Zbrani zapiski za 1. letnik magistrskega študija

Patrik Žnidaršič

13. marec 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 Teorija grafov

1.1 Matchings

Definition 1.1.1. 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 1.1.2. 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 1.1.3. 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 1.1.4. 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 1.1.5 (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 1.1.6. 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 1.1.7. *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 1.1.8. A KÖNIG-EGERVÁRY graph is a graph G with $\alpha'(G) = \beta(G)$.

Theorem 1.1.9 (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. Our 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 1.1.10. *If G is a bipartite graph, then $\alpha(G) = \beta'(G)$.*

Definition 1.1.11. 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 1.1.12 (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 1.1.13. A matching M is a PERFECT MATCHING if it covers all vertices.

Corollary 1.1.14. *In a bipartite graph G, there is a perfect matching if and only if $|A| = |B|$ and A satisfies Hall's condition.*

Definition 1.1.15. 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 1.1.16. *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 1.1.17. *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 1.1.18. *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. □

Question 7. Show that any non-maximum matching in a graph admits an augmenting path.

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.

1.1.1 Tutte's theorem

Definition 1.1.19. 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 1.1.20 (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. □

Question 8. State and prove Tutte's theorem.

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.

1.1.2 Factors

Definition 1.1.21. A FACTOR is a spanning subgraph (it contains all vertices). A k-FACTOR is a k-regular spanning subgraph.

Definition 1.1.22. A CUBIC GRAPH is a 3-regular graph.

Theorem 1.1.23 (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.

Question 9. State and prove Petersen's theorem for cubic graphs.

Theorem 1.1.24. *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. □

Question 10. Show that every even-regular graph has a 2-factor.

1.2 Connectivity

Definition 1.2.1. 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 1.2.2. 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 $\kappa(G) \leq n - \alpha(G) = \beta(G)$.

Question 11. Define k-connected graphs. Explain why $\kappa(G) \leq \beta(G)$.

Theorem 1.2.3. *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. □

Question 12. What is the minimum number of edges in a k-connected graph on n vertices? Prove the bound.

Definition 1.2.4. A set $F \subseteq E(G)$ is a DISCONNECTING SET if $G - F$ is disconnected.

Definition 1.2.5. 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 1.2.6. 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 1.2.7. *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 1.2.8. *If G is k-connected, then G is k-edge-connected.*

Question 13. Show that a k-connected graph is always k-edge-connected.

Corollary 1.2.9. *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 1.2.10 (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

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

□

Question 14. State and prove Whitney's theorem.

Theorem 1.2.11 (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 $|S| \geq k$. □

Question 15. State and prove the expansion lemma.

Theorem 1.2.12. *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. □

Question 16. Show that a graph is 2-connected if and only if it has no isolated vertices and there is a cycle through any pair of edges.

Proposition 1.2.13 (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. □

Question 17. State the subdivision lemma.

1.2.1 Ear decomposition of a graph

Definition 1.2.14. 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 1.2.15. 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 1.2.16. *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}. □

Question 18. What is an open ear decomposition? Show that a graph is 2-connected if and only if it has one.

Proposition 1.2.17. *A graph G is 2-edge-connected if and only if G is connected and there is no cut edge in G.*

Definition 1.2.18. 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 1.2.19. 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 1.2.20. *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. □

Question 19. What is a closed ear decomposition? Show that a graph is 2-edge-connected if and only if it has one.

Definition 1.2.21. 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 1.2.22 (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 of $P_0, \ldots, P_k$. 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. □

Question 20. State and prove Robbins' theorem.

1.2.2 Cuts

Definition 1.2.23. 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 1.2.24. The maximum number of internally vertex-disjoint x, y-paths in a graph G is labeled with $\lambda_G(x, y)$.

Theorem 1.2.25 (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. □

Question 21. State and prove Menger's theorem for vertex cuts.

Definition 1.2.26. 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 1.2.27 (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. □

Question 22. State and prove Menger's theorem for edge cuts.

Lemma 1.2.28. *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 $\kappa(G - e) + 1 \geq \kappa(G)$. □

Question 23. Show that for any $e \in E(G)$, we have $\kappa(G) - 1 \leq \kappa(G - e) \leq \kappa(G)$.

Theorem 1.2.29 (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

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

Question 24. State Menger's theorem.

1.3 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 1.3.1 (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 1.3.2 (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. □

Question 25. State and prove Brooks' theorem.

1.3.1 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 1.3.3. *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 $\chi(G) = k$. $\multimap$ □

Question 26. Show that the difference between $\chi(G)$ and $\omega(G)$ can be arbitrarily large.

Theorem 1.3.4. *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. □

Question 27. State and prove a lower bound for $m(G)$ in a graph with $\chi(G) = k$.

1.3.2 Turàn's theorem

Definition 1.3.5. 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 1.3.6. G is a COMPLETE k-PARTITE GRAPH if every edge between the partite classes is present in G.

Definition 1.3.7. 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 1.3.8 (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. □

Question 28. What is a Turàn graph? State and prove Turàn's theorem.

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 $\chi(G) = r$.

1.3.3 Chordal graphs

Definition 1.3.9. A graph G is CHORDAL if there is no induced subgraph that is isomorphic to a cycle C_n for $n \geq 4$.

Definition 1.3.10. A vertex v is a SIMPLICIAL vertex in G if $N_G[v]$ induces a clique.

Lemma 1.3.11 (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. □

Question 29. State and prove Voloshin's lemma.

Theorem 1.3.12. *A graph G is chordal if and only if there is a simplicial elimination ordering of the vertices of G.*

Proof. Right-to-left: If G is not chordal, then we will show there is no simplicial elimination ordering. Suppose there is one, $v_1, v_2, \ldots, v_n$. Since G is not chordal, it has a cycle C_k with $k \geq 4$. Let v_i be the first vertex of this cycle in the ordering. Then v_i has two neighbours along the cycle with no edge between them, which contradicts the definition of the simplicial elimination ordering.

Left-to-right: By Voloshin's lemma, there is a simplicial vertex v_1 in G. Then $G - v_1$ is also a chordal graph, so it has a simplicial vertex v_2. Then $G - v_1 - v_2$ is a chordal graph, and we keep going. We get a simplicial elimination ordering $v_1, \ldots, v_n$. □

Question 30. Characterize chordal graphs and prove the characterization.

Theorem 1.3.13. *If G is chordal, then $\chi(G) = \omega(G)$.*

Proof. In general $\omega(G) \leq \chi(G)$. Let $v_1, \ldots, v_n$ be a simplicial elimination ordering in G. Consider the greedy coloring in the reverse of that order. When we color a vertex v_i, then the neighbours that are already covered are $R_i = N(v_i) \cap \{v_{i+1}, \ldots, v_n\}$. As it is a simplicial elimination ordering, R_i is a clique, and $R_i \cup \{v_i\}$ is also a clique. Since $|R_i \cup \{v_i\}| \leq \omega(G)$, at most $\omega(G) - 1$ neighbours of v_i are colored before v_i, so we can color v_i with one of $\{1, 2, \ldots, \omega(G)\}$. Since this is true for any v_i, we have $\chi(G) \leq \omega(G)$. □

1.3.4 Perfect graphs

Definition 1.3.14. A graph G is a PERFECT GRAPH if $\chi(H) = \omega(H)$ holds for every induced subgraph H of G.

Theorem 1.3.15. *Chordal graphs are perfect.*

Proof. An induced subgraph of a chordal graph is chordal. □

Question 31. Show that chordal graphs are perfect.

Theorem 1.3.16. *Bipartite graphs are perfect.*

Proof. If G is bipartite, then either $E(G) = \varnothing$, in which case $\omega(G) = \chi(G) = 1$, or $E(G) \neq \varnothing$, in which case $\chi(G) = \omega(G) = 2$. Every subgraph of a bipartite graph is bipartite. □

Theorem 1.3.17 (Vizing)**.** *For every graph G, we have $\Delta(G) \leq \chi'(G) \leq \Delta(G) + 1$.*

Theorem 1.3.18. *If G is bipartite, then $\chi'(G) = \Delta(G)$.*

Proof. Let $k = \Delta(G)$. Consider the following two cases. First, if G is k-regular, there exists a perfect matching M_1 in G. The graph $G - M_1$ is $(k-1)$-regular and bipartite, so there exists a perfect matching M_2 in $G - M_1$. We can continue this procedure to find matchings $M_1, \ldots, M_k$. Now assign color i to M_i.

In the second case, if G is not k-regular, we will construct a k-regular bipartite graph which has G as a subgraph. We add edges to G such that the maximum degree remains k and the graph remains bipartite. If this is not possible but there are still vertices of degree $< k$, then we can follow the following construction.

- Create a copy G' of G, except it has partite classes switched.
- Create a new graph by combining G and G', and adding connections $u \sim u'$ for $u \in G$ with $\deg(u) < k$.

This increases the minimum degree of G by at least one, and we still have a bipartite graph. If we continue this process, we will get a k-regular bipartite graph. □

Question 32. Show that bipartite graphs are Vizing class one.

Theorem 1.3.19. *If G is bipartite, then its line graph $L(G)$ is perfect.*

Proof. For every graph, $\chi'(G) = \chi(L(G))$, as every edge coloring of G corresponds to a vertex coloring of $L(G)$. If G is bipartite, $\chi'(G) = \Delta(G)$. The maximal cliques in $L(G)$ corresponds to the vertices of G (since there are no triangles). If a vertex is of degree k, then the corresponding clique in $L(G)$ is also of size k, so $\Delta(G) = \omega(L(G))$. This means $\chi(L(G)) = \omega(L(G))$.

The induced subgraphs of $L(G)$ are still line graphs of bipartite graphs, as if $H = L(G)[S]$, then H is the line graph of the subgraph of G obtained by only keeping the edges corresponding to the vertices in S. □

Question 33. Show that the line graph of a bipartite graph is perfect.

Theorem 1.3.20 (Perfect graph theorem)**.** *A graph G is perfect if and only if its complement $\overline{G}$ is perfect.*

Theorem 1.3.21 (Strong perfect graph theorem)**.** *A graph G is perfect if and only if neither G nor $\overline{G}$ have an induced odd cycle of size at least 5.*

Question 34. State the weak and strong perfect graph theorems.

Definition 1.3.22. A (β, α')-PERFECT GRAPH is a graph for which $\beta(H) = \alpha'(H)$ for every induced subgraph H.

Theorem 1.3.23. *A graph is (β, α')-perfect if and only if it is bipartite.*

1.3.5 Gallai-Roy-Vitaver theorem

Theorem 1.3.24. *If G is a simple graph and $\vec{D}$ is an orientation of G, and $l(\vec{D})$ is the longest path in $\vec{D}$, then $\chi(G) \leq l(\vec{D}) + 1$. Further, there is an orientation for which the equality holds.*

Proof. Consider an arbitrary orientation $\vec{D}$ of G. Let $\vec{D}'$ be a maximal subdigraph of $\vec{D}$ that does not contain any directed cycles. Define a vertex coloring c for G with $c(v)$ being one plus the length of the longest directed path in $\vec{D}'$ that ends in v. This is a proper vertex coloring; consider an edge $\overrightarrow{uv} \in \vec{D}$. Consider two cases.

- If $\overrightarrow{uv} \in \vec{D}'$, then $l'(u) < l'(v)$, since v cannot be on the longest path ending at u (as there are no directed cycles).
- If $\overrightarrow{uv} \notin \vec{D}'$, then there is a directed path in $\vec{D}'$ between v and u (which forms a cycle with $\overrightarrow{uv}$). Then the coloring c gives increasing colors to the vertices along this path, and $c(u) > c(v)$.

We skip the proof of the second statement. □

Question 35. State the Gallai-Roy-Vitaver theorem. Prove the $\chi(G) \leq$ inequality.

1.4 Planar graphs

Definition 1.4.1. If G is a graph, then the DRAWING of G into the plane is a function $h(x)$ defined on $V(G) \cup E(G)$ such that for every vertex v, $h(v)$ is a point and for every edge uv, $h(uv)$ is a $h(u), h(v)$-curve.

Remark. We can assume without loss of generality that the image of an edge is a polygonal curve, so composed of finitely many line segments.

Definition 1.4.2. The PLANAR EMBEDDING of a graph G is a drawing where the curves corresponding to the edges of G do not intersect, except in common end vertices.

Definition 1.4.3. A PLANE GRAPH is a particular embedding of a planar graph.

Theorem 1.4.4 (Jordan)**.** *Every closed simple curve in the plane divides it into exactly two regions, a bounded inside region and an unbounded outside region.*

Definition 1.4.5. If G is a plane graph, then a FACE of G is a maximal region that does not contain any points from the image of the embedding function f.

Definition 1.4.6. The DUAL GRAPH G^* of a plane graph G is obtained by switching the role of the vertices and faces of G, with two faces being connected if and only if they share a common border. If there are multiple common edges on the boundary between two faces, make that many edges. If there is an edge which borders twice on the same face, add a loop.

Remark. Two different embeddings of a planar graph may have non-isomorphic dual graphs.

Proposition 1.4.7. *Dual graphs are always connected.*

Proposition 1.4.8. *If G is a connected plane graph, then $(G^*)^* \cong G$.*

Remark. Note that dual graphs are always planar.

Definition 1.4.9. The LENGTH $l(F)$ of a face F in a plane graph G is the total length of the walk(s) along the boundary of F, in units of number of edges.

Remark. With this definition, cut edges are counted twice.

Remark. For any plane graph,

$$\sum_{F \text{ face}} l(F) = 2m(G).$$

Theorem 1.4.10. *Suppose G is a plane graph. The following statements are equivalent:*

- *G is bipartite,*
- *every face of G has an even length,*
- *G^* is Eulerian (connected and all vertices are of even degree).*

Proof. 1 to 2: The length of a face is the length of a cycle in G, and possibly some cut edges, which are counted twice.

2 to 1: Suppose G is not bipartite, so there is an odd cycle C. Consider the sum of the lengths of faces inside C. Every edge inside C (but not part of C) is counted twice, but every edge of C is counted only once. So the sum is odd. Then there must be a face F in the cycle with an odd length, which is a contradiction.

Equivalence between 2 and 3: Note that G^* is connected and $\deg_{G^*}(x) = l(x)$. □

Question 36. When is a plane graph bipartite? State two equivalences.

Theorem 1.4.11. *If G is a plane graph and $D \subseteq E(G)$, then D is the set of edges of a cycle if and only if the corresponding dual edge set D^* is a minimal edge cut in G^*.*

Proof. Consider three cases.

- If D is exactly the edge set of a cycle C, then D^* contains all edges between the inside and outside faces of C, so it is an edge cut in G^* and it is minimal, as the faces inside C are all connected in G^*.
- If D is a proper subset of the set of edges of a cycle, we can prove that D^* is an edge cut, but not minimal.
- If D does not contain the edge set of a cycle, we can show that D^* is not an edge cut. □

Definition 1.4.12. A planar graph G is OUTERPLANAR if there exists an embedding such that the boundary of the outer face contains all vertices.

Theorem 1.4.13. *If G is a simple outerplanar graph, then $\delta(G) \leq 2$.*

Proof. If $n \leq 4$, then this is simple casework. For $n \geq 4$, we will prove that additionally, there are at least two nonadjacent such vertices. Induction on $n = n(G)$. We covered the base case $n = 4$ before, just do the induction step, in which we consider two cases.

In the first case, if there is a cut vertex v, then $G - v$ splits into two graphs G_1, G_2. Both $G_1 + v$ and $G_2 + v$, with the original edges from v, are outerplanar graphs. By the induction hypothesis, there are two nonadjacent vertices in $G_1 + v$ of degree at most 2, so there is a vertex $z_1 \neq v$ in $G_1 + v$ with degree at most 2 in $G_1 + v$. Similarly, there is a vertex z_2 in $G_2 + v$ for which the same holds. These vertices are of the same degree in G.

In the second case, if there is no cut vertex, then there is no cut edge, and the boundary of the outer face is a Hamiltonian cycle. If there is no chord in this cycle, the statement is proven. Otherwise, let xy be a chord. Consider the two cycles H_1, H_2 that the chord divides our Hamiltonian cycle into. Both are outerplanar, so using the induction hypothesis on each, we can find two nonadjacent vertices of degree at most 2. They can't be adjacent to a vertex in the other part of the Hamiltonian cycle, as that edge would cross xy. □

Question 37. Show that a simple outerplanar graph has $\delta(G) \leq 2$.

Theorem 1.4.14 (Euler)**.** *If G is a connected plane graph, then*

$$n(G) + f(G) - m(G) = 2.$$

Definition 1.4.15. A SUBDIVISION of G is obtained by replacing some edges of G with internally vertex-disjoint paths.

Definition 1.4.16. A KURATOWSKI GRAPH is one which is a subdivision of K_5 or $K_{3,3}$.

Lemma 1.4.17. *If G is a planar graph and $e \in E(G)$, then there is an embedding of G such that e is on the boundary of the outer face.*

Proof. Consider an embedding of G and place a sphere on that plane, touching one of the faces adjacent to e. Project the entire drawing onto the sphere by taking the intersection of a line from a given point to the highest point on the sphere. Every face on the plane will correspond to a region of the sphere with the same set of boundary edges. Now project the surface of the sphere onto another plane above the sphere, by drawing a line through the bottom-most point of the sphere and another point of the sphere, and taking the intersection of that line with the new plane. □

Question 38. Show that for any edge e of a planar graph, there is an embedding with e on the boundary of the outer face.

Lemma 1.4.18. *If G is a minimal nonplanar graph, then it is 2-connected.*

Proof. If G is disconnected, it cannot be minimal nonplanar, as components are proper subgraphs, which would all be planar if G was minimal. If G has a cut vertex v, then consider the v-lobes (the subgraphs of G induced by each of the components of $G - v$, combined with v). By the minimality of G, all v-lobes are planar subgraphs. Lemma 1.4.17 states that these graphs can be embedded in such a way that v is on the outer face. We can transform each picture such that they all fit in a portion of the plane, so we can combine the drawings to find G is planar. □

Question 39. Show that a minimal nonplanar graph is 2-connected.

Lemma 1.4.19. *If $S = \{x, y\}$ is a minimum vertex cut in G and G is nonplanar, then $G - S$ contains a component G_i such that the S-lobe H_i, along with the edge xy, is nonplanar.*

Proof. Suppose to the contrary, that each $H_i + xy$ is planar. By the first lemma, we can embed all these graphs such that xy is on the outer face. We can construct an embedding for $G + xy$ in the following way:

- Start with the embedding of $H_1 + xy$
- Choose a face in $H_1 + xy$ that borders xy. Then we may attach the embedding of $H_2 + xy$ onto that face without intersecting $H_1 + xy$.
- Continue this process.

So $G + xy$ is planar, and G must be too. □

Question 40. Show that in a nonplanar graph with a cut set $S = \{x, y\}$, there is a lobe H in $G - S$ such that $H + xy$ is nonplanar.

Lemma 1.4.20. *If G is a nonplanar graph without a Kuratowski subgraph and G has the minimum number of edges under this condition, then it is 3-connected.*

Proof. We know by lemma 1.4.18 that G is 2-connected. Suppose that $S = \{x, y\}$ is a vertex cut. Again as before consider the S-lobes $H_1, H_2, \ldots$ with an added edge xy. By lemma 1.4.19, we know there is a graph $H_i + xy$ which is not planar. By the minimality of $m(G)$, we know $H_i + xy$ has a Kuratowski subgraph, as it has fewer edges than G (note that S is a minimal vertex cut). Label that subgraph F. As it's not present in G by assumption, it must include the edge xy.

Consider another component G_j of $G - S$. Since S is a minimum vertex cut, both x and y have neighbours in G_j. Since G_j is connected, there is an x, y-path in H_j (not considering xy). We can replace xy in $H_i + xy$ with that path, and get a Kuratowski graph. But this is a subgraph of G. □

Question 41. Suppose that G is the graph with the smallest number of edges among all nonplanar graphs without a Kuratowski subgraph. Show that G is 3-connected.

Theorem 1.4.21. *If G is a 3-connected graph with $n(G) \geq 5$, then there exists an edge $e \in E(G)$ such that $G \cdot e$ (G with e contracted) is 3-connected.*

Proof. Suppose there is no such edge. Let $e = xy \in E(G)$. Since $G \cdot e$ is not 3-connected, there is a vertex cut S with 2 vertices. Let w be the contracted vertex of $G \cdot e$. If $w \notin S$, then replacing w with the original edge still leaves $G - S$ disconnected, which can't happen. So $w \in S$. Let $S = \{w, z\}$. Then $S' = \{x, y, z\}$ is a vertex cut in G.

We have shown that every edge xy in G has a mate, so a vertex w such that $\{x, y, w\}$ is a disconnecting set. Let $f = uv$ be an edge with mate z such that $G - \{u, v, z\}$ has the largest component among all edge and mate vertex cuts. Let G_i be that largest component, and G_j be some other component in $G - \{u, v, z\}$. Since $\{u, v, z\}$ is minimal, there is a vertex z' of G_j which is adjacent to z. Let z^* be a mate for zz' and let H be the subgraph induced by $G_i \cup \{u, v\}$. Consider three cases.

- If $z^* \in V(H)$ and $H - z^*$ is disconnected, then $G - \{z, z^*\}$ is disconnected, as at least one component of $H - z^*$ has edges only to z, not u or v. But G is 3-connected, so this can't happen.
- If $z^* \in V(H)$, but $H - z^*$ is connected, then we can similarly show that $\{z, z'\}$ is a disconnecting set.
- If $z^* \notin V(H)$, then $G - \{z, z', z^*\}$ has a component which includes the entire H, but $n(H) > n(G_i)$, so our choice for $\{u, v, z\}$ wasn't optimal.

In each case, we got a contradiction. □

Question 42. Show that in any 3-connected graph on at least 5 vertices, there is an edge which we can contract and still have a 3-connected graph.

Lemma 1.4.22. *If G contains no Kuratowski subgraphs and $e \in E(G)$, then $G \cdot e$ contains no Kuratowski subgraph.*

Proof. Let $e = xy$ and let w be the contracted vertex in $G \cdot e$. Suppose that $G \cdot e$ contains a Kuratowski subgraph F.

- If F doesn't contain w, then F is present in G.
- If $w \in V(F)$ and $\deg_F(w) = 2$, then we just uncontract the edge and find a Kuratowski subgraph in G.
- If $w \in V(F)$ and $\deg_F(w) \geq 3$, consider three subcases.
  - If $N_F(w) \subseteq N_G(x)$ or symmetrically for y, then we may replace w by x and find a Kuratowski subgraph in G.
  - If all but one neighbour of w is a neighbour of x in G, then we can subdivide the edge going into the neighbour of y and find a Kuratowski subgraph of G.
  - Otherwise, since F is Kuratowski, $\Delta(F) \leq 4$. In this case, we therefore have precisely two x-neighbours and two y-neighbours. Label them x_1, x_2, y_1, y_2. In this case, F must be a subdivision of K_5.

    Consider the first branch vertices (those of degree at least 3) along the paths $wx_1 \dots x_1'$, $wx_2 \dots x_2'$, $wy_1 \dots y_1'$, $wy_2 \dots y_2'$. These paths are unique as possible intermediate vertices are of degree 2. Since they were part of a subdivision of K_5 in F, there are disjoint paths between any two of these vertices. We can use them to find a subdivision of $K_{3,3}$, with the partite classes being (x, y_1', y_2') and (y, x_1', x_2'). □

Question 43. Show that edge contraction preserves the property that a graph has no Kuratowski subgraphs.

Theorem 1.4.23. *If G is a 3-connected graph without Kuratowski subgraphs, then there exists a convex embedding of G such that no 3 vertices are on a line.*

Proof. Use induction on $n(G)$. The base case is $n(G) = 4$, which is just K_4. For the induction step, since G is 3-connected, there is an edge $e \in E(G)$ such that $G \cdot e$ is 3-connected. Let $e = xy$ and let z be the contracted vertex. Since G has no Kuratowski subgraph, $G \cdot e$ has no Kuratowski subgraph. By the induction hypothesis, there is a convex embedding for $G \cdot e$ with no 3 vertices on a line.

Since $G \cdot e$ is 3-connected, $G \cdot e - z$ is 2-connected, so if we remove the edges incident to z from $G \cdot e$, the region containing z must be a cycle. Let $x_1, x_2, \dots, x_k$ be the neighbours of x on the cycle, labeled in cyclic order. Consider the following cases for the neighbours of y.

- If all neighbours of y are between x_i and x_{i+1}, then we can find a convex embedding for G by drawing y inside the $x_i x x_{i+1}$ angle, and very near x, while drawing x in z's place. We can ensure that the resulting drawing has convex faces and no three points in a line.

- If x and y have at least 3 common neighbours, then they form a subdivision of K_5 along with x and y.
- If there are two neighbours x_1', x_2' of x and two neighbours of y, y_1, y_2, which are four different vertices such that the cyclic order is $x_1' y_1 x_2' y_2$, this induces a subdivision of $K_{3,3}$, which is a Kuratowski subgraph. □

Question 44. Show that a 3-connected graph without Kuratowski subgraphs can be embedded into the plane such that all interior faces are convex and no 3 vertices are on a line.

Theorem 1.4.24 (Kuratowski)**.** *A graph is planar if and only if it contains no Kuratowski subgraph.*

Proof. Note that a subdivision of G is planar if and only if G is planar, which proves the left-to-right implication. For the other direction, suppose there exist nonplanar graphs without Kuratowski subgraphs. We've shown that the minimal such graph must be 3-connected, so by the previous theorem, it is planar. This is a contradiction. □

Question 45. State and prove Kuratowski's theorem.

Definition 1.4.25. A graph H is a MINOR in G if we can obtain H after deleting and contracting edges in G.

Proposition 1.4.26. *If G contains a subdivision of a graph F as a subgraph, then F is a minor in G.*

Proof. We can contract the edges of between two subdivision vertices or between a subdivision vertex and a non-subdivision vertex. □

Remark. The opposite implication is not true. The Petersen graph has K_5 as a minor, but no subgraph is a subdivision of K_5.

Question 46. Show that if G contains a subdivision of F as a subgraph, then F is a minor in G, but that the reverse doesn't hold.

Theorem 1.4.27 (Wagner)**.** *A graph G is planar if and only if neither K_5 nor $K_{3,3}$ are a minor of G.*

Question 47. State Wagner's theorem.

Theorem 1.4.28 (four-colour theorem)**.** *If G is a planar graph, then $\chi(G) \leq 4$.*

Theorem 1.4.29. *If G is a planar graph, then $\chi(G) \leq 5$.*

Proof. If G is a planar graph, then by Euler's formula, we get $m(G) \leq 3n(G) - 6$, so $\delta(G) \leq 5$. Induction on $n(G)$. Consider a vertex $v \in V(G)$ of degree at most 5. By the induction hypothesis, there is a 5-colouring of $G - v$. If $N(v)$ uses fewer than 5 colours,

we can use a missing colour for v. Otherwise, $\deg_G(v) = 5$ and the neighbours $v_1, \ldots, v_5$, numbered cyclically, have colours $1, \ldots, 5$.

Let $G_{1,3}$ be the subgraph of G induced by the vertices with colours 1 and 3. Consider two cases.

- If v_1 and v_3 are in different components of $G_{1,3}$, then we may switch the colours in one of the two components, and we get a 5-colouring of G where $N(v)$ is coloured with 4 colours.
- If v_1 and v_3 belong to the same component, there is a v_1, v_3-path in $G_{1,3}$, where vertices alternate between colours 1 and 3. In the original graph, this path forms a cycle with v. One of v_2, v_4 must be inside this cycle, and the other outside. This means v_2, v_4 are in different components of $G_{2,4}$, since any v_2, v_4-path would need to cross the cycle, so share a common vertex with it, which can't happen, as $G_{1,3}$ and $G_{2,4}$ have no common vertices. This means we can use the previous construction for $G_{2,4}$. □

Question 48. Prove the 5-color theorem.

Definition 1.4.30. A graph G is a MAXIMAL PLANAR GRAPH if G is planar and for any $e \in E(\overline{G})$, $G + e$ is not a planar graph.

Definition 1.4.31. An embedding of G is a TRIANGULATION if every face has a boundary that is a 3-cycle.

Proposition 1.4.32. *Let G be a planar graph with $n(G) \geq 3$. Then the following statements are equivalent.*

- $m(G) = 3n(G) - 6$,
- *every embedding of G is a triangulation,*
- *there is an embedding of G that is a triangulation,*
- *G is a maximal planar graph.*

Proof. Note that

$$m = 3n - 6 \Leftrightarrow \sum_{F \text{ face}} l(F) = 3f \Leftrightarrow \text{every face has length 3.}$$

Since the embedding is arbitrary, this proves the equivalence of the first three statements.

1 to 4: If we add an edge, we will have more edges than we can in a planar graph.

4 to 2: Suppose there is an embedding that isn't a triangulation. Then we may add an edge to a face which is a cycle with at least 4 vertices (it may be a degenerate cycle). □

Question 49. Characterize maximal planar graphs.

2 Teorija izračunljivosti

2.1 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 2.1.1. 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 2.1.2. 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 2.1.3. 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 2.1.4. *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 2.1.5. *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 2.1.6. *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 2.1.7. *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 2.1.8. *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 2.1.9. A k-tape Turing machine COMPUTES $f : \Sigma_1^* \times \cdots \times \Sigma_k^* \rightharpoonup \Sigma^*$ 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 2.1.10. 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 2.1.11. 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 $\gamma_2(g(w)) \simeq f(\gamma_1(w))$.

Question 7. What is a representation? What does it mean for a function to be $(\gamma_1 \to \gamma_2)$-computable?

Definition 2.1.12. 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 2.1.13. 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 2.1.14. 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 2.1.15. *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 $1 \mapsto 1$.

2.1.1 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.

2.2 Computability of natural numbers

Definition 2.2.1. The COMPUTABLE PARTIAL FUNCTIONS is the set

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

Definition 2.2.2. 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.

Question 8. What are the primitive recursive functions? Why are they total?

Definition 2.2.3. 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.

Question 9. Define partial recursive functions.

Proposition 2.2.4. *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 2.2.5. *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

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

□

Question 10. Show that the partial recursive functions coincide with the computable partial functions.

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.

Question 11. How is $\phi_e^n(x_1, \dots, x_n)$ defined?

Proposition 2.2.6. *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 2.2.7 (the universal function)**.** *For any $n \geq 1$, the $(n+1)$-arity function*

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

exists and is computable.

Proposition 2.2.8. *The unary total function below is not computable.*

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

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. □

Question 12. Give two examples of non-computable total functions and show they aren't computable.

Theorem 2.2.9 (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 2.2.10 (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 $e, x_1, \ldots, x_n$:*

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

Question 13. State the S-M-N and Kleene's normal form theorems.

2.3 Computable and computably enunmerable sets

Definition 2.3.1. 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 2.3.2. KLEENE'S HALTING SET is the set

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

Proposition 2.3.3. *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. □

Question 14. What are computable and what are computably enumerable sets? Define Kleene's halting set and show it is not computable.

Proposition 2.3.4. *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 computably 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

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

□

Question 15. Show that a set is computably enumerable if and only if it is the domain of some computable partial function.

Proposition 2.3.5. *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 2.3.6. *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 2.3.7. *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. □

Question 16. Show that a set is computable if and only if both it and its complement are computably enumerable.

Corollary 2.3.8. *The set $\overline{K}$ is not computably enumerable.*

Theorem 2.3.9. *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 $s(w, y) = 0$. □

Question 17. Characterize computably enumerable sets and prove the characterization.

Theorem 2.3.10. *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. □

Question 18. Characterize the computable sets and prove the characterization.

Corollary 2.3.11. *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 2.3.12. 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 2.3.13 (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.*

Question 19. When is a subset of $\mathcal{C}$ decidable or semidecidable? State the reduction lemma.

Theorem 2.3.14 (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. □

Question 20. State and prove Rice's theorem.

Definition 2.3.15. 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 2.3.16 (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. □

Question 21. State and prove the Rice-Shapiro theorem.

2.3.1 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 2.3.17. *Every productive set is not computably enumerable.*

Proposition 2.3.18. *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 $g(e) \in \overline{K} \setminus W_e$. □

Question 22. Define productive sets. Show that $\overline{K}$ is productive.

Lemma 2.3.19. *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 $g(h(e)) \in A \setminus W_{h(e)}$. □

Question 23. Show that if a productive set A reduces to B, then B is also productive.

Theorem 2.3.20. *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. □

Definition 2.3.21. A set A is CREATIVE if it is computably enumerable and its complement is productive.

Theorem 2.3.22. *If $\varnothing \subsetneq B \subsetneq \mathcal{C}$ is such that I_B is computably enumerable, then is it creative.*

Theorem 2.3.23. *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. □

Question 24. Show that every productive set has an infinite computably enumerable subset.

Definition 2.3.24. 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 2.3.25. A set $A \subseteq \mathbb{N}$ is SIMPLE if it is computably enumerable and its complement is immune.

Theorem 2.3.26 (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 $B \not\subseteq A$. □

Question 25. State and prove Post's theorem.

2.4 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.

Question 26. Define type 2 Turing machines.

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

Theorem 2.4.1. *The following are equivalent:*

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

Proof. 1 to 2: Suppose that there is a T2M M computing p. To decide whether w is in $\mathrm{Prefix}(p)$ 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 $\mathrm{Prefix}(p)$. 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. □

2.4.1 Topological aspects of computing with infinite words

Theorem 2.4.2. *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. □

Question 27. Show that a semidecidable language is an open set.

Theorem 2.4.3. *If a partial function $f : \Sigma^\omega \rightharpoonup \Sigma^\omega$ is computable, then for every $p \in \mathrm{dom}\, f$ and for every $n \geq 0$ there exists an $m \geq 0$ such that for all $q \in \mathrm{dom}\, f$, if $q|_m = p|_m$, then $f(q)|_n = f(p)|_n$.*

Proof. Let M be a T2M that computes f. Consider any $p \in \mathrm{dom}\, f$ 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. □

Question 28. Show that computable functions are continuous on their domain.

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 2.4.2 states: A semidecidable language is an open set.

Remark. Theorem 2.4.3 states: Computable functions are continuous with respect to the subspace topology on $\operatorname{dom} f$.

Proposition 2.4.4. *If $L \subseteq \Sigma^\omega$ is decidable, it is clopen.*

Proof. The complement is semidecidable. □

Theorem 2.4.5. *An ω-language L is decidable if and only if it is clopen.*

Definition 2.4.6. A subset $Z \subseteq \Sigma^\omega$ is G_δ if it is a countable intersection of open sets.

Theorem 2.4.7. *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. □

Question 29. Show that the domains of computable partial functions are G_δ.

Theorem 2.4.8. *The topological space Σ^ω is compact.*

Proof. Tychonoff. □

Corollary 2.4.9. *The compact subsets of Σ^ω are exactly the closed sets.*

Lemma 2.4.10. *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 $w_1, \ldots, w_k$. □

Question 30. Show that clopen sets are decidable.

2.4.2 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 2.4.11. 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 2.4.12. *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 2.4.13. *A real number is γ_c-computable if and only if it is computable with respect to the decimal or binary representation.*

Question 31. What is the Cauchy representation?

Definition 2.4.14. 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 2.4.15. *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 2.4.16 (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 $|f(x) - f(x')| < \varepsilon$. □

Question 32. State and prove the continuity theorem.

2.5 Algorithmic information theory

A computable partial function $u : \{0,1\}^* \rightharpoonup \{0,1\}^*$ is UNIVERSAL if for any computable partial function $f : \{0,1\}^* \rightharpoonup \{0,1\}^*$ there exists a word $v_f \in \{0,1\}^*$ such that for all $w \in \{0,1\}^*$ we have $f(w) \simeq u(v_f w)$.

Proposition 2.5.1. *A universal computable partial function exists.*

Given a universal function u, the KOLMOGOROV COMPLEXITY $C_u(w)$ of a word w is

$$C_u(w) = \min\{|v| \mid u(v) = w\}.$$

Proposition 2.5.2. *Suppose u, u' are universal computable functions. Then for all words w, $C_{u'}(w) \leq C_u(w) + C$ for some constant C, equal for all words.*

Proposition 2.5.3. *For any $n \geq 0$, there exists a word w of length n such that $C_u(w) \geq n$.*

Proof. There are only at most 2^{n-1} words with complexity $< n$, but there are 2^n words of length n. □

Definition 2.5.4. An infinite word $p \in \{0,1\}^\omega$ is K-INCOMPRESSIBLE if for all $n \geq 0$, we have $C(p|_n) \geq n - O(1)$.

Theorem 2.5.5. *No infinite sequence is K-incompressible.*

Proof. Define a map

$$\lceil w \rceil = \sum_{i=0}^{|w|-1} w_i 2^i + 2^{|w|} - 1$$

which has an inverse word : $\mathbb{N} \to \{0,1\}^*$. Let $p \in \{0,1\}^\omega$. Consider any $d \geq 0$ and let $m = \lceil p|_d \rceil$. Define $h : \{0,1\}^* \to \{0,1\}^*$ by $h(w) = \text{word}(|w|)w$. Note that

$$p_0 p_1 \dots p_{d-1} p_d \dots p_{d+m-1} = p|_{d+m} = h(p_d \dots p_{d+m-1}) = u(v_h p_d \dots p_{d+m-1}),$$

so $C_u(p|_{d+m}) \leq |v_h| + \lceil p|_d \rceil$. □

Definition 2.5.6. A partial function $f : \{0,1\}^* \rightharpoonup \{0,1\}^*$ is PREFIX-FREE if $\text{dom}\, f \subseteq \{0,1\}^*$ satisfies the following property: for all $w, w' \in \text{dom}\, f$, w is not a proper prefix of w'.

Definition 2.5.7. A computable prefix-free function $u : \{0,1\}^* \rightharpoonup \{0,1\}^*$ is UNIVERSAL if for any computable prefix-free function $f : \{0,1\}^* \rightharpoonup \{0,1\}^*$ there exists $v_f \in \{0,1\}^*$ such that for all w, $f(w) \simeq u(v_f w)$.

