\theta(\|T(X) - e(\theta)\|^2)$. Velja

$$\begin{aligned}\mathrm{SKN}(\theta) &= E_\theta(\|T(X) - E_\theta(T(X)) + E_\theta(T(X)) - e(\theta)\|^2)\\ &= E_\theta(\|T(X) - E_\theta(T(X))\|^2)\\ &+ E_\theta(\langle T(X) - E_\theta(T(X)), E_\theta(T(X)) - e(\theta)\rangle)\\ &+ E_\theta(\|E_\theta(T(X) - e(\theta))\|^2).\end{aligned}$$

Ker je $E_\theta(T(X) - E_\theta(T(X))) = 0$, sledi

$$\begin{aligned}\mathrm{SKN}(\theta) &= E_\theta(\|T(X) - E_\theta(T(X))\|^2) + E_\theta(\|E_\theta(T(X) - e(\theta))\|^2) \\ &= \mathrm{sl\,var}_\theta(T(X)) + \|b_\theta(T)\|^2 .\end{aligned}$$

Definicija 5.2.14. Privzemimo model $(\mathcal{P}_n)_n$ za $x_1, x_2, \ldots$ in naj bo $e_n : \mathcal{P}_n \to \mathbb{R}^r$ zaporedje ocenjevanih funkcij ter $T_n : \mathbb{R}^n \to \mathbb{R}^r$ zaporedje cenilk za e_n. To zaporedje je NEPRISTRANSKO ZA e_n V LIMITI, če velja

$$\lim_{n\to\infty} E_{P_n}(T_n - e_n(P_n)) = 0$$

za vsako dopustno porazdelitev $(P_n)_n \in (\mathcal{P}_n)_n$.

Definicija 5.2.15. Naj bo sedaj $Y_{(\mathcal{P}_n)_n}$ družina porazdelitev. Dalje naj bo $(a_n)_n$ zaporedje pozitivnih realnih števil in $T_n : \mathbb{R}^n \to \mathbb{R}^r$ zaporedje statistik. Naj bo $e_n : \mathcal{P}_n \to \mathbb{R}^r$ zaporedje ocenjevanih funkcij. Če velja $\lim a_n \in (0, \infty]$ in

$$a_n(T_n(X^{(n)}) - e_n(P_n)) \xrightarrow[n\to\infty]{d} Y_{(P_n)_n}$$

za vse $(P_n)_n \in (\mathcal{P}_n)_n$ in vsak X z $X^{(n)} \sim P_n$, potem pravimo, da je T_n ASIMPTOTIČNO NEPRISTRANSKO ZAPOREDJE CENILK za $(e_n)_n$.

Primer. Če so T_n dosledne, velja $T_n(X^{(n)}) - e_n(P_n) \to 0$.

Primer. Iz CLI dobimo $\sqrt{n}(\overline{X} - \mu) \to N(0, \Sigma)$. V tem primeru so porazdelitve res odvisne od $(P_n)_n$ (zaradi Σ).

5.2.3 Asimptotična normalnost

V zgornjem kontekstu (definicija 5.2.14) pravimo, da je zaporedje $(T_n)_n$ ASIMPTOTIČNO NORMALNO zaporedje cenilk za $(e_n)_n$, če obstaja zaporedje funkcij $V_n : \mathcal{P}_n \to \mathrm{SPD}(r)$ (simetrične pozitivno definitne matrike $r \times r$), za katerega velja

$$V_n(P_n)^{-1/2}(T_n(X^{(n)} - e_n(P_n))) \xrightarrow[n\to\infty]{d} N(0, I)$$

za vsak $(P_n)_n \in (\mathcal{P}_n)_n$ in X s to porazdelitvijo.

Rečemo, da je $V_n(P_n)$ ASIMPTOTIČNA VARIANCA za $T_n(X^{(n)})$. Če je $V_n(P_n) = \frac{1}{n}A((P_n)_n)$ za družino simetričnih pozitivno definitnih matrik $A : (\mathcal{P}_n)_n \to \mathbb{R}^{r\times r}$, je to ASIMPTOTIČNA VARIANCA V OŽJEM SMISLU.