Definition 2.5.8. The PREFIX-FREE COMPLEXITY $K(w)$ is defined as

$$K(w) = \min\{|v| \mid u(v) = w\}$$

for a universal prefix-free function u.

Proposition 2.5.9. *For any $d > 1$, there exists a constant C such that $K(w) \leq d\,|w| + C$ for any word w.*

Proof. Outline. For any large N, we work with an encoding of $\{0,1\}^*$ as words Σ^* where $|\Sigma| = 2^N - 1$. We map w to $\lceil w \rceil$, then represent Σ with binary words of length N. We are left with one non-represented word, which is 0^N without loss of generality. Then we assign to w the representation $\langle \lceil w \rceil \rangle$ followed by N zeros. □

Definition 2.5.10. An infinite word p is PREFIX-FREE INCOMPRESSIBLE if there exists a constant c such that for all n, $K(p|_n) \geq n - c$.

Definition 2.5.11. Chaitin's halting probability Ω is the probability that u halts if given an input generated randomly by throwing coins, so

$$\Omega = \sum_{w \in \operatorname{dom} u} 2^{-|w|}.$$

Then take p^Ω as the binary representation of Ω. If there are multiple options, choose the one with infinite ones.

Theorem 2.5.12. *Chaitin's halting probability p^Ω is prefix-free incompressible.*

Proof. Note that

$$\Omega = \sum_{i=0}^{\infty} p_i^\Omega 2^{-(i+1)}.$$

We will define a computable function $f : \{0,1\}^* \rightharpoonup \{0,1\}^*$ satisfying $f(p^\Omega\big|_n) \downarrow$ and $K(f(p^\Omega\big|_n)) > n$ for all n. If we do, then there is a number c such that for all $n \geq 0$,

$$K(f(p^\Omega\big|_n)) \leq K(p^\Omega\big|_n) + c,$$

so $K(p^\Omega\big|_n) > n - c$.

Let $w_0, w_1, \ldots$ be a computable enumeration of $\operatorname{dom} u$ without repetitions. Given an input word $p_0 p_1 \ldots p_{n-1}$, the algorithm of f finds the smallest $N \geq 0$ (if it exists) with the property

$$\sum_{i=0}^{N-1} 2^{-|w_i|} \geq \sum_{i=0}^{n-1} p_i 2^{-(i+1)}.$$

Then the algorithm returns the first $v \in \{0,1\}^*$ such that $v \notin \{u(w_0), \ldots, u(w_{N-1})\}$. Define $f(p_0 \ldots p_{n-1}) = v$.

Clearly $f(p^\Omega\big|_n)$ is defined for any $n \in \mathbb{N}$, as we have chosen a representation of p^Ω with infinite ones. From

$$\sum_{i=0}^{n-1} p_i^\Omega 2^{-(i+1)} \leq \sum_{i=0}^{N-1} 2^{-|w_i|} < \Omega < \sum_{i=0}^{n-1} p_i^\Omega 2^{-(i+1)} + 2^{-n}$$

we have

$$\Omega - \sum_{i=0}^{n-1} 2^{-|w_i|} < 2^{-n},$$

so every word w_i for $i \geq N$ is such that $|w_i| > n$. This means that for every finite word $v \notin \{u(w_0), \ldots, u(w_{N-1})\}$, the smallest $w \in \operatorname{dom} u$ with $u(w) = v$ has $|w| > n$ or $K(v) > n$. □

2.6 Algorithmic randomness

The law of large numbers states that for a sequence, generated randomly by coin tosses,

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}q_i=\frac{1}{2}\right)=1.$$

We say that a sequence q satisfies the law of large numbers if

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}q_i=\frac{1}{2}.$$

If q does not satisfy the law, then there exists a $\delta>0$ such that $\left|m_n(q)-\frac{1}{2}\right|<\delta$ for infinitely many n, where

$$m_n(q)=\frac{1}{n}\sum_{i=0}^{n-1}q_i.$$

Given δ, we then have the following non-randomness test. Define T_n as the set of all sequences $q\in\{0,1\}^\omega$ for which there are at least n different values of i where $\left|m_n(q)-\frac{1}{2}\right|>\delta$. Note that we can semidecide T_n.

Definition 2.6.1. A NAIVE NON-RANDOMNESS TEST is a sequence $(T_n)_n$ of subsets of $\{0,1\}^\omega$ such that

- every T_n is open,
- $\lim\lambda(T_n)=0$ where $\lambda(T_n)$ is the probability that a fair randomly generated sequence lands in T_n.

Then $q\in 2^\omega$ SATISFIES $(T_n)_n$ if $q\in\bigcap T_n$.

Definition 2.6.2. A MARTIN-LÖF NON-RANDOMNESS TEST is a sequence $(T_n)_n$ of subsets of 2^ω such that

- $(T_n)_n$ is a computable sequence of computably open sets,
- $\lim\lambda(T_n)=0$ with a computable rate of convergence.

A sequence q is MARTIN-LÖF RANDOM if it fails every ML test.

Question 33. What is a naive non-randomness test and what is a Martin-Löf non-randomness test?

Proposition 2.6.3. *Computable sequences are not ML random.*

Proof. Let p be a computable sequence. Then $(\langle p|_n\rangle)_n$ is an ML test which p satisfies. □

Question 34. Show that computable sequences are not ML random.

Definition 2.6.4. An ML test $(T^u_n)_n$ is UNIVERSAL if any non-ML random p satisfies $(T^u_n)_n$.

It holds that every open set $U \subseteq 2^\omega$ can be written as a disjoint union of cylinder sets $\langle w_i \rangle$ for some $(w_i)_{i \in I}$. Define $\lambda(\langle w \rangle) = 2^{-|w|}$- Then the probability of an open set is

$$\lambda(U) = \sum_{i \in I} \lambda(w_i)$$

for the disjoint union above. We can prove that $\lambda(U)$ is well-defined.

We represent open sets as infinite sequences $w_0, w_1, \ldots$ where every w_i is either a word in 2^* or the symbol $\varnothing$. We say that an open set is COMPUTABLE if it is represented by some computable infinite sequence $w_0, w_1, \ldots$. Note that this definition coincides with the semidecidable sets.

Then a sequence of open sets $(T_n)_n$ is represented by a sequence of sequences, $T_i \sim (w_{ij})_j$. We say that $(T_n)_n$ is a COMPUTABLE SEQUENCE OF COMPUTABLE OPENS if the single sequence representing w_{ij} above, using the pairing function, is computable.

A computable total function $r : \mathbb{N} \to \mathbb{N}$ is a COMPUTABLE RATE OF CONVERGENCE for a Cauchy sequence $(x_n)_n \subseteq [0, 1]$ converging to 0 if for any n and $m \geq r(n)$, we have $x_m \leq 2^{-n}$.

Theorem 2.6.5 (Martin-Löf)**.** *There exists a universal ML test.*

Proof. Outline. We can show the following:

- There is a computable function from $\mathbb{N}$ to the representations of computable sequences of computably open sets, and it finds some representation of any sequence.
- There is a function from the computable sequences of computably open sets to ML tests, which preserves rapid ML tests in its domain (an ML test is rapid if $\lambda(T_n) \leq 2^{-n}$).
- There is a function from ML tests to rapid decreasing ML tests which preserves the set of sequences that satisfy the test (an ML test is decreasing if $T_{n+1} \subseteq T_n$ for all n).

So a sequence p is ML random if and only if it fails every rapid decreasing test. From all three points above, we have a computable enumeration of the rapid decreasing ML tests. It is surjective in the sense that for every rapid decreasing ML test, at least one representation is found. Let $(T^0_n)_n, (T^1_n)_n, \ldots$ be this enumeration. Take

$$T^u_n = \bigcup_{i=0}^{\infty} T^i_{n+i+1}.$$

This is clearly a rapid decreasing sequence, and it is computable. □

Question 35. Show that there exists a universal ML test.

Corollary 2.6.6. *The set R_{ML} of all ML random sequences is a countable union of closed sets (F_σ). Additionally, $\lambda(R_{ML}) = 1$ and $|R_{ML}| \geq 2^{\aleph_0}$.*

Proof. The complement of R_{ML} is exactly the intersection of all T_n^u. □

Theorem 2.6.7. *The following are equivalent for any infinite sequence $q \in 2^\omega$.*

- *q is ML random,*
- *q is prefix-free incompressible.*

3 Uvod v funkcionalno analizo

3.1 Normirani in Banachovi prostori

Definicija 3.1.1. 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 3.1.2. Normiran prostor je BANACHOV, če je poln za inducirano metriko.

Trditev 3.1.3. *Seštevanje in množenje vektorjev s skalarjem sta zvezni operaciji.*

Definicija 3.1.4. 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 3.1.5. *Če je X kompakten Hausdorffov prostor, je $\mathcal{C}(X)$ Banachova algebra.*

Trditev 3.1.6. *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. □

Vprašanje 1. Dokaži: če je X Banachov prostor in $Y \leq X$, je Y Banachov natanko tedaj, ko je zaprt v X.

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.

3.1.1 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 3.1.7. *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

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

Vprašanje 2. Definiraj napolnitev normiranega prostora. Pokaži, da napolnitev vsebuje originalen prostor kot gost podprostor.

Posledica 3.1.8. *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. $\square$

3.1.2 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 3.1.9. Normirana prostora X in Y sta IZOMORFNA, če obstaja linearni homeomorfizem med njima.

Vprašanje 3. Kdaj sta normi ekvivalentni? Kdaj sta normirana prostora izomorfna?

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 3.1.10. *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 3.1.11. *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

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

Vprašanje 4. Definiraj normo na kvocientnem prostoru. Kaj mora veljati, da bo res norma? Kdaj je kvocient Banachov? Dokaži.

Trditev 3.1.12. *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 $\lim x_n = x - y$. $\square$

Vprašanje 5. Pokaži: če sta Y in X/Y Banachova, je X Banachov.

Posledica 3.1.13. *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. $\square$

Vprašanje 6. Pokaži: vsak končnorazsežen normiran prostor je Banachov.

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.

3.2 Linearni funkcionali

Izrek 3.2.1. *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*

Vprašanje 7. Karakteriziraj zveznost linearnih preslikav med normiranima prostoroma.

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

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

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.

Vprašanje 8. Kako definiraš normo na (omejenem) linearnem operatorju? Povej še vsaj eno izražavo.

Trditev 3.2.2. *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 3.2.3. DUALNI PROSTOR prostora X je $X^* = B(X, \mathbb{F})$.

Izrek 3.2.4. *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 $T_m \to T$. $\square$

Vprašanje 9. Pokaži: če je Y Banachov, je $B(X,Y)$ Banachov.

Posledica 3.2.5. *Dualni prostor je vedno Banachov.*

Izrek 3.2.6. *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 $Sx = \lim T x_n$.

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

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

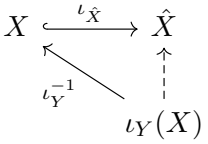
torej je S omejen in $\|S\| \leq \|T\|$. Obrat je očiten. □

Vprašanje 10. Dokaži: linearen operator $T : Y \to Z$, ki slika v poln prostor, lahko enolično razširimo do operatorja $\overline{Y} \to Z$.

Posledica 3.2.7. *Naj bosta $S, T : X \to Y$ omejena operatorja, ki se ujemata na gostem podprostoru. Potem je $S = T$.*

Posledica 3.2.8. *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 3.2.9. Naj bo $T : X \to X$ omejen operator med normiranima prostoroma. Tedaj je T OBRNLJIV, če je bijektiven in T^{-1} omejen.

Definicija 3.2.10. 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 3.2.11. *Naj bo $T \in B(X, Y)$. Naslednji trditvi sta ekvivalentni.*

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

3.2.1 Banachov izrek

Definicija 3.2.12. 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 $\lambda \geq 0$.

Vprašanje 11. Kaj je sublinearni funkcional?

Primer. Vsaka polnorma je sublinearni funkcional.

Izrek 3.2.13 (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 $X = \tilde{Y}$. □

Vprašanje 12. Povej in dokaži realni Hahn-Banachov izrek.

Lema 3.2.14. *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 $\|f\| \leq \left\|\tilde{f}\right\|$. □

Vprašanje 13. Pokaži, da lahko $\mathbb{R}$-linearni funkcional iz kompleksnega vektorskega prostora X enolično razširimo do $\mathbb{C}$-linearnega funkcionala, in da razširitev ohranja normo.

Izrek 3.2.15 (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 $\left|\tilde{F}(x)\right| \leq p(x)$. □

Vprašanje 14. Povej in dokaži kompleksni Hahn-Banachov izrek.

Izrek 3.2.16 (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 $\|F\| = \|f\|$.*

Vprašanje 15. Povej Hahn-Banachov izrek za normirane prostore.

Posledica 3.2.17. *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 3.2.18. *Naj bo X normiran in $x \in X$. Tedaj je*

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

Vprašanje 16. Pokaži, da je $\|x\| = \max\{|f(x)| \mid f \in X^*, \|f\| = 1\}$.

Posledica 3.2.19. *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 $\frac{1}{d} f$. □

Vprašanje 17. Pokaži, da lahko z linearnim funkcionalom ločiš točko od zaprtega podprostora.

Izrek 3.2.20. *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 $f(z) = 1$. □

Vprašanje 18. Kako lahko izraziš zaprtje podprostora z jedri linearnih funkcionalov? Dokaži enakost.

3.2.2 Adjungirani operator in drugi dual

Definicija 3.2.21. Za normiran prostor X in $f \in X^*$ definiramo $\hat{x}(f) = f(x)$.

Trditev 3.2.22. *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 obstaja 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 3.2.23. *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 3.2.24. *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

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

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

Vprašanje 19. Kako je definiran adjungiran operator v smislu Banachovih prostorov? Pokaži, da adjungiranje ohranja normo operatorja.

Definicija 3.2.25. Normiran prostor je REFLEKSIVEN, če je i_X surjekcija.

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

3.3 Temeljni izreki funkcionalne analize

Izrek 3.3.1 (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 $x \in X$ in $r > 0$ velja

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

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. □

Vprašanje 20. Povej in dokaži Bairov izrek.

Posledica 3.3.2. *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 $\bigcup_j A_j = \left(\bigcap_j A_j^{\mathsf{c}}\right)^{\mathsf{c}} \neq X$. □

Vprašanje 21. Povej in dokaži posledico Bairovega izreka z zaprtimi množicami.

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 3.3.3 (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 $T \in \mathcal{F}$. □

Vprašanje 22. Povej in dokaži princip enakomerne omejenosti.

Izrek 3.3.4 (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 $x \in A$. □

Vprašanje 23. Povej in dokaži izrek o šibki omejenosti.

Posledica 3.3.5. *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 3.3.6. *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 3.3.7 (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 TU vsebuje odprto okolico 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 $\mathring{B}(0,\eta_i) \subseteq T(\mathring{B}(0,\varepsilon_i))$. 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 $y \in \mathring{B}(0,\eta_0) \subseteq T(\mathring{B}(0,\varepsilon_0))$. 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 $y - Tx_0 \in \mathring{B}(0,\eta_1) \subseteq T(\mathring{B}(0,\varepsilon_1))$, 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-slika 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. □

Vprašanje 24. Povej in dokaži izrek o odprti preslikavi.

Posledica 3.3.8. *Naj bo T omejen linearen bijektiven operator med Banachovima prostoroma. Potem je njegov inverz tudi omejen.*

Vprašanje 25. Povej posledico izreka o odprti preslikavi.

Posledica 3.3.9. *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. □

Vprašanje 26. Dokaži: dve normi, za kateri je X Banachov prostor, sta bodisi ekvivalentni bodisi neprimerljivi.

Lema 3.3.10. *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 3.3.11 (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.

$$\begin{array}{ccc} X & \xrightarrow{(x,Tx)} & \Gamma_T \\ & \searrow^{T} & \downarrow \mathrm{pr}_2 \\ & & Y \end{array}$$

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 $T = \mathrm{pr}_2 \circ \mathrm{pr}_1^{-1}$. □

Vprašanje 27. Povej in dokaži izrek o zaprtem grafu.

3.4 Hilbertovi prostori

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

Definicija 3.4.1. 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 3.4.2 (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 3.4.3. *Skalarni produkt je zvezna preslikava.*

Izrek 3.4.4 (Jordan, von Neumann)**.** *Če v normiranem prostoru velja paralelogramska enakost, je norma porojena s skalarnim produktom.*

Definicija 3.4.5. 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 3.4.6. Vektorja x in y sta PRAVOKOTNA, če $\langle x, y\rangle = 0$. Označimo $x \perp y$.

Definicija 3.4.7. Množici A in B sta PRAVOKOTNI, če je $\langle a, b\rangle = 0$ za vsak $a \in A$ ter $b \in B$.

Izrek 3.4.8 (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 3.4.9. *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 $k, k' \in K$. □

Vprašanje 28. Pokaži: v Hilbertovem prostoru obstaja enolična projekcija točke na konveksno zaprto množico.

Izrek 3.4.10. *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

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

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

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

Vprašanje 29. Pokaži: Če je M zaprt podprostor v Hilbertovem prostoru in $x_0 \in M$, potem je $x - x_0$ pravokoten na M natanko tedaj, ko je $d(x, M) = \|x - x_0\|$.

Definicija 3.4.11. Naj bo X prostor s skalarnim produktom. Za $x \in X$ definiramo $\{x\}^\perp = \{y \in X \,|\, x \perp y\}$, za $A \subseteq X$ pa

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

Lema 3.4.12. *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 $y \in \{x\}^\perp$. $\square$

Vprašanje 30. Pokaži: $A^\perp$ je vedno zaprt podprostor.

Izrek 3.4.13. *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 $M^{\perp\perp} = \ker(I - P) = \operatorname{im} P = M$. □

Vprašanje 31. Definiraj ortogonalni projektor na zaprt podprostor M v Hilbertovem prostoru H. Dokaži, da je linearen, idempotent, da velja $\|Px\| \leq \|x\|$ in da velja $H = M \oplus M^\perp$.

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

Posledica 3.4.14. *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 $A^{\perp\perp} = [A]^{\perp\perp}$. □

Vprašanje 32. Kaj je $A^{\perp\perp}$ za $A \subseteq H$?

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 3.4.15. *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 3.4.16 (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 $f(x) = \left\langle x, z/\|z\|^2 \right\rangle$. □

Vprašanje 33. Povej in dokaži Rieszov izrek.

Posledica 3.4.17. *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 3.4.18. *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 3.4.19. *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 3.4.20. *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

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

Torej $\|y' - y\| = 0$. □

Vprašanje 34. Pokaži, da so Hahn-Banachove razširitve v Hilbertovem prostoru enolične.

Posledica 3.4.21. *Vsak Hilbertov prostor je refleksiven (vložitev v H^{**} je surjektivna).*

3.4.1 Ortonormirani sistemi

Definicija 3.4.22. 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 3.4.23. *Vsaka ortogonalna množica je linearno neodvisna.*

Definicija 3.4.24. 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 $\subseteq$).

Vprašanje 35. Definiraj kompleten ortonormiran sistem.

Trditev 3.4.25. *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 3.4.26. *Vsak Hilbertov prostor ima bazo.*

Trditev 3.4.27. *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 3.4.28 (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. □

Vprašanje 36. Povej in dokaži Besselovo neenakost.

Posledica 3.4.29. *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. □

Vprašanje 37. Naj bo $\{e_i\}_i$ ONS in $x \in X$. Pokaži, da je kvečjemu števno mnogo vektorjev e_i pravokotnih na x.

Posledica 3.4.30. *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 števno mnogo $e \in E$. Na njih uporabimo Besselovo neenakost. □

Izrek 3.4.31. *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 števno 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
$$\|s_m - s_n\|^2 = \left\| \sum_{k=n+1}^{m} \langle x, e_k \rangle \, e_k \right\|^2 = \sum_{k=n+1}^{m} |\langle x, e_k \rangle|^2 \leq \sum_{k=n+1}^{\infty} |\langle x, e_k \rangle|^2 \,.$$
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. $\rightarrow\!\!\!\!\!\!\times\leftarrow$ □

Vprašanje 38. Povej 5 trditev, ekvivalentnih dejstvu, da je $E \subseteq H$ KONS. Dokaži ekvivalence. Kaj je Parsevalova enakost?

Trditev 3.4.32. *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 števno 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. □

Vprašanje 39. Pokaži: če sta E, F neskončna KONS-a, imata isto kardinalnost.

Definicija 3.4.33. Dimenzija Hilbertovega prostora je enaka kardinalnosti katerekoli njene baze.

Lema 3.4.34. *V separabilnem metričnem prostoru je vsaka družina paroma disjunktnih odprtih krogel kvečjemu števno neskončna.*

Dokaz. Naj bo $\{\mathring{B}(x_i, \varepsilon_i) \mid i \in I\}$ družina paroma disjunktnih odprtih krogel. Naj bo S števna 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 3.4.35. *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

$$\left\|e - e'\right\|^2 = \left\langle e - e', e - e'\right\rangle = \|e\|^2 - 2\,\mathrm{Re}\left\langle e, e'\right\rangle + \left\|e'\right\|^2 = 2.$$

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

$$\left\{ \sum_{k=1}^{n} \lambda_k e_k \,|\, n \in \mathbb{N}, \lambda_k \in \mathbb{Q} + i\mathbb{Q} \right\}. \qquad \square$$

Vprašanje 40. Karakteriziraj separabilnost neskončnorazsežnega Hilbertovega prostora in dokaži karakterizacijo.

Definicija 3.4.36. 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 3.4.37. *Naj bo $U : X \to Y$ linearna izometrija med prostoroma s skalarnim produktom. Tedaj U ohranja skalarni produkt.*

Dokaz. Računamo

$$\|U(x+y)\|^2 = \|Ux\|^2 + \|Uy\|^2 + 2\operatorname{Re}\langle Ux, Uy\rangle$$
$$\|U(x+y)\|^2 = \|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\operatorname{Re}\langle x, y\rangle$$

torej $\operatorname{Re}\langle x, y\rangle = \operatorname{Re}\langle Ux, Uy\rangle$. Če namesto x pišemo ix, dobimo še $-\operatorname{Im}\langle x, y\rangle = -\operatorname{Im}\langle Ux, Uy\rangle$. $\square$

Vprašanje 41. Pokaži: linearna izometrija ohranja skalarni produkt.

Lema 3.4.38. *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 $0 = \langle y, Ue_i\rangle = \langle U^{-1}y, e_i\rangle$, torej $y = 0$. $\square$

Izrek 3.4.39. *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. □

Vprašanje 42. Pokaži: Hilbertova prostora sta izomorfna natanko tedaj, ko imata isto dimenzijo.

Posledica 3.4.40. *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.

Vprašanje 43. Definiraj neskončno direktno vsoto Hilbertovih prostorov.

3.4.2 Stone-Weierstrassov izrek

Lema 3.4.41. *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 3.4.42. *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 3.4.43 (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 $f(y) \leq f_{x_m}(y) < g(y) + \varepsilon$. □

Vprašanje 44. Povej in dokaži mrežni Stone-Weierstrassov izrek.

Lema 3.4.44. *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

$$S_n(1) = \lim_{x \to 1} S_n(x) \leq 1,$$

torej je zaporedje $(S_n(1))_n$ naraščajoče in navzgor omejeno. Sedaj 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 3.4.45 (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 je 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 $|f| \in \overline{A}$. □

Vprašanje 45. Povej in pokaži realni Stone-Weierstrassov izrek.

Izrek 3.4.46 (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 $x \in K$.*

Vprašanje 46. Povej Weierstrassov izrek.

Izrek 3.4.47 (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_0 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 ravno dokazanem obstajata $g_1, h_1 \in A_0$, da je $\|g - g_1\| < \varepsilon/2$ in $\|h - h_1\| < \varepsilon/2$. Potem je $\|f - (g_1 + ih_1)\| < \varepsilon$. □

Vprašanje 47. Povej in dokaži kompleksni Stone-Weierstrassov izrek.

3.5 Omejeni operatorji med Hilbertovimi prostori

Definicija 3.5.1. 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 3.5.2. Seskvilinearna forma je ZVEZNA (ali OMEJENA), če obstaja $M \geq 0$, da $|u(x, y)| \leq M \|x\| \|y\|$.

Vprašanje 48. Kaj je seskvilinearna forma? Kdaj je omejena?

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 3.5.3. *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 $\overline{u(x, y)} = \langle y, Ax \rangle = \overline{\langle Ax, y \rangle}$. 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. □

Vprašanje 49. Pokaži, da lahko vsako omejeno seskvilinearno formo izrazimo s skalarnim produktom.

Definicija 3.5.4. 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 3.5.5. *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 $U^* = U^{-1}$. □

Vprašanje 50. Pokaži: U je izomorfizem Hilbertovih prostorov natanko tedaj, ko je obrnljiv in $U^* = U^{-1}$.

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 3.5.6. *Naj bosta $A \in B(H, K)$ in $B \in B(K, L)$ omejena operatorja. Tedaj $(BA)^* = A^*B^*$.*

Dokaz. Preprost račun. □

Posledica 3.5.7. *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 3.5.8. *Če je $A \in B(H, K)$, je* $\ker A^* = (\operatorname{im} A)^\perp$.

Dokaz. Velja

$$x \in \ker A^* \Leftrightarrow A^*x = 0 \Leftrightarrow \forall y.\, \langle A^*x, y\rangle = 0 \Leftrightarrow \forall y.\, \langle x, Ay\rangle = 0 \Leftrightarrow x \in (\operatorname{im} A)^\perp.$$ □

Vprašanje 51. Kaj je $\ker A^*$?

Posledica 3.5.9. *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 3.5.10. 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 3.5.11. *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 $\gamma = \alpha$. □

Vprašanje 52. Pokaži: če je $\langle Ax, x\rangle \in \mathbb{R}$ za vsak x, je A sebi adjungiran.

Izrek 3.5.12. *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

$$\begin{aligned} 4\left|\operatorname{Re}\langle Ax, y\rangle\right| &\leq |\langle A(x+y), x+y\rangle| + |\langle A(x-y), x-y\rangle| \\ &\leq w(A)\left(\|x+y\|^2 + \|x-y\|^2\right) \\ &= 2w(A)\left(\|x\|^2 + \|y\|^2\right) \end{aligned}$$

oziroma

$$\left|\operatorname{Re}\langle Ax, y\rangle\right| \leq \frac{1}{2} w(A)\left(\|x\|^2 + \|y\|^2\right).$$

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

$$\left|\operatorname{Re}\left\langle Ax, \frac{Ax}{\|Ax\|}\right\rangle\right| \leq w(A),$$

zato $\|Ax\| \leq w(A)$ za vsak x z $\|x\| = 1$ in $Ax \neq 0$. Sedaj lahko naredimo supremum po enotski sferi. □

Vprašanje 53. Kako lahko izraziš normo sebi adjungiranega operatorja? Dokaži.

Trditev 3.5.13. *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 3.5.14. *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.

Vprašanje 54. Pokaži, da v $B(H)$ velja C^*-aksiom.

Trditev 3.5.15. *Za $A \in B(H)$ so naslednje trditve ekvivalentne:*

- *A je normalen,*
- *$\|A^*x\| = \|Ax\|$ za vsak $x \in H$,*
- *če je $\mathbb{F} = \mathbb{C}$, potem je $\operatorname{Re} A \operatorname{Im} A = \operatorname{Im} A \operatorname{Re} A$.*

Dokaz. Ekvivalenca 1 in 2: Velja

$$\|A^*x\| = \|Ax\| \Leftrightarrow \langle A^*x, A^*x\rangle = \langle Ax, Ax\rangle \Leftrightarrow \langle AA^*x, x\rangle = \langle A^*Ax, x\rangle \Leftrightarrow \langle Bx, x\rangle = 0$$

za $B = AA^* - A^*A$. Zgornje velja za vse $x \in H$ natanko tedaj, ko je $B = 0$.

Ekvivalenca 1 in 3: Pišimo $A = H + iK$ za $H = \operatorname{Re} A$ in $K = \operatorname{Im} A$. Potem je A normalen natanko tedaj, ko je $(H - iK)(H + iK) = (H + iK)(H - iK)$, oziroma $HK = KH$. □

Vprašanje 55. Karakteriziraj normalnost operatorja in dokaži karakterizacijo.

Posledica 3.5.16. *Naj bo $A \in B(H)$ normalen in $\lambda \in \mathbb{F}$. Tedaj je* $\ker(A^* - \lambda I) = \ker(A - \overline{\lambda} I)$.

Definicija 3.5.17. Operator $A \in B(H)$ je POZITIVNO SEMIDEFINITEN, če je sebi adjungiran in $\langle Ax, x\rangle \geq 0$ za vsak $x \in H$.

Opomba. Če je A pozitivno definiten, je injektiven in zato

$$\ker A = \ker A^* = (\operatorname{im} A)^\perp,$$

torej $\overline{\operatorname{im} A} = (\operatorname{im} A)^{\perp\perp} = (\ker A)^\perp = H$, torej je $\operatorname{im} A$ gost v H.

Lema 3.5.18. *Naj bosta $A, B \in B(H)$ pozitivno semidefinitna. Potem je $A + B \geq 0$. Če je $A > 0$, je $A + B > 0$.*

Definiramo množici

- $B(H)_{\text{sa}} = \{A \in B(H) \mid A = A^*\}$,
- $B(H)^+_{\text{sa}} = \{A \in B(H)_{\text{sa}} \mid A \geq 0\}$.

Trditev 3.5.19. *Relacija $A \geq B \Leftrightarrow A - B \geq 0$ je delna urejenost na $B(H)_{sa}$.*

Definicija 3.5.20. Podmnožica $C \subseteq X$ realnega vektorskega prostora X je STOŽEC, če velja

- $C + C \subseteq C$,
- $\lambda C \subseteq C$ za vsak $\lambda \geq 0$,
- $C \cap (-C) = \{0\}$.

Vprašanje 56. Kaj je stožec v realnem vektorskem prostoru?

Trditev 3.5.21. *Množica $B(H)^+_{sa}$ je stožec v $B(H)_{sa}$.*

Trditev 3.5.22. *Naj bo $A \in B(H)$ sebi adjungiran in $B \in B(K, H)$.*

- *če $A \geq 0$, potem $B^*AB \in B(K)^+_{sa}$,*
- *velja $-\|A\| I \leq A \leq \|A\| I$,*
- *$B^*B \in B(K)^+_{sa}$ in $BB^* \in B(H)^+_{sa}$.*

Izrek 3.5.23. *Naj bo $A \in B(H)$. Tedaj obstaja natanko en $B \in B(H)^+_{sa}$, da je $A = B^2$.*

3.5.1 Idempotenti in invariantni podprostori

Če je P idempotent, velja

$$x = \underbrace{x - Px}_{\in \ker P} + \underbrace{Px}_{\in \operatorname{im} P}.$$

Ker sta $\ker P$ in $\operatorname{im} P$ zaprta v H, velja $H = \operatorname{im} P \oplus \ker P$.

Definicija 3.5.24. Idempotent $P \in B(H)$ je ORTOGONALNI PROJEKTOR, če je $\ker P = (\operatorname{im} P)^\perp$.

Izrek 3.5.25. *Naj bo $P \in B(H)$ neničelni idempontent. Naslednje trditve so ekvivalentne:*

- *P je ortogonalni projektor,*
- *$\|P\| = 1$,*
- *$P^* = P$,*
- *P je normalen,*
- *$\langle Px, x\rangle \geq 0$ za vsak $x \in H$.*

Dokaz. 1 v 2: Vemo že, da je $\|Px\| \leq \|x\|$ za vsak $x \in H$. Če izberemo $x \in \operatorname{im} P$, je $Px = x$, torej $\|P\| = 1$.

2 v 1: Vzemimo $x \in (\ker P)^\perp$, da je $x - Px \in \ker P$. Sledi $\langle x - Px, x\rangle = 0$, oziroma $\langle Px, x\rangle = \|x\|^2$. Torej

$$\|x\|^2 = \langle Px, x\rangle \leq \|Px\|\,\|x\| \leq \|P\|\,\|x\|^2 = \|x\|^2,$$

iz česar sledi $\|x\| = \|Px\| = \sqrt{\langle Px, x\rangle}$. Potem je

$$\|x - Px\|^2 = \|x\|^2 + \|Px\|^2 - 2\operatorname{Re}\langle x, Px\rangle = \|x\|^2 + \|Px\|^2 - 2\langle Px, x\rangle = 0.$$

Torej $x = Px \in \operatorname{im} P$, oziroma $(\ker P)^\perp \subseteq \operatorname{im} P$.

Za drugo inkluzijo vzemimo $y \in \operatorname{im} P$ in zapišimo $y = y_1 + y_2$ za $y_1 \in \ker P$, $y_2 \in (\ker P)^\perp$. Potem je

$$y = Py = P(y_1 + y_2) = Py_2 = y_2 \in (\ker P)^\perp,$$

torej $\operatorname{im} P \subseteq (\ker P)^\perp$.

1 v 3: Za $x, y \in H$ zapišimo $x = x_1 + x_2$, $y = y_1 + y_2$, kjer sta $x_1, y_1 \in \ker P$ ter $x_2, y_2 \in (\ker P)^\perp$. Računamo

$$\begin{aligned}\langle P^*x, y\rangle = \langle x, Py\rangle = \langle x_1 + x_2, Py_2\rangle = \langle x_1 + x_2, y_2\rangle = \langle x_2, y_2\rangle\\ \langle Px, y\rangle = \langle Px_2, y_1 + y_2\rangle = \langle x_2, y_1 + y_2\rangle = \langle x_2, y_2\rangle.\end{aligned}$$

3 v 4 je očitno. 4 v 1: Operator P je normalen, torej je $(\operatorname{im} P)^\perp = (\ker P)^* = \ker P$. Torej $\operatorname{im} P = P^\perp$.

3 v 5: Računamo

$$\langle Px, x\rangle = \left\langle P^2x, x\right\rangle = \langle Px, P^*x\rangle = \langle Px, Px\rangle \geq 0.$$

5 v 1 izpustimo. □

Vprašanje 57. Kdaj je idempotent $P \in B(H)$ ortogonalni projektor? Karakteriziraj in dokaži.

Definicija 3.5.26. Zaprt podprostor M normiranega prostora X je INVARIANTEN za $A \in B(X)$, če je $AM \subseteq M$.

Trditev 3.5.27. *Zaprt podprostor M je invarianten za $A \in B(H)$ natanko tedaj, ko je $M^\perp$ invarianten za A^*.*

Dokaz. Naj bo M invarianten za A. Če je $y \in M^\perp$, potem je $\langle A^*y, x\rangle = \langle y, Ax\rangle = 0$, in zato $A^*y \in M^\perp$. Drugo implikacijo dobimo, ko zamenjamo vlogi M in $M^\perp$ ter A in A^*. □

Posledica 3.5.28. *Naj bo $A \in B(H)_{sa}$. Potem je zaprt podprostor M invarianten za A natanko tedaj, ko je $M^\perp$ invarianten za A.*

3.5.2 Kompaktni operatorji

Definicija 3.5.29. Naj bo $T : X \to Y$ linearna preslikava med normiranima prostoroma. Operator T je KOMPAKTEN, če je $\overline{T(B_X)}$ kompakt v Y.

Opomba. $B_X = \{x \in X \mid \|x\| = 1\}$

Opomba. Vsak kompakten operator je omejen.

Označimo $K(X, Y) = \{T : X \to Y \mid T \text{ kompakten}\}$.

Izrek 3.5.30. *Za metrični prostor M so naslednje trditve ekvivalentne:*

- *M je kompakten,*
- *vsako neskončno zaporedje v M ima stekališče,*
- *M je poln in povsem omejen (za vsak $\varepsilon > 0$ lahko M pokrijemo s končno mnogo kroglami polmera ε).*

Izrek 3.5.31. *Naj bo $T : X \to Y$ linearen operator. Naslednje trditve so ekvivalentne:*

- *T je kompakten,*
- *T preslika omejene množice v relativno kompaktne množice,*
- *če je $(x_m)_m$ omejeno zaporedje v X, potem ima $(Tx_m)_m$ stekališče v Y.*

Vprašanje 58. Definiraj in karakteriziraj kompaktnost linearnega operatorja.

Trditev 3.5.32. *Naj bosta X, Y Banachova prostora. Potem velja:*

- *$K(X,Y)$ je zaprt podprostor v $B(X,Y)$,*
- *za vsake $K \in K(X,Y)$, $A \in B(X)$ ter $B \in B(Y)$ sta $KA, BK \in K(X,Y)$.*

Dokaz. Samo prva točka. Naj bosta $K_1, K_2 \in K(X,Y)$ ter $\lambda, \mu \in \mathbb{F}$. Naj bo $(x_m)_m$ omejeno zaporedje v X. Ker je K_1 kompakten, obstaja podzaporedje $(x_{n_k})_k$, da $(K_1 x_{n_k})_k$ konvergira. Podobno imamo podzaporedje $(x_{n_{k_j}})_j$, da je $(K_2 x_{n_{k_j}})_j$ konvergentno. Potem $(\mu K_1 x_{n_{k_j}} + \lambda K_2 x_{n_{k_j}})_j$ konvergira v Y.

Naj bo sedaj $(K_n)_n$ zaporedje kompaktnih operatorjev, ki konvergira k nekemu K. Pokažimo, da je K kompakten. Ker je Y Banachov, je $\overline{K(B_X)}$ kompakt natanko tedaj, ko je $\overline{K(B_X)}$ povsem omejena, oziroma ko je $K(B_X)$ povsem omejena. Naj bo $\varepsilon > 0$. Obstaja $n \in \mathbb{N}$, da je $\|K - K_n\| < \varepsilon/3$. Ker je K_n kompakten, obstajajo $x_1, \ldots, x_m \in B_X$, da je

$$K_n(B_X) \subseteq \bigcup_{i=1}^{m} \mathring{B}(Kx_i, \varepsilon/3).$$

Pokažimo še, da je

$$K(B_X) \subseteq \bigcup_{i=1}^{m} \mathring{B}(x_i, \varepsilon).$$

Če je $x \in B_X$, potem obstaja j, da je $\|K_n x_j - K_n x\| < \varepsilon/3$. Velja

$$\begin{aligned} \|Kx_j - Kx\| &\leq \|Kx_j - K_n x_j\| + \|K_n x_j - K_n x\| + \|K_n x - Kx\| \\ &\leq \|K - K_n\| \left(\|x_j\| + \|x\|\right) + \|K_n(x_j - x)\| \\ &< \varepsilon. \end{aligned}$$

□

Vprašanje 59. Pokaži, da je $K(X,Y)$ zaprt podprostor $B(X,Y)$ za Banachova X, Y.

Trditev 3.5.33. *Naj bo $A \in B(X)$ operator z $\dim \operatorname{im} A < \infty$. Potem je A kompakten.*

Dokaz. Velja $A(B_X) \subseteq \operatorname{im} A$. Ker je množica omejena v končnorazsežnem prostoru, je relativno kompaktna. □

Definicija 3.5.34. Operator $A \in B(X,Y)$ je KONČNEGA RANGA, če je $\dim \operatorname{im} A < \infty$. V tem primeru je $\operatorname{rang} A = \dim \operatorname{im} A$.

Definicija 3.5.35. Množico vseh operatorjev končnega ranga $X \to Y$ označimo z $F(X,Y)$.

Vprašanje 60. Pokaži, da je $F(X,Y) \subseteq K(X,Y)$.

Izrek 3.5.36. *Naj bo X normiran prostor. Tedaj je B_X kompakt natanko tedaj, ko je $\dim X < \infty$.*

Posledica 3.5.37. *Identiteta* $\mathrm{id} : X \to X$ *je kompakten operator natanko tedaj, ko je* $\dim X < \infty$.

Naj bo X neskončnorazsežen Banachov prostor. Tedaj je $\mathrm{id} \in B(X) \setminus K(X)$. Tvorimo prostor $B(X)/K(X)$. Ker je X Banachov, je $B(X)/K(X)$ Banachov, ker pa je $K(X)$ zaprti ideal v Banachovi algebri $B(X)$, je tudi $B(X)/K(X)$ Banachova algebra. Pravimo ji CALKINOVA ALGEBRA.

Trditev 3.5.38. *Naj bo X normiran in Y Banachov. Naj bo $A \in K(X,Y)$ ter $\overline{A}$ enolična razširitev A na $\overline{X}$ po zveznosti. Tedaj je $\overline{A} \in K(\overline{X}, \overline{Y})$.*

3.5.3 Izrek Arzela-Ascoli

Definicija 3.5.39. Naj bo K kompakten Hausdorffov prostor. Množica $H \subseteq \mathcal{C}(K)$ je ENAKOZVEZNA, če za vsaka $x \in K$ in $\varepsilon > 0$ obstaja odprta množica $U_x \ni x$, da je $|f(y) - f(x)| < \varepsilon$ za vse $y \in U_x$ in $f \in H$.

Izrek 3.5.40 (Arzela-Ascoli)**.** *Naj bo K kompakten Hausdorffov prostor in $H \subseteq \mathcal{C}(K)$ družina funkcij. Tedaj je H relativno kompaktna natanko tedaj, ko je enakozvezna in po točkah omejena.*

Vprašanje 61. Povej izrek Arzela-Ascoli.

Izrek 3.5.41. *Naj bo $(f_n)_n \subseteq \mathcal{C}([a,b])$ tako zaporedje, da za vse $\varepsilon > 0$ obstaja $\delta > 0$, da za vsak n velja*

$$|y - x| < \delta \implies |f_n(x) - f_n(y)| < \varepsilon.$$

Če je $(f_n)_n$ omejeno zaporedje, je relativno kompaktno.

Dokaz. Naj bo $(f_n)_n$ kot v izreku. Naj bo $(x_n)_n$ množica racionalnih števil na $[a,b]$, postavljena v zaporedje. Zaporedje $(f_n(x_1))_n$ je omejeno v $\mathbb{C}$, torej ima konvergentno podzaporedje $(f_n^{(1)}(x_1))_n$. Zaporedje $(f_n^{(1)}(x_2))_n$ je omejeno, torej ima konvergentno podzaporedje $(f_n^{(2)}(x_2))_n$. Postopek ponavljamo, da dobimo neskončno zaporedij $(f_n^{(k)}(x_k))_n$. Tvorimo $\tilde{f}_n = f_n^{(n)}$. Po konstrukciji $(\tilde{f}_n(x_j))_n$ konvergira za poljuben $j \in \mathbb{N}$.

Naj bo $\varepsilon > 0$. Obstaja $\delta > 0$, da velja

$$|y - x| < \delta \implies |f_n(y) - f_n(x)| < \frac{\varepsilon}{3}.$$

Ker je $[a,b]$ kompakt, obstaja $p \in \mathbb{N}$, da je

$$[a,b] \subseteq \bigcup_{j=1}^{p} (x_j - \delta, x_j + \delta).$$

Obstaja $n_\varepsilon \in \mathbb{N}$, da za vse $m, n \geq n_\varepsilon$ velja $\left|\tilde{f}_n(x_j) - \tilde{f}_m(x_j)\right| < \varepsilon/3$ za vse $j = 1, \ldots, p$. Za $x \in [a,b]$ potem obstaja j, da $|x_j - x| < \delta$, torej

$$\left|\tilde{f}_n(x) - \tilde{f}_m(x)\right| \leq \left|\tilde{f}_n(x) - \tilde{f}_n(x_j)\right| + \left|\tilde{f}_m(x) - \tilde{f}_m(x_j)\right| + \left|\tilde{f}_n(x_j) - \tilde{f}_m(x_j)\right| < \varepsilon,$$

torej je $\left\|\tilde{f}_n - \tilde{f}_m\right\| \leq \varepsilon$ za $m, n \geq n_\varepsilon$. Torej je zaporedje $(\tilde{f}_n)_n$ Cauchyjevo in zato konvergira. Limita je stekališče $(f_n)_n$. □

Vprašanje 62. Pokaži: če za omejeno zaporedje $(f_n)_n$ zveznih funkcij na kompaktnem intervalu za poljuben $\varepsilon > 0$ obstaja $\delta > 0$, da za vsak n iz $|x - y| < \delta$ sledi $|f_n(x) - f_n(y)| < \varepsilon$, potem ima zaporedje stekališče.

Trditev 3.5.42. *Naj bo $H \subseteq \mathcal{C}([a,b])$ družina zvezno odvedljivih funkcij. Če je supremum množice $\{\|f\|_\infty \mid f \in H\}$ končen, potem je H enakozvezna.*

Dokaz. Za $f \in H$ je

$$|f(y) - f(x)| = \left|f'(\xi)(y - x)\right| \leq \sup_{g \in H} \|g\|_\infty |y - x|$$

po Lagrangeovem izreku, za neki ξ. □

Izrek 3.5.43. *Naj bo $k \in \mathcal{C}([a,b] \times [a,b])$ in K integralski operator, dan s predpisom*

$$(Kf)(x) = \int_a^b k(x,y) f(y) dy.$$

Tedaj sta $K : (\mathcal{C}([a,b]), \|\cdot\|_\infty) \to (\mathcal{C}([a,b]), \|\cdot\|_\infty)$ in $K : (\mathcal{C}([a,b]), \|\cdot\|_2) \to (\mathcal{C}([a,b]), \|\cdot\|_2)$ kompaktna operatorja.

Dokaz. Ker je k enakomerno zvezna, za vsak $\varepsilon > 0$ obstaja $\delta > 0$, da iz $|(x,y) - (z,w)| < \delta$ sledi $|k(x,y) - k(z,w)| < \varepsilon$. Potem je

$$\begin{aligned} |Kf(x) - Kf(y)| &\leq \int_a^b |k(x,z) - k(y,z)|\,|f(z)|\,dz \\ &< \varepsilon \int_a^b |f(z)|\,dz \\ &= \varepsilon \langle 1, f \rangle_{L^2} \\ &\leq \varepsilon \sqrt{b-a}\,\|f\|_2\,, \end{aligned}$$

torej je $K(B_{\mathcal{C}([a,b])})$ je enakozvezna. Je tudi omejena:

$$|Kf(x)| \leq \int_a^b |k(x,y) f(y)|\,dy \leq \|k\|_\infty \|f\|_\infty (b-a).$$

Po izreku Arzela-Ascoli je slika krogle relativno kompaktna glede na $\|\cdot\|_\infty$ normo. Za drugo normo vzemimo omejeno zaporedje $(f_n)_n \subseteq \mathcal{C}([a,b])$. Od prej vemo, da je $(Kf_n)_n$ enakozvezna, podobno pokažemo tudi, da je omejena. □

Vprašanje 63. Pokaži, da je integralski operator z zveznih jedrom kompakten na $\mathcal{C}([a,b])$ za drugo ali neskončno normo.

3.5.4 Kompaktnost adjungiranega operatorja

Izrek 3.5.44. *Naj bo $T \in B(H,K)$. Naslednje trditve so ekvivalentne:*

- *T je kompakten,*
- *T^* je kompakten,*
- *obstaja zaporedje $(T_n)_n \subseteq F(H,K)$, da $T_n \to T$.*

Dokaz. 3 v 1 je jasno. 1 v 3: Množica $\overline{T(B_H)}$ je kompaktna v K, torej je separabilna in velja

$$\overline{\operatorname{im} T} \subseteq \bigcup_{n\in\mathbb{N}} n\overline{T(B_H)}.$$

Torej je tudi $\overline{\operatorname{im} T}$ separabilen Hilbertov prostor, in ima števen KONS $(e_n)_n$. Naj bo P_n ortogonalni projektor na linearno ogrinjačo prvih n vektorjev e_i. Projektor je končnega ranga in zato kompakten. Definiramo $T_n = P_nT$.

Če je $y \in \overline{\operatorname{im} T}$, potem je $y = \sum_m \langle y, e_m\rangle e_m$. Velja $P_ny \xrightarrow[n\to\infty]{} y$, torej za $x \in H$ velja $Tx \in \operatorname{im} T$, torej $T_nx \xrightarrow[n\to\infty]{} Tx$. Torej $T_n \to T$ po točkah.

Pokazati moramo še, da $T_n \to T$ v $B(H,K)$. Ker je T kompakten, za vsak $\varepsilon > 0$ obstajajo $x_1, \ldots, x_m \in B_H$. da je

$$T(B_H) \subseteq \bigcup_{j=1}^{m} \mathring{B}(Tx_j, \frac{\varepsilon}{3}).$$

Za $x \in B_H$ obstaja x_j, da je $\|Tx - Tx_j\| < \varepsilon/3$. Potem je

$$\|Tx - T_nx\| \le \|Tx - Tx_j\| + \|Tx_j - T_nx_j\| + \|T_nx_j - T_nx\| < \varepsilon,$$

saj

$$\|T_nx_j - T_nx\| \le \|P_n\| \, \|Tx_j - Tx\| < \frac{\varepsilon}{3}.$$

1 v 2: Najprej izračunamo

$$\|T^*x\|^2 = \langle T^*x, T^*x\rangle = \langle TT^*x, x\rangle \le \|TT^*x\| \, \|x\| .$$

Ker je T kompakten, je TT^* kompakten. Naj bo $(x_n)_n$ omejeno v K, torej obstaja konvergentno podzaporedje $(x_{n_k})_k$. Velja

$$\left\|T^*x_{n_k} - T^*x_{n_j}\right\|^2 \le \left\|TT^*(x_{n_k} - x_{n_j})\right\| \left\|x_{n_k} - x_{n_j}\right\| .$$

Prvi člen konvergira k 0 za $k, j \to \infty$, drugi člen pa je omejen. Torej je zaporedje $(T^*x_{n_k})_k$ Cauchyjevo v H, ki je poln, zato konvergira, in je T^* kompakten.

2 v 1: Velja $T^{**} = T$. □

Vprašanje 64. Pokaži: omejen operator T je kompakten natanko tedaj, ko je T^* kompakten, kar je natanko tedaj, ko obstaja zaporedje $(T_n)_n$ operatorjev končnega ranga, ki konvergira k T.

3.6 Spektralna teorija

Naj bo A kompleksna Banachova algebra z enoto. Naj bo $a \in A$. Definiramo

$$\rho(a) = \{\lambda \in \mathbb{C} \mid \lambda - a \text{ obrnljiv v } A\}.$$

Opomba. Identificiramo $\lambda - a = \lambda \cdot 1 - a$.

Pravimo, da je $\rho(a)$ RESOLVENTA elementa a. Označimo

$$R(\lambda, a) = (\lambda - a)^{-1}.$$

Preslikava $\lambda \mapsto R(\lambda, a)$ je RESOLVENTNA FUNKCIJA. Množici $\sigma(a) = \mathbb{C} \setminus \rho(a)$ pravimo SPEKTER elementa a.

Trditev 3.6.1. *Naj bo $A \in B(H)$. Tedaj je $\sigma(A^*) = \{\overline{\lambda} \mid \lambda \in \sigma(A)\}$.*

Dokaz. Naj bo B tak, da je $(\lambda I - A)B = B(\lambda I - A) = I$. Če obe strani adjungiramo, dobimo $B^*(\overline{\lambda} I - A^*) = (\overline{\lambda} I - A^*)B^* = I$. □

Definicija 3.6.2. Naj bo X kompleksen Banachov prostor in A omejen operator na X. Potem je λ v ZVEZNEM delu spektra ($\lambda \in \sigma_c(A)$), če je $\lambda I - A$ injektiven in $\overline{\mathrm{im}(\lambda I - A)} = X$ ter $\mathrm{im}(\lambda I - A) \neq X$. Pravimo, da je λ v RESIDUALNEM delu spektra ($\lambda \in \sigma_r(A)$), če je $\lambda I - A$ injektiven in $\mathrm{im}(\lambda I - A)$ ni gosta. Pravimo, da je λ v TOČKASTEM delu spektra ($\lambda \in \sigma_p(A)$), če je λ lastna vrednost za A.

Vprašanje 65. Kaj je resolventa in kaj je spekter? Na katere tri dele se razdeli spekter omejenega operatorja na kompleksnem Banachovem prostoru?

Trditev 3.6.3. *Naj bo $A \in B(H)$. Če je $A = A^*$, je $\sigma_p(A) \subseteq \mathbb{R}$. Če je A normalen, so lastni vektorji medseboj pravokotni.*

Trditev 3.6.4. *Naj bo H Hilbertov prostor in $A \in B(H)$. Če je $\lambda \in \sigma_r(A)$, je $\overline{\lambda} \in \sigma_p(A^*)$. Če je $\lambda \in \sigma_p(A)$, je $\overline{\lambda} \in \sigma_p(A^*) \cup \sigma_r(A^*)$.*

Dokaz. Če $\mathrm{im}(A - \lambda I)$ ni gost v H, potem je

$$\ker(A^* - \overline{\lambda} I) = \ker(A - \lambda I)^* = \mathrm{im}(A - \lambda I)^{\perp} \neq \{0\},$$

torej obstaja lastni vektor za $\overline{\lambda}$ za A^*.

Za drugo trditev velja $\ker(\lambda I - A) \neq \{0\}$, če vzamemo pravokotni komplement obeh strani, dobimo $\overline{\mathrm{im}(\overline{\lambda} I - A^*)} \neq H$. □

Vprašanje 66. Kako sta povezana točkast in residualni spekter operatorja s spektrom adjungiranega operatorja? Dokaži.

Posledica 3.6.5. *Če je $A \in B(H)$ normalen, je $\sigma_r(A) = \varnothing$.*

Dokaz. Če je $\lambda \in \sigma_r(A)$, potem je $\overline{\lambda} \in \sigma_p(A^*)$ in zato

$$\ker(A - \lambda I) = \ker(A - \lambda I)^* = \ker(A^* - \overline{\lambda} I) \neq \{0\},$$

kjer smo v prvem koraku uporabili normalnost. Sledi $\lambda \in \sigma_p(A)$. $\rightarrow\!\leftarrow$ □

Vprašanje 67. Pokaži: normalen operator ima prazen residualni spekter.

Lema 3.6.6. *Naj bo X kompleksen Banachov prostor in $A \in B(X)$. Naj bo $\lambda \in \sigma_c(A)$. Tedaj obstaja $(x_n)_n$ z $\|x_n\| = 1$, da $Ax_n - \lambda x_n \to 0$.*

Izrek 3.6.7. *Naj bo $A \in B(H)$ sebi adjungiran operator. Tedaj $\sigma(A) \subseteq \mathbb{R}$.*

Dokaz. Ker je A sebi adjungiran, je normalen, zato $\sigma_r(A) = \varnothing$. Dokazati moramo torej le še $\sigma_c(A) \subseteq \mathbb{R}$. Če je $\lambda \in \sigma_c(A)$, obstaja zaporedje $(x_n)_n$ z $\|x_n\| = 1$ in $Ax_n - \lambda x_n \to 0$. Pišimo $\lambda = \alpha + i\beta$. Potem velja

$$|\langle(\lambda I - A)x_n, x_n\rangle| \leq \|x_n\| \, \|(\lambda I - A)x_n\| \xrightarrow[n\to\infty]{} 0,$$

skalarni produkt pa je enak

$$\langle(\lambda I - A)x_n, x_n\rangle = \langle \alpha x_n - Ax_n, x_n\rangle + i\beta \langle x_n, x_n\rangle = \langle \alpha x_n - Ax_n, x_n\rangle + i\beta,$$

torej $\beta = 0$. □

Vprašanje 68. Pokaži: spekter sebi adjungiranega operatorja je realen.

Izrek 3.6.8. *Naj bo A Banachova algebra z enico in naj bo $\|a\| < 1$ za nek $a \in A$. Tedaj je $1 - a$ obrnljiv in*

$$(1 - a)^{-1} = \sum_{n=0}^{\infty} a^n.$$

Dokaz. Označimo delne vsote z s_n. Če je $n > m$, velja

$$\|s_n - s_m\| = \left\|a^{m+1} + \cdots + a^n\right\| \leq \|a\|^{m+1} + \cdot + \|a\|^n \leq \|a\|^{m+1} \sum_{k=0}^{\infty} \|a\|^k = \frac{\|a\|^{m+1}}{1 - \|a\|},$$

kar konvergira k 0 za $n, m \to \infty$. Torej je zaporedje Cauchyjevo, in obstaja limita s. Velja $(1 - a)s_n = 1 - a^{n+1}$. Leva stran konvergira k $(1 - a)s$, desna pa k 1. □

Vprašanje 69. Pokaži: če je a element Banachove algebre in $\|a\| < 1$, je $1 - a$ obrnljiv.

Lema 3.6.9. *Naj bosta $\lambda, \mu \in \rho(a)$. Velja $R(\mu, a) - R(\lambda, a) = (\lambda - \mu)R(\lambda, a)R(\mu, a) = (\lambda - \mu)R(\mu, a)R(\lambda, a)$.*

Dokaz. Ker $(\lambda - a)$ in $(\mu - a)$ komutirata, komutirata tudi njuna inverza. Preostalo je preprost račun. □

Trditev 3.6.10. *Naj bo $a \in A$, kjer je A kompleksna Banachova algebra z enico. Potem je $\rho(a)$ odprta v $\mathbb{C}$, $\sigma(a)$ pa kompaktna in vsebovana v $B(0, \|a\|)$.*

Dokaz. Naj bo $\lambda_0 \in \rho(a)$. Iščemo $\varepsilon > 0$, da iz $|\lambda - \lambda_0| < \varepsilon$ sledi $\lambda \in \rho(a)$. Izrazimo lahko

$$\lambda - a = (\lambda_0 - a)(1 + (\lambda - \lambda_0)(\lambda_0 - a)^{-1}).$$

Če je $\left\|(\lambda - \lambda_0)(\lambda_0 - a)^{-1}\right\| \leq \varepsilon \left\|(\lambda_0 - a)^{-1}\right\| < 1$, potem je

$$1 + (\lambda - \lambda_0)(\lambda_0 - a)^{-1}$$

obrnljiv, torej $\lambda - a$ produkt dveh obrnljivih elementov.

Za drugo trditev je $\sigma(a) = \mathbb{C} \setminus \rho(a)$ zaprta množica. Če je $|\lambda| > \|a\|$, je

$$\lambda - a = \lambda(1 - \frac{a}{\lambda}),$$

ampak potem $\|a/\lambda\| < 1$, in je $\lambda \in \rho(a)$. Torej je $\sigma(a) \subseteq B(0, \|a\|)$. □

Vprašanje 70. Pokaži: če je a element Banachove algebre, je $\sigma(a)$ kompakt, vsebovan v $B(0, \|a\|)$.

Opomba. Če je $|\lambda| > \|a\|$, je $\lambda \in \rho(a)$ in

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

Opomba. Če je $\lambda_0 \in \rho(a)$ in $|\lambda - \lambda_0| < \left\|(\lambda_0 - a)^{-1}\right\|^{-1}$, potem je $\lambda \in \rho(a)$ in velja

$$(\lambda - a)^{-1} = (\lambda_0 - a)^{-1} \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n (\lambda_0 - a)^{-n}.$$

Trditev 3.6.11. *Naj bo $\mu \in \rho(a)$. Tedaj velja*

$$\lim_{\lambda \to \mu} \frac{(\lambda - a)^{-1} - (\mu - a)^{-1}}{\lambda - \mu} = -(\mu - a)^{-2}.$$

Trditev 3.6.12. *Resolventna funkcija je zvezna.*

Dokaz. Naj bo $\lambda_0 \in \rho(a)$ in $|\lambda - \lambda_0| < \|R(\lambda_0, a)\|^{-1}$. Potem velja

$$R(\lambda, a) = \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n R(\lambda_0, a)^{n+1},$$

torej je

$$R(\lambda, a) - R(\lambda_0, a) = \sum_{n=1}^{\infty} (\lambda_0 - \lambda)^n R(\lambda_0, a)^{n+1}.$$

Iz tega sledi

$$\begin{aligned} \|R(\lambda, a) - R(\lambda_0, a)\| &\leq |\lambda - \lambda_0| \, \|R(\lambda_0, a)\|^2 \sum_{n=0}^{\infty} |\lambda - \lambda_0|^n \, \|R(\lambda_0, a)\|^n \\ &= \frac{|\lambda_0 - \lambda| \, \|R(\lambda_0, a)\|^2}{1 - |\lambda_0 - \lambda| \, \|R(\lambda_0, a)\|} \\ &\xrightarrow[\lambda \to \lambda_0]{} 0. \end{aligned}$$

□

Vprašanje 71. Pokaži: resolventna funkcija je zvezna.

Izrek 3.6.13. *Naj bo $a \in A$ in A kompleksna Banachova algebra z enico. Tedaj je $\sigma(a) \neq \varnothing$.*

Dokaz. Recimo $\sigma(a) = \varnothing$. Potem je $\rho(a) = \mathbb{C}$, ker pa velja

$$\lim_{\lambda \to \mu} \frac{R(\lambda, a) - R(\mu, a)}{\lambda - \mu} = -R(\mu, a)^2,$$

imamo za vsak $\varphi \in A^*$

$$\varphi\left(\lim_{\lambda \to \mu} \frac{R(\lambda, a) - R(\mu, a)}{\lambda - \mu} \right) = \lim_{\lambda \to \mu} \frac{\varphi(R(\lambda, a)) - \varphi(R(\mu, a))}{\lambda - \mu} = -\varphi(R(\mu, a)^2),$$

torej je $\varphi \circ R(\cdot, a)$ cela holomorfna funkcija.

Če je $|\lambda| > \|a\|$, je

$$\|R(\lambda, a)\| \leq \sum_{n=0}^{\infty} \frac{\|a\|^n}{|\lambda|^{n+1}} = \frac{1}{|\lambda|} \frac{1}{1 - \|a\|/|\lambda|},$$

to pa konvergira k 0 za $|\lambda| \to \infty$. Torej je $R(\cdot, a)$ omejena, zato je tudi $\varphi \circ R(\cdot, a)$ omejena. Ker je cela holomorfna, je konstantna in velja $\varphi(R(\lambda, a)) = 0$ za vsak $\lambda \in \mathbb{C}$. Za $\lambda \in \rho(a)$ velja $(\lambda - a)R(\lambda, a) = 1$, torej $R(\lambda, a) \neq 0$, in smo prišli v protislovje s Hahn-Banachom. □

Vprašanje 72. Dokaži: spekter elementa Banachove algebre z enico je vedno neprazen.

Definicija 3.6.14. Naj bo $a \in A$. Definirajmo

$$r(a) = \sup_{\lambda \in \sigma(a)} |\lambda| = \max_{\lambda \in \sigma(a)} |\lambda| .$$

To število imenujemo SPEKTRALNI RADIJ elementa a.

Lema 3.6.15. *če je $\lambda \in \sigma(a)$, je $\lambda^n \in \sigma(a^n)$.*

Dokaz. Dokažemo obratno; če je $\lambda^n \in \rho(a^n)$, je $\lambda \in \rho(a)$. Računamo

$$\lambda^n - a^n = (\lambda - a)(\lambda^{n-1} + \cdots + a^{n-1})$$

in desni člen produkta označimo z b. Obstaja $c \in A$, da je $(\lambda^n - a^n)c = c(\lambda^n - a^n) = 1$, torej $(\lambda - a)bc = 1$ in ima $\lambda - a$ desni inverz. Podobno pokažemo, da ima levi inverz. □

Trditev 3.6.16. *Naj bo $a \in A$ in p polinom. Tedaj je $\sigma(p(a)) = p(\sigma(a))$.*

Izrek 3.6.17 (Gelfandova formula)**.** *Naj bo $a \in A$. Potem je*

$$r(a) = \lim_{n\to\infty} \|a^n\|^{1/n} = \liminf_{n\to\infty} \|a^n\|^{1/n} = \inf_{n\in\mathbb{N}} \|a^n\|^{1/n}.$$

Dokaz. Naj bo $\lambda \in \sigma(a)$, da je $|\lambda| = r(a)$. Potem je $\lambda^n \in \sigma(a^n)$, torej $|\lambda|^n \leq \|a^n\|$ in velja $r(a) = |\lambda| \leq \|a^n\|^{1/n}$ za vsak $n \in \mathbb{N}$. Zaporedje je navzdol omejeno, torej ima infimum, ki je

$$r(a) \leq \inf_{n\in\mathbb{N}} \|a^n\|^{1/n} \leq \liminf_{n\to\infty} \|a^n\|^{1/n}.$$

Če pokažemo, da je $\limsup \|a^n\|^{1/n} \leq r(a)$, potem bo dokaz končan.

Naj bo $f \in A^*$. Definiramo $\tilde{f}(\lambda) = f(R(\lambda, a))$ kot preslikavo $\rho(a) \to \mathbb{C}$. To je holomorfna funkcija. Če je $|\lambda| > \|a\|$, lahko $R(\lambda, a)$ eksplicitno izrazimo in zato velja

$$\tilde{f}(\lambda) = \sum_{n=0}^{\infty} \frac{f(a^n)}{\lambda^{n+1}}.$$

To je Laurantov razvoj $\tilde{f}$ v okolici 0 na komplement krogle $B(0, \|a\|)$. Razvoj lahko naredimo tudi na $B(0, r(a))^{\mathsf{c}}$. Ker je enoličen, zgornji predpis velja za vsak $\lambda > r(a)$. Za $\varepsilon > 0$ in $\lambda = r(a) + \varepsilon$ torej konvergira vrsta

$$\sum_{n=0}^{\infty} \frac{f(a^n)}{(r(a)+\varepsilon)^{n+1}},$$

torej obstaja $M > 0$, da je $\left|f(a^n)/(r(a)+\varepsilon)^{n+1}\right| \leq M$ za vse $n \in \mathbb{N}$. Po principu šibke omejenosti je potem množica

$$\left\{ \frac{a^n}{(r(a)+\varepsilon)^{n+1}} \mid n \in \mathbb{N} \right\}$$

omejena in obstaja $\tilde{M} > 0$, da

$$\left\| \frac{a^n}{(r(a)+\varepsilon)^{n+1}} \right\| \leq \tilde{M}$$

za vse $n \in \mathbb{N}$. Sledi

$$\limsup_{n\to\infty} \|a^n\|^{1/n} \leq \lim_{n\to\infty} (\tilde{M}(r(a)+\varepsilon))^{1/n}(r(a)+\varepsilon) = r(a)+\varepsilon,$$

saj prvi člen produkta konvergira k 1. Torej je $\limsup \|a^n\|^{1/n} \leq r(a)$. □

Vprašanje 73. Povej in dokaži Gelfandovo formulo.

Posledica 3.6.18. *Naj bo A sebi adjungiran operator na Hilbertovem prostoru. Tedaj velja $r(A) = \|A\|$.*

Dokaz. Po izreku velja $r(A) = \lim \|A^n\|^{1/n}$. Ker je $\left\|A^2\right\| = \|A^*A\| = \|A\|^2$, lahko z indukcijo pokažemo, da velja $\left\|A^{2^n}\right\| = \|A\|^{2^n}$. S tem dobimo podzaporedje zaporedja v limiti, ki bo konvergiralo k $\|A\|$. □

Vprašanje 74. Pokaži, da za sebi adjungiran operator velja $\|A\| = r(A)$.

3.6.1 Spekter kompaktnega operatorja

Primer. Naj bo $H = l^2$ in $(e_n)_n$ neka ortonormirana baza. Naj bo $(d_n)_n$ zaporedje omejenih števil. Definiramo diagonalni operator $D : l^2 \to l^2$ z

$$Dx = \sum_{n=0}^{\infty} d_n \langle x, e_n \rangle e_n.$$

Preverimo lahko naslednja dejstva:

- $\|D\| = \sup_n |d_n|$,
- D je sebi adjungiran natanko tedaj, ko je $d_n \in \mathbb{R}$ za vsak n,
- D je normalen,
- D je unitaren natanko tedaj, ko je $|d_n| = 1$ za vsak n,
- D je kompakten natanko tedaj, ko je $d_n \to 0$.

Lema 3.6.19 (Rieszova lema o pravokotnici)**.** *Naj bo X normiran prostor in Y zaprt podprostor v X. Tedaj za vsak $\varepsilon \in (0,1)$ obstaja $x \in X$, da velja $\|x\| = 1$ in $d(x, Y) \geq \varepsilon$.*

Vprašanje 75. Povej Rieszovo lemo o pravokotnici.

Opomba. Če je X končnorazsežen, obstaja $x \in X$ z $\|x\| = 1$ in $d(x, Y) = 1$. Skličemo se lahko na kompaktnost.

Trditev 3.6.20. *Naj bo Y končnorazsežen podprostor v normiranem prostoru X. Tedaj obstaja zaprt podprostor $Z \subseteq X$, da je $X = Y \oplus Z$.*

Dokaz. Naj bo $\{e_1, \dots, e_n\}$ baza za Y in $\{f_1, \dots, f_n\}$ dualna baza. Ker je dimenzija Y končna, je f_i omejen za vse $i = 1, \dots, n$. Naj bo F_i katerakoli Hahn-Banachova razširitev funkcije f_i. Definiramo

$$Z = \bigcap_{n=1}^{n} \ker F_i.$$

Ker je F_i omejen, je $\ker F_i$ zaprta množica, torej je presek zaprt.

Če je $x \in X$, lahko izrazimo

$$x = \underbrace{F_1(x)e_1 + \dots + F_n(x)e_n}_{\in Y} + \underbrace{x - F_1(x)e_1 - \dots - F_n(x)e_n}_{\in Z}.$$

Če je $x \in Y \cap Z$, velja $x = \alpha_1 e_1 + \dots + \alpha_n e_n$, ampak tudi $F_i(x) = 0 = \alpha_i$ za vse i, torej $x = 0$. □

Vprašanje 76. Dokaži: za končnorazsežen podprostor $Y \leq X$ obstaja zaprt podprostor $Z \leq X$, da je $X = Y \oplus Z$.

Trditev 3.6.21. *Naj bo X Banachov prostor in $K \in K(X)$. Če je $\lambda \neq 0$, je $\dim \ker(K - \lambda I) < \infty$. Dodatno je $\operatorname{im}(K - \lambda I)$ zaprta v X.*

Dokaz. Izrazimo lahko $K - \lambda I = \lambda(\frac{1}{\lambda}K - I)$, torej je $\ker(K - \lambda I) = \ker(\frac{1}{\lambda}K - I)$. Označimo drugo preslikavo z S. Velja $x \in \ker(S - I) \Leftrightarrow Sx = x$, torej je $S|_{\ker(S-I)}$ identiteta, in hkrati kompakten operator. Torej je jedro končnorazsežno.

Za drugo trditev zapišemo $X = \ker(K - \lambda I) \oplus Y$, kjer je Y zaprt podprostor X. Velja

$$\operatorname{im}(K - \lambda I) = (K - \lambda I)(\ker(K - \lambda I) \oplus Y) = (K - \lambda I)Y = \operatorname{im}(K - \lambda I|_Y).$$

Označimo $T = K - \lambda I$.

Dovolj je dokazati, da je $T|_Y$ navzdol omejen, saj bo potem topološki izomorfizem $Y \to \operatorname{im} T$. Ker je Y zaprt, je Banachov, zato bo $\operatorname{im} T$ Banachov, polni prostori pa so vedno zaprti.

Recimo, da $T|_Y$ ni navzdol omejen. Tedaj za vsak $n \in \mathbb{N}$ obstaja $y_n \in Y$, da $\|y_n\| = 1$ in $\|Ty_n\| \leq 1/n$. Ker je K kompakten, obstaja podzaporedje $(y_{n_k})_k$, da $(Ky_{n_k})_k$ konvergira. Velja

$$y_{n_k} = \lambda \frac{1}{\lambda} y_{n_k} = \frac{1}{\lambda}\left(\lambda y_{n_k} - Ky_{n_k} + Ky_{n_k}\right) \to y$$

za nek $y \in Y$ z $\|y\| = 1$ saj sta prva dva člena vsote del zaporedja Ty_n, ki konvergira. Enostavno lahko pokažemo, da velja $Ky = \lambda y$. To pomeni, da je $y \in \ker(K - \lambda I) \cap Y = \{0\}$, ampak $\|y\| = 1$. $\rightarrow\!\!\!\!\!\!\times\!\!\!\!\!\!\leftarrow$ □

Vprašanje 77. Pokaži: če je K kompakten operator in $\lambda \neq 0$, je jedro $K - \lambda I$ končnorazsežno, slika $K - \lambda I$ pa je zaprta.

Lema 3.6.22. *Naj bo K kompakten operator na kompleksnem Banachovem prostoru X. Tedaj ima K za vsak $\varepsilon > 0$ le končno mnogo linearno neodvisnih lastnih vektorjev za lastne vrednosti λ z $|\lambda| \geq \varepsilon$.*

Dokaz. Recimo, da imamo $\varepsilon > 0$ in zaporedje $(x_n)_n$ linearno neodvisnih lastnih vektorjev, ki zaporedoma pripadajo lastnim vrednostim λ_n z $|\lambda_n| \geq \varepsilon$. Definirajmo X_n kot linearno ogrinjačo prvih n lastnih vektorjev. Velja $\dim X_n < \infty$, torej so ti prostori zaprti, hkrati pa velja

$$\{0\} \subsetneq X_1 \subsetneq X_2 \subsetneq X_3 \subsetneq \cdots$$

Po Rieszovi lemi o pravokotnici obstaja $y_n \in X_n \setminus X_{n-1}$ z $\|y_n\| = 1$, da je $d(y_n, X_{n-1}) \geq \frac{1}{2}$ za vse $n \geq 2$. Pišimo $y_n = \alpha_1 x_1 + \cdots + \alpha_n x_n$ in $z_n = y_n/\lambda_n$. Potem velja $\|z_n\| = |\lambda_n|^{-1} \leq \frac{1}{\varepsilon}$ ter

$$Kz_n - y_n = \alpha_1 \left(\frac{\lambda_1}{\lambda_n} - 1 \right) x_1 + \cdots + \alpha_{n-1} \left(\frac{\lambda_{n-1}}{\lambda_n} - 1 \right) x_{n-1} \in X_{n-1}.$$

Za $m < n$ izračunamo

$$\|Kz_n - Kz_m\| = \| \underbrace{y_n}_{\in X_n} - \underbrace{(y_n - Kz_n + Kz_m)}_{\in X_{n-1}} \| \geq \frac{1}{2}$$

torej $(Kz_n)_n$ nima Cauchyjevega podzaporedja, zato nima stekališča. To je protislovje s kompaktnostjo K. □

Vprašanje 78. Pokaži, da ima kompakten operator le končno mnogo linearno neodvisnih lastnih vektorjev, ki pripadajo lastnim vrednostim z $|\lambda| > \varepsilon$.

Izrek 3.6.23. *Za kompakten operator K na kompleksnem Banachovem prostoru X veljajo naslednje trditve:*

- *če je $\dim X = \infty$, je $0 \in \sigma(K)$,*
- *če je $\lambda \in \sigma(K) \setminus \{0\}$, je λ lastna vrednost K,*
- *$\sigma(K)$ je končen ali številno neskončen,*
- *če je $\sigma(K)$ neskončen in so $(\lambda_n)_n$ lastne vrednosti, štete z geometrijsko večkratnostjo, potem $\lambda_n \to 0$.*

Dokaz. Prva točka: Če $0 \notin \sigma(K)$, je K obrnljiv in je zato $\mathrm{id} = KK^{-1}$ kompaktna, torej je $\dim X < \infty$.

Tretja točka: Računamo

$$\sigma(K) \setminus \{0\} = \bigcup_{n \in \mathbb{N}} \sigma(K) \cap \left\{ \lambda \in \mathbb{C} \mid |\lambda| \geq \frac{1}{n} \right\}.$$

To je številna unija končnih množic.

Četrta točka: Iz leme dobimo, da za vsak $\varepsilon > 0$ obstaja n_ε, za katerega za poljuben $n \geq n_\varepsilon$ velja $|\lambda_n| < \varepsilon$. Torej $\lambda_n \to 0$.