Opomba. Asimptotična varianca ni enolična, lahko jo npr. pomnožimo s poljubnim zaporedjem, ki konvergira k 1.

Izrek 5.2.16. *Privzamemo model linearne regresije $\mathcal{P}'_n$ od prej, $X = Z\beta + \varepsilon$. Privzemimo, da $\frac{1}{v}Z^TZ$ konvergira k neki simetrični pozitivno definitni matriki razsežnosti $d \times d$. Tedaj imajo $Z^{(n)}$ poln rang za dovolj velika števila n in za $\hat\beta = (Z^TZ)^{-1}Z^TX$ velja*

$$\frac{1}{\sigma}(Z^TZ)(\hat\beta - \beta) \xrightarrow[n\to\infty]{d} N(0, I).$$

5.2.4 Konstrukcija cenilk

Obravnavamo neodvisne in enako porazdeljene slučajne spremenljivke X, s pripadajočim modelom, parametriziranim z $\Theta^{\text{odp}} \subseteq \mathbb{R}^d$. Privzemimo, da obstajajo momenti $\mu_j = \mu_j(\theta) = e_j(\theta) = E_\theta(X^j)$ za $1 \leq j \leq d$, in da je funkcija $e = (e_1, \ldots, e_d) : \Theta \to \mathbb{R}^d$ obrnljiva z inverzom $g : \operatorname{im} e \to \Theta$. Za momente imamo standardne cenilke

$$\hat{\mu}_j = \frac{1}{n} \sum_{i=1}^{n} X_k^j,$$

ki so krepko dosledne cenilke za momente po krepkem zakonu velikih števil. Če je g zvezna, je $g(\hat{\mu}_1, \ldots, \hat{\mu}_d)$ tudi krepko dosledna cenilka za θ.

Dodatno privzemimo obstoj momentov $E_\theta(X^j)$ za $j \leq 2d$. Vektorji $(X_i, X_i^2, \ldots, X_i^d)$ so neodvisni in enako porazdeljeni s pričakovano vrednostjo $(\mu_1, \ldots, \mu_d)$ in variančno matriko $\Sigma = [\mu_{k+l} - \mu_k \mu_l]_{k,l}$. Privzemimo, da je Σ neizrojena, da po CLI velja

$$\sqrt{n}\,((\hat{\mu}_1, \ldots, \hat{\mu}_d) - (\mu_1, \ldots, \mu_d)) \xrightarrow[n\to\infty]{d} N(0, \Sigma).$$

Če je g diferenciabilna, potem

$$\sqrt{n}\,(g(\hat{\mu}_1, \ldots, \hat{\mu}_d) - g(\mu_1, \ldots, \mu_d)) \xrightarrow[n\to\infty]{d} N(0, Dg(\mu)\Sigma Dg^T(\mu)).$$

To pomeni, da je $g(\hat{\mu}_1, \ldots, \hat{\mu}_d)$ asimptotično normalno zaporedje cenilk.

Alternativno lahko poiščemo cenilko po metodi največjega verjetja. Naj bo $X : \Omega \to \mathbb{R}^m$ proučevani slučajni vektor z modelom, parametriziranim z odprto množico $\Theta \subseteq \mathbb{R}^d$. Privzemimo gostote $f(\cdot, \theta)$, da je

$$P_\theta(X \in B) = \int_B f(x, \theta) d\nu(x).$$

Tu je ν neka σ-končna mera, ki dominira model $\{P_\theta \mid \theta \in \Theta\} \ll \nu$. Funkciji $L : \mathbb{R}^n \times \Theta \to [0, \infty)$, definirani z $L(x, \theta) = f(x, \theta)$, pravimo VERJETJE. Če za dano realizacijo x vektorja X obstaja $\hat{\theta} \in \mathbb{R}^d$, za katerega je $L(x, \hat{\theta})$ maksimum vrednosti $\{L(x, \theta) \mid \theta \in \overline{\Theta}\}$, mu pravimo OCENA PO MNV za x. Če $\hat{\theta} = \hat{\theta}(x)$ obstaja za ν-skoraj vse x, funkciji $\hat{\theta} : \mathbb{R}^d \to \overline{\Theta}$ pravimo CENILKA NAJVEČJEGA VERJETJA ZA θ.