Druga točka: Recimo, da je $0 \neq \lambda \in \sigma(K)$. Recimo, da je $K - \lambda I$ injektiven. Če dokažemo, da je $K - \lambda I$ surjektiven, bo sledilo, da je obrnljiv z omejenim inverzom, oziroma $\lambda \in \rho(K)$.

Pokažimo prvo, da je $\operatorname{im}(K-\lambda I)^n$ zaprta za vsak n. Trditev pove, da ima $K-\lambda I$ zaprto sliko. Izrazimo lahko

$$(K - \lambda I)^n = \sum_{n=0}^{\infty} \binom{n}{k} (-\lambda)^k K^{n-k} =: \tilde{K} + (-\lambda)^n I.$$

Operator $\tilde{K}$ je kompakten, torej ima $(K - \lambda I)^n$ zaprto sliko.

Definirajmo $X_n = \ker(K - \lambda I)^n$ za $n \in \mathbb{N}_0$. Velja $X_0 = \{0\} \subseteq X_1 \subseteq X_2 \subseteq \cdots$. Recimo, da so vse inkluzije stroge. Po Rieszovi lemi o pravokotnici za vsak $n \in \mathbb{N}$ obstaja $x_n \in X_n \setminus X_{n-1}$ z $\|x_n\| = 1$ in $d(x_n, X_{n-1}) \geq \frac{1}{2}$. Za $m < n$ je

$$\|Kx_m - Kx_n\| = \|\lambda x_n + Kx_n - \lambda x_n - Kx_m\| = |\lambda| \left\| x_n + \frac{1}{\lambda}(K - \lambda I)x_n - \frac{1}{\lambda} K x_m \right\|.$$

Velja $x_m \in X_m$, torej $Kx_m \in X_m$ (saj K in $K - \lambda I$ komutirata), iz česar sledi $Kx_m \in X_{n-1}$, saj je $X_m \subseteq X_{n-1}$. Hkrati je $(K - \lambda I)x_n \in X_{n-1}$, torej $\frac{1}{\lambda}((K - \lambda I)x_n - Kx_m) \in X_{n-1}$. Velja torej $\|Kx_n - Kx_m\| \geq \frac{1}{2}|\lambda|$. Kot prej pridemo v protislovje s kompaktnostjo operatorja K. Torej mora obstajati n, za katerega je $X_n = X_{n+1}$.

Za konec recimo, da $K-\lambda I$ ni surjektiven. Potem $X \neq (K-\lambda I)X$, velja pa $(K-\lambda I)^n X = (K - \lambda I)^{n+1} X$. To je protislovje z injektivnostjo $(K - \lambda I)^n$. □

Vprašanje 79. Kaj lahko poveš o spektru kompaktnega operatorja na kompleksnem Banachovem prostoru? Dokaži.

Izrek 3.6.24. *Naj bo $K : H \to H$ kompakten sebi adjungiran operator na kompleksnem Hilbertovem prostoru H. Potem obstajata zaporedje $(\lambda_n)_n \subseteq \mathbb{R}$ in ortonormiran sistem $(e_n)_n$ (lahko sta oba končna) z*

- $|\lambda_1| \geq |\lambda_2| \geq \cdots$, $\lambda_n \neq 0$, *in če je zaporedje neskončno, velja* $\lambda_n \to 0$,
- $Ke_n = \lambda_n e_n$,
- *če je* $\lambda \in \sigma_p(K) \setminus \{0\}$, *se* λ *pojavi v* $(\lambda_n)_n$ *natanko tolikokrat, kot je* $\dim \ker(K-\lambda I)$,
- $Kx = \sum \lambda_n \langle x, e_n \rangle e_n$.

Dokaz. Ker je K sebi adjungiran, velja $r(K) = \|K\|$ in $\sigma(K) \subseteq [-\|K\|, \|K\|]$. Če je $K = 0$, nimamo nič za dokazati. Sicer lahko razvrstimo lastne vrednosti K v padajoče zaporedje po absolutni vrednosti, pri čemer jih štejemo glede na geometrijske večkratnosti. Če je $\lambda \in \sigma_p(K) \setminus \{0\}$, lahko izberemo ortonormirano bazo za $\ker(K - \lambda I)$. Če je

$\mu \neq \lambda$ tudi neničelna lastna vrednost, bo njena ONB pravokotna na ONB za λ. Postopek ponovimo za vsako lastno vrednost v zaporedju. Dobljene vektorje zložimo v $(e_n)_n$, da je $Ke_n = \lambda_n e_n$. Definiramo Y kot zaprtje linearne ogrinjače teh vektorjev. Potem je $(e_n)_n$ ONB za Y in $H = Y \oplus Y^\perp$.

Ker je $Ke_n = \lambda_n e_n$, je linearna ogrinjača $(e_n)_n$ invariantna na K, ki je zvezen, torej enako velja za Y. Iz sebi adjungiranosti sledi $K(Y^\perp) \subseteq Y^\perp$. Če je $K|_{Y^\perp} \neq 0$, je njegova norma neničelna lastna vrednost za K, kar je nemogoče, ker smo že našli vse lastne vrednosti. Torej za $x \in H$ velja $Kx = \sum \lambda_n \langle x, e_n \rangle e_n$. □

Vprašanje 80. Kaj lahko poveš o spektru kompaktnega sebi adjungiranega operatorja na kompleksnem Hilbertovem prostoru? Dokaži.

4 Statistika 2

4.1 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 4.1.1. *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 4.1.2 (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 $\|S_i\| \in \{0, 1\}$. □

4.1.1 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 4.1.3. *Statistike, ki so od vzorca odvisne le preko T, so avtomatično nepristranske cenilke z enakomerno najmanjšo disperzijo za svojo pričakovano vrednost.*

Izrek 4.1.4. *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. □

4.2 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 4.2.1. *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 4.2.2. *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 4.2.3. *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 4.2.4 (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 $X_n + Y_n \xrightarrow{d} X + c$, $X_n Y_n \xrightarrow{d} cX$ 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 4.2.5. *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 4.2.6 (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 4.2.7 (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 4.2.8 (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*

$$\sqrt{n}\left(\frac{1}{n}\sum_i X_i - \mu\right) \xrightarrow[n\to\infty]{d} N(0,\Sigma).$$

4.2.1 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 4.2.9 (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 4.2.10. 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 4.2.11. *Naj λ_{min} označuje najmanjšo lastno vrednost matrike. Če $\lambda_{min}(Z^TZ)$ v porazdelitvi konvergira k ∞, je zaporedje cenilk za regresijski parameter β SKN dosledno.*

Trditev 4.2.12. *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 4.2.13. *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. □

4.2.2 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, ki je tu enaka $\mathrm{SKN}(\theta) = E_\theta(\|T(X) - e(\theta)\|^2)$. 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 4.2.14. 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 4.2.15. 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).

4.2.3 Asimptotična normalnost

V zgornjem kontekstu (definicija 4.2.14) 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 4.2.16. *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*

$$\frac{1}{\sigma}(Z^TZ)(\hat\beta - \beta) \xrightarrow[n\to\infty]{d} N(0, I).$$

4.2.4 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 4.2.17. *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 $\sqrt{n}(\hat{\theta}^{(n)} - \theta) \to N(0, \text{FI}(\theta)^{-1})$.

4.3 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 $\alpha \in \{0.05, 0.1, 0.01\}$.

4.3.1 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 4.3.1 (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*

$$-2 \log \lambda(X_1, \ldots, X_n) \xrightarrow[n \to \infty]{d} \chi^2(\dim \Theta - \dim H).$$

5 Simetrije grafov

5.1 Introduction

Definition 5.1.1. Let Γ be a graph with vertex set Ω and adjacency relation $\sim$. A permutation g of Ω is an AUTOMORPHISM or a SYMMETRY of Γ if for any $u, v \in \Omega$, u and v are adjacent if and only if u^g and v^g are adjacent.

Example. Consider the symmetries of the Petersen graph. The standard picture is invariant under rotations by a multiple of $2\pi/5$, and by the 5 reflections along a line through the origin and an outer vertex. We can describe these symmetries by permutations (labeling the vertices as in figure 5.1a):

- A one-step rotation in the negative direction is represented by

  $$\rho = (1\,2\,3\,4\,5)(6\,7\,8\,9\,10).$$

  Then the other rotations are ρ^2, ρ^3, ρ^4 and $\rho^5 = \mathrm{id}$.

- The reflection around the vertical line is

  $$\tau = (1)(2\,5)(3\,5)(6)(7\,10)(8\,9).$$

  We have more reflections, for example

  $$\tau' = (2)(1\,3)(4\,5)(7)(6\,8)(9\,10).$$

The orbits of all the above described permutations are $\{1, 2, 3, 4, 5\}$ and $\{6, 7, 8, 9, 10\}$. But there is another automorphism,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 6 & 8 & 10 & 7 & 9 & 1 & 3 & 5 & 2 & 4 \end{pmatrix} = (1\,6)(2\,8\,5\,9)(3\,10\,4\,7)$$

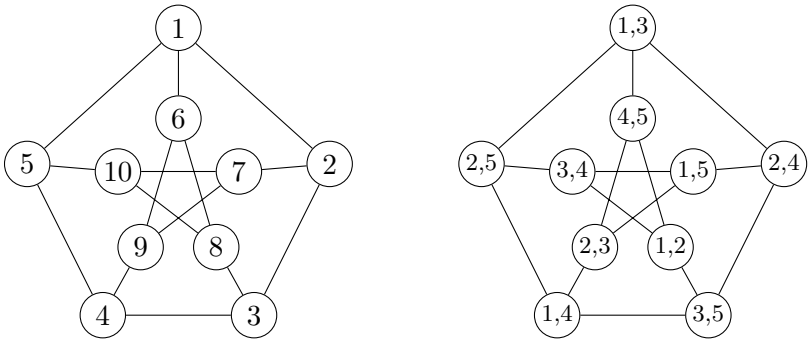
This permutation is of degree 4. It in interesting to consider σ^2; it must map the outer cycle to the inner cycle, and vice versa, so it must be in $\langle \rho, \tau \rangle$. We can compute $\sigma^2 = \tau$. As it turns out, there is no symmetry of this graph which maps the outer cycle to the inner cycle, and vice versa, but is of order 2.

Consider also the conjugation $\rho^\sigma = \sigma^{-1}\rho\sigma$. We can compute $\rho^\sigma = \rho^2$. Since $\tau^\sigma = (\sigma^2)^\sigma = \tau$,

$$\langle \tau, \rho \rangle^\sigma = \langle \tau^\sigma, \rho^\sigma \rangle = \langle \tau, \rho^2 \rangle = \langle \tau, \rho \rangle$$

and we have $\langle \tau, \rho \rangle \trianglelefteq H := \langle \rho, \tau, \sigma \rangle$ and $\sigma^2 \in \langle \tau, \rho \rangle$. So $|H| = 2\,|\langle \tau, \rho \rangle| = 20$. But there are more automorphisms.

We can relabel the vertices if the Petersen graph with elements of the set $\binom{[5]}{2}$ (note that for a set S, $\binom{S}{k}$ denotes the set of all k-subsets of S), as in figure 5.1b. Note that $U, V \in \binom{[5]}{2}$ are connected if and only if they are disjoint (as sets). For any permutation $g \in S_5$, we can find a permutation $\alpha_g \in \mathrm{Sym}\binom{[5]}{2}$ which permutes the pairs by permuting their elements as g would. So there are at least $|S_5| = 120$ automorphisms of the Petersen graph. Note that $g \mapsto \alpha_g$ is a group homomorphism.



(a) The first labeling. (b) The second ordering

Figure 5.1: The Petersen graph

Definition 5.1.2. Let Ω be a nonempty finite set. A PERMUTATION on Ω is a bijection $\Omega \to \Omega$. We denote the set of all permutations of Ω with $\operatorname{Sym}\Omega$, or S_n, if $\Omega = [n]$.

Definition 5.1.3. If $g \in \operatorname{Sym}\Omega$ and $\omega \in \Omega$, then $\omega^g = g(\omega)$ denotes the application of g on ω.

Definition 5.1.4. The composition of two permutations $g, h \in \operatorname{Sym}\Omega$ is defined as the permutation which acts as

$$g \circ h : \omega \to g(h(\omega)) = (\omega^h)^g.$$

We introduce the inverse composition as our product operation on $\operatorname{Sym}\Omega$, defined as

$$gh = h \circ g.$$

Then $\omega^{(gh)} = (\omega^g)^h =: \omega^{gh}$. Note that both $(\operatorname{Sym}\Omega, \circ)$ and $(\operatorname{Sym}\Omega, \cdot)$ are groups, and they are different groups, but they are isomorphic.

Definition 5.1.5. We denote the set of all transpositions of Ω with T_Ω.

Remark. Note that $\langle T_\Omega \rangle = \operatorname{Sym}\Omega$.

Definition 5.1.6. Denote by $\operatorname{Alt}\Omega$ the set of all even permutations of Ω.

Remark. This is an index-2 normal subgroup in $\operatorname{Sym}\Omega$.

Definition 5.1.7. The CYCLIC group of order n is $\operatorname{Cyc}(n) = \langle (1\,2\,3\,\ldots\,n) \rangle \leq S_n$.

Definition 5.1.8. The DIHEDRAL group of order n is

$$\operatorname{Dih}(n) = \begin{cases} \left\langle (1\,2\,3\,\ldots\,n), (1)(2,n)(3,n-1)\ldots(\frac{n}{2},\frac{n+2}{2}) \right\rangle & n \text{ even} \\ \left\langle (1\,2\,3\,\ldots\,n), (1)(2,n-1)\ldots(\frac{n+1}{2},\frac{n+3}{2}) \right\rangle & n \text{ odd} \end{cases}$$

as a subgroup of S_n.

Remark. This is the automorphism group of a cyclic graph on n vertices.

Definition 5.1.9. Let G be an abstract group and Ω a set. A GROUP ACTION is a group homomorphism $\rho : G \to \operatorname{Sym}\Omega$.

Remark. We often write ω^g when we really mean $\omega^{\rho(g)}$.

Remark. As a shorthand, we can write $\rho(g) = g^\Omega$, and $G^\Omega = \rho(G) = \{g^\Omega \mid g \in G\}$.

Definition 5.1.10. An action is FAITHFUL if the kernel of the action is trivial.

Remark. Note that $G^\Omega = \operatorname{im}\rho \cong G/\ker\rho$.

Remark. Any faithful action is an isomorphism $G \to G^\Omega$, so we may view G as a subgroup in $\operatorname{Sym}\Omega$. So G becomes a permutation group.

Definition 5.1.11. The POINT STABILISER of an element $\omega \in \Omega$ is the subgroup $G_\omega = \{g \in G \mid \omega^g = \omega\}$.

Definition 5.1.12. The SET STABILISER G_Δ of a subset $\Delta \subseteq \Omega$ is the subgroup $G_\Delta = \{g \in G \mid \Delta^g = \Delta\}$.

Definition 5.1.13. The POINT-WISE STABILISER $G_{(\Delta)}$ of a subset $\Delta \subseteq \Omega$ is the subgroup $G_{(\Delta)} = \bigcap_{\omega\in\Delta} G_\omega$.

Remark. Note that G_Δ has an induced action on Δ. But even if the action of G on Δ is faithful, G_Δ might act on Δ unfaithfully. If we label the induced permutation group with G^Δ_Δ, then we have

$$G^\Delta_\Delta \cong G_\Delta/G_{(\Delta)}.$$

Definition 5.1.14. The ORBIT of a point $\omega \in \Omega$ is the set $\omega^G = \{\omega^g \mid g \in G\}$. We label $\Omega/G = \{\omega^g \mid \omega \in \Omega\}$.

Remark. The set Ω/G is a partition of Ω into disjoint sets.

Definition 5.1.15. A group G acts TRANSITIVELY on Ω if $|\Omega/G| = 1$.

Lemma 5.1.16. *Let G act on Ω faithfully and let $\Omega_1, \ldots, \Omega_r$ be the orbits of this action. Then we can embed G into the group $G^{\Omega_1} \times \cdots \times G^{\Omega_r}$. Additionally, if $\pi_i : G^{\Omega_1} \times \cdots \times G^{\Omega_r} \to G^{\Omega_i}$ is the canonical projection, then the composition of π_i and the embedding is surjective.*

Remark. We call $G^{\Omega_1}, \ldots, G^{\Omega_r}$ the TRANSITIVE CONSTITUENTS.

Remark. In this case, we say that G is a SUBDIRECT PRODUCT of $G^{\Omega_1}, \ldots, G^{\Omega_r}$ (it is a subset of a direct product).

Lemma 5.1.17 (orbit-stabiliser)**.** *If G acts on Ω and $\omega \in \Omega$, then $|G| = |G_\omega| \cdot |\omega^G|$.*

Definition 5.1.18. Let $g \in G$. Then $\operatorname{Fix}_\Omega(g) = \{\omega \in \Omega \mid \omega^g = \omega\}$ is the set of fixed points of g, and $\operatorname{Supp}_\Omega(g) = \Omega \setminus \operatorname{Fix}_\Omega(g)$.

Lemma 5.1.19 (Cauchy-Frobenius / the lemma that is not Burnside's)**.** *Let G act on Ω. Then*

$$|\Omega/G| = \frac{1}{|G|}\sum_{g\in G}|\mathrm{Fix}_\Omega(g)|\,.$$

Corollary 5.1.20. *If G acts transitively on Ω, then it contains an element that fixes no points.*

Remark. We call such elements DERANGEMENTS.

Definition 5.1.21. An action ρ is REGULAR if it is transitive and the stabiliser of a point is trivial.

Remark. If an action is regular, then the stabiliser of every point is trivial.

Example (action by right multiplication). Let G be a group. Define $\Omega = G$ and for $g, h \in G$, $h^g := hg$. This is a regular action.

Example (action by left multiplication). Let G be a group. Define $\Omega = G$ and for $g, h \in G$, $h^g := g^{-1}h$. This is also a regular action.

Remark. All regular actions are isomorphic to one another.

Example (action by conjugation). Again set $\Omega = G$ and for $g, h \in G$, define $h^g = g^{-1}hg$. The stabiliser of a point h is precisely the centraliser of h in the group, i.e. the elements of G which commute with h. The kernel of this action is

$$\bigcap_{h\in G} G_h = \bigcap_{h\in G} C_G(h) = Z(G).$$

This action is then faithful if and only if the center is trivial.

Example (action on subgroups). Let $\Omega = S(G) = \{H \mid H \leq G\}$ be the set of all subgroups of G. Then define $H^g = g^{-1}Hg$. The centraliser of H is precisely the normaliser of H in G.

Example (action on Sylow subgroups). Let G be a group with $|G| = p^e m$ where $p \in \mathbb{P}$, $e \in \mathbb{N}$ and p does not divide m. Let $\mathrm{Syl}_p(G)$ be the set of all Sylow p-subgroups of G. If we take $\Omega = \mathrm{Syl}_p(G)$, then we can introduce $P^g = g^{-1}Pg$. This action is transitive.

Example (action on cosets). Let $H \leq G$ be a subgroup of G. Take $\Omega = G/H$ be the set of left cosets of H, and introduce the action $(Hh)^g = H(hg)$. If $Hh = Hk$, then $hk^{-1} \in H$, so also $hgg^{-1}k \in H$, which means this action is well-defined. We can compute that the stabiliser of a coset Hx is precisely $x^{-1}Hx$.

The kernel of the action is the intersection of all the stabilisers, which is the largest subgroup of H that is normal in G. This is called the CORE of H. So the action is faithful if and only if the core of H is trivial, so if and only if H contains no nontrivial normal subgroups of G.

Lemma 5.1.22. *Let G act on Ω and let $\omega \in \Omega, x \in G$. Then*

$$G_{\omega^x} = (G_\omega)^x,$$

where the superscript on the left denotes the action on Ω, and the superscript on the right is the conjugation by x.

Proof. Let $g \in G_{\omega^x}$. Then $(\omega^x)^g = \omega^x$, which is equivalent to $\omega^{xg} = \omega^x$, which is again equivalent to $\omega^{xgx^{-1}} = \omega$, so $xgx^{-1} \in G_\omega$ or $g \in (G_\omega)^x$. □

Example. Let $k \in \mathbb{N}$. We label $\Omega^{\{k\}} = \binom{\Omega}{k}$. If G acts on Ω, then we have an induced action of G on $\Omega^{\{k\}}$, by $X^g = \{x^g \mid x \in X\}$. More generally, G acts on the power set $P(\Omega)$ by the same rule. This action preserves cardinality.

5.2 Graphs

It is implied that every graph is finite, simple and undirected.

Definition 5.2.1. The ARC SET of a graph Γ is $A(\Gamma) = \{(u, v) \mid uv \in E(\Gamma)\}$.

Definition 5.2.2. A MORPHISM from a graph Γ to another graph Λ is a mapping $\varphi : V(\Gamma) \to V(\Lambda)$ such that for each edge $uv \in E(\Gamma)$, there is an edge $\varphi(u)\varphi(v) \in E(\Lambda)$.

Remark. A morphism $\varphi : \Gamma \to \Lambda$ induces a map $E(\Gamma) \to E(\Lambda)$ with $uv \in E(\Gamma) \mapsto \varphi(u)\varphi(v) \in E(\Lambda)$.

Definition 5.2.3. If φ is injective, it is a MONOMORPHISM. If it is surjective both as a vertex map and an edge map, then it is an EPIMORPHISM.

Definition 5.2.4. If φ is bijective as a vertex map, and φ^{-1} is also a graph morphism, then φ is an ISOMORPHISM.

Definition 5.2.5. If φ is an isomorphism $\Gamma \to \Gamma$, then φ is an AUTOMORPHISM of Γ. We denote the set of all automorphisms of Γ with $\operatorname{Aut}\Gamma$.

Lemma 5.2.6. *Let Γ and Λ be graphs and let $\varphi : V(\Gamma) \to V(\Lambda)$ be a map. Then φ is an isomorphism if and only if it is bijective and $\phi(E(\Gamma)) = E(\Lambda)$.*

Remark. If g is a permutation of $V(\Gamma)$, it is enough to check that $u \sim v$ implies $u^g \sim v^g$, since the sets $E(\Gamma)$ and $E(\Gamma)^g$ are of the same cardinality.

Remark. A permutation g of $V(\Gamma)$ is an automorphism if and only if g is in the stabilizer of $E(\Gamma)$ in the action of $\operatorname{Sym}(V(\Gamma))$ on the power set of $V(\Gamma)^{\{2\}}$. So $\operatorname{Aut}\Gamma = \operatorname{Sym}(V(\Gamma))_{E(\Gamma)}$ in this particular action. The orbit-stabiliser theorem then gives us $|\operatorname{Sym}(V)| = |\operatorname{Aut}\Gamma| \left|E(\Gamma)^{\operatorname{Sym}(V)}\right|$, or for $n = |V(\Gamma)|$,

$$|\operatorname{Aut}\Gamma| = \frac{n!}{|E(\Gamma)^{\operatorname{Sym} V}|}.$$

The number in the denominator is then precisely the number of graphs on vertex set V that are isomorphic to Γ.

5.2.1 Finding graphs with prescribed automorphisms

Consider the following problem. Given a permutation group $G \leq \operatorname{Sym}\Omega$, determine for which sets $E \subseteq \binom{\Omega}{2}$ does $G \leq \operatorname{Aut}(\Omega, E)$ hold. Or, in particular, how many such sets are there. Label this number with $\gamma(G)$.

Lemma 5.2.7. *Let Γ be a graph, $V = V(\Gamma)$, $E = E(\Gamma)$ and let $G \leq \operatorname{Sym}\Omega$. Then $G \leq \operatorname{Aut}\Gamma$ if and only if E is a union of some orbits of G in its induced action on $\binom{V}{2}$.*

Corollary 5.2.8. *Let $G \leq \operatorname{Sym}\Omega$. Then*

$$\gamma(G) = 2^{\operatorname{orb}_2(G)},$$

where $\operatorname{orb}_2(G)$ is the number of orbits of G on its induced action on $\binom{\Omega}{2}$.

By the Cauchy-Frobenius lemma, we have

$$\operatorname{orb}_2(G) = \frac{1}{|G|} \sum_{g \in G} \left|\operatorname{Fix}_2(g)\right|,$$

where $\operatorname{Fix}_2(g)$ is the number of unordered pairs of elements of Ω fixed by g.

Suppose that $G = \langle g \rangle$ is cyclic. Then the above reduces to

$$\operatorname{orb}_2(G) = \frac{1}{|g|} \sum_{i=0}^{|g|-1} \left|\operatorname{Fix}_2(g^i)\right|,$$

where $|g|$ is the order of g. Now consider several cases. For the first case, suppose $\Omega = \{1, 2, \ldots, m\}$ and $g = (1\,2\,\ldots\,m)$. Then a pair $e = \{x, y\} \subseteq \Omega$ is fixed by g^i if either g^i fixes both x and y, which happens only in the case $i = 0$, or if g^i maps x to y and vice versa, but this happens only if m is even, $i = \frac{m}{2}$ and x and y are opposite. We thus have

$$\left|\operatorname{Fix}_2(g^i)\right| = \begin{cases} \binom{m}{2} & i = 0 \\ \frac{m}{2} & m \text{ even}, i = \frac{m}{2} \\ 0 & \text{otherwise} \end{cases}$$

so

$$\operatorname{orb}_2(\langle g \rangle) = \frac{1}{m}\left(\binom{m}{2} + \operatorname{even}(m)\,\frac{m}{2}\right) = \left\lfloor \frac{m}{2} \right\rfloor,$$

where $\operatorname{even}(m)$ is either 1 or 0 if m is even or odd, respectively.

For the second case, suppose $\Omega = \{1, 2, \ldots, m, 1', 2', \ldots, k'\}$, so $|\Omega| = m + k'$, and $g = (1\,2\,\ldots\,m)(1'\,2'\,\ldots\,k')$. Note that $|g| = \operatorname{lcm}(m, k)$. Now choose some $e = \{i, j'\}$ for $1 \leq i \leq m$ and $1 \leq j \leq k$. We can estimate

$$\left|e^{\langle g \rangle}\right| = \frac{|\langle g \rangle|}{|\langle g \rangle_e|} = \operatorname{lcm}(m, k)$$

since if g^r fixes e, then it also fixes i and j' (since g cannot map a non-primed element to a primed element), so $g^r \in \langle g \rangle_e$ implies $g^r \in \langle g \rangle_i$. Then by the following lemma, since a cyclic group is Abelian, g^r fixes all points $1, \ldots, m$. Similarly, it fixes all points $1', \ldots, k'$. So g^r is the identity (since a permutation group always acts faithfully).

Lemma 5.2.9. *Suppose G is an Abelian group acting on Ω. Then every $g \in G_\omega$ fixes all points in ω^G. In other words, G_ω acts trivially on ω^G.*

Proof. Take $g \in G_\omega$ and $\delta \in \omega^G$. Then $\delta = \omega$ for some $h \in G$, so

$$\delta^g = \omega^{hg} = \omega^{gh} = \omega^h = \delta$$

which we were trying to prove. □

Remark. If G is Abelian and acts transitively and faithfully, then it acts regularly. Since all regular actions are isomorphic, this means Abelian groups are boring.

Continuing the previous discussion, this implies we have precisely

$$\frac{mk}{\operatorname{lcm}(m,k)} = \gcd(m,k)$$

orbits containing a pair of the form $\{i, j'\}$.

For the general case, suppose g consists of k fixed points $K = \operatorname{Fix}(g)$ and r cycles $M_1, \ldots, M_r$ of lengths $m_1, \ldots, m_r \geq 2$. Label $n = k + m_1 + \cdots + m_k$. If $g \neq \mathrm{id}$, then $k < n$, $r \geq 1$ and $r \leq \frac{n}{2}$. Let E be an orbit of $\langle g \rangle$ on $\binom{\Omega}{2}$. Then E is of one of the following types:

1. E contains a pair $e \subseteq \binom{K}{2}$. Then $E = \{e\}$. There are $\binom{k}{2}$ orbits of this type.
2. E contains a pair $e = \{x, y\}$, where $x \in K$ and $y \in M_i$ for some i. Then $E = \{\{x, z\} \mid z \in M_i\}$, and there are rk orbits of this type.
3. E contains a pair $e \subseteq \binom{M_i}{2}$ for some i. Then by the first case, there are

   $$\left\lfloor \frac{m_1}{2} \right\rfloor + \left\lfloor \frac{m_2}{2} \right\rfloor + \cdots + \left\lfloor \frac{m_r}{2} \right\rfloor$$

   such orbits.
4. E contains a pair $e = \{x, y\}$ with $x \in M_i, y \in M_j$ and $i \neq j$. By the second case above, there are $\gcd(m_i, m_j)$ of these orbits for fixed i, j, so in total

   $$\sum_{1 \leq i < j \leq r} \gcd(m_i, m_j)$$

   orbits of this type.

We counted

$$\operatorname{orb}_2(\langle g\rangle) = \binom{k}{2} + kr + \sum_{i=1}^{r} \left\lfloor \frac{m_i}{2} \right\rfloor + \sum_{1\le i<j\le r} \gcd(m_i, m_j).$$

This is an exact formula, but it's horrendous, so we will use it to derive some less exact bounds. We can estimate

$$\sum_{i=1}^{r} \left\lfloor \frac{m_i}{2} \right\rfloor \le \sum_{i=1}^{r} \frac{m_i}{2} = \frac{n-k}{2}$$

and, since the minimum is always less than the average,

$$\begin{aligned}\sum_{i<j} \gcd(m_i, m_J) \le \sum_{i<j} \min(m_i, m_j) &\le \sum_{i<j} \frac{m_i + m_j}{2} = \frac{1}{2}(r-1)\sum_{i=0}^{r} m_i \\ &= \frac{(r-1)(n-k)}{2}.\end{aligned}$$

We thus see that

$$\operatorname{orb}_2(\langle g\rangle) \le \binom{k}{2} + kr + \frac{n-k}{2} + \frac{(r-1)(n-k)}{2} = \binom{k}{2} + \frac{r(n+k)}{2}.$$

This is further equal to

$$\begin{aligned}\operatorname{orb}_2(\langle g\rangle) &= \binom{k}{2} + \frac{r(n+k)}{2} \\ &\le \binom{n}{2} + \frac{k(k-1) - n(n-1)}{2} + \frac{r(n+k)}{2} \\ &= \binom{n}{2} - \frac{(n-r-1)(n-k)}{2}.\end{aligned}$$

Now let us find some useful lower bounds for $(n-r-1)(n-k)$. Since $n = k+m_1+\cdots+m_r$, we have $1 \le r \le \frac{n}{2}$ and $n-k \ge 2r$. Plugging this into the expression above yields us

$$\operatorname{orb}_2(\langle g\rangle) \le \binom{n}{2} - r(n-r-1).$$

The second term on the right is a quadratic function $r \mapsto -r(n-r-1)$, with a minimum at $\frac{n-1}{2}$. In the bounds $1 \le r \le \frac{n}{2}$, the maximum value of this quadratic function is achieved at $r = 1$, and it is equal to $-(n-2)$. So

$$\operatorname{orb}_2(\langle g\rangle) \le \binom{n}{2} - (n-2).$$

We can similarly show an alternate bound

$$\operatorname{orb}_2(\langle g\rangle) \le \binom{n}{2} - \frac{(n-2)(n-k)}{4}$$

by considering $(n-r-1)(n-k)$ and the inequality $n-r \ge \frac{n}{2}$.

5.2.2 Graph counting

Consider the following question: How many non-isomorphic graphs on n vertices are there? First, how many graphs with vertex set $[n]$ can we find? The answer is of course precisely the size of the power set of $\binom{[n]}{2}$, so 2 to that power. But some of those graphs are isomorphic, so we've counted them too many times.

Let $F(n)$ denote the number of non-isomorphic graphs on n vertices. We're interested in

$$\lim_{n\to\infty} \frac{F(n)}{2^{\binom{n}{2}}},$$

if the limit even exists.

Let Γ, Δ be graphs with vertex set $[n]$. A permutation $g \in \mathrm{Sym}(n)$ is an isomorphism between Γ and Δ if and only if $E(\Gamma)$ and $E(\Delta)$ are in the same orbit of $\langle g\rangle$ in its action on $P(\binom{[n]}{2})$. Then the number of all graphs $[n]$, isomorphic to Γ is

$$|[\Gamma]| = \left|E(\Gamma)^{\mathrm{Sym}(n)}\right| = \frac{|\mathrm{Sym}(n)|}{\left|\mathrm{Sym}(n)_{E(\Gamma)}\right|} = \frac{n!}{|\mathrm{Aut}\,\Gamma|}.$$

So, since $\mathrm{Aut}\,\Gamma \geq 1$, we find

$$\frac{2^{\binom{n}{2}}}{n!} \leq F(n) \leq 2^{\binom{n}{2}}.$$

By the Cauchy-Frobenius lemma, we have

$$F(n) = \frac{1}{n!} \sum_{g\in\mathrm{Sym}(n)} \left|\mathrm{Fix}_2(g)\right|,$$

where $|\mathrm{Fix}_2(g)|$ is the number of fixed points of g in the action on $P(\binom{[n]}{2})$. Note that g fixes a point $E \in P(\binom{[n]}{2})$ if and only if the graph $([n], E)$ has g in its automorphism group. We know that

$$|\mathrm{Fix}_2(g)| = 2^{\mathrm{orb}_2(\langle g\rangle)},$$

which gives us a formula for $F(n)$:

$$F(n) = \frac{1}{n!} \sum_{g\in\mathrm{Sym}(n)} 2^{\mathrm{orb}_2(\langle g\rangle)}.$$

Theorem 5.2.10 (Erdös, Renyi)**.** *There exists a function $f : \mathbb{N} \to \mathbb{N}$ with $\lim f(n) = 0$ such that*

$$F(n) = (1 + f(n))\frac{2^{\binom{n}{2}}}{n!},$$

i.e. $F(n) = (1 + o(1))2^{\binom{n}{2}}/n!$

Proof. We estimate $\mathrm{orb}_2(\langle g\rangle)$ for a given $g \in \mathrm{Sym}(n)$. Fix some $m \in \{1, \dots, n\}$. Then split the set $\mathrm{Sym}(n)$ into the following classes:

$$A = \{\mathrm{id}\}$$
$$B = \{g \in \mathrm{Sym}(n) \mid g \text{ moves more than } m \text{ points}\}$$
$$C = \mathrm{Sym}(n) \setminus A \setminus B.$$

We've shown that

$$n!F(n) = \sum_{g \in \mathrm{Sym}(n)} 2^{\mathrm{orb}_2(\langle g\rangle)}.$$

If $g = \mathrm{id}$, then $\mathrm{orb}_2(\{\mathrm{id}\}) = \binom{n}{2}$. For $g \in B$, we use

$$\mathrm{orb}_2(\langle g\rangle) \le \binom{n}{2} - \frac{(n-2)(n-k)}{4} \le \binom{n}{2} - \frac{(n-2)m}{4}.$$

Since $|B| \le n! \le n^n$, the contribution of all elements of B to the sum is at most

$$n^n 2^{\binom{n}{2} - \frac{(n-2)m}{4}}.$$

Similarly, for $g \in C$, we use the estimates $\mathrm{orb}_2(\langle g\rangle) \le \binom{n}{2} - (n-2)$ and $|C| \le \binom{n}{m} m! \le n^m$. We then have

$$n!F(n) \le 2^{\binom{n}{2}} \left(1 + \underbrace{n^n 2^{-(n-2)m/4} + n^m 2^{-(n-2)}}_{f_m(n)}\right).$$

Note that

$$f_m(n) = 2^{n \log_2 n - (n-2)m/4} + 2^{m \log_2 n - (n-2)}$$

We can choose $m = \lfloor 6 \log_2(n) \rfloor$. Then for large enough n, we have $5 \log_2(n) \le m \le 6 \log_2(m)$, so

$$f(n) = f_m(n) \le 2^{n \log_2(n) - 5(n-2)\log_2(n)/4} + 2^{6 \log_2^2(n) - (n-2)}$$

but this converges to 0 for $n \to \infty$. □

Fix a number $n \in \mathbb{N}$. Then let $\mathcal{G}(n)$ be the set of isomorphism classes of graphs on n vertices, so $|\mathcal{G}(n)| = F(n)$. Also introduce $\mathcal{A}(n)$ as the set of all isomorphisms classes for which the isomorphism group is trivial. Denote $A(n) = |\mathcal{A}(n)|$. Finally, introduce $\mathcal{B}(n) = \mathcal{G}(n) \setminus \mathcal{A}(n)$ and $B(n) = |\mathcal{B}(n)|$, so $A(n) + B(n) = F(n)$.

Theorem 5.2.11. *The limit*

$$\lim_{n \to \infty} \frac{A(n)}{F(n)} = 1,$$

so most graphs have trivial isomorphism groups.

Proof. Recall $F(n) = (1 + f(n))2^{\binom{n}{2}}/n!$ where $f(n) \to 0$. Observe

$$2^{\binom{n}{2}} = \sum_{[\Gamma]\in\mathcal{G}(n)} |[\Gamma]| = \sum_{[\Gamma]\in\mathcal{A}(n)} |[\Gamma]| + \sum_{[\Gamma]\in\mathcal{B}(n)} |[\Gamma]|$$

but $|[\Gamma]| = |\mathrm{Sym}(n)| \,/\, |\mathrm{Aut}\,\Gamma| = n!/\,|\mathrm{Aut}\,\Gamma|$. Then

$$2^{\binom{n}{2}} \leq \sum_{[\Gamma]\in\mathcal{A}(n)} n! + \sum_{[\Gamma]\in\mathcal{B}(n)} \frac{n!}{2} = n!\left(A(n) + \frac{1}{2}B(n)\right) = n!\left(F(n) - \frac{1}{2}B(n)\right)$$

so, dividing this by $n!$ and $F(n)$,

$$\frac{1}{1+f(n)} \leq 1 - \frac{1}{2}\frac{B(n)}{F(n)} < 1.$$

By the sandwich theorem, $\frac{B(n)}{F(n)} \to 0$. □

5.3 Vertex transitive graphs

Definition 5.3.1. Let P be a graph and $G \leq \mathrm{Aut}\,\Gamma$. If G acts transitively on $V(\Gamma)$, then we say Γ is G-VERTEX-TRANSITIVE. If it is $(\mathrm{Aut}\,\Gamma)$-vertex transitive, then it is VERTEX-TRANSITIVE.

Example. The complete and empty graphs are vertex-transitive.

Example. Cycles are vertex-transitive. They are even G-vertex-transitive for $G = \langle(1\,2\,\ldots\,n)\rangle$.

5.3.1 Cayley graphs

Let G be a group and $S \subseteq G$ be a set which is closed under inverting and which doesn't include 1. We say that S is a CAYLEY SUBSET of G.

Definition 5.3.2. Let S be a Cayley subset of G, and let Γ be a graph for which the following holds:

- $V(\Gamma) = G$,
- $E(\Gamma) = \{\{g, sg\} \,|\, g \in G, s\ inS\}$.

Then Γ is the CAYLEY GRAPH of G with respect to S, and we label $\mathrm{Cay}(G, S) = \Gamma$.

Example. Let $G = (\mathbb{Z}_6, +)$ and $S = \{1, 3, 5\}$. Then its Cayley graph is shown in figure 5.2, and is isomorphic to $K_{3,3}$.

Note that two elements $x, y \in G$ are adjacent in $\mathrm{Cay}(G, S)$ if and only if $yx^{-1} \in S$. The neighbourhood of a vertex is then

$$\mathrm{Cay}(G, S)(x) = \{sx \,|\, s \in S\},$$

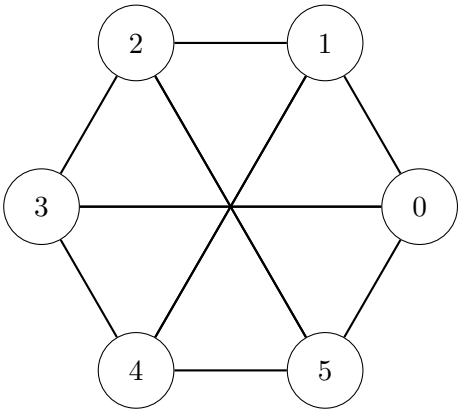


Figure 5.2: Cayley graph of $\mathbb{Z}_6$

where we used the notation $\Gamma(x) = N_\Gamma(x)$. In particular, $\mathrm{Cay}(G, S)$ is an $|S|$-regular graph.

6 Komutativna algebra

6.1 Introduction

6.1.1 Rings

Definition 6.1.1. A RING is a set A, combined with operations $+$ and $\cdot$, such that $(A, +)$ is an Abelian group, $(A, \cdot)$ is a monoid, and distributivity holds: $a(b + c) = ab + ac$, $(b + c)a = ba + ca$ for any $a, b, c \in A$.

Remark. We will assume that all rings are commutative.

Example. We know of a few simple rings:

- $\mathbb{Z}$,
- $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$,
- the ZERO RING $\underline{0} = \{0\}$,
- if A is a ring, then $A[X], A[X_1, \ldots, X_n]$ are also rings,
- if M is an Abelian group, $\mathrm{End}(M)$ is (usually) a non-commutative ring,
- if K is a field and V is a K-vector space of finite dimension $d > 1$, then $\mathrm{End}(V) \cong M_d(K)$ is a non-commutative ring,
- if X is a topological space, $\mathcal{C}(X, \mathbb{R})$ is a ring with pointwise operations,
- if $\{A_i\}_{i \in I}$ is a family of rings, then
$$\prod_{i \in I} A_i$$
is a ring,
- if A is a ring, the set of formal power series' $A[\![X]\!]$ consists of sequences $f : \mathbb{N}_0 \to A$ with operations $(f + g)(n) = f(n) + g(n)$ and
$$(fg)(n) = \sum_{k=0}^{n} f(k)g(n - k).$$
- We can also introduce other products for sequences:
$$(fg)(n) = f(n)g(n) \qquad \text{(pointwise product)}$$
$$(f * g)(n) = \sum_{d|n} f(d)g(n/d) \qquad \text{(Dirichlet's convolution)}$$
- if $\{A_i\}_{i \in I}$ is a family of subrings of A, then their intersection is a subring of A.

Definition 6.1.2. Let A be a ring and $a \in A$. Then

- a is INVERTIBLE (or a UNIT) if there exists $b \in A$ such that $ab = 1$,

- a is a ZERO DIVISOR if there exists an element $b \in A \setminus \{0\}$ such that $ab = 0$,
- a is NILPOTENT if $a^n = 0$ for some n.

Definition 6.1.3. We label the group of units of A by A^* or A^x, and we label the set of non-zero-divisors by $A^\cdot$.

Definition 6.1.4. A ring A is an (INTEGRAL) DOMAIN if $A \neq \underline{0}$ and $ab = 0$ implies $a = 0$ or $b = 0$ for any $a, b \in A$.

Remark. This is equivalent to the statement that 0 is the only zero-divisor of A.

Example. Let $f = \sum f(n)x^n \in A[\![x]\!]$. Then f is invertible if and only if $f(0) \in A^x$.

Definition 6.1.5. If $A \leq B$ are rings and $S \subseteq B$ is a subset, then

$$A[S] = \bigcap_{A \leq A' \leq B, S \subseteq A'} A'$$

is the subring of B obtained by adjoining S to A.

Remark. Note that

$$A[S] = \left\{ \sum_{i=1}^{m} a_i s_i \,|\, a_i \in A, s_i \in S, m \in \mathbb{N} \right\}.$$

Definition 6.1.6. A subset $I \subseteq A$ is an IDEAL if $I \neq \varnothing$ and I is closed under addition and (left) multiplication with elements of A. We denote $I \trianglelefteq A$.

Remark. Given an ideal I of A, the quotient $A/I = \{a + I \,|\, a \in A\}$ is a ring. We sometimes also label $a + I = \bar{a} = [a] = [a]_I$.

Definition 6.1.7. If $X \subseteq A$, then (X) (or $\langle X \rangle$) is the ideal generated by X, so the smallest ideal of A which includes X.

Definition 6.1.8. The PRINCIPAL IDEALS are those generated by a single element, $(a) = Aa$.

Example. The ideals of $\mathbb{Z}$ are $\{n\mathbb{Z} \,|\, n \in \mathbb{N}_0\}$.

Example. The ideal $(x, y) \trianglelefteq K[x, y]$ is not principal.

Remark. If $I \trianglelefteq A$, then $A = I$ if and only if $1 \in I$. This is also equivalent to $I \cap A^x \neq \varnothing$.

Definition 6.1.9. An ideal $I \trianglelefteq A$ is

- a PRIME IDEAL if I is proper and for any $a, b \in A$, if $ab \in I$, then $a \in I$ or $b \in I$,
- a MAXIMAL IDEAL if I is proper and if for all $J \trianglelefteq A$, $I \subseteq J$ implies $J = A$.

The spectrum $\mathrm{Spec}(A)$ is the set of all prime ideals. The set of all maximal ideals is denoted by $\mathrm{Max}(A)$.

Theorem 6.1.10. *An ideal I of A is prime if and only if A/I is a domain. It is maximal if and only if A/I is a field.*

Corollary 6.1.11. *All maximal ideals are prime.*

Remark. Every ideal is contained in a maximal ideal (if Zorn's lemma is true).

Example. The prime ideals of $\mathbb{Z}$ are $\{p\mathbb{Z} \mid p \text{ prime}\} \cup \{\{0\}\}$. Each of $p\mathbb{Z}$ is also maximal, as $\mathbb{Z}/p\mathbb{Z} = \mathbb{Z}_p$ is a field.

Example. In $K[x, y]$, we have $\underline{0} \subsetneq (x) \subsetneq (x, y)$. Each of these are prime ideals, and (x, y) is maximal.

Remark. The trivial ideal $\underline{0}$ is prime if ad only if A is a domain.

Definition 6.1.12. If $I, J \trianglelefteq A$, then the following are also ideals:

- $I + J = \{a + b \mid a \in I, b \in J\} = (I, J)$,
- $I \cap J$,
- $I \cdot J = (\{ab \mid a \in I, b \in J\})$ (ideal generated by the products),
- radial ideal: $\sqrt{I} = \{a \in A \mid \exists n \geq 1. a^n \in I\}$.

Remark. The non-trivial step in proving that the radical ideal really is an ideal is that it is closed under addition. Let $a, b \in \sqrt{I}$. Then let n be large enough that $a^n, b^n \in I$. Compute

$$(a + b)^{2n} = \sum_{i=0}^{2n} a^i b^{2n-i} \binom{2n}{i}.$$

If $i \geq n$, then $a^i \in I$, but if $i \leq n$, then $b^{2n-i} \in I$. So in every case, the terms of the sum are in I, which means the entire sum is in I.

Remark. Note that $\sqrt{I}$ is not necessarily an ideal in a non-commutative ring.

Definition 6.1.13. The NILRADICAL of A is $N(A) = \sqrt{\underline{0}}$ (the set of all nilpotents). The JACOBSON RADICAL $J(A)$ is the intersection of all maximal ideals.

Remark. If $A \neq \underline{0}$, then $N(A), J(A)$ are proper ideals.

Lemma 6.1.14. *For any ring A, $N(A) \subseteq J(A)$. Also, $J(A) = \{a \in A \mid \forall b \in A. 1 - ba \in A^x\}$.*

Proof. For the first point, take $a \in N(A)$, so $a^n = 0$ for some $n \geq 1$. If $M \in \mathrm{Max}(A)$, then $a^n \in M$, as 0 is in every ideal. Since M is prime, this implies $a \in M$.

For the second point, let $a \in J(A)$ and $b \in A$. For any $M \in \mathrm{Max}(A)$, $1 - ab \notin M$, as since $a \in M$, also $ab \in M$, but then if $1 - ab \in M$, this implies $1 \in M$. So $1 - ab$ cannot be in any maximal ideal, which means $(1 - ab) = A$. But then $1 - ab \in A^x$, as also $1 \in (1 - ab)A = A$.

For the other inclusion, let $a \in A$ be such that for every $b \in A$, we have $1 - ab \in A^x$. Let $M \in \mathrm{Max}(A)$ and suppose $a \notin M$. Then $(M, a) = A$, so $1 = m + xa$ for some $m \in M, x \in A$. This means $m = 1 - xa \in M$, but this element is invertible by assumption, so $M = A$. $\rightarrow\!\!\!\!\!\!\times\!\!\!\!\!\!\leftarrow$ □

Lemma 6.1.15 (prime avoidance)**.** *Let I be an ideal of A and let $P_1, \ldots, P_n \in \mathrm{Spec}(A)$. If $I \subseteq P_1 \cup \ldots \cup P_n$, then there exists some k such that $I \subseteq P_k$.*

Proof. Induction on n. If $n = 1$, there is nothing to do. Let $n > 1$. Suppose that for any i, there is an element $a_i \in I \setminus \bigcup_{j \neq i} P_j$, so $a_i \in P_i$. Consider

$$a = \sum_{j=1}^{n} a_1 \ldots \hat{a}_j \ldots a_n = \underbrace{a_1 \ldots \hat{a}_i \ldots a_n}_{\notin P_i} + \underbrace{\sum_{j \neq i} a_1 \ldots \hat{a}_j \ldots a_n}_{\in P_i}$$

for some arbitrary i, because otherwise there would exist some $j \neq i$ such that $a_j \in P_i$ by primality of P_i, contradicting the choice of a_j.

So $a \notin P_i$ for all i. This is a contradiction, as $a \in I \subseteq P_1 \cup \ldots \cup P_n$. □

Lemma 6.1.16. *Let $I_1, \ldots, I_n$ be ideals of A and let $P \in \mathrm{Spec}(A)$. If $I_1 \cap \ldots \cap I_n \subseteq P$, then there exists some k such that $I_k \subseteq P$.*

Proof. Suppose that for every j, $I_j \nsubseteq P$, so there exist $a_j \in I_j \setminus P$. Then $a_1 \ldots a_n \in I_1 \ldots I_n \subseteq I_1 \cap \ldots \cap I_n \subseteq P$. Since P is prime, this implies there exists a j such that $a_j \in P$, which is a contradiction. □

Remark. If $f : A \to B$ is a ring homomorphism, then $\ker f \trianglelefteq A$. More generally, if $I \trianglelefteq B$, then $f^{-1}(I) \trianglelefteq A$. Also, if $P \in \mathrm{Spec}(B)$, then $f^{-1}(P) \in \mathrm{Spec}(A)$.

Proposition 6.1.17 (universal property for quotients)**.** *Let $I \trianglelefteq A$ and $\pi : A \to A/I$ be the canonical epimorphism. For every ring homomorphism $f : A \to B$ with $I \subseteq \ker f$, there exists a unique ring homomorphism $\hat{f} : A/I \to B$ such that $f = \hat{f} \circ \pi$.*

Corollary 6.1.18 (first isomorphism theorem)**.** *If $f : A \to B$ is a ring homomorphism, then $A/\ker f \cong f(A)$.*

Theorem 6.1.19 (isomorphism theorems)**.** *Let $I \trianglelefteq A$. Then*

- *$\{J \trianglelefteq A \,|\, I \subseteq J \subseteq A\}$ is in a bijective correspondence with the ideals of A/I, with the map*
$$J \mapsto J/I = \{a + I \,|\, a \in J\},$$
- *if $J \trianglelefteq A$ is such that $I \subseteq J \subseteq A$,*
$$A/J \cong \frac{A/I}{J/I},$$

- *if $B \subseteq A$ is a subring, then $B + I$ is a subring of A, $B \cap I \trianglelefteq B$, and*

$$\frac{B+I}{I} \cong \frac{B}{B \cap I}.$$

Theorem 6.1.20 (Chinese remainder theorem)**.** *If $I_1, \ldots, I_n \trianglelefteq A$ are pairwise comaximal (so $I_i + I_j = A$ for any i, j), then*

$$A/I_1 \cap \ldots \cap I_n \cong A/I_1 \times \cdots \times A/I_n$$

and $I_1 \cap \ldots \cap I_n = I_1 \cdot \cdots \cdot I_n$.

Example. $\mathbb{Z}/p^n q^m \mathbb{Z} \cong \mathbb{Z}/p^n\mathbb{Z} \times \mathbb{Z}/q^m\mathbb{Z}$ for primes p, q.

6.1.2 Modules

Definition 6.1.21. Let A be a ring. A MODULE M is an additive Abelian group with a scalar multiplication operation $\cdot : A \times M \to M$, such that

- $1m = m$,
- $(ab)m = a(bm)$,
- $(a + b)m = am + bm$,
- $a(m + n) = am + an$.

If M is an Abelian group, an A-module structure can equivalently be described by a ring homomorphism $\varepsilon : A \to \mathrm{End}(M)$, called the STRUCTURE HOMOMORPHISM. Given a module, we can define

$$\varepsilon(a)(m) = am,$$

and in the other direction, the scalar product is given by $am := \varepsilon(a)(m)$.

Example. The $\mathbb{Z}$-modules are precisely the Abelian groups.

Example. If K is a field, then K-modules are precisely the K-vector spaces.

Remark. The submodules of A are precisely the ideals of A.

Example. If V is a K-vector space, and $\varphi \in \mathrm{End}_K(V)$, then we can construct a ring homomorphism $\phi : K[x] \to \mathrm{End}_K(V)$ by

$$\sum_{i=0}^{n} a_i x^i \mapsto \sum_{i=0}^{n} a_i \varphi^i.$$

This induces a $K[x]$-module structure on V, with x acting as φ, $x \cdot v = \varphi(v)$. Since $K[x]$ is a principal ideal domain, $\ker \phi = (m_\varphi)$ for some m_φ, which turns out to be the minimal polynomial of φ.

Let M be an A-module and $E \subseteq M$. Then we define

$$\langle E \rangle_A = \left\{ \sum_{i=1}^{n} a_i m_i \,|\, a_i \in A, m_i \in E \right\}$$

as the A-module generated by E (it is of course the smallest submodule that includes E).

If $\{N_i\}_{i \in I}$ is a family of submodules, then the intersection of all N_i is a submodule, as is

$$\sum_{i \in I} M_i = \left\langle \bigcup_{i \in I} M_i \right\rangle.$$

If $I \trianglelefteq A$, then

$$IM = \left\{ \sum_{i=1}^{n} a_i m_i \,|\, a_i \in I, m_i \in M \right\}$$

is also a submodule of M.

Remark. Let $\varphi : A \to B$ be a ring homomorphism. If N is a B-module, then it is also an A-module via $an := \varphi(a)n$. Similarly, if φ is an epimorphism and M is an A-module such that $IM = 0$ for $I = \ker \varphi$, then M is also a B-module via $bm := am$ for any $a \in \varphi^{-1}(b)$.

Remark. In particular, if M is an A-module and $I \trianglelefteq A$, then M/IM is an A/I-module via $(a + I)(m + IM) = am + IM$.

Remark. This construction gives a category equivalence between the category of A-modules M with $IM = 0$ and the category of A/I-modules.

Example. If M is an Abelian group (so $\mathbb{Z}$-module) and p is prime, then M/pM is a $\mathbb{Z}/p\mathbb{Z}$-vector space.

Theorem 6.1.22 (Isomorphism theorems)**.** *Let M be a module and N, N' submodules.*

- $\operatorname{im} f \cong M/\ker f$ *for any module homomorphism f,*
- *there is a bijection between $\{X \,|\, N \leq X \leq M\}$ and the submodules of M/N,*
- *if $N \leq X \leq M$, then*
$$M/N \cong \frac{M/N}{X/N},$$
- $(N + N')/N \cong N'/(N \cap N')$.

Let $\{M_i\}_{i \in I}$ be a family of A-modules. Then the product is defined as

$$\prod_{i \in I} M_i = \{\{m_i\}_{i \in I} \,|\, m_i \in M_i\}$$

and the direct sum (coproduct) as

$$\bigoplus_{i\in I} M_i = \{\{m_i\}_{i\in I} \mid m_i \in M_i, m_i \neq 0 \text{ for finitely many } i\}.$$

These have the usual (universal) properties.

Remark. If I is finite, then the direct sum and direct product over I are isomorphic.

Remark. If all M_i are submodules of some module M, then

$$M' = \sum_{i\in I} M_i$$

(the smallest submodule of M which includes all M_i) is an internal direct sum if and only if the natural epimorphism $\oplus_i M_i \to \sum_i M_i$ is an isomorphism. This is equivalent to the condition that every $m \in M'$ has a unique representation

$$m = \sum_{i\in I'} m_i$$

for $m_i \in M_i$ and $I' \subset I$ some finite subset. This is again equivalent to the condition that for any M_j,

$$M_j \cap \sum_{i\neq j} M_i = \varnothing.$$

Definition 6.1.23. A module M_A is FREE if $M_A \cong A_A^{(I)} = \oplus_{i\in I} A$ for some I. A BASIS of M_A is a family $\{m_i\}_{i\in I}$ such that

$$\{a_i\}_{i\in I} \mapsto \sum_{i\in I} a_i m_i$$

is an $A^{(I)} \to M$ isomorphism.

Example. If K is a field, then every K-module is free.

Example. As a $\mathbb{Z}$-module, $\mathbb{Z}^n$ is free.

Example. As a $\mathbb{Z}$-module, $\mathbb{Z}/n\mathbb{Z}$ is not free (for $n \neq 0$).

Example. As a $(\mathbb{Z}/n\mathbb{Z})$-module, $\mathbb{Z}/n\mathbb{Z}$ is free.

Example. As a $\mathbb{Z}$-module, $\mathbb{Q}$ is not free.

Remark. Every module M_A is a quotient of a free module.

Definition 6.1.24. A module M_A is FINITELY GENERATED if it is generated by a finite number of elements.

Remark. This is equivalent to the existence of an epimorphism $A_A^k \to M$ for some $k \geq 0$.

Example. As a $\mathbb{Z}$-module, $\mathbb{Q}$ is not finitely generated.

Lemma 6.1.25 (Nakayama)**.** *Let M be a finitely generated A-module.*

- *If $J(A)M = M$, then $M = 0$.*
- *If $N \leq M$ such that $M = N + J(A)M$, then $M = N$.*

Proof. Suppose $M \neq 0$. Since M is finitely generated, there exists a minimal generating set $m_1, \ldots, m_r$. By assumption, since $M = J(A)M$, we have

$$m_r = a_1 m_1 + \ldots + a_r m_r$$

where all $a_i \in J(A)$. This implies $(1 - a_r)m_r = a_1 m_1 + \ldots + a_{r-1} m_{r-1}$, so $(1 - a_r)m_r \in \langle m_1, \ldots, m_{r-1} \rangle$. But by lemma 6.1.14, $1 - a_r$ is invertible, so $m_r \in \langle m_1, \ldots, m_{r-1} \rangle$, which contradicts the minimality of r.

For the second part, apply the first statement to M/N, observing that $J(A)(M/N) = (J(A)M + N)/N = M/N$ by assumption, so $M/N = 0$ and $M = N$. □

Definition 6.1.26. A ring A is LOCAL if $A \neq \underline{0}$ and A has a unique maximal ideal $\mathfrak{m}$.

Example. The ring $\mathbb{Z}_{(2)} = \{\frac{a}{b} \in \mathbb{Q} \mid 2 \nmid b\}$ has $J(\mathbb{Z}_{(2)}) = 2\mathbb{Z}_{(2)}$. This is the unique maximal ideal, as any element not in $2\mathbb{Z}_{(2)}$ is invertible.

Remark. If $(A, \mathfrak{m})$ is local, then $A/\mathfrak{m}$ is a field, and $\mathfrak{m} = J(A)$.

Corollary 6.1.27. *Let $(A, \mathfrak{m})$ be local and M_A a finitely generated A-module. If $x_1, \ldots, x_r \in M$ are such that $x_1 + \mathfrak{m}M, \ldots, x_r + \mathfrak{m}M$ is a basis of the $(A/\mathfrak{m})$-vector space $M/\mathfrak{m}M$, then $x_1, \ldots, x_r$ generate M as an A-module.*

Definition 6.1.28. A COMPLEX $(M_\cdot, f_\cdot)$ is a sequence of modules $(M_n)_n$ together with homomorphisms $f_n : M_n \to M_{n+1}$ such that $f_{n+1} \circ f_n = 0$. We say that a sequence $(M_n)_n$ is EXACT if $\operatorname{im} f_n = \ker f_{n+1}$ for all n.

Definition 6.1.29. A SHORT exact sequence is an exact sequence of the form

$$0 \longrightarrow M \xrightarrow{f} N \xrightarrow{g} P \longrightarrow 0$$

Remark. In the sequence above, f is injective and g is surjective.

Example. Let $N \leq M$. Then

$$0 \longrightarrow N \hookrightarrow M \xrightarrow{\pi} M/N \longrightarrow 0$$

is a short exact sequence.

Example.

$$0 \longrightarrow \mathbb{Z} \xrightarrow{f} \mathbb{Z} \xrightarrow{g} \mathbb{Z}/n\mathbb{Z} \longrightarrow 0$$

is a short exact sequence for $f(x) = nx$ and $g(y) = y + n\mathbb{Z}$.

Remark. In the short exact sequence

$$0 \longrightarrow M \xrightarrow{f} N \xrightarrow{g} P \longrightarrow 0$$

we have $M \cong \ker g$ and $P \cong N/\operatorname{im} f$. Then the following diagram commutes and has exact rows:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M & \xrightarrow{f} & N & \xrightarrow{g} & P & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle \mathrm{id}} & & \downarrow{\scriptstyle (\hat{g})^{-1}} & & \\ 0 & \longrightarrow & K & \hookrightarrow & N & \xrightarrow{\pi} & N/K & \longrightarrow & 0 \end{array}$$

for $K = \ker g = \operatorname{im} f$ and $\hat{g}(n + K) = g(n)$.

Lemma 6.1.30. *For a short exact sequence $0 \to M \xrightarrow{f} N \xrightarrow{g} P \to 0$ the following are equivalent:*

- *There exists a commutative diagram with exact rows*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M & \xrightarrow{f} & N & \xrightarrow{g} & P & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \mathrm{id}} & & \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle \mathrm{id}} & & \\ 0 & \longrightarrow & M & \xrightarrow{\hat{f}} & M \oplus P & \xrightarrow{\hat{g}} & P & \longrightarrow & 0 \end{array}$$

  for $\hat{f}(m) = (m, 0)$ and $\hat{g}(m, p) = p$.

- *there exists a homomorphism $s : P \to N$ such that $g \circ s = \mathrm{id}_P$,*
- *there exists a homomorphism $r : M \to N$ such that $r \circ f = \mathrm{id}_M$.*

In this case $h : N \to M \oplus P$ is an isomorphism. We say that the short exact sequence SPLITS.

For A-modules M, N, the set $\operatorname{Hom}(M, N)$ is itself an A-module for pointwise operations. Then we have two functors:

- the covariant functor $\operatorname{Hom}(M, \cdot) : X \mapsto \operatorname{Hom}(M, X)$ which maps a morphism $X \xrightarrow{f} Y$ to $f_*(g) = f \circ g$,
- the contravariant functor $\operatorname{Hom}(\cdot, N) : X \mapsto \operatorname{Hom}(X, N)$, which maps from the opposite category of the category of A-modules, into the category of A-modules. It maps a homomorphism $X \xrightarrow{f} Y$ to $f^*(g) = g \circ f$.

Lemma 6.1.31. *A sequence $0 \to N \xrightarrow{f} M \xrightarrow{g} P$ is exact if and only if for any A-module X, the sequence $0 \to \operatorname{Hom}(X, N) \xrightarrow{f_*} \operatorname{Hom}(X, M) \xrightarrow{g_*} \operatorname{Hom}(X, P)$ is exact. A sequence $N \xrightarrow{f} M \xrightarrow{g} P \to 0$ is exact if and only if for any A-module X, $0 \to \operatorname{Hom}(P, X) \xrightarrow{g^*} \operatorname{Hom}(M, X) \xrightarrow{f^*} \operatorname{Hom}(N, X)$ is exact.*

Remark. The Hom-functors are left exact.

Definition 6.1.32. Let $M_1, \dots, M_n$ and P be A-modules. A map $f : M_1 \times \cdots \times M_n \to P$ is A-MULTILINEAR if for every index i,

$$f(m_1, \dots, m_{i-1}, am_i + bm_i', m_{i+1}, \dots, m_n) = af(m_1, \dots, m_{i-1}, m_i, m_{i+1}, \dots, m_n) + bf(m_1, \dots, m_{i-1}, m_i', m_{i+1}, \dots, m_n).$$

The TENSOR PRODUCT $M_1 \otimes \cdots \otimes M_n$ is the A-module together with the multilinear map $\otimes : M_1 \times \cdots \times M_n \to M_1 \otimes \cdots \otimes M_n$, defined by the following universal property: For every A-module P and every multilinear map $f : M_1 \times \cdots \times M_n \to P$ there exists a unique A-homomorphism $\hat{f} : M_1 \otimes \cdots \otimes M_n \to P$ such that $\hat{f} \circ \otimes = f$, so that the diagram below commutes.

$$\begin{array}{ccc} M_1 \times \cdots \times M_n & \xrightarrow{\otimes} & M_1 \otimes \cdots \otimes M_n \\ & \searrow^{f} & \downarrow \hat{f} \\ & & P \end{array}$$

Remark. The tensor product can also be written in terms of elementary tensors,

$$M \otimes_A N = \left\{ \sum_{i=1}^{s} m_i \otimes n_i \,|\, m_i \in M, n_i \in N \right\}.$$

Example. In $\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/2\mathbb{Z}$, we have $2 \otimes 1 = 1 \otimes (2 \cdot 1) = 0$. But in $2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/2\mathbb{Z}$, this doesn't hold. The map $(2x, y) \mapsto xy$ is bilinear, and it maps $2 \otimes 1$ to a nonzero element.

Remark. Note that $(M_1 \otimes M_2) \otimes M_3 \cong M_1 \otimes (M_2 \otimes M_3) \cong M_1 \otimes M_2 \otimes M_3$. Also, $M_1 \otimes M_2 \cong M_2 \otimes M_1$.

Let M be an A-module. Then we have a functor $M \otimes \cdot : X \mapsto M \otimes X$, which maps $X \xrightarrow{f} Y$ to $\mathrm{id} \otimes f : m \otimes x \mapsto m \otimes f(x)$.

Theorem 6.1.33 (Hom-$\otimes$ adjunction)**.** *Let M be an A-module. Then $\cdot \otimes M$ is left adjoint to* $\mathrm{Hom}(M, \cdot)$*. Equivalently, for all A-modules M, N, P, there is an A-isomorphism*

$$\begin{aligned} \mathrm{Hom}(N \otimes M, P) &\to \mathrm{Hom}(N, \mathrm{Hom}(M, P)) \\ f &\mapsto (n \mapsto (m \mapsto f(n \otimes m))) \end{aligned}$$

with inverse $\rho \mapsto (n \otimes m \mapsto \rho(n)(m))$.

Corollary 6.1.34. *The functor $M \otimes \cdot$ is right exact, i.e. for every exact sequence $N \to P \to Q \to 0$, also $M \otimes N \to M \otimes P \to M \otimes Q \to 0$ is exact. Also,*

$$M \otimes \left(\bigoplus_{i \in I} N_i \right) \cong \bigoplus_{i \in I} M \otimes N_i$$

for all families $\{N_i\}_{i \in I}$.

Proof. For the first claim, we have $\mathrm{Hom}(M \otimes X, Y) \cong \mathrm{Hom}(M, \mathrm{Hom}(X, Y))$ for all A-modules X, Y. The sequence $0 \to \mathrm{Hom}(Q, X) \to \mathrm{Hom}(P, X) \to \mathrm{Hom}(N, X)$ is also exact by lemma 6.1.31, and using the same lemma again, $0 \to \mathrm{Hom}(M, \mathrm{Hom}(Q, X)) \to \mathrm{Hom}(M, \mathrm{Hom}(P, X)) \to \mathrm{Hom}(M, \mathrm{Hom}(N, X))$ is also exact. These terms are isomorphic to $0 \to \mathrm{Hom}(M \otimes Q, X) \to \mathrm{Hom}(M \otimes P, X) \to \mathrm{Hom}(M \otimes N, X)$. If we use lemma 6.1.31 backwards, we get that $M \otimes N \to M \otimes P \to M \otimes Q \to 0$ is exact.

We skip the proof of the second claim, it's not that hard. We can construct an isomorphism using the universal properties of direct sums and tensor products. □

Example. If $I \trianglelefteq A$ and M is an A-module, then $M/IM \cong M \otimes_A A/I$. We have a short exact sequence $0 \to I \to A \to A/I \to 0$ with the inclusion and projection maps, so we get $I \otimes M \to A \otimes M \to A/I \otimes M \to 0$. Now we take the map $\mu : A \otimes M \to M$ defined by $a \otimes m \mapsto am$, which has inverse $m \mapsto 1 \otimes m$. Since $\mathrm{im}\, \mu|_{I \otimes M} = IM$, this gives us

$$M/IM \cong (A \otimes M)/(I \otimes M) \cong A/(I \otimes M).$$

Proposition 6.1.35. *If $A \neq \underline{0}$ and $A^{(I)} \cong A^{(J)}$ for sets I, J, then $|I| = |J|$.*

Proof. Let $M \in \mathrm{Max}(A)$, then A/M is a field, call it K. Then

$$A^{(I)} \otimes_A A/M \cong (A \otimes A/M)^{(I)} \cong (A/M)^{(I)} = K^{(I)},$$

so these are K-vector spaces. From $A^{(I)} \cong A^{(J)}$ we get $K^{(I)} \cong K^{(J)}$, and since the dimension of a vector space is well-defined, this concludes the proof. □

Definition 6.1.36. Let $A \neq \underline{0}$. If M is a finitely generated free A-module, its RANK is the unique $n \in \mathbb{N}_0$ such that $M \cong A^n$.

6.1.3 Projective, injective and flat modules

Let M be a module over some ring A. Note the following:

- $\cdot \otimes M$ is not exact, as it doesn't preserve the short exact sequence $0 \to 2\mathbb{Z} \to \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \to 0$ for $M = \mathbb{Z}$,
- $\mathrm{Hom}(M, \cdot)$ and $\mathrm{Hom}(\cdot, M)$ are not right exact, as they don't preserve that sequence for $M = \mathbb{Z}/2\mathbb{Z}$.

Definition 6.1.37. A module M is called PROJECTIVE if $\mathrm{Hom}(M, \cdot)$ is exact. It is INJECTIVE if $\mathrm{Hom}(\cdot, M)$ is exact. It is FLAT if $M \otimes \cdot$ is exact.

Theorem 6.1.38. *For an A-module P, the following are equivalent:*

1. *P is projective,*
2. *for every epimorphism $g : M_A \to N_A$, the induced $g_* : \mathrm{Hom}(P, M) \to \mathrm{Hom}(P, N)$ is an epimorphism,*

3. *for every diagram with an exact row of the form below, there exists a ψ such that the diagram commutes:*

$$\begin{array}{ccccc} & & & & P \\ & & \overset{\exists\psi}{\swarrow} & & \downarrow{\scriptstyle\varphi} \\ M & \xrightarrow[f]{} & N & \longrightarrow & 0 \end{array}$$

4. *every epimorphism $g : M \to P$ splits, i.e. there exists $s : P \to M$ such that $g \circ s = \mathrm{id}_P$,*

5. *there exists an A-module M such that $M \oplus P$ is free.*

Proof. 1 to 2 is clear. 2 to 3: By assumption, $f_* : \mathrm{Hom}(P, M) \to \mathrm{Hom}(P, N)$ is an epimorphism. Since f_* is surjective, there exists $\psi \in \mathrm{Hom}(P, M)$ such that $\varphi = f_*(\psi) = f \circ \psi$. This is precisely the commutativity of the diagram.

3 to 4: Consider the following diagram.

$$\begin{array}{ccccc} & & & & P \\ & & \overset{\exists s}{\swarrow} & & \downarrow{\scriptstyle\mathrm{id}} \\ M & \xrightarrow[g]{} & P & \longrightarrow & 0 \end{array}$$

Since g is surjective, by assumption, there exists a map s for which the diagram commutes. This is the splitting map.

4 to 5: Let $g : A^{(I)} \to P$ be an epimorphism for some index set I. This gives us a short exact sequence

$$0 \to \ker g \to A^{(I)} \xrightarrow{g} P \to 0.$$

By assumption, this sequence splits, so $A^{(I)} \cong P \oplus \ker g$ by lemma 6.1.30.

5 to 1: Let

$$0 \to M \xrightarrow{f} N \xrightarrow{g} Q \to 0$$

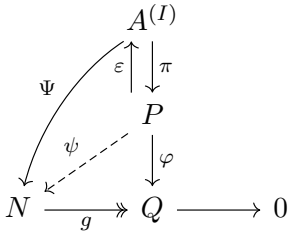
be exact. Consider the sequence

$$0 \to \mathrm{Hom}(P, M) \xrightarrow{f_*} \mathrm{Hom}(P, N) \xrightarrow{g_*} \mathrm{Hom}(Q, N) \to 0.$$

By left exactness of $\mathrm{Hom}(P, \cdot)$, we only have to check surjectivity of g_*. Let $\varphi \in \mathrm{Hom}(P, Q)$. We need to show that there exists $\psi : P \to N$ such that $g_*(\psi) = g \circ \psi = \varphi$.

Fix a module C_A and an index set I such that $P \oplus C \cong A^{(I)}$, let $\pi : A^{(I)} \to P$ be the canonical projection and $\varepsilon : P \to A^{(I)}$ the embedding such that $\pi \circ \varepsilon = \mathrm{id}_P$. For each standard basis vector $e_i \in A^{(I)}$, choose $n_i \in N$ such that $g(n_i) = \varphi \circ \pi(e_i)$ (we can do this since g is surjective). Then there exists a map $\Psi \in \mathrm{Hom}(A^{(I)}, N)$ such that $\Psi(e_i) = n_i$

for each $i \in I$. By construction, $g \circ \Psi = \varphi \circ \pi$. We define $\psi = \Psi \circ \varepsilon$. Then the diagram below commutes.



□

Example. Free modules are projective.

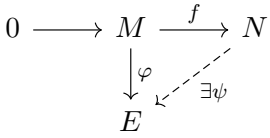
Example. For $A = \mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$. Then $\mathbb{Z}/2\mathbb{Z}$ is a projective A-module, but not free.

Corollary 6.1.39. *Let $\{P_i\}_{i \in I}$ be a family of A-modules. Then all P_i are projective if and only if their direct sum $\bigoplus_i P_i$ is projective.*

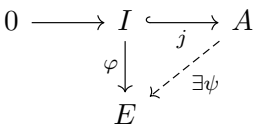
Definition 6.1.40. An A-module E is INJECTIVE if $\mathrm{Hom}(\cdot, E)$ is exact.

Theorem 6.1.41. *For an A-module E, the following are equivalent:*

1. *P is injective,*
2. *for every monomorphism $f : M \to N$, f_* is an epimorphism,*
3. *for every diagram with exact row of the below form, there exists a homomorphism ψ such that $\varphi = \psi \circ f$,*



4. *every monomorphism $f : E \to M$ splits, i.e. there exists $r : M \to N$ such that $r \circ f = \mathrm{id}_E$,*
5. *(Baer's criterion) for every $I \trianglelefteq A$, the inclusion map $j : I \to A$, and every diagram of the below form, there exists a ψ such that $\varphi = \psi \circ j$,*



Proof. The equivalence of 1 and 2 is clear, since $\mathrm{Hom}(\cdot, E)$ is left-exact in any case, so we need only check right exactness. 2 and 3 are equivalent, since f_* is surjective if and only if for every $\varphi \in \mathrm{Hom}(M, E)$ there is a homomorphism ψ such that $\varphi = \psi \circ f$, but this is exactly 3.