Trditev 5.2.17. *Če obstaja enolična $\hat{\theta}$ in je $T : \mathbb{R}^n \to \mathbb{R}^m$ zadostna statistika, velja $\hat{\theta}(x) = \hat{\theta}(Tx)$.*

Dokaz. Po Fisher-Neymannu je $f(x, \theta) = g(Tx, \theta)h(x)$. Brez škode za splošnost je $h(x) > 0$, sicer $\hat{\theta}$ ni enolična. Maksimizacija se reducira na maksimizacijo funkcije $\theta \mapsto g(Tx, \theta)$. □

Naj bo $\mathcal{P}$ parametrični model z gostotami (glede na neko σ-končno mero) $\{f(\cdot,\theta)\,|\,\theta \in \Theta\}$, kjer je $\Theta^{\text{odp}} \subseteq \mathbb{R}^d$. Privzemimo dodatne regularnostne privzetke, med drugim da je množica $S = \{x\,|\,f(x,\theta) > 0\}$ neodvisna od θ. Velja

$$0 = \partial_{\theta_i} 1 = \partial_{\theta_i} \int f(x,\theta) d\nu(x) = \int \frac{\partial(\log f)}{\partial\theta_i} f(x,\theta) d\nu(x) = E_\theta(\partial_{\theta_i}(\log f(X,\theta))).$$

Funkciji

$$V_\theta(x) = \text{grad}_\theta(\log L)(x,\theta)$$

pravimo FUNKCIJA ZBIRA. Ker je $E_\theta(V_\theta(X)) = 0$ za vse θ, velja

$$0 \leq \text{var}_\theta(V_\theta(X)) = E_\theta(V_\theta(X)V_\theta(X)^T),$$

to matriko imenujemo FISHERJEVA INFORMACIJA in označimo s $\text{FI}(\theta)$. Če zgornje še enkrat odvajamo po θ_j, dobimo

$$0 = \int \frac{\partial^2(\log f)}{\partial\theta_i\partial\theta_j} f(x,\theta) d\nu(x) + \int \frac{\partial(\log f)}{\partial\theta_i} \frac{\partial(\log f)}{\partial\theta_j} f(x,\theta) d\nu(x)$$

kar je natanko

$$\text{FI}(\theta) = -E_\theta(H(\log L)(X,\theta)).$$

Če nam gostote f dopuščajo faktorizacijo

$$f(x,\theta) = h(x)\exp(-\psi(\theta) + \langle Q(\theta), T(x)\rangle)\,,$$

kjer so $\psi : \Theta \to \mathbb{R}$, $Q : \Theta \to \mathbb{R}^m$, $T : \mathbb{R}^n \to \mathbb{R}^m$ in $h : \mathbb{R}^m \to [0,\infty)$ primerne funkcije, potem pravimo, da je model EKSPONENTEN. Pravimo, da je model NARAVNO PARAMETRIZIRAN, če je $Q = \text{id}$.

Pod nekaterimi regularnostnimi privzetki lahko zagotovimo obstoj zaporedja slučajnih vektorjev $\hat{\theta}^{(n)}$ z naslednjimi lastnostmi:

- verjetnost, da $\hat{\theta}^{(n)}$ reši logaritemsko enačbo verjetja, konvergira k 1,
- zaporedje je dosledno za θ,
- velja $\sqrt{n}(\hat{\theta}^{(n)} - \theta) \to N(0, \text{FI}(\theta)^{-1})$.