3 to 4: Consider

$$\begin{array}{ccccc} 0 & \longrightarrow & E & \xrightarrow{f} & M \\ & & {\scriptstyle \mathrm{id}_E}\downarrow & \swarrow_{\exists r} & \\ & & E & & \end{array}$$

Then by 3, we have the existence of r.

4 to 3: Let $K = \{(\varphi(m), -f(m)) \in E \times N \,|\, m \in M\}$ and $Q = E \times N/K$. There are homomorphisms $j_E : E \to Q$, which maps $e \mapsto (e, 0) + K$ and $j_N : N \to Q$, which maps $n \mapsto (0, n) + K$. We can check that j_E is a monomorphism: if $e \in E$ such that $j_E(e) = 0$, then $(e, 0) \in K$, so there exists $m \in M$ such that $e = \varphi(m)$ and $0 = f(m)$, but since f is a monomorphism, $m = 0$ and $e = 0$.

$$\begin{array}{ccccc} 0 & \longrightarrow & M & \xrightarrow{f} & N \\ & & {\scriptstyle \varphi}\downarrow & & \downarrow{\scriptstyle j_N} \\ & & E & \underset{r}{\overset{j_E}{\rightleftarrows}} & Q \end{array}$$

Then by assumption, there exists $r \in \mathrm{Hom}(Q, E)$ such that $r \circ j_E = \mathrm{id}_E$. Define $\psi = r \circ j_N$. Then for $m \in M$, we have

$$\begin{aligned} \varphi(m) &= r \circ j_E \circ \varphi(m) = r((\varphi(m), 0) + K) = r((\varphi(m), 0) - (\varphi(m), -f(m)) + K) \\ &= r((0, f(m)) + K) = r \circ j_N(f(m)) = \psi \circ f(m). \end{aligned}$$

Which proves the implication.

3 to 5 is easy. 5 to 3: Let $\Omega := \{(N', \psi') \,|\, f(M) \leq N' \leq N, \psi \in \mathrm{Hom}(N', E), \psi \circ f = \varphi\}$. There is a partial order on Ω, given by

$$(N', \psi') \leq (N'', \psi'') \Leftrightarrow N' \subseteq N'' \wedge \psi''\big|_{N'} = \psi'.$$

Clearly $\Omega \neq \varnothing$, as $(f(M), \varphi \circ f^{-1}) \in \Omega$. If $\Omega_0 \subseteq \Omega$ is a chain, then

$$\left(\bigcup_{(N', \psi') \in \Omega_0} N' \,,\, \psi_0 \right)$$

is an upper bound for Ω_0 for ψ_0 defined by $\psi_0|_{N'} = \psi'$ for all $(N', \psi') \in \Omega_0$. Then by Zorn's lemma, there exists a maximal $(N', \psi') \in \Omega$.

We claim that $N' = N$. Suppose that $N' \subsetneq N$ and let $x \in N \backslash N'$. Then $I = \{a \in A \,|\, ax \in N'\}$ is an ideal of A. Define $\mu_I : I \to E$ by $a \mapsto \psi'(ax)$. This is an A-homomorphism, so by assumption, we can extend it to $\mu : A_A \to E_A$. Then $N'' := N' + Ax$ is a module with $N'' \supsetneq N'$, and there is a homomorphism $\psi'' : N'' \to E$, defined with $n + ax \mapsto \psi'(n) + \mu(a)$.

We need to check ψ'' is well defined. Suppose that $n + ax = n' + a'x$ with $n, n' \in N'$ and $a, a' \in A$. Then $n - n' = (a' - a)x$, so $a' - a \in I$ by definition of I, so

$$\mu(a') - \mu(a) = \mu(a' - a) = \mu_I(a' - a) = \psi'((a' - a)x) = \psi'(n - n') = \psi'(n) - \psi'(n')$$

meaning $\psi'(n) + \mu(a) = \psi'(n') + \mu(a')$, so ψ'' is well-defined. Then $(N'', \psi'') \in \Omega$, which contradicts the maximality of (N', ψ'). □

Example. For $A = \mathbb{Z}$, we see that $\mathbb{Z}_{\mathbb{Z}}$ is not injective, as for $n \neq \pm 1$ and $f(x + n\mathbb{Z}) = 1$, the map does not extend to $\mathbb{Z}$.

$$\begin{array}{ccccc} 0 & \longrightarrow & n\mathbb{Z} & \hookrightarrow & \mathbb{Z} \\ & & {\scriptstyle f}\downarrow & \swarrow_{\nexists} & \\ & & \mathbb{Z} & & \end{array}$$

Example. The module $\mathbb{Q}_{\mathbb{Z}}$ is injective, as we can extend $f : n \mapsto q$ with $g : 1 \mapsto f/q$.

$$\begin{array}{ccccc} 0 & \longrightarrow & n\mathbb{Z} & \hookrightarrow & \mathbb{Z} \\ & & {\scriptstyle f}\downarrow & \swarrow_{g} & \\ & & \mathbb{Q} & & \end{array}$$

Definition 6.1.42. Let A be a domain. An A-module M is DIVISIBLE if for all $m \in M$ and for all $a \in A^*$, there exists an element $m_0 \in M$ such that $m = am_0$.

Example. The $\mathbb{Z}$-modules $\mathbb{Q}$ and $\mathbb{Q}/\mathbb{Z}$ are divisible, but $\mathbb{Z}/n\mathbb{Z}$ is not.

Proposition 6.1.43. *Let A be a principal ideal domain. Then M_A is injective if and only if M_A is divisible.*

Corollary 6.1.44. *Let $\{E_i\}_{i \in I}$ be a family of modules. Then all E_i are injective if and only if their product $\prod_i E_i$ is injective.*

Lemma 6.1.45 (snake lemma)**.** *Given a diagram of A-modules*

$$\begin{array}{ccccccccc} & & M & \xrightarrow{i} & N & \longrightarrow & P & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle h} & & \\ 0 & \longrightarrow & M' & \longrightarrow & N' & \xrightarrow{p} & P' & & \end{array}$$

with exact rows, there is a an exact sequence

$$\ker f \to \ker g \to \ker h \to \operatorname{coker} f \to \operatorname{coker} g \to \operatorname{coker} h.$$

If i is mono, so is $i|_{\ker f}$. If p is an epi, then so is $\bar{p} : \operatorname{coker} g \to \operatorname{coker} h$.

Lemma 6.1.46. *Let E be an A-module and $0 \to M \xrightarrow{f} N \xrightarrow{g} P \to 0$ be a split short exact sequence. Then $0 \to E \otimes M \xrightarrow{E\otimes f} E \otimes N \xrightarrow{E\otimes g} E \otimes P \to 0$ is also a short exact sequence.*

Proof. Sketch. Apply $E \otimes \cdot$ to

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M & \xrightarrow{f} & N & \xrightarrow{g} & P & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \mathrm{id}_M} & & \downarrow{\scriptstyle \sim} & & \downarrow{\scriptstyle \mathrm{id}_P} & & \\ 0 & \longrightarrow & M & \longrightarrow & M \oplus P & \longrightarrow & P & \longrightarrow & 0 \end{array}$$

Then note that $E \otimes \cdot$ preserves direct sums. □

Remark. The functors $\mathrm{Hom}(M, \cdot)$ and $\mathrm{Hom}(\cdot, M)$ have the same property.

Theorem 6.1.47. *For an A-module E, the following are equivalent:*

1. *E is flat (i.e. $E \otimes_A \cdot$ is exact),*
2. *for every monomorphism $f : M \to N$, also $E \otimes f : e \otimes m \mapsto e \otimes f(m)$ is a monomorphism,*
3. *for every finitely generated ideal $I \trianglelefteq A$, the homomorphism $\mu_I : I \otimes E \to E$, mapping $x \otimes e \mapsto xe$, is a monomorphism, and hence induces an isomorphism $I \otimes E \to IE$.*

Proof. The equivalence of 1 and 2 is easy, as we know $E \otimes \cdot$ is left-exact. 2 to 3: Apply the assumption to the inclusion $I \to A$ and use $A \otimes_A E \cong E$ by $a \otimes e = ae$.

3 to 2: Let $f : M \to N$ be a monomorphism. Observe that if $\varphi : N \to N'$ is an isomorphism, then it suffices to show the claim for the inclusion map $j : M' := \varphi(f(M)) \to N' := \varphi(N)$. If we apply $E \otimes \cdot$ to

$$\begin{array}{ccccc} 0 & \longrightarrow & M & \xhookrightarrow{f} & N \\ & & {\scriptstyle \varphi\circ f}\downarrow & & \downarrow{\scriptstyle \varphi} \\ & & M' & \xhookrightarrow[j]{} & N' \end{array}$$

then we get

$$\begin{array}{ccc} E \otimes M & \xrightarrow{E\otimes f} & E \otimes N \\ {\scriptstyle E\otimes(\varphi\circ f)}\downarrow & & \downarrow{\scriptstyle E\otimes\varphi} \\ E \otimes M' & \xrightarrow[E\otimes j]{} & E \otimes N' \end{array}$$

so $\ker(E \otimes M) \cong \ker(E \otimes M')$ via $E \otimes (\varphi \circ f)$.

Also observe the following: Let $M \leq N$ and $j : M \to N$ be the inclusion map. If for every finitely generated $M' \leq M$ and the inclusion map $j' : M' \to N$, we have that $E \otimes j'$ is mono, then also $E \otimes j$ is a mono. Let

$$x = \sum_{i=1}^{s} e_i \otimes m_i \in \ker(E \otimes j)$$

with $s \geq 0$, $e_i \in E$ and $m_i \in M$. Let $M' = \langle m_1, \ldots, m_s \rangle \leq M$ be a finitely generated submodule of M, and let $j' : M' \to N, \varepsilon : M' \to N$ be inclusions. Then the following diagram commutes:

$$E \otimes M' \xrightarrow{E\otimes\varepsilon} E \otimes M \xrightarrow{E\otimes j} E \otimes N, \quad E \otimes M' \xrightarrow[E\otimes j']{} E \otimes N$$

Let

$$x' = \sum_{i=1}^{s} e_i \otimes m_i \in E \otimes M'.$$

We can see that $(E \otimes \varepsilon)(x') = x$, and by assumption, $(E \otimes j) \circ (E \otimes \varepsilon)(x') = 0$, but then $(E \otimes j')(x') = 0$, which means $x' = 0$, since $E \otimes j'$ is a monomorphism. This means $x = 0$, so $E \otimes j$ is injective.

Let $f : M \to N$ be a monomorphism. Consider the following special case: $M \leq N$ and N is a finitely generated free module. Let $j : M \to N$ be the inclusion. Use induction on $r = \dim N$. If $r = 0$, there is nothing to check. If $r = 1$, then $N \cong A_A$, so the claim holds by assumption and the above observations. For the induction step, suppose $r \geq 2$. Then write $N = F_1 \oplus F_2$ with $\dim F_1 = r-1$ and $\dim F_2 = 1$. We have canonical epimorphisms $\pi_i : N \to F_i$ and canonical embeddings $\varepsilon_i : F_i \to N$. Let $M_1 = M \cap F_1 = \varepsilon_1^{-1}(M)$ and $M_2 = \pi_2(M) \cong M/M_1$. Then the following diagram has exact rows, with the bottom row being split exact:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & M_1 & \xrightarrow{\varepsilon_1|_{M_1}} & M & \xrightarrow{\pi_2|_M} & M_2 & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle j_1} & & \downarrow{\scriptstyle j} & & \downarrow{\scriptstyle j_2} & & \\
0 & \longrightarrow & F_1 & \xrightarrow{\varepsilon_1} & N & \xrightarrow{\pi_2} & F_2 & \longrightarrow & 0
\end{array}$$

We obtained $j_1 = j|_{M_1}$ and j_2 by factoring $\pi_2 \circ j$ through M_1. We then use $E \otimes \cdot$ on the diagram above, and use the resulting diagram in the snake lemma. We can use the induction hypothesis on F_1 and F_2, so $\ker(E \otimes j_1) = 0$ and $\ker(E \otimes j_2) = 0$. Then $\ker(E \otimes j) = 0$, which is exactly what we wanted to prove.

Now consider the case where N is free, but not necessarily finitely generated. By the second observation, we can assume without loss of generality that M is finitely generated. Write $N = \bigoplus_i Ae_i$, where $\{e_i\}_{i \in I}$ is an A-basis for N. Since M is finitely generated, there exists a finite $I_0 \subseteq I$ such that $M \subseteq \bigoplus_{i \in I_0} Ae_i := N_0$. Let $\varepsilon_M : M \to N_0$ and $\varepsilon_0 : N_0 \to N$ be the inclusions. Now $j = \varepsilon_0 \circ \varepsilon_M$. Then $E \otimes j = (E \otimes \varepsilon_0) \circ (E \otimes \varepsilon_M)$.

The right term is a monomorphism by case 1, and the left is a monomorphism since ε_0 is split. Then $E \otimes j$ is a monomorphism.

In the third case, suppose $j : M \to N$ is an inclusion and N is arbitrary. Let $\pi : F \to N$ be an epimorphism with F free. For $K = \ker \pi$, the following diagram commutes and has exact rows:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & K & \longrightarrow & \pi^{-1}(M) & \xrightarrow{\pi|_{\pi^{-1}(M)}} & M & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \mathrm{id}_K} & & \downarrow & & \downarrow{\scriptstyle j} & & \\ 0 & \longrightarrow & K & \hookrightarrow & F & \longrightarrow & N & \longrightarrow & 0 \end{array}$$

Using $E \otimes \cdot$, we find

$$\begin{array}{ccccccccc} & & & & 0 & & & & \\ & & & & \downarrow & & & & \\ & & E \otimes K & \longrightarrow & E \otimes \pi^{-1}(M) & \longrightarrow & E \otimes M & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle E \otimes \mathrm{id}} & & \downarrow & & \downarrow & & \\ & & E \otimes K & \longrightarrow & E \otimes F & \longrightarrow & E \otimes j & \longrightarrow & 0 \end{array}$$

where the second column is exact by the second case, as is the sequence $0 \to E \otimes K \to E \otimes F$. By the snake lemma, then $\ker E \otimes j = 0$. □

Corollary 6.1.48. *Let $\{E_i\}_{i \in I}$ be a family of modules. Then they are all flat if and only if their direct sum is flat.*

Proof. Easy exercise. Use that the tensor product distributes over direct sums. □

Corollary 6.1.49. *Projective modules are flat.*

Proof. Projective modules are direct summands of free modules. By the previous corollary, flat modules are closed under direct sums and summands, so it suffices to show that A_A is flat. Note that for any A-module M, we have $A \otimes M \cong M$. If $f : M \to N$ is a monomorphism, then

$$\begin{array}{ccc} A \otimes M & \xrightarrow{A \otimes f} & A \otimes N \\ \downarrow{\scriptstyle \sim} & & \downarrow{\scriptstyle \sim} \\ M & \xrightarrow[f]{} & N \end{array}$$

commutes and $A \otimes f$ is a monomorphism. □

Example. The module $\mathbb{Q}_{\mathbb{Z}}$ is flat but not projective. It is not projective since it is injective (divisible, but $\mathbb{Z}$ is not a field). It is flat: Take the map $\mu_n : n\mathbb{Z} \otimes \mathbb{Q} \to \mathbb{Q}$,

$$\mu_n : nx \otimes q \mapsto nxq.$$

It has inverse $\mu_n^{-1}(q) = 1 \otimes q$. ⊡

6.2 Localisations

Given a ring A, we wish to adjoin some inverses (fractions), but not all of them.

Definition 6.2.1. A subset $S \subseteq A$ is a MULTIPLICATIVELY CLOSED SET (or MC SET) if $1 \in S$ and for any $s, s' \in S$, their product $ss' \in S$.

Example. For $a \in A$, $S = \{1, a, a^2, \ldots\}$ is an mc set.

Example. The set $S = A \setminus P$, where $P \in \operatorname{Spec}(A)$ is multiplicatively closed.

Example. The set of all non-zero divisors of A is multiplicatively closed.

Example. If $I \trianglelefteq A$ is an ideal, then $S = 1 + I$ is multiplicatively closed.

Definition 6.2.2. Given an mc set $S \subseteq A$, the LOCALISATION $S^{-1}A$ of S is defined as follows:

- As a set, $S^{-1}A = A \times S/\sim$, where
$$(a, s) \sim (b, t) :\Leftrightarrow \exists u \in S. atu = bsu.$$
- We introduce the notation
$$\frac{a}{s} = [(a, s)]_\sim.$$
- We introduce the operations
$$\frac{a}{s} \cdot \frac{a'}{s'} = \frac{aa'}{ss'}, \qquad \frac{a}{s} + \frac{a'}{s'} = \frac{as' + sa'}{ss'},$$
which makes $S^{-1}A$ a ring with zero $\frac{0}{1}$ and one $\frac{1}{1}$.

We need to check that $\sim$ defined above is an equivalence relation. It is clearly reflexive and symmetric. If $(a, s) \sim (a', s')$ and $(a', s') \sim (a'', s'')$. Then there exist $u, v \in S$ such that $as'u = a'su$ and $a's''v = a''s'v$, but then

$$(as'')s'uv = as'us''v = a'sus''v = a''ss'vu = (a''s)s'uv,$$

so $(a, s) \sim (a'', s'')$ and $\sim$ is an equivalence relation.

We introduce the ring homomorphism $j = j_S : A \to S^{-1}A$, defined with $a \mapsto \frac{a}{1}$. Note that $j(S) \subseteq (S^{-1}A)^x$. Then we have the following universal property: For any $\varphi : A \to B$ with $\varphi(S) \subseteq B^x$, there exists a unique ring homomorphism $\bar{\varphi} : S^{-1}A \to B$ such that $\varphi = \bar{\varphi} \circ j$.

Sketch of the proof. For uniqueness, note that

$$\bar{\varphi}\left(\frac{a}{s}\right) = \bar{\varphi}\left(\frac{a}{1}\right)\varphi\left(\frac{1}{s}\right).$$

We can multiply both sides with $\bar{\varphi}(s)$ to find $\varphi(a) = \bar{\varphi}\left(\frac{a}{s}\right)\varphi(s)$. This also leads us to define $\bar{\varphi}(a/s) = \varphi(a)\varphi(s)^{-1}$. □

Note that $\varphi(a)\varphi(s)^{-1} = 0$ if and only if $\varphi(a) = 0$, so

$$\ker \bar{\varphi} = \{\frac{a}{s} \mid a \in \ker \varphi, s \in S\}.$$

Similarly, $j(a) = 0$ if and only if $\frac{a}{1} = 0$, or if there exists an $s \in S$ for which $as = 0$. Then

$$\ker j = \{a \in A \mid \exists s \in S \,.\, as = 0\}$$

can be nontrivial.

Definition 6.2.3. If S is the set of non-zero-divisors of A, then $S^{-1}A = \mathcal{F}(A)$ is the TOTAL RING OF FRACTIONS of A. If A is a domain, we also call it the FIELD OF FRACTIONS or QUOTIENT FIELD.

If S does not contain zero-divisors, then $S^{-1}A \to \mathcal{F}(A)$ embeds canonically by the universal property.

Definition 6.2.4. If $S = A \setminus P$ for a prime ideal P, write $A_P = S^{-1}A$. If $S = \{1, a, a^2, \ldots\}$ for some $a \in A$, we write $A_a = S^{-1}A$.

Example. If p is a prime number, then

$$\mathbb{Z}_{(p)} = \left\{\frac{a}{b} \mid a \in \mathbb{Z}, p \nmid b\right\}$$

but

$$\mathbb{Z}_p = \left\{\frac{a}{p^k} \in \mathbb{Q} \mid a \in \mathbb{Z}, k \geq 0\right\}$$

Recall: An A-algebra is a ring B that is simultaneously an A-module (with the same addition), for which there is a ring homomorphism $A \to B$.

Example. The polynomial ring $A[X]$ is an A-algebra.

Example. Rings are precisely $\mathbb{Z}$-algebras.

Example. If $I \trianglelefteq A$, then $A \to A/I$ gives an A-algebra.

Example. If $S \subseteq A$ is an mc set, then $j : A \to S^{-1}A$ gives an A-algebra.

Remark. We say that the homomorphism *is* an algebra, even though is only gives us an algebra.

If $f : A \to B$ is an A-algebra, we can

- extend ideals: $I \trianglelefteq A \mapsto f(I)B = \langle f(I) \rangle_B$,
- contract ideals: $J \trianglelefteq B \mapsto f(J)^{-1} \trianglelefteq A$.

For the localisation $j : A \to S^{-1}A$ and an ideal $I \trianglelefteq A$, we write

$$I \cdot S^{-1}A = j(I)S^{-1}A = \left\{\frac{a}{s} \mid s \in S, a \in I\right\}.$$

Definition 6.2.5. Let $I \trianglelefteq A$ and $X \subseteq A$, define the COLON IDEAL

$$(I : X) = \{a \in A \mid aX \subseteq I\}.$$

Remark. It is easy to check that this is an ideal.

Proposition 6.2.6. *Let $S \subseteq A$ be an mc set and $j : A \to S^{-1}A$ the localisation. Then*

- *For all $J \trianglelefteq S^{-1}A$, we have $J = j(j^{-1}(J))S^{-1}A$. In particular, every ideal of $S^{-1}A$ is an extension of an ideal of A.*
- *For any $I \trianglelefteq A$,*

$$j^{-1}(j(I)S^{-1}A) = j^{-1}(IS^{-1}A) = \bigcup_{s \in S}(I : s) \supseteq I.$$

- *$I \cdot S^{-1}A = S^{-1}A$ if and only if $I \cap S^{-1}A \neq \varnothing$.*

Proof. For the first point, the inclusion $\supseteq$ is clear. The second inclusion: If $a/s \in J$, then $j(a) = \frac{a}{1} \in J$, meaning $a \in j^{-1}(J)$, so $\frac{a}{1} \in j(j^{-1}(J))$ and $\frac{a}{s} \in j(j^{-1}(J))S^{-1}A$.

For the second point, consider

$$\begin{aligned} a \in j^{-1}(I \cdot S^{-1}A) &\Leftrightarrow \exists x \in I, s \in S \,.\, \frac{a}{1} = \frac{x}{s} \\ &\Leftrightarrow \exists x \in I, s, t \in S \,.\, ast = xt \\ &\Leftrightarrow \exists s' \in S \,.\, as' \in I \\ &\Leftrightarrow \exists s' \in S \,.\, a \in (I : s') \end{aligned}$$

and the inclusion holds since $1 \in S$ and $(I : 1) = I$.

For the third point, the right-to-left implication holds because $j(S) \subseteq (S^{-1}A)^x$, and for the left-to-right implication, $\frac{1}{1} = \frac{x}{s}$ with $x \in I$, $s \in S$, so there exists $t \in S$ such that $st = xt \in S \cap I$. □

Corollary 6.2.7. *If $P \in \operatorname{Spec}(A)$, then A_P is a local ring with unique maximal ideal PA_P.*

Proof. By the first and third point of the proposition, with $S = A \setminus P$. □

Example. Let $p \in \mathbb{P}$. Then $\mathbb{Z}_{(p)}$ has a maximal ideal

$$\left\{\frac{a}{b} \in \mathbb{Q} \mid p \nmid b, p \mid a\right\}$$

Example. If K is a field and $P = (x, y) \trianglelefteq K[x, y]$, then

$$K[x, y]_{(x,y)} = \left\{ \frac{f}{g} \mid f, g \in K[x, y], g(0, 0) \neq 0 \right\}$$

is local, and has maximal ideal

$$(x, y)K[x, y]_{(x,y)} = \left\{ \frac{f}{g} \mid g(0, 0) \neq 0, f(0, 0) = 0 \right\}.$$

Example. Different ideals can localise to the same ideal. For example, $2\mathbb{Z}_{(2)} = 6\mathbb{Z}_{(2)}$, because $3 \in \mathbb{Z}^x_{(2)}$.

Corollary 6.2.8. *There is a bijection*

$$\{P \in \operatorname{Spec}(A) \mid P \cap S = \varnothing\} \to \operatorname{Spec}(S^{-1}A),$$

given with $P \mapsto P \cdot S^{-1}A$.

Proof. The inverse of this map is $Q \mapsto j^{-1}(Q)$. We need to check that this inverse actually maps into $\operatorname{Spec} A$, which is true since the preimages of prime ideals under ring homomorphisms are prime, and by the previous proposition, $j^{-1}(Q)S^{-1}A = Q$ and $j^{-1}(Q) \cap S = \varnothing$. For the other direction, let $P \in \operatorname{Spec} A$ be such that $P \cap S = \varnothing$. Then $PS^{-1}A$ is prime: It is equal to

$$PS^{-1}A = \left\{ \frac{x}{s} \mid x \in P, s \in S \right\}.$$

Suppose that $\frac{a}{s} \cdot \frac{b}{t} \in PS^{-1}A$, so $\frac{ab}{st} = \frac{x}{s'}$ with $x \in P$, $s' \in S$. Then there exists $u \in S$ such that $abs'u = xstu \in P$, but since $s'u \notin P$, we have $ab \in P$, so $a \in P$ or $b \in P$. Therefore $\frac{a}{s} \in PS^{-1}A$ or $\frac{b}{t} \in PS^{-1}A$.

Now,

$$j^{-1}(PS^{-1}A) = \bigcup_{s \in S} (P : s) = P$$

since $(P : s) = P$ for any s, as $s \notin P$. □

Corollary 6.2.9. *The following statements hold:*

- *the nilradical is equal to*
$$N(A) = \bigcap_{P \in \operatorname{Spec}(A)} P,$$
- *if* $I \trianglelefteq A$, *then*
$$\sqrt{I} = \bigcap_{P \in \operatorname{Spec}(A), I \subseteq P} P.$$

Proof. For the first point, the $\subseteq$ inclusion holds since for $P \in \operatorname{Spec}(A)$ and $a \in N(A)$, so there exists $n \in \mathbb{N}$ such that $a^n = 0 \in P$ and $a \in P$.

The other inclusion: Let a be in the intersection. Then A_a is a ring with $\operatorname{Spec}(A) = \varnothing$, meaning $A_a = \underline{0}$ (since otherwise, we have maximal ideals), so $1 \in \ker(j : A \to A_a)$, meaning there exists $n \in \mathbb{N}_0$ for which $a^n 1 = 0$, so $a \in N(A)$.

For the second point, apply the first to A/I, noting that $N(A/I) = \sqrt{I}/I$. □

6.2.1 Localisation of modules

Let M be an A-module and $S \subseteq A$ an mc set. Define

$$(m, s) \sim (m', s') \Leftrightarrow \exists t \in S \,.\, ms't = m'st.$$

This is an equivalence relation on $M \times S$, which induces the set

$$S^{-1}M = \left\{ \frac{m}{s} := [(m, s)]_\sim \,|\, m \in M, s \in S \right\}.$$

This is an $S^{-1}A$-module via

$$\frac{a}{s}\frac{m}{s'} = \frac{am}{ss'}.$$

As before, we have an A-homomorphism $j : M \to S^{-1}M$, defined with $m \mapsto \frac{m}{1}$. It has kernel $\ker j = \{m \in M \,|\, \exists s \in S \,.\, sm = 0\}$.

If $f : M \to N$ is an A-homomorphism, then $S^{-1}f : S^{-1}M \to S^{-1}N$ is an $S^{-1}A$-homomorphism for $\frac{m}{s} \mapsto \frac{f(m)}{s}$. This means that localisation is a functor from the category of A-modules to the category of $S^{-1}A$-modules.

Proposition 6.2.10. *Localisation is an exact functor. If $M \xrightarrow{f} N \xrightarrow{g} P$ is exact, then so is $S^{-1}M \xrightarrow{S^{-1}f} S^{-1}N \xrightarrow{S^{-1}g} S^{-1}P$.*

Proof. Since $g \circ f = 0$, we have $0 = S^{-1}(g \circ f) = S^{-1}g \circ S^{-1}f$, so $\operatorname{im} S^{-1}f \subseteq \ker S^{-1}g$. Let $\frac{n}{s} \in \ker S^{-1}g$. Then $\frac{g(n)}{s} = 0$ in $S^{-1}P$, so there exists $t \in S$ for which $0 = g(n)t = g(nt)$, meaning $nt \in \ker g = \operatorname{im} f$. So we have $nt = f(m)$ for some $m \in M$. Then

$$\frac{n}{s} = \frac{nt}{st} = \frac{f(m)}{st} = S^{-1}f(\frac{m}{st}).$$

This shows the exactness of the sequence. □

In particular, if $M \leq N$, then without loss of generality, $S^{-1}M \leq S^{-1}N$ and we can think of it as

$$S^{-1}M = \left\{ \frac{m}{s} \in S^{-1}N \,|\, m \in M \right\}.$$

Also, if $I \trianglelefteq A$, then $S^{-1}I = IS^{-1}A \trianglelefteq S^{-1}A$.

Corollary 6.2.11. *If N is an A-module and $M, M' \leq N$, then*

- $S^{-1}(M + M') = S^{-1}M + S^{-1}M'$,
- $S^{-1}(M \cap M') = S^{-1}M \cap S^{-1}M'$,
- $S^{-1}(M/N) \cong S^{-1}N/S^{-1}M$ *as* $S^{-1}A$*-modules.*

Proof. The first point is trivial. For the second point, $\subseteq$ is easy, and for the other inclusion: Let $\frac{m}{s} = \frac{m'}{t}$ with $s, t \in S$, $m \in M$ and $m' \in M'$. There exists $u \in S$ for which $mtu = m'su \in M \cap M'$, so $\frac{m}{s} = \frac{mtu}{stu} qin S^{-1}(M \cap M')$.

For the third point, the exactness of $0 \to M \to N \to N/M \to 0$ induces the short exact sequence $0 \to S^{-1}M \to S^{-1}N \to S^{-1}(N/M) \to 0$, meaning $S^{-1}(N/M) \cong S^{-1}N/S^{-1}M$. □

Remark. If $I \trianglelefteq A$, then $S^{-1}\sqrt{I} = \sqrt{S^{-1}I} \trianglelefteq S^{-1}A$ and in particular $S^{-1}N(A) = N(S^{-1}A)$.

Lemma 6.2.12. *Let $I \trianglelefteq A$ and $\pi : A \to A/I$ the canonical epimorphism. Let $S \subseteq A$ be an mc set. If M is an A-module with $IM = 0$, then M is an A/I-module, $T = \pi(S)$ is an mc set and $S^{-1}M \cong T^{-1}M$ canonically.*

Proof. Define $\varphi : S^{-1}M \to T^{-1}M$ mapping $\frac{m}{s} \mapsto \frac{m}{\pi(s)}$. We can easily show that this is an A-module epimorphism. It is also injective: Suppose $\frac{m}{\pi(s)} = \frac{0}{1} \in T^{-1}M$, so there exists $t \in T$ such that $mt = 0$, and $s' \in S$ for which $t = \pi(s')$. This means $0 = m\pi(s') = m(s' + I) = ms' + MI$, so $ms' \in MI = 0$. So $\frac{m}{s} = 0$. □

Corollary 6.2.13. *Let $I \trianglelefteq A$, $S \subseteq A$ an mc set, $\pi : A \to A/I$ the canonical epimorphism and $j : A \to S^{-1}A$. Then $T = \pi(S)$ is an mc set and $S^{-1}A/S^{-1}I \cong T^{-1}(A/I)$ as an A-module, and also as a ring.*

Proof. Note that $S^{-1}A/S^{-1}I \cong S^{-1}(A/I)$, and by the lemma, $S^{-1}(A/I) \cong T^{-1}(A/I)$. □

Corollary 6.2.14. *Let $P \in \operatorname{Spec}(A)$. Then $\mathcal{F}(A/P) \cong A_P/PA_P$.*

Proof. Apply the previous corollary with $S = A \setminus P$:

$$\frac{A_P}{PA_P} = \frac{S^{-1}A}{S^{-1}P} \cong T^{-1}(A/P) = \mathcal{F}(A/P)$$

for $T = (A/P) \setminus \underline{0}$. □

Remark. This field is called the RESIDUE FIELD of A of P. We denote it with $k(P)$.

Example. Let $p \in \mathbb{P}$. Then $\mathbb{Z}_{(p)}/p\mathbb{Z}_{(p)} \cong \mathbb{Z}/p\mathbb{Z}$.

Example. Let K be a field. Then $K[x,y]_{(x,y)}/(x,y)_{(x,y)} \cong K$. Similarly, $K[x,y]_{(x)}/(x)_{(x)} \cong \mathcal{F}(K[x,y]/(x)) = \mathcal{F}(K[y])$, which is just the set of rational functions $K(y)$ in y.

6.2.2 Local properties

For an A-module M, $P \in \operatorname{Spec}(A)$ and the localisation $j_P : M \to M_P$, we often write $x_P := j_P(x)$. Similarly, if $f \in \operatorname{Hom}(M_A, N_A)$, then $f_P \in \operatorname{Hom}(M_P, N_P)$ is the localised homomorphism.

Proposition 6.2.15. *Let M be an A-module. Then for any $x \in M$, $x = 0$ if and only if for all $P \in \operatorname{Spec}(A)$, $x_P = 0$. This holds if and only if for all $P \in \operatorname{Max}(A)$, $x_P = 0$. Also, $M = 0$ if and only if for all $P \in \operatorname{Spec}(A)$, $M_P = 0$, which holds if and only if for all $P \in \operatorname{Max}(A)$, $M_P = 0$.*

Proof. The implications from 1 to 2 and from 2 to 3 are clear. 3 to 1: Suppose that $x_P = 0$ for all $P \in \operatorname{Max}(A)$. Let $I := \operatorname{ann}(x) := \{a \in A \mid ax = 0\} \trianglelefteq A$. Since $x_P = 0$, it follows that $x \in \ker j_P$, so there exists $s \in A \setminus P$ for which $sx = 0$. Then $I \nsubseteq P$ for all maximal ideals P, so $I = A$ as any proper ideal is contained in a maximal ideal. But then $x = 1 \cdot x = 0$ as $1 \in \operatorname{ann}(x)$.

For the second claim, $M = 0$ if and only if all elements of M are 0. But by the first claim, this is equivalent to the statement that for all $x \in M$ and $P \in \operatorname{Max}(A)$, we have $x_P = 0$. This is again equivalent to the statement that all M_P are zero for $P \in \operatorname{Max}(A)$. □

Proposition 6.2.16. *Let M, N be two A-modules and $f \in \operatorname{Hom}(M, A)$. Then f is a monomorphism if and only if for all $P \in \operatorname{Spec}(A)$, f_P is a monomorphism, which holds if and only if for all $P \in \operatorname{Max}(A)$, f_P is a monomorphism. These equivalences also hold for epimorphisms and isomorphisms.*

Proof. Consider the natural exact sequence

$$0 \longrightarrow \ker f \longrightarrow M \xrightarrow{f} N \longrightarrow \operatorname{coker} f \longrightarrow 0$$

Then also the sequence

$$0 \longrightarrow (\ker f)_P \longrightarrow M_P \xrightarrow{f_P} N_P \longrightarrow (\operatorname{coker} f)_P \longrightarrow 0$$

is exact, so $\ker f_P = (\ker f)_P$ and $\operatorname{coker} f_P = (\operatorname{coker} f)_P$. Since f is mono if and only if $\ker f = 0$ and f is epi if and only if $\operatorname{coker} f = 0$, the claim follows form the previous proposition. □

Properties of a module or homomorphism of such form are called LOCAL PROPERTIES.

6.2.3 Scalar extension and restriction

Suppose $f : A \to B$ is an A-algebra (i.e. f is a ring homomorphism). If M is a B-module, then $a \cdot m := f(a)m$ makes M into an A-module. This is called RESTRICTION OF SCALARS. It gives a functor from the category of B-modules to the categories of A-modules.

As an A-algebra, the ring B is an A-module via $a \cdot b = f(a)b$. If N is an A-module, then $B \otimes_A N$ is a B-module via $b \cdot (x \otimes n) = (bx) \otimes N$. This gives us a functor from the category of A-modules to the category of B-modules, called the EXTENSION OF SCALARS.

Note that while these functors go in the opposite direction, they are not inverse.

Example. Consider $f : \mathbb{Z} \to \mathbb{Q}$. Then $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}^n = \mathbb{Q}^n$ and $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}/n\mathbb{Z} = \underline{0}$.

Example. Let $I \trianglelefteq A$ and $\pi : A \to A/I$ be the canonical epimorphism. Then $A/I \otimes_A M \cong M/IM$.

Proposition 6.2.17. *Let $S \subseteq A$ be an mc set. For an A-module M, $S^{-1}A \otimes_A M \cong S^{-1}M$ as $S^{-1}A$-modules, by $\frac{a}{s} \otimes m \mapsto \frac{am}{s}$.*

Proof. The $S^{-1}A \times M \to S^{-1}M$ map $(\frac{a}{s}, m) \mapsto \frac{am}{s}$ is A-bilinear. Then there exists an A-homomorphism $\varphi : S^{-1}A \otimes_A M \to S^{-1}M$ with $\varphi(\frac{a}{s} \otimes m) = \frac{am}{s}$. It is also a $S^{-1}A$-homomorphism:

$$\varphi(\frac{a'}{s'}(\frac{a}{s} \otimes m)) = \varphi(\frac{aa'}{ss'} \otimes m) = \frac{aa'm}{ss'} = \frac{a'}{s'}\varphi(\frac{a}{s} \otimes m).$$

It is clearly surjective. For injectivity, let

$$x = \sum_{i=1}^{n} \frac{a_i}{s_i} \otimes m_i \in \ker \varphi.$$

Take a common denominator $\frac{a_i}{s_i} = \frac{a_i'}{s}$, so

$$x = \sum_{i=1}^{n} \frac{1}{s} \otimes (a_i' m_i) = \frac{1}{s} \otimes \sum_{i=1}^{n} (a_i' m_i) = \frac{1}{s} \otimes m$$

where m denotes the sum on the right. If we map this with φ, we get $\frac{m}{s} = 0$, so there exists $t \in S$ for which $mt = 0$, so $x = \frac{1}{st} \otimes mt = 0$.

TODO *There is something we still need to check. Look at the lecture notes.* □

7 Diferencialna geometrija

7.1 Osnovni pojmi

Naj bo X topološka mnogoterost (2-števen lokalno evklidski Hausdorffov prostor).

Definicija 7.1.1. GLADEK ATLAS na X je družina $U = \{(U_\alpha, \varphi_\alpha) \,|\, \alpha \in A\}$ kjer je A neka indeksna množica, $\varphi_\alpha : U_\alpha \to V_\alpha \subseteq \mathbb{R}^n$ homomorfizmi, $\bigcup_\alpha U_\alpha = X$ in kjer je za poljubna α, β z $U_\alpha \cap U_\beta \neq \varnothing$ preslikava $\varphi_\beta \circ \varphi_\alpha^{-1}$ difeomorfizem $\varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$.

Definicija 7.1.2. Naj bosta $U = \{(U_\alpha, \varphi_\alpha)\}_{\alpha \in A}$ ter $V = \{(V_\beta, \psi_\beta)\}_{\beta \in B}$ dva atlasa na X. Atlasa sta EKVIVALENTNA, če je $\psi_\beta \circ \varphi_\alpha^{-1}$ difeomorfizem $\varphi_\alpha(U_\alpha \cap V_\beta) \to \psi_\beta(U_\alpha \cap V_\beta)$ za vsak par α, β, za katera je $U_\alpha \cap V_\beta$.

Definicija 7.1.3. Naj bo U atlas na X. Ekvivalenčni razred $[U]$ glede na relacijo ekvivalentnosti atlasov se imenuje GLADKA STRUKTURA na X. Mnogoterost, opremljena z gladko strukturo, je GLADKA MNOGOTEROST.

Opomba. Obstajajo topološke mnogoterosti, na katerih je več neekvivalentnih gladkih struktur.

Naj bo sedaj $X \subseteq \mathbb{R}^N$ (gladka) mnogoterost, vložena v $\mathbb{R}^N$, za katero velja $\dim X = n < N$. Dodatno naj bo $m \in X$ ter $\gamma : (-\varepsilon, \varepsilon) \to X$ gladka krivulja z $\gamma(0) = m$. Potem je

$$V = \left.\frac{d}{dt}\right|_{t=0} \gamma(t)$$

eden od tangentnih vektorjev na X v točki m.

Definicija 7.1.4. TANGENTNI PROSTOR T_mX na X v točki m je množica

$$T_mX = \left\{ \left.\frac{d}{dt}\right|_{t=0} \gamma(t) \,|\, \gamma : (-\varepsilon, \varepsilon) \to X, \gamma(0) = m \right\}.$$

To je vektorski podprostor dimenzije n v $\mathbb{R}^N$.

Naj bo sedaj X ponovno abstraktna gladka mnogoterost.

Definicija 7.1.5. Krivulja $\gamma : (a, b) \to X$ je GLADKA, če je vsak kompozitum $\varphi_\alpha \circ \gamma$ gladka preslikava za vsak α (na primernem definicijskem območju).

Definicija 7.1.6. Krivulji $\gamma_1, \gamma_2 : (-\varepsilon, \varepsilon) \to X$, za kateri velja $\gamma_1(0) = \gamma_2(0) = m$ sta EKVIVALENTNI, če velja

$$\left.\frac{d}{dt}\right|_{t=0} \varphi_\alpha(\gamma_1(t)) = \left.\frac{d}{dt}\right|_{t=0} \varphi_\alpha(\gamma_2(t))$$

za poljuben α, za katerega je $m \in U_\alpha$. Označimo $\gamma_1 \sim \gamma_2$.

Definicija 7.1.7. TANGENTNI VEKTOR na X v točki m je eden od ekvivalenčnih razredov za zgornjo relacijo. TANGENTNI PROSTOR T_mX je prostor vseh tangentnih vektorjev na X v točki m.

Opomba. Izkaže se, da je ekvivalenčna relacija $\sim$ neodvisna od karte.

Predstavnik tangentnega vektorja $[\gamma]$ označimo z

$$\underline{\varphi_\alpha\dot{(}\gamma)} = \left.\frac{d}{dt}\right|_{t=0} \varphi_\alpha(\gamma(t)).$$

To je vektor v $\mathbb{R}^n$. Poglejmo si, kaj se zgodi z $\underline{\varphi_\alpha\dot{(}\gamma)}$, če zamenjamo karto. Naj bo (U_β, φ_β) še ena karta, ki vsebuje m. Potem je

$$\underline{\varphi_\beta\dot{(}\gamma)} = \left.\frac{d}{dt}\right|_{t=0} \varphi_\beta(\gamma(t)) = \left.\frac{d}{dt}\right|_{t=0} (\varphi_\beta\varphi_\alpha^{-1}\varphi_\alpha)(\gamma(t)) = D_{\varphi_\alpha(m)}(\varphi_\beta\varphi_\alpha^{-1}) \left.\frac{d}{dt}\right|_{t=0} \varphi_\alpha(\gamma(t))$$

po verižnem pravilu za odvajanje, kar pa je enako

$$\underline{\varphi_\beta\dot{(}\gamma)} = D_{\varphi_\alpha(m)}(\varphi_\beta\varphi_\alpha^{-1})\underline{\varphi_\alpha\dot{(}\gamma)}$$

Preslikava $D_{\varphi_\alpha(m)}$ je linearni izomorfizem $\mathbb{R}^n \to \mathbb{R}^n$.

V tangentnem prostoru so nam na voljo običajni operaciji vektorskih prostorov:

- Seštevanje: $[\gamma_1] + [\gamma_2] = [t \mapsto \varphi_\alpha^{-1}(\varphi_\alpha(m) + t(\underline{\varphi_\alpha\dot{(}\gamma_1)} + \underline{\varphi_\alpha\dot{(}\gamma_2)}))]$.
- Množenje s skalarjem: $a[\gamma] = [\gamma^a]$ za $\gamma^a(t) = \gamma(at)$.

Vsakemu tangentnemu vektorju lahko priredimo diferencialni operator prvega reda. Naj bo $f : X \to \mathbb{R}$ gladka funkcija. V smeri $[\gamma]$ ga odvajamo s smernim odvodom

$$(Vf)_{(m)}[\gamma] = \left.\frac{d}{dt}\right|_{t=0} (f \circ \gamma)(t).$$

To je neodvisno od izbire predstavnika γ: če sta $\gamma_1, \gamma_2 \in [\gamma]$, velja

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=0} (f \circ \gamma_i)(t) &= \left.\frac{d}{dt}\right|_{t=0} (f\varphi_\alpha^{-1}\varphi_\alpha\gamma_i)(t) \\
&= \vec{\nabla}.(f\varphi_\alpha^{-1}) \left.\frac{d}{dt}\right|_{t=0} (\varphi_\alpha\gamma_i)(t) \\
&= \vec{\nabla}.(f\varphi_\alpha^{-1})\underline{\varphi_\alpha\dot{(}\gamma_i)}.
\end{aligned}$$

Ker sta γ_1 in γ_2 ekvivalentna, je dobljeni izraz neodvisen od izbire i.

Tangentne vektorje lahko obravnavamo kot parcialne diferencialne operatorje prvega reda. Naj bo $[\gamma] = V_m \in T_mX$ tangentni vektor. PDO, ki ga pripišemo temu vektorju, je smerni odvod; če je $f : X \to \mathbb{R}$ gladka funkcija, definiramo

$$V_m(f) = \left.\frac{d}{dt}\right|_{t=0} f(\gamma(t)).$$

V lokalnih koordinatah, danih z atlasom $(U_\alpha, \varphi_\alpha)$, lahko zapišemo

$$V_m(f) = \frac{d}{dt}\Big|_{t=0} f(\gamma(t)) = \frac{d}{dt}\Big|_{t=0} f \circ \varphi_\alpha^{-1} \circ \varphi_\alpha \circ \gamma(t) = \sum_{i=1}^{n} \frac{\partial(f\varphi_\alpha^{-1})}{\partial x_i} v_i$$

kjer je $\frac{d}{dt}(\varphi_\alpha\gamma)(0) = (v_1, \ldots, v_n)$ in $x_1, \ldots, x_n$ izbrane koordinate v $\mathbb{R}^n$. To je očitno linearen operator, velja $V(fg) = V(f)g + fV(g)$.

7.1.1 Vektorski svežnji

Oglejmo si navadno diferencialno enačbo $\dot{\vec{x}} = V(\vec{x})$, kjer je V vektorsko polje na neki podmnožici $\mathbb{R}^n$. Vemo: Za dan Cauchyjev problem lahko najdemo krivuljo γ, za katero velja $\dot{\gamma} = V(\gamma)$. Naj bo sedaj X gladka mnogoterost. Vektorsko polje na X je definirano kot preslikava

$$V : X \to \bigsqcup_{m \in X} T_m X,$$

ki slika v disjunktno unijo tangentnih prostorov na X.

Definicija 7.1.8. Naj bosta X in E taki gladki mnogoterosti, da obstaja projekcija $\pi : E \to X$ z naslednjimi lastnostmi:

- Za vsak $b \in X$ je $\pi^{-1}(b) \subseteq E$ vektorski prostor, izomorfen $\mathbb{F}^k$.
- Za vsak $b \in X$ obstaja odprta okolica $b \in U \subseteq X$ in difeomorfizem $T_U : \pi^{-1}(U) \to U \times \mathbb{F}^k$, za katerega je vsaka skrčitev T_U na $\pi^{-1}(a)$ (lokalen) linearen izomorfizem $\pi^{-1}(a) \to \{a\} \times \mathbb{F}^k$ za poljuben $a \in U$. Potem je T_U LOKALNA TRIVIALIZACIJA.

Potem je (E, X, π) VEKTORSKI SVEŽENJ ranga k.

Definicija 7.1.9. Sveženj $\pi : E \to X$ je TRIVIALEN, če je difeomorfen $E = X \times V$.

Vpeljemo še naslednjo terminologijo:

- π je PROJEKCIJA NA SVEŽENJ,
- $V(\mathbb{F}^k)$ je VLAKNO,
- X je OSNOVNI PROSTOR,
- E je CELOTEN PROSTOR,
- T_U je TRIVIALIZACIJA.

Na E lahko konstruiramo gladek atlas. Naj bo $U = \{(U_\alpha, \varphi_\alpha)\}_{\alpha \in A}$ tak gladek atlas na X, da je restringiran sveženj $E/U_\alpha = \pi^{-1}(U_\alpha)$ trivialen. Za odprte množice v novem atlasu $\tilde{U}$ vzamemo množice $\pi^{-1}(U_\alpha)$ za $\alpha \in A$, karte pa so dane s preslikavami

$$\tilde{\tau}_\alpha : \pi^{-1}(U_\alpha) \to \mathbb{F}^{n+k}$$
$$\tilde{\tau}_\alpha(v) = (\varphi_\alpha(\pi(v)), v_1, \ldots, v_k) = (\varphi_\alpha(\pi(v)), \mathrm{pr}_2(\tau_\alpha(v))),$$

kjer je τ_α difeomorfizem $\pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{F}^k$, ki slika $v \mapsto (\pi(v), (v_1, \ldots, v_k))$. Včasih pišemo tudi

$$\tau_\alpha(m) = (\pi(m), v_1(\pi(m)), \ldots, v_k(\pi(m))) = (b, v_1(b), \ldots, v_k(b)).$$

Oglejmo si prehodno preslikavo

$$\tau_\beta \circ \tau_\alpha^{-1} : (b, v_1(b), \ldots, v_k(b)) \mapsto (b, w_1(b), \ldots, w_k(b)).$$

Zapišemo jo lahko v obliki

$$\tau_\beta \circ \tau_\alpha^{-1} \left(b, \begin{bmatrix} v_1(b) \\ \vdots \\ v_k(b) \end{bmatrix} \right) = \left(b, g_{\beta\alpha}(b) \begin{bmatrix} v_1(b) \\ \vdots \\ v_k(b) \end{bmatrix} \right)$$

oziroma $\tau_\beta \circ \tau_\alpha^{-1} = (\mathrm{id}, g_{\beta\alpha}(b))$. Za vsak par α, β, za katera je $U_\alpha \cap U_\beta \neq \varnothing$, torej dobimo gladko preslikavo $g_{\beta\alpha} \in \mathrm{GL}(k, \mathbb{F})$. Dodano velja $g_{\gamma\beta} g_{\beta\alpha} = g_{\gamma\alpha}$.

Definicija 7.1.10. Kocikel, prirejen atlasu $U = \{(U_\alpha, \varphi_\alpha)\}_\alpha$, z vrednostmi v $\mathrm{GL}(k, \mathbb{F})$, je družina preslikav $g_{\beta\alpha} : U_\alpha \cap U_\beta \to \mathrm{GL}(k, \mathbb{F})$, za katere za vsak primeren b velja

- $g_{\alpha\gamma}(b) g_{\gamma\beta}(b) g_{\beta\alpha}(b) = I$,
- $g_{\alpha\beta}(b) = g_{\beta\alpha}(b)^{-1}$.

Opomba. Če imamo dan X in kocikel, lahko do izomorfizma natančno konstruiramo pripadajoči sveženj.

Definicija 7.1.11. Naj bo $\pi : E \to X$ sveženj. Gladek lokalni prerez svežnja E nad U_α je gladka preslikava $s : U_\alpha \to E/U_\alpha$.

Opomba. Ker je $E/U_\alpha = \pi^{-1}(U_\alpha)$, za primeren b velja $\pi \circ s(b) = b$.

Definicija 7.1.12. Lokalno ogrodje svežnja $\pi : E \to X$ nad U_α je k-terica prerezov $(s_1, \ldots, s_k) : U_\alpha \to E/U_\alpha$, ki so med seboj linearno neodvisni, torej da so za vsak $b \in U_\alpha$ vektorji $s_1(b), \ldots, s_k(b)$ linearno neodvisni v $\pi^{-1}(b) \cong \mathbb{F}^k$.

Definicija 7.1.13. Tangentni sveženj TX gladke mnogoterosti X je disjunktna unija

$$TX = \bigsqcup_{m \in X} T_m X.$$

Opremimo TX z gladko strukturo. Naj bo $U = \{(U_\alpha, \varphi_\alpha)\}_\alpha$ atlas za X. Potem definiramo

$$TU_\alpha = \bigsqcup_{b \in U_\alpha} T_b X := TX/U_\alpha.$$

Naj bo U_α kontraktibilna. Vpeljemo lokalno trivializacijo $\tau_\alpha : TU_\alpha \to V_\alpha \times \mathbb{R}^n$,

$$\tau_\alpha(v(b)) = (\varphi_\alpha(b), D_b\varphi_\alpha(v(b))) = (\varphi_\alpha(b), v_1(b), \ldots, v_n(b)).$$

Torej imamo atlas $TU = \{(TU_\alpha, \tau_\alpha)\}_\alpha$. Za bazo lokalnih prerezov v točki $b_0 = \varphi_\alpha^{-1}(x_0)$ vzamemo vektorje $[\gamma_i]_{b_0}$, pri katerih so reprezentativne krivulje koordinatne premice $t \mapsto \varphi_\alpha^{-1}(x_1^0, \ldots, x_i^0 + t, \ldots, x_n^0)$ za $x_0 = (x_1^0, \ldots, x_n^0)$. Tak tangentni vektor označimo z $\frac{\partial}{\partial x_i^\alpha}(b)$.

7.1.2 Liejev odvod

Spomnimo se: Nad lokalno karto $(U_\alpha, \varphi_\alpha)$, nad katero je $TU_\alpha \subseteq TX$ trivialen, imamo odlikovana bazo prerezov TU_α, ki jo označimo z

$$\frac{\partial}{\partial x_1^\alpha}, \frac{\partial}{\partial x_2^\alpha}, \ldots, \frac{\partial}{\partial x_n^\alpha}.$$

Velja $\frac{\partial}{\partial x_i^\alpha} = [c_i^\alpha]_m$, kjer je c_i^α i-ta koordinatna krivulja na U_α. Vsako vektorsko polje V lahko lokalno zapišemo v obliki

$$V(m) = \sum_{i=1}^{n} a_i^\alpha(m) \frac{\partial}{\partial x_i^\alpha}(m)$$

za neke a_i^α.

Definicija 7.1.14. Smerni odvod funkcije $f : X \to \mathbb{R}$ v smeri vektorskega polja V je funkcija $V(f) : X \to \mathbb{R}$, podana s predpisom

$$V(f)(m) = \left.\frac{d}{dt}\right|_{t=0} f(\gamma(t)),$$

kjer je γ integralska krivulja V skozi točko $\gamma(0) = m$.

Velja potem

$$V(f)(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i^\alpha}(x) a_i^\alpha(x).$$

Naj bosta V in W gladki vektorski polji na $U \subseteq \mathbb{R}^n$. Opazujemo vrednosti W vzdolž integralske krivulje $\gamma^V(t)$ polja V. Naj bo $\phi_t^V : U_\alpha \to \phi_t^V(U_\alpha)$ tok vektorskega polja V. Definiramo

$$\mathcal{L}_V W(m) = \left.\frac{d}{dt}\right|_{t=0} \left(D_{\phi_t^V(m)} \phi_t^V\right)^{-1} W(\gamma^V(t))$$

kjer je $\gamma^V(0) = m$.

Izračunajmo $\mathcal{L}_V W$ v koordinatah. Če označimo $\phi_t^V(x) = ((\phi_t^V)_1, \ldots, (\phi_t^V)_n)(x_1, \ldots, x_n)$, potem je

$$D\phi_t^V = \begin{bmatrix} \frac{\partial (\phi_t^V)_1}{\partial x_1} & \cdots & \frac{\partial (\phi_t^V)_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial (\phi_t^V)_n}{\partial x_1} & \cdots & \frac{\partial (\phi_t^V)_n}{\partial x_n} \end{bmatrix}.$$

Če na definiciji $\mathcal{L}_V W$ uporabimo produktno pravilo za odvajanje, dobimo

$$\mathcal{L}_V W(m) = \left(\frac{d}{dt}\Big|_{t=0} \left(D\phi_t^V\right)^{-1} \right) W(\gamma^V(0)) + \left(D\phi_t^V\right)^{-1}\Big|_{t=0} \frac{d}{dt}\Big|_{t=0} W(\gamma^V(t))$$

V našem primeru je $\phi_t^V = \mathrm{id}$ pri $t = 0$. Potem je

$$\frac{d}{dt}\Big|_{t=0} D\phi_t^V = \begin{bmatrix} \frac{\partial V_1}{\partial x_1} & \cdots & \frac{\partial V_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial V_n}{\partial x_1} & \cdots & \frac{\partial V_n}{\partial x_n} \end{bmatrix} (m),$$

saj odvod po času in po x_i komutirata. Potem imamo, upoštevaje $\partial_t(A^{-1}) = -A^{-1}\dot{A}A^{-1}$,

$$\mathcal{L}_V W = -\begin{bmatrix} \frac{\partial V_1}{\partial x_1} & \cdots & \frac{\partial V_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial V_n}{\partial x_1} & \cdots & \frac{\partial V_n}{\partial x_n} \end{bmatrix} \cdot \begin{bmatrix} W_1 \\ \vdots \\ W_n \end{bmatrix} + \begin{bmatrix} \sum_{j=1}^n \frac{\partial W_1}{\partial x_j} V_j \\ \vdots \\ \sum_{j=1}^n \frac{\partial W_n}{\partial x_j} V_j \end{bmatrix},$$

kjer so vsa polja evaluirana v točki m. Potem je i-ta komponenta tega vektorja enaka

$$-\sum_{j=1}^n \frac{\partial V_i}{\partial x_j} W_j + \sum_{j=1}^n \frac{\partial W_i}{\partial x_j} V_j,$$

torej je zapis $\mathcal{L}_V W$ antisimetričen v V in W.

Rezultat nas navede na drugo definicijo Liejevega odvoda. Naj bo $f : X \to \mathbb{R}$ poljubna funkcija. Potem

$$(\mathcal{L}_V W)(f) = [V, W](f) = V(W(f)) - W(V(f)) = \sum_{i,j} \left(V_j \frac{\partial W_i}{\partial x_j} - W_j \frac{\partial V_i}{x_j} \right) \left(\frac{\partial f}{\partial x_i} \right).$$

Naj bosta ϕ_t^V in ϕ_t^W tokova V in W. Pričakovali bi, da mora veljati $\phi_s^W \circ \phi_t^V = \phi_t^V \circ \phi_s^W$, vendar to ni nujno res. Izračunamo lahko

$$\frac{d^2}{dtds}\Big|_{s=t=0} f(\phi_t^V(\phi_s^W(m))) - f(\phi_s^W(\phi_t^V(m))) = [W, V](f)(m).$$

Če tokova komutirata, potem torej velja $\mathcal{L}_V W = 0$.

7.1.3 Frobeniusov izrek

Naj bo X gladka mnogoterost in TX njen tangentni sveženj. Naj bodo $V_1, \ldots, V_r$ lokalna gladka vektorska polja v TX in $r < n = \dim TX$. Označimo z $D_V X$ lokalni vektorski podsveženj v TX. Polja $V_1, \ldots, V_r$ naj bodo v vsaki točki, kjer so definirana, linearno neodvisna. Zanimalo nas bo, kdaj obstajajo integralske podmnogoterosti v X, ki "pointegrirajo" $D_V X$.

Definicija 7.1.15. Podmnogoterost $N \subseteq X$ dimenzije r je INTEGRALSKA PODMNOGOTEROST za distribucijo $D_V X$, če za vsako točko $m = (b, \ldots) \in D_V(X)$ velja $T_b N = \mathrm{Lin}\{V_1(b), \ldots, V_r(b)\}$.

Izrek 7.1.16 (Frobenius)**.** *Naj bo X mnogoterost dimenzije n in $V_1, \ldots, V_r$ vektorska polja, ki so v vsaki točki linearno neodvisna. Naj za ta polja velja*

$$[V_i, V_j](m) = \sum_{k=1}^{n} c_{i,j}^k(m) V_k,$$

kjer so $C_{i,j}^k : X \to \mathbb{R}$ gladke funkcije. Potem za vsak $p \in X$ obstaja podmnogoterost $N \subseteq X$, da je $p \in N$ in da za vsak $q \in N$ velja $T_q N = \mathrm{Lin}\{V_1(q), \ldots, V_r(q)\}$.

Proof. Ideja. Z Gaussovim postopkom nadomestimo polja $V_1, \ldots, V_r$ s polji $\tilde{V}_1, \ldots, \tilde{V}_r$, za katera velja $\left[\tilde{V}_i, \tilde{V}_j\right] = 0$. Definiramo preslikavo

$$\varphi(t_1, \ldots, t_r) = \phi_{t_r}^{\tilde{V}_r} \circ \phi_{t_{r-1}}^{\tilde{V}_{r-1}} \circ \cdots \circ \phi_{t_2}^{\tilde{V}_2} \circ \phi_{t_1}^{\tilde{V}_1}(p).$$

Potem moramo pokazati, da za vsako točko $q \in N$ velja $T_q N = \mathrm{Lin}\{\tilde{V}_1(q), \ldots, \tilde{V}_r(q)\}$. Izberimo $i = \{1, \ldots, r\}$ in si oglejmo

$$\left.\frac{d}{ds}\right|_{s=0} \varphi(t_1^0, \ldots, t_{i-1}^0, t_i^0 + s, t_{i+1}^0, \ldots, t_r^0) = \left.\frac{d}{ds}\right|_{s=0} \phi_s^{\tilde{V}_i} \circ \phi_{t_r^0}^{\tilde{V}_r} \circ \cdots \circ \phi_{t_1^0}^{\tilde{V}_1}(p) = \tilde{V}_1(q).$$

□

7.1.4 Liejeve grupe

Če je G hkrati mnogoterost in grupa, ter sta množenje in invertiranje gladki preslikavi, je G Liejeva grupa. Grupna struktura nam poda veliko poceni difeomorfizmov, na primer leve in desne translacije ter adjungiranje

$$L_g : h \mapsto gh$$
$$R_g : h \mapsto hg$$
$$\tilde{A}_g : h \mapsto ghg^{-1}$$

Definicija 7.1.17. Vektorsko polje X_ξ je LEVO INVARIANTNO VEKTORSKO POLJE, če je podano s predpisi

$$X_\xi(g) = D_e L_g(\xi),$$

kjer je $\xi \in T_e G$ in e enota grupe.

Naj bo $\{\xi_1, \ldots, \xi_n\}$ neka baza $T_e G$. Ker je za vsak $g \in G$ preslikava $D_e L_g : T_e G \to T_g G$ linearni izomorfizem, je $\{D_e L_g(\xi_1), \ldots, D_e L_g(\xi_n)\} = \{g\xi_1, \ldots, g\xi_n\}$ baza $T_g G$. Iz tega sledi, da je tangentni sveženj TG trivialen.

Za $x \in T_hG$ pišemo $g \cdot x := D_hL_g(x)$ in podobno $x \cdot g$.

Oglejmo si preslikavo $\tilde{A} : (g, h) \mapsto ghg^{-1}$. Če nadomestimo h s krivuljo $h(t)$, za katero je $h(0) = e$ in posledično

$$\left.\frac{d}{dt}\right|_{t=0} h(t) = \xi \in T_eG,$$

lahko definiramo preslikavo $\mathrm{Ad}_g : T_eG \to T_eG$ s predpisom

$$\xi \mapsto \left.\frac{d}{dt}\right|_{t=0} \tilde{A}_g(h(t)).$$

Opazimo

$$\mathrm{Ad}_g(\xi) = (D_eR_g)^{-1}(D_eL_g)(\xi) = g \cdot \xi \cdot g^{-1}.$$

Integralske krivulje levo invariantnih vektorskih polj so rešitve začetnih problemov $\dot{\gamma}(t) = X_\xi(\gamma(t)) = \gamma(t) \cdot \xi$ z $\gamma(0) = e$. Podobno za desno invariantna polja rešujemo $\dot{\gamma}(t) = \xi \cdot \gamma(t)$. Če je G matrična Liejeva grupa (torej ξ matrika), so rešitve teh enačb oblike

$$\gamma(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!}\xi^n = \exp(t\xi).$$

Običajno pišemo $\varphi_\xi(t) := \gamma(t)$. Potem so preslikave $\varphi_\xi : (-\varepsilon, \varepsilon) \mapsto G$ homomorfizmi, kar sledi iz izrekov o tokovih za sisteme NDE. Slike $\operatorname{im}\varphi_\xi$ so torej enodimenzionalne podgrupe v G.

Definicija 7.1.18. Splošna eksponentna preslikava $\exp : T_eG \to G$ je podana s predpisom $\exp(\xi) = \varphi_\xi(1)$.

Opomba. Velja $\exp(t\xi) = \varphi_{t\xi}(1) = \varphi_\xi(t)$.

V prostor $\mathfrak{g} = T_eG$ bomo vpeljali operacijo

$$[\![\xi, \eta]\!] = \mathcal{L}_{X_\xi}(X_\eta)(e).$$

Trditev 7.1.19. *Naj bosta $\xi, \eta \in \mathfrak{g}$ poljubna. Potem velja*

$$[\![\xi, \eta]\!] = \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{\varphi_\eta(t)}(\xi).$$

Proof. Integralska krivulja polja X_ξ skozi g je podana s predpisom

$$\gamma_g^{X_\xi}(t) = g \cdot \varphi_\xi(t), \quad \gamma^{X_\xi}(0) = g.$$

Definirajmo $Y_\eta = X_\eta$. Potem je $Y_\eta(g \cdot \varphi_\xi(t)) = g \cdot \varphi_\xi(t) \cdot \eta$ in

$$\phi_t^{X_\xi}(h) = \gamma_h^{X_\xi}(t) = h \cdot \varphi_\xi(t),$$

torej

$$(\phi_t^{X_\xi})^{-1}(gh) = gh \cdot \varphi_\xi(-t) = gh\varphi_\xi(t)^{-1}.$$

Sedaj lahko izračunamo

$$\begin{aligned}\mathcal{L}_{X_\xi}(Y_\eta)(g) &= \left.\frac{d}{dt}\right|_{t=0} \left(g \cdot \varphi_\xi(t)\eta\varphi_\xi(t)^{-1}\right) \\ &= \left.\frac{d}{dt}\right|_{t=0} g\,\mathrm{Ad}_{\varphi_\xi(t)}(\eta).\end{aligned}$$

Vstavimo $g = e$ in dobimo

$$[\![\xi, \eta]\!] = \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{\varphi_\xi(t)}(\eta). \qquad \square$$

Zgornji dokaz nam tudi pove

$$[X_\xi, Y_\eta](g) = g \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{\varphi_\xi(t)}(\eta) = g[\![\xi, \eta]\!].$$

Spotoma smo torej dokazali, da je Liejev produkt levo invariantnih vektorskih polj levo invariantno vektorsko polje.

Kako pa ta formula izgleda v primeru matričnih Liejevih grup? Za matrične grupe velja

$$\varphi_\xi(t) = \exp(t\xi) = I + t\xi + \cdots + \frac{t^n}{n!}\xi^n + \cdots,$$

iz česar sledi

$$[\![\xi, \eta]\!] = [X_\xi, X_\eta](e) = \left.\frac{d}{dt}\right|_{t=0} \exp(t\xi)\eta\exp(-t\xi) = \left.\frac{d}{dt}\right|_{t=0} (I + t\xi)\eta(I - t\xi),$$

to pa je enako

$$[\![\xi, \eta]\!] = \xi\eta - \eta\xi = [\xi, \eta].$$

Primer. Če je $G = SU(n)$, je tangentni prostor enak

$$\mathfrak{su}(n) = \{\xi \in \mathfrak{gl}(n, \mathbb{C}) \,|\, \xi^* = -\xi, \mathrm{sl}(\xi) = 0\}.$$

Res; naj bo $g(t)$ krivulja z $g(0) = I$. Potem za vsak t velja $g^*(t)g(t) = I$, kar odvajamo v

$$\frac{d}{dt}(g^*(t))g(t) + (g(t))^* \frac{d}{dt}(g(t)) = 0,$$

kar pa še evaluiramo v 0, in dobimo.

7.2 Glavni svežnji

Glavni svežnji so objekti, sorodne vektorskim svežnjem. Vlakna v glavnih svežnjih so Liejeve grupe in ne vektorski prostori.

Definicija 7.2.1. Gladka mnogoterost P, skupaj s preslikavo $\pi : P \to M$, je GLAVNI G-SVEŽENJ, če velja:

- grupa G gladko deluje na P z desne z $\rho : P \times G \to P$, $\rho(p, g) = p \cdot g = \rho_g(p)$,
- prostor orbit $M = P/G$ je gladka mnogoterost in naravna projekcija
$$\pi : P \to P/G = M$$
je gladka,
- sveženj P je lokalno trivialen. To pomeni, za vsako točko $m \in M$ obstaja odprta okolica $m \in U \subseteq M$, da obstaja difeomorfizem $\varphi : \pi^{-1}(U) \subseteq P \to U \times G$ oblike $\varphi(p) = (\pi(p), s(p))$, kjer je s ekvivariantna preslikava; $s(\rho_g(p)) = s(p) \cdot g$.

Primer. Naj bo M mnogoterost in G Liejeva grupa. Potem je $P = M \times G$ glavni sveženj.

Primer. Naj bo $P = S^3 = \{(z_1, z_2) \in \mathbb{C}^2 \mid |z_1|^2 + |z_2|^2 = 1\}$. Liejeva grupa $U(1) = \{e^{i\varphi} \mid \varphi \in [0, 2\pi]\}$ deluje na S^3 z
$$\rho_{e^{i\varphi}}(z_1, z_2) = (z_1 e^{i\varphi}, z_2 e^{i\varphi}).$$
Prostor orbit $S^3/U(1)$ je tedaj
$$M = \{[z_1, z_2] \mid (z_1, z_2) \in S^3\} = \mathbb{CP}^1 \approx S^2.$$

Opomba. Velja $S^3 \approx SU(2)$.

Naj bo $\{U_\alpha\}_\alpha$ atlas na M in naj bo $\pi^{-1}(U_\alpha)$ za vsak α trivialen. Potem je $\varphi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times G$
$$p \mapsto (\pi(p), s_\alpha(p))$$
trivializacija. Recimo, da je $\pi(p) \in U_\alpha \cap U_\beta$. Definiramo lahko preslikavo
$$g_{\beta\alpha} : U_\alpha \cap U_\beta \to G,$$
podano s predpisom $g_{\beta\alpha}(m) = s_\beta(p) s_\alpha^{-1}(p)$ za poljuben $p \in \pi^{-1}(m)$. Ta je res neodvisna od izbire p: Če je tudi $q \in \pi^{-1}(m)$, obstaja $g \in G$, da je $q = p \cdot g$, zaradi ekvivariantnosti s pa je potem $s_\beta(p) s_\alpha^{-1}(p) = s_\beta(q) s_\alpha^{-1}(q)$. Očitno za $g_{\beta\alpha}$ veljata lastnosti kocikla.

Kakor pri vektorskih svežnjih tudi tu velja: Če imamo gladko mnogoterost M z atlasom $\{(U_\alpha, \varphi_\alpha)\}_\alpha$, Liejevo grupo G kocikel $\{g_{\beta\alpha}\}_{\beta,\alpha}$, lahko konstruiramo glavni sveženj. Definiramo lahko namreč
$$\tilde{P} = \prod_{\alpha \in A} (U_\alpha \times G)$$
in $P = \tilde{P}/\sim$ za
$$(\alpha, m_1, h) \sim (\beta, m_2, g) \Leftrightarrow m_1 = m_2 \wedge g_{\beta\alpha}(m_1)h = g.$$

8 Logika

8.1 Introduction

First-order logic is a formal language for expressing properties of mathematical statements. The language is determined by a "signature."

Example. Consider the following arithmetic signature, consisting of:

- two constants 0, 1,
- two binary function symbols $+$, $\cdot$.

Given the signature, we have a derived notions of "terms" and "formulas." In this case, terms represent numbers, for example $0, 1, x, x \cdot (y + z), x \cdot x \cdot x + 1 \cdot 0, \ldots$ are all terms of this language. On the other hand, formulas represent properties of their free variables. As an example, the formula $\exists z \;\; y = x + z$ expresses the property $x \leq y$. Here, z is a bound variable, and x, y are free variables. ⊡

A formula with free variables is called a "sentence". When a sentence is interpreted in a structure, it is either true or false.

Example. Fermat's last theorem can be expressed as

$$\forall n \;\; n > 2 \implies \neg (\exists x, y, z \;\; x > 0 \wedge y > 0 \wedge z > 0 \wedge x^n + y^n = z^n)$$

This uses exponentiation, which we can expresses in the arithmetic language, as was shown by Gödel. The sentence above is true, and the proof is left as an exercise for the reader. ⊡

Example. The sentence

$$x > 1 \wedge \forall y, z \;\; x = y \cdot z \implies x = y \vee x = z$$

encodes "x is prime." We can use this for example to express some open problems, such as the twin prime conjecture,

$$\forall x \exists y \;\; y > x \wedge \text{Prime}(y) \wedge \text{Prime}(y + 2).$$

I need some text to put a boxdot here. ⊡

The richness of the examples suggests that the following is a difficult question:

> Given an input sentence A, decide whether A is true or false in $\mathbb{N}$, i.e. if $\mathbb{N} \models A$.

Theorem 8.1.1 (Church, Turing)**.** *The above question is undecidable.*

We can interpret our arithmetic signature in any ring with a unit, such as in the real numbers. Then, a formula with n free variables describes an n-dimensional geometric shape (actually, a subset of $\mathbb{R}^n$). This signature, interpreted in $\mathbb{R}$, gives a language for Euclidean geometry. But more important is the following.

Theorem 8.1.2 (Tarski)**.** *The problem, "given a sentence A, decide if $\mathbb{R} \models A$," is decidable.*

Definition 8.1.3. A SIGNATURE is a combination of a set of predicate (or relation) symbols $\mathcal{P}$ and function symbols $\mathcal{F}$, together with an arity function $\mathrm{Ar} : \mathcal{P} \amalg \mathcal{F} \to \mathbb{N}$, which returns the number of arguments that a symbol takes.

Example. The signature discussed above has $\mathcal{P} = \varnothing$, $\mathcal{F} = \{0, 1, \cdot, +\}$, and Ar defined with

$$0 \mapsto 0 \qquad 1 \mapsto 0 \qquad + \mapsto 2 \qquad \cdot \mapsto 2$$

Definition 8.1.4. A function symbol of arity 0 is called a CONSTANT.

Example. Another important signature is $(\{E\}, \varnothing)$ with $\mathrm{Ar}(E) = 2$. This is the signature for interpreting in graphs.

Example. Yet another important signature is $(\{\in\}, \varnothing)$ with $\mathrm{Ar}(\in) = 2$. This is the signature of set theory. Note that this is just the same signature as above.

Definition 8.1.5. TERMS are defined by the following two rules:

- a variable is a term,
- if $t_1, \ldots, t_n$ are terms and $f \in \mathcal{F}$ is a function symbol of arity n, then $f(t_1, \ldots, t_n)$ is a term.

Definition 8.1.6. FORMULAS are defined by the following rules:

- if t_1, t_2 are terms, then $t_1 = t_2$ is a formula,
- if $t_1, \ldots, t_n$ are terms and $p \in \mathcal{P}$ is a predicate symbol of arity n, then $p(t_1, \ldots, t_n)$ is a formula,
- if A, B are formulas, then so are $\neg A$, $A \wedge B$, $A \vee B$ and $A \implies B$,
- if A is a formula, then so are $\exists x \;\; A$ and $\forall x \;\; A$.

Formulas of the first two types are called ATOMIC FORMULAS.