5.3 Preizkušanje domnev

Obravnavamo model $\mathcal{P}$ za slučajni vektor X. Privzemimo, da obstaja dekompozicija $\mathcal{P} = \mathcal{H} \cup \mathcal{A}$ na neprazni disjunktni množici. NERANDOMIZIRAN POIZKUS DOMNEVE $\mathcal{H}$ proti $\mathcal{A}$ je odločitveno pravilo $\phi : \mathbb{R}^n \to \{0,1\}$, ki ga izvedemo tako:

- če je $\phi(x) = 1$, domnevo $\mathcal{H}$ zavrnemo,
- če je $\phi(x) = 0$, domneve $\mathcal{H}$ ne zavrnemo.

Če je $\mathcal{P}$ parametriziran, torej v bijektivni korespondenci z množico Θ, potem označimo sliki $\mathcal{H}$ in $\mathcal{A}$ pod to bijekcijo s H in A. Pravimo, da preizkušamo H proti A.

Poznamo dve vrsti napake. Če je v resnici $P_X \in H$, mi pa domnevo vseeno zavržemo, temu pravimo NAPAKA PRVE VRSTE, če pa $P_X \notin H$, a mi domneve ne zavržemo, je to NAPAKA DRUGE VRSTE. Pri primernih zveznostnih predpostavkah tipično velja

$$\sup\{P_\theta(\text{napaka prve vrste}) \mid \theta \in H\} + \sup\{P_\theta(\text{napaka druge vrste}) \mid \theta \in A\} = 1,$$

torej popolnega preizkusa ni. V praksi odločitev o zavrnitvi napravimo na podlagi neke testne statistike T v smislu

$$\phi(x) = \begin{cases} 1, & T(x) \in B \\ 0, & T(X) \notin B \end{cases}$$

za neko zavrnitveno območje B.

Zaradi komplementarnosti maksimalnih vrednosti napak obeh vrst izberemo pomembnejšo in jo skušamo omejiti. Po potrebi zamenjamo H in A, da je to napaka prve vrste. VELIKOST PREIZKUSA je potem

$$\sup_{\theta \in H} P_\theta(\text{napaka prve vrste}).$$

Pravimo, da ima preizkus STOPNJO ZNAČILNOSTI ali STOPNJO TVEGANJA α, če je njegova velikost največ α. Standardne izbire so $\alpha \in \{0.05, 0.1, 0.01\}$.

5.3.1 Preizkušanje na podlagi razmerja verjetij

Privzamemo parametrični model s prostorom parametrov $\Theta \subseteq \mathbb{R}^d$ in gladka verjetja $L(x, \theta)$. Naj bo $H \subseteq \Theta$ preizkušana domneva. RAZMERJE VERJETIJ je H je funkcija $\lambda : \mathbb{R}^n \to [0, 1]$, definirana z

$$\lambda(x) = \frac{\sup\{L(x, \theta) \mid \theta \in H\}}{\sup\{L(x, \theta) \mid \theta \in \Theta\}}.$$

Vedno lahko potem konstruiramo preizkus oblike

$$\phi(x) = \begin{cases} 1, & \lambda(x) < D \\ 0, & \lambda(x) \geq D \end{cases}$$

za neko primerno konstanto D.

Izrek 5.3.1 (Wilksov izrek o asimptotični porazdelitvi razmerja verjetij)**.** *Naj bodo $X_1, X_2, \ldots$ neodvisni enako porazdeljeni slučajni vektorji z gostoto $f(x, \theta)$, kjer je $\theta \in \Theta$ in je Θ gladka končnorazsežna mnogoterost. Dalje privzemimo, da je H zaprta gladka podmnogoterost brez roba, tako da je $H \subseteq \Theta$ prava vložitev. Naj veljajo primerni gladkostni privzetki na gostote, ki zagotavljajo asimptotično normalnost cenilk največjega verjetja. Če dejanski parameter θ pripada H, potem*

$$-2 \log \lambda(X_1, \ldots, X_n) \xrightarrow[n \to \infty]{d} \chi^2(\dim \Theta - \dim H).$$