Definition 8.1.7. A STRUCTURE for a signature $(\mathcal{P}, \mathcal{F})$ is a triple

$$\mathcal{M} = \left(M, (p^{\mathcal{M}})_{p \in \mathcal{P}}, (f^{\mathcal{M}})_{f \in \mathcal{F}}\right),$$

where M is the underlying set, $p^{\mathcal{M}} : M^{\mathrm{Ar}\, p} \to \{\top, \bot\}$ is the interpretation of predicate p, and $f^{\mathcal{M}} : M^{\mathrm{Ar}\, f} \to M$ is the interpretation of function symbol f.

A term t is interpreted in a structure $\mathcal{M}$ as follows. We assume an ENVIRONMENT $\rho : \mathrm{Vars} \to M$, then define $t^{\rho}_{\mathcal{M}}$ by

- a variable is interpreted as $x^{\rho} = \rho(x)$,
- a function symbol is interpreted as $(f(t_1, \ldots, t_n))^{\rho} = f^{\mathcal{M}}(t_1^{\rho}, \ldots, t_n^{\rho})$.

The SATISFACTION RELATION $\mathcal{M} \models_\rho A$ then states that formula A is true in $\mathcal{M}$ under environment ρ. We define it with

- $\mathcal{M} \models_\rho t_1 = t_2$ if and only if $t_1^\rho = t_2^\rho$,
- for a predicate p, $\mathcal{M} \models_\rho p(t_1, \ldots, t_n)$ if and only if $p^{\mathcal{M}}(t_1^\rho, \ldots, t_n^\rho) = \top$,
- the logical operations are interpreted as usual, so for example

$$(\mathcal{M} \models_\rho A \implies B) :\Leftrightarrow (\mathcal{M} \not\models_\rho A \vee \mathcal{M} \models_\rho B)$$

- The universal quantifier is interpreted as

$$(\mathcal{M} \models_\rho \forall x \;\; A) :\Leftrightarrow \forall d \in M \,.\, \mathcal{M} \models_{\rho[x \mapsto d]} A$$

  where $\rho[x \mapsto d]$ means that we use the environment ρ, except we replace the value of x with d, so

$$\rho[x \mapsto d](y) = \begin{cases} \rho(y) & \text{if } y \text{ is not } x, \\ d & \text{if } y \text{ is } x \end{cases}$$

  We define the interpretation of the existential quantifier similarly.

Example. The structure discussed above is $\mathcal{N} = (\mathbb{N}, 0, +, 1, \cdot)$. We also have a related structure, $\mathcal{R} = (\mathbb{R}, 0, +, 1, \cdot)$.

Lemma 8.1.8. *Let A be formula in structure $\mathcal{M}$ and $\mathrm{FV}(A)$ the set of free variables in A. If $\rho|_{\mathrm{FV}(A)} = \rho'|_{\mathrm{FV}(A)}$, then $\mathcal{M} \models_\rho A$ if and only if $\mathcal{M} \models_{\rho'} A$.*

Corollary 8.1.9. *If A is a sentence, then $\mathcal{M} \models_\rho A$ is independent of ρ. In this case, we write $\mathcal{M} \models A$.*

Definition 8.1.10. The THEORY $\mathrm{Th}(\mathcal{M})$ of a structure $\mathcal{M}$ is the set of all sentences that $\mathcal{M}$ satisfies.

Remark. We can restate the Church-Turing theorem to "$\mathrm{Th}(\mathcal{N})$ is undecidable," and Tarski's theorem to "$\mathrm{Th}(\mathcal{R})$ is decidable."

Proposition 8.1.11 (law of the excluded middle)**.** *For any sentence A, either $M \models A$ or $M \models \neg A$.*

Remark. This says that the theory of an individual structure is COMPLETE.

Example. Consider the signature (e, m), where e is a constant and m is a function symbol of arity 2. We have several interesting classes of structures on this signature: Group (the class of groups), AbGroup (the class of Abelian groups), …

Example. For the signature $(\leq)$, where $\leq$ is a binary relation, we also have several classes of structures: PartOrd (the partial orderings), LinOrd (the linear orderings), …

Definition 8.1.12. Let M be a class of structures. Then $\mathrm{M} \models A$ if and only if for every $\mathcal{M} \in \mathrm{M}$, we have $\mathcal{M} \models A$. We say that A is VALID in M.

Definition 8.1.13. The THEORY $\mathrm{Th}(\underline{\mathrm{M}})$ of a collection of structures $\underline{\mathrm{M}}$ is the set of all valid sentences in $\underline{\mathrm{M}}$.

Example. For $\underline{\mathrm{Group}}$, the sentence $\forall x \exists y \;\; m(x,y) = e \wedge m(y,x) = e$ is valid, but the sentence $\exists x \;\; x \neq e \wedge m(x,x) = e$ is not valid, and neither is its negation.

So for the theory of a collection, it is in general not (necessarily) true that either A is valid or $\neg A$ is valid. The theory of a collection is not complete.

Theorem 8.1.14 (Szmielew)**.** *The theory* $\mathrm{Th}(\underline{AbGroup})$ *is decidable.*

Theorem 8.1.15 (Novikov)**.** *The theory* $\mathrm{Th}(\underline{Group})$ *is undecidable.*

Remark. Novikov proved that even a simple case, the "word problem for groups" is undecidable.

A set of AXIOMS is a set S of sentences which determine a class of structures

$$\underline{\mathrm{Mod}}(S) = \{\mathcal{M} \,|\, \forall A \in S \,.\, \mathcal{M} \models A\}$$

called the COLLECTION OF MODELS of S.

Example. If

$$S_G = \left\{ \begin{array}{c} \forall x \;\; m(x,e) = x \wedge m(e,x) = x, \\ \forall x,y,z \;\; m(x,m(y,z)) = m(m(x,y),z), \\ \forall x \exists y \;\; m(x,y) = e \wedge m(y,x) = e \end{array} \right\},$$

then $\underline{\mathrm{Mod}}(S) = \underline{\mathrm{Group}}$.

Definition 8.1.16. A sentence A is a LOGICAL CONSEQUENCE if

$$S \models A \Leftrightarrow \forall \mathcal{M} \in \underline{\mathrm{Mod}}(S) \,.\, \mathcal{M} \models A$$

holds. We say that S ENTAILS A.

Remark. A sentence A is a logical consequence of S if and only if $A \in \mathrm{Th}(\underline{\mathrm{Mod}}(S))$.

Remark. For a set of axioms S, we define $\mathrm{Th}(S)$ to be $\mathrm{Th}(\underline{\mathrm{Mod}}(S))$.

Definition 8.1.17. A sentence A is LOGICALLY VALID if it satisfies the empty set of axioms, $\varnothing \models A$, so if $\mathcal{M} \models A$ for every $\mathcal{M}$. We label this with $\models A$.

Definition 8.1.18. A set of sentences T is a THEORY if it is closed under logical consequence, i.e. if for all sentences A, if $T \models A$, then $A \in T$.

Definition 8.1.19. A theory T is COMPLETE if for every sentence A either $A \in T$ or $\neg A \in T$.

Definition 8.1.20. A theory T is CONSISTENT if for every sentence A at most one of $A, \neg A$ belong to T.

Remark. A theory T is inconsistent if and only if T is the set of all sentences.

A great contribution of logic is the notion of formal proof. Using this we write precise finite proofs showing that a sentence A follows from an agreed set of axioms. We define the provability relation $S \vdash A$ which states that "A is derivable from the axioms S."

Theorem 8.1.21 (soundness theorem)**.** *For all axioms S and sentences A, if $S \vdash A$, then $S \models A$.*

Theorem 8.1.22 (completeness theorem)**.** *For all axioms S and sentences A, if $S \models A$, then $S \vdash A$.*

Remark. This is the completeness of the proof system; it means that we haven't forgotten any means of proof. It is different than the incompleteness theorem (both are Gödel's).

Remark. As an application, to verify a sentence A for $\mathrm{Th}(\underline{\mathrm{Group}})$, we just need to construct a formal proof showing that $S_G \vdash A$. We know that such a proof exists.

We've seen that the theory of natural numbers is undecidable. Can we at least find a complete set of axioms like we did for groups? There is a natural set of axioms PA (Peano's) for $\mathrm{Th}(\mathcal{N})$, which axiomatises Peano arithmetic. In practice, many known theorems of number theory are provable in PA, for example that there are infinitely many primes, or the prime number theorem.

Theorem 8.1.23 (Gödel's first incompleteness theorem)**.** *One can find a sentence G such that*

- $\mathcal{N} \models G$,
- $PA \nvdash G$,
- $PA \nvdash \neg G$.

Remark. The theory $\mathrm{Th}(PA)$ is incomplete.

One possible interpretation of this is that perhaps we've just missed some necessary axioms. But as it turns out, whatever axioms S we end up with, if S is decidable, then we can repeat the theorem.

8.2 Syntactic Manipulations

Consider the language of natural numbers. We use the following abbreviations:

- $\exists x > t \;\; A$ abbreviates $\exists x \;\; x > t \wedge A$,
- $\forall x > t \;\; A$ abbreviates $\forall x \;\; x > t \implies A$

Here, t is not a logical variable, but rather a meta-variable; that is, we can substitute it for any term. You can of course use meta-variables in different situations, for example, to be substituted by formulas or just regular variables. We will use $s, t, u, \ldots$ as meta-variables for terms, $A, B, C, \ldots$ for formulas, $x, y, z, \ldots$ for variables, $P, Q, R, \ldots$

for relation symbols, $f, g, \ldots$ for function symbols etc. Note that $y > x$ is itself an abbreviation of $\exists z \;\; y = x + z + 1$.

Let A be a formula, t a term and x a free variable. By $A[t/x]$, we denote the formula we get from substituting t for x in A. Similarly, if s is a term, then $s[t/x]$ denotes the term we get by substituting t for the variable x in s.

Example. The abbreviation Prime(y) means $\neg(y = 1) \wedge \forall x \forall z \;\; y = x \cdot z \implies (x = 1 \vee z = 1)$. We can substitute Prime$(y)[y + 2/y]$, but we must be careful to rename the bound variable x when substituting Prime$(y)[x/y]$.

More formally, we define $s[t/x]$ as the term we get when we literally replace all occurrences of the variable x with the term t. For $A[t/x]$, we rename all bound variables in A so that no bound variable occurs in the set $\text{Vars}(t) \cup \{x\}$. Then we replace all remaining occurrences of x by t.

Even more formally, for a variable y, we define

$$y[t/x] := \begin{cases} t & \text{if } y \text{ is syntactically equal to } x \\ y & \text{otherwise} \end{cases}$$

and for a function symbol,

$$(f(s_1, \ldots, s_k))[t/x] := f(s_1[t/x], \ldots, s_k[t/x]),$$

which covers $s[t/x]$ for any term s. For a formula A, we can write a similar recursive definition of $A[t/x]$. The only interesting part of the definition are the quantifiers,

$$(\forall y \;\; A)[t/x] := \forall y \;\; A[t/x]$$

assuming we've already renamed the bound variables. If we haven't, we can use the following definition:

$$(\forall y \;\; A)[t/x] := \begin{cases} \forall y \;\; A & \text{if } y \text{ is literally } x \\ \forall y \;\; A[t/x] & \text{if } y \notin \text{Vars}(t) \text{ and } y \text{ is not } x \\ \forall w \;\; A[w/y][t/x] & \text{otherwise} \end{cases}$$

where w is some chosen fresh w (that isn't x or used in A or t).

Substitution has a semantic interpretation,

$$(S[t/x])^{\rho}_{\mathcal{M}} = S^{\rho[x \mapsto t^{\rho}]}_{\mathcal{M}},$$

and

$$\mathcal{M} \models_{\rho} A[t/x] \Leftrightarrow \mathcal{M} \models_{\rho[x \mapsto t^{\rho}]} A.$$

Definition 8.2.1. Two formulas A, B are LOGICALLY EQUIVALENT ($A \equiv B$) if the formula $A \Leftrightarrow B$ is valid. Equivalently, if for all $\mathcal{M}$ and ρ, $\mathcal{M} \models_{\rho} A$ if and only if $\mathcal{M} \models_{\rho} B$.

Example. For any formula A, $\top \equiv A \vee \neg A$ and $\bot \equiv A \wedge \neg A$.

Example. For any formulas A, B, we have $A \Leftrightarrow B \equiv (A \implies B) \wedge (B \implies A)$.

Example. If we have negation, then we need only one of $\wedge, \vee, \implies$, we can express one with another. We also only need one quantifier.

As discussed above, combining $\neg$ with one of $\wedge, \vee, \implies$ and one of $\forall, \exists$, will give us a SUFFICIENT SET OF CONNECTIVES AND QUANTIFIERS.

8.2.1 Normal forms

A formula is in PRENEX NORMAL FORM (PNF) if it has the form $Q_1 x_1 Q_2 x_2 \dots Q_k x_k B$, where Q_i are quantifiers and B is quantifier free. We include the possibility of $k = 0$.

Proposition 8.2.2. *Every formula A is equivalent to a formula A^{PNF} in prenex normal form.*

Proof. We restrict to formulas over the sufficient set $\neg, \vee, \exists$ (for the input only). We will describe a recursive algorithm that computes an equivalent formula in prenex normal form.

For $(A \vee A')^{\text{PNF}}$, first recursively compute $A^{\text{PNF}} = Q_1 x_1 \dots Q_k x_K B$ and $A'^{\text{PNF}} = Q'_1 x'_1 \dots Q'_{k'} x'_{k'} B'$, then return $Q_1 x_1 \dots Q_k x_k Q'_1 x'_1 \dots Q'_{k'} x'_{k'} (B \vee B')$. We of course assume there is no overlap in bound variables.

The PNF of $\exists x \ \ A$ is $\exists x \ \ A^{\text{PNF}}$. Finally, for negation $\neg A$, recursively compute $A^{\text{PNF}} = Q_1 x_1 \dots Q_k x_k B$, then switch each quantifier to $\bar{Q}_1 x_1 \dots \bar{Q}_k x_k (\neg B)$. □

For quantifier free formulas, we consider normal forms with connectives $\wedge, \top, \vee, \bot, \neg$. Recall that an atomic formula is one of $t_1 = t_2$ or $P(t_1, \dots, t_n)$. We say that a LITERAL is either an atomic formula or its negation.

A formula is in NEGATION NORMAL FORM if it is built from literals using only conjunctions and disjunctions, i.e. if $\neg$ occurs only before an atomic formula. We consider $\top$ as a conjunction and $\bot$ as a disjunction.

Proposition 8.2.3. *Every quantifier free formula A is equivalent to a formula A^{NNF} in negation normal form.*

Proof. Use De Morgan's laws. □

A formula is in CONJUNCTIVE NORMAL FORM if it is a conjunction of disjunctions of literals, i.e. of the form $D_1 \wedge \dots \wedge D_k$, where $k \geq 0$ and each D_i is of the form $L_{i1} \vee \dots \vee L_{il_i}$, with L_{ij} being literals. Similarly, a formula is in DISJUNCTIVE NORMAL FORM if it is a disjunction of conjunctions of literals.

Proposition 8.2.4. *Every quantifier free formula A is equivalent to formulas A^{DNF} and A^{CNF} in disjunctive and conjunctive normal forms, respectively.*

Remark. While the negation normal form can be found efficiently, finding conjunctive and disjunctive normal forms is an NP-complete problem.

8.3 Quantifier elimination

Definition 8.3.1. A theory T is DECIDABLE if there exists an algorithm that takes as input a sequence A and returns true if $T \models A$ (equivalently, if $A \in T$), and it returns false if $T \not\models A$.

Definition 8.3.2. We say that quantifiers are ELIMINABLE from a formula $A(x_1, \ldots, x_k)$ relative to a theory T if there exists a quantifier-free formula $A^{\mathrm{QF}}(x_1, \ldots, x_k)$ such that

$$T \models \forall x_1, \ldots, x_k \;\; A \Leftrightarrow A^{\mathrm{QF}}.$$

We denote this condition with $A \equiv A^{\mathrm{QF}}$.

Remark. It is important that A^{QF} uses the same (or a subset of) free variables as A does.

Definition 8.3.3. A theory T ENJOYS QUANTIFIER ELIMINATION if quantifiers are eliminable relative to T from every formula.

Up to logical equivalence, only $\top$ and $\bot$ are quantifier-free sentences in the theory of valid sentences $V_\varnothing$ over the empty signature. So since we have sentences that are not equivalent to either (such as $\forall x \forall y \;\; x = y$), this theory does not enjoy quantifier elimination.

For every $n \in \mathbb{N}$, consider the sentence

$$\mathrm{Card}_{\geq n} := \exists x_1, \ldots, x_n \;\; \bigwedge_{i=1}^{n} \bigwedge_{j=i+1}^{n} x_i \neq x_j$$

We will consider a new theory, called FINITE CARDINAL BOUNDS, with signature $(C_{\geq n})_{n \in \mathbb{N}}$, where all C_n are propositional constants (0-ary logical predicates), and the theory given by axioms $C_n \Leftrightarrow \mathrm{Card}_{\geq n}$ for all $n \in \mathbb{N}$. We claim the following.

Theorem 8.3.4. *FCB enjoys quantifier elimination.*

Proposition 8.3.5. *A theory T enjoys quantifier elimination if and only if quantifiers are eliminable from formulas of the form $\exists x \;\; (L_1 \wedge \ldots \wedge L_k)$ where $k \geq 0$ and all L_i are literals.*

Proof. The left-to-right implication is trivial. Right-to-left: We will first show that quantifiers are eliminable from formulas of the form $\exists x \;\; A$, where A is quantifier-free.

We can write A in disjunctive normal form to get $\exists x \;\; C_1 \vee \ldots \vee C_n$, where C_i are conjunctions of literals. This is equivalent to

$$(\exists x \;\; C_1) \vee (\exists x \;\; C_2) \vee \ldots \vee (\exists x \;\; C_n),$$

where we can use the assumption. For a general formula, replace the universal quantifiers with existential ones, then eliminate them one by one from the inside out. □

Remark. It is enough to consider formulas of the form in the theorem, but with the additional requirement that each L_i contains x, as we can simply move every other literal before the quantifier.

Proof of Theorem 8.3.4. Consider a formula $\exists x \;\; L_1 \wedge \ldots \wedge L_k$. Each L_i can be of the form $x = y_i$, $x \neq y_i$, $x = x$ or $x \neq x$. There's no need to consider the last two forms, as they are equivalent to $\top$ and $\bot$, respectively. So we can assume y_i are different from x and pairwise distinct.

Suppose that one of L_i is an equality, say L_k. Then we have an equivalent formula $(L_1 \wedge \ldots \wedge L_k)[y_k/x]$, as we require that $x = y_k$. We're left with the case of $\exists x \;\; x \neq y_1 \wedge \ldots \wedge x \neq y_k$, so all literals are negations of equality. The equivalent formula is then

$$\bigvee_{P \in \mathcal{P}_n} \left(\left(\bigwedge_{X \in P} \bigwedge_{i \in X} \bigwedge_{j \in X, j > i} y_i = y_j \right) \wedge C_{|P|+1} \right),$$

where $\mathcal{P}(n)$ is the set of all partitions of the set $\{1, \ldots, n\}$. □

Implicit in the proof above is an algorithm that given $A(x_1, \ldots, x_k)$, computes its quantifier-free equivalent $A^{\mathrm{QF}}(x_1, \ldots, x_k)$. As an additional remark, this algorithm has a horrendously high time complexity.

Theorem 8.3.6. *The theory of valid sentences over the empty signature is decidable.*

Proof. Given an input sentence A, compute A^{QF} in FCB as above. We then convert A^{QF} to conjunctive normal form. We get $D_1 \wedge \ldots \wedge D_n$, where D_i are disjunctions of literals. Note that the formula is valid if and only if each D_i is valid.

A disjunction of literals has the form

$$C_{m_1} \vee \ldots \vee C_{m_l} \vee \neg C_{n_1} \vee \ldots \vee \neg C_{n_k},$$

which is the same as

$$C_{\min(m_1, \ldots, m_l)} \vee \neg C_{\max(n_1, \ldots, n_k)}.$$

This is valid if and only if $\min(m_1, \ldots, m_l) \leq \max(n_1, \ldots, n_k)$. If this test passes for all D_i, we return true; otherwise, we return false. □

TODO *Include section 4.4 from the published lecture notes*

9 Dinamični sistemi

Opomba. Okolica = odprta okolica.

9.1 Uvod

Dinamični sistem je kombinacija množice možnih stanj in evolucijskega pravila. Obravnavamo delec v množici možnih stanj, pravilo pa nam pove, kaj se z njim dogaja "v naslednjem trenutku." Pravilo je DETERMINISTIČNO, kar pomeni, da je naslednji "korak" odvisen le od trenutnega stanja delca.

Dinamične sisteme delimo na DISKRETNE in ZVEZNE. Prvi so rekurzivne zveze $x_{n+1} = F(x_n)$ za neko funkcijo $F : S \to S$ in evklidski prostor S, drugi pa so (nelinearni) sistemi navadnih diferencialnih enačb $\dot{x} = F(x)$. V obeh primerih gre za t.i. AVTONOMNE SISTEME, v katerih pravilo ni odvisno od časa.

Opomba. Vsak sistem lahko prevedemo na avtonomen sistem, če zapakiramo čas kot dodatno spremenljivko.

Primer. Sherlock Holmes najde truplo v jezeru, in izmeri njegovo temperaturo. Ob prihodu izmeri 9 °C, eno uro kasneje 7 °C, temperatura jezera pa je 5 °C. Vprašanje je, kdaj je prišlo do umora.

Fizikalni zakon pravi, da telesna temperatura pada sorazmerno z razliko do temperature jezera. Uporabimo lahko diskretni model,

$$T_{n+1} = T_n - k(T_n - T_J),$$

kjer je T_J temperatura jezera, k konstanta, T_n pa temperatura ob n-tem trenutku. Če vstavimo meritve v to enačbo, dobimo $k = \frac{1}{2}$. Sistem lahko potem rešimo eksplicitno, in sicer dobimo

$$T_n = \frac{C}{2^n} + 5\,^\circ\mathrm{C}$$

za neko konstanto C. Če je $T_0 = 37\,^\circ$C, izračunamo $C = 32\,^\circ$C, iz česar dobimo, da je $T_3 = 9\,^\circ$C, torej se je umor zgodil pred tremi urami. ⊡

Primer. Zgornji problem lahko rešujemo tudi z zveznim modelom,

$$\dot{T} = -k(T - T_J)$$

za nek drug k kot prej. Rešitev te diferencialne enačbe je $T = T_J + (T_0 - T_J)\exp(-kt)$, kjer smo že upoštevali začetni pogoj. Uporabiti moramo še meritve, za katere dobimo

$$T(n) = (37 - 5)e^{-kn} + 5 = 9,$$
$$T(n+1) = (37 - 5)e^{-k(n+1)} + 5 = 7,$$

oziroma $k = \log 2$ in $n = 3$. Enako kot prej torej sklepamo, da se je umor zgodil 3 ure pred začetkom opazovanja. ⊡

9.2 Dinamika realnih funkcij

9.2.1 Osnovni pojmi

Za začetek se bomo omejili na funkcije $f : I \to I$, kjer je I interval v $\mathbb{R}$. Obravnavali bomo zaporedja $x_{n+1} = f(x_n)$ pri različnih začetnih pogojih $x_0 \in I$. Uvedimo oznako

$$f^n = \underbrace{f \circ f \circ \cdots \circ f}_{n \text{ pojavitev } f}.$$

Definicija 9.2.1. ORBITA točke x_0 pri funkciji f je množica $O_f(x_0) = \{f^n(x_0) \mid n \in \mathbb{N}_0\}$. Množici vseh orbit za $x_0 \in I$ pravimo DINAMIKA FUNKCIJE f.

Primer. Za $x_{n+1} = x_n^2 + 1$ je orbita točke 1 enaka $\{1, 2, 5, 26, \ldots\}$.

Orbite lahko (za funkcije ene spremenljivke) vizualiziramo s t.i. PAJČEVINASTIM DIAGRAMOM (angl. *cobweb diagram*), prikazanim na sliki 9.1.

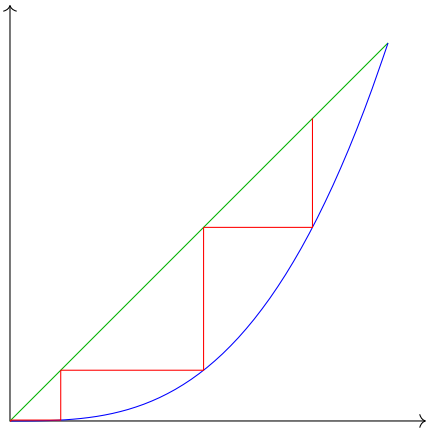


Figure 9.1: Pajčevinasti diagram za $f(x) = x^3$ z začetno točko $x_0 = 0.8$.[1]

Definicija 9.2.2. Točka x_0 je PERIODIČNA s periodo $n \in \mathbb{N}$ za f, če zanjo velja $f^n(x_0) = x_0$ in je n najmanjše število s to lastnostjo. Če je $n = 1$, taki točki pravimo FIKSNA TOČKA.

Primer. Funkcija $f(x) = -x^3$ ima le eno fiksno točko, 0. Izračunamo lahko, da imamo tudi dve točki periode 2, to sta 1 in -1.

Definicija 9.2.3. Orbiti n-periodične točke pravimo n-CIKEL.

Definicija 9.2.4. Naj bo $x_0 \in I$ fiksna točka za $f : I \to I$.

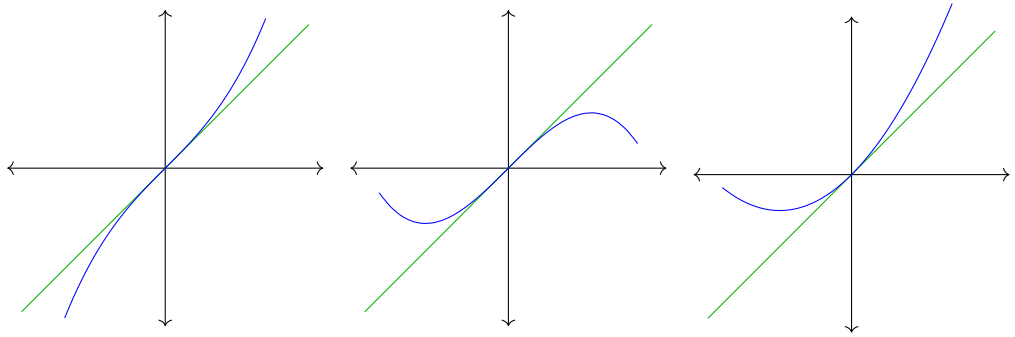
- Točka x_0 je ŠIBKO PRIVLAČNA, če obstaja okolica $U \ni x$, da za vsak $y_0 \in U$ velja $y_n = f^n(x_0) \to x_0$. Okolici U pravimo OBMOČJE PRIVLAKA za x_0, največjemu intervalu znotraj U pa NEPOSREDNO OBMOČJE PRIVLAKA za x_0.

[1] Prirejeno po: stackexchange

- Točka x_0 je ŠIBKO ODBOJNA, če obstaja okolica $U \ni x$, da za vsak $y_0 \in U \setminus \{x_0\}$ obstaja $m \in \mathbb{N}$, da $f^m(y_0) \notin U$.

Definicija 9.2.5. Če je f zvezno odvedljiva, uvedemo dodatne pojme. Točka x_0 je PRIVLAČNA, če je $|f'(x_0)| < 1$, ODBOJNA, če je $|f'(x_0)| > 1$, oziroma NEVTRALNA, če je $|f'(x_0)| = 1$.

Primer. Funkcije $f_1(x) = x + x^3$, $f_2(x) = x - x^3$ in $f(x) = x + x^2$ imajo vse fiksno točko v 0, a se obnaša drugače. V prvem primeru je 0 šibko odbojna točka, v drugem šibko privlačna, v zadnjem pa niti ena niti druga. Preverimo lahko, da je v vseh primerih $|f_i'(0)| = 1$, torej so vse točke odbojne.



(a) $f_1(x) = x + x^3$ (b) $f_2(x) = x - x^3$ (c) $f_3(x) = x + x^2$

Figure 9.2: Funkcije iz primera

Izrek 9.2.6. *Naj bo $f \in \mathcal{C}^1(I)$ in $x_0 \in I$ fiksna točka. Potem velja:*

1. *Če je $|f'(x_0)| < 1$, je x_0 šibko privlačna.*
2. *Če je $|f'(x_0)| > 1$, je x_0 šibko odbojna.*

Opomba. Če je točka privlačna, je šibko privlačna. Če je odbojna, je šibko odbojna.

Proof. Naj bo λ tak, da je $|f'(x_0)| < \lambda < 1$. Potem obstaja $\delta > 0$, da je $|f'(x)| < \lambda$ za vse $x \in (x_0 - \delta, x_0 + \delta)$. Po Lagrangeovem izreku obstaja $c \in (x, x_0)$, da velja

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(c).$$

Velja $f(x_0) = x_0$, torej je $|f(x) - x_0| = |f'(c)|\,|x - x_0| < \lambda\,|x - x_0|$ in posledično $f(x) \in (x_0 - \delta, x_0 + \delta)$. Induktivno potem $|f^n(x) - f(x_0)| < \lambda^n\,|x - x_0|$, kar konvergira k 0 za $n \to \infty$.

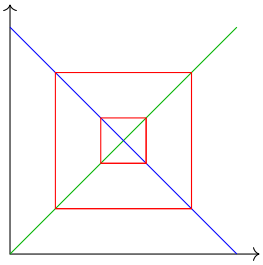
Podobno za drugo točko. □

Definition 9.2.7. Fiksna točka $x_0 \in I$ je STABILNA za $f : I \to I$, če za vsako njeno okolico $U \subseteq I$ obstaja manjša okolica $U' \subseteq U$, da za vsak $x \in U'$ velja $O_f(x) \subseteq U$.

Opomba. Vse privlačne točke so stabilne (glej zgornji dokaz).

Opomba. Stabilne šibke privlačne točke imenujemo ASIMPTOTSKO STABILNE.

Primer. Poglejmo si $f(x) = 1 - x$. Ker je $f^2(x) = x$, ima f fiksno točko v $x_0 = \frac{1}{2}$, vse ostale točke pa so periodične s periodo 2. Točka $\frac{1}{2}$ je nevtralna (ni privlačna/odbojna), je pa stabilna.



Opomba. Če je f zvezna in gre $f^n(x) \to x_0$, je x_0 nujno fiksna točka.

Definicija 9.2.8. Naj bo x_0 n-periodična točka zvezne funkcije f. Pripadajoči n-cikel je ŠIBKO PRIVLAČEN, če je vsak njegov element šibko privlačna točka za f^n in ŠIBKO OBOJEN, če je vsak njegov element šibko odbojna točka za f^n. Podobno definiramo privlačnost, odbojnost in nevtralnost cikla za zvezno odvedljivo f.

Izrek 9.2.9. *Če je $f \in \mathcal{C}^1(I)$, so vse periodične točke n-cikla istega tipa za f^n.*

Proof. Naj bo $\{x_1, \ldots, x_n\}$ cikel funkcije f. Recimo, da je x_n šibko privlačna za f^n, torej obstaja okolica $U_n \ni x_n$, da za $x \in U$ velja $f^{nk}(x) \xrightarrow[k\to\infty]{} x_n$. Definiramo U_{n-j} kot tisto povezano komponento praslike $f^{-j}(U_n)$, ki vsebuje x_j.

Vemo, da za vsak $x \in U_{n-j}$ po konstrukciji velja $f^j(x) \in U_n$, torej $f^{nk+j}(x) \xrightarrow[k\to\infty]{} x_n$, iz česar sledi $f^{nk}(x) \xrightarrow[k\to\infty]{} x_{n-j}$ (uporabimo f^{n-j} na prejšnji limiti).

Dokaz za šibko odbojnost izpustimo. Za preostale tri lastnosti računamo

$$\begin{aligned}\left|(f^n)'(x_n)\right| &= \left|(f \circ f \circ \cdots \circ f)'(x_n)\right| \\ &= \left|f'(f^{n-1}(x_n))\right| \left|f'(f^{n-2}(x_n))\right| \cdots \left|f'(x_n)\right| \\ &= \left|f'(x_1)\right| \left|f'(x_2)\right| \cdots \left|f'(x_n)\right|.\end{aligned}$$

Podoben razvoj lahko naredimo v poljubni točki cikla, in dobimo enak rezultat. □

9.2.2 Definicija kaosa

Definicija 9.2.10 (Devaney)**.** Dinamični sistem, podan z $f : I \to I$ je KAOTIČEN, če zanj veljajo naslednje lastnosti:

(C1) Množica periodičnih točk je gosta v I,

(C2) Tranzitivnost: za poljubna odprta intervala $U_1, U_2 \subseteq I$ obstajata $x_0 \in U_1$ ter $n \in \mathbb{N}$, da je $f^n(x_0) \in U_2$.

(C3) Občutljivostna konstanta: Obstaja $\beta > 0$, da v poljubni okolici U poljubne točke x_0 najdemo tudi točko $y_0 \in U$, za katero je $|f^n(x_0) - f^n(y_0)| > \beta$ za nek $n \in \mathbb{N}$.

Opomba. Tretja točka definicije je zanimiva pri izbiri majhnih okolic U, saj pomeni, da lahko vsaki točki poljubno blizu najdemo točko s popolnoma drugačno dinamiko. Temu pravimo "metuljev pojav." Pomeni, da je orbita občutljiva na začetne pogoje.

Opomba. Za drugo točko je zadosti pokazati, da obstaja gosta orbita. Če je f zvezna, je to tudi potreben pogoj po Bairovem izreku.

Opomba. Če je f zvezna, se izkaže, da prvi dve točki implicirata tretjo.

Primer (Podvojitvena preslikava). Definiramo $D : [0, 1) \to [0, 1)$ z

$$D(x) = 2x - \lfloor 2x \rfloor = \begin{cases} 2x & x < \frac{1}{2} \\ 2x - 1 & x \geq \frac{1}{2} \end{cases}$$

Za dokaz kaotičnosti uporabimo dvojiški zapis, kjer dodatno prepovemo neskončen niz enic. Naša preslikava potem število $0.x_1x_2x_3\ldots$ slika v $0.x_2x_3x_4\ldots$, zato tej preslikavi pravimo tudi "operator zamika."

Naj bo $x \in [0, 1)$ poljubna. Potem je za dovolj velik $N \in \mathbb{N}$ točka

$$\tilde{x} = 0.x_1x_2\ldots x_Nx_1x_2\ldots x_N\ldots$$

periodična in se nahaja v majhni okolici točke x.

Za točko (C2) je dovolj pokazati, da obstaja gosta orbita. To bo natanko orbita točke

$$x = 0.0100011011000001010011100\ldots$$

To število vsebuje vse končne dvojiške zapise. Posledično je orbita gosta, saj za poljuben $\tilde{x} \in [0, 1]$, ki ima dvojiški zapis $\tilde{x} = 0.\tilde{x}_1\tilde{x}_2\ldots$. Potem za $N \in \mathbb{N}$ obstaja $m \in \mathbb{N}$, da je $f^m(x) = 0.\tilde{x}_1\ldots\tilde{x}_N$.

Za točko (C3) bomo pokazali, da za $x \in [0, 1]$ poljubno blizu obstaja $\tilde{x}$, da za nek m velja $|f^m(x) - f^m(\tilde{x})| = \frac{1}{2}$. Konkretno za $x = 0.x_1x_2\ldots$ vzamemo

$$\tilde{x} = 0.x_1x_2\ldots x_N\tilde{x}_{N+1}x_{N+2}\ldots,$$

kjer je $\tilde{x}_{N+1} \neq x_{N+1}$. Potem je $|x - \tilde{x}| < 2^{-N}$, ampak $\left|f^N(x) - f^N(\tilde{x})\right| = \frac{1}{2}$. ⊡

Primer (Šotorska preslikava). Tudi preslikava

$$T(x) = \begin{cases} 2x & x \leq \frac{1}{2} \\ 2 - 2x & x \geq \frac{1}{2} \end{cases}$$

je kaotična. Graf T^n ima 2^n enako razporejenih šotorov, vsi od katerih so visoki 1, torej imamo 2^n fiksnih točk, torej imamo na vsakem intervalu oblike $[\frac{j}{2^n}, \frac{j+1}{2^n}]$ periodično točko. Potem (C1) očitno velja. Sledi tudi (C2), saj se pri uporabi T vsak interval zgornje oblike slika v dvakrat večji interval, torej bomo po nekem številu korakov prekrili celoten $[0, 1]$. Točka (C3) sledi iz prvih dveh, saj je $[0, 1]$ kompakten.

9.2.3 Konjugacije in semikonjugacije

Definicija 9.2.11. Pravimo, da sta $f : I \to I$ in $g : J \to J$ KONJUGIRANI, če obstaja homeomorfizem $h : I \to J$, da spodnji diagram komutira:

$$\begin{array}{ccc} I & \xrightarrow{f} & I \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle h} \\ J & \xrightarrow{g} & J \end{array}$$

Opomba. Že iz pogoja $h \circ f = g \circ h$ sledi $h(O_f(x)) = O_g(h(x))$ za poljuben $x \in I$. Iz zahteve, da je h homeomorfizem, dobimo še $f^n = h^{-1} \circ g^n \circ h$, in obstaja homeomorfna korespondenca med orbitama f in g.

Primer. Oglejmo si $f(x) = x^2 - 2x + 2$ na $[1, \infty)$ ter $g(x) = x^2$ na $[0, \infty)$. Potem za $h(x) = x - 1$ dobimo $h(f(x)) = g(h(x))$ za poljuben x.

Opomba. Vsaka kvadratna funkcija $f(x) = ax^2 + bx + c$ je konjugirana neki funkciji $q_C(x) = x^2 + C$ za $C \in \mathbb{R}$.

Opomba. V primeru se je ohranil tudi karakter fiksnih točk. Še več; če sta h, h^{-1} tudi odvedljivi, za fiksno točko x velja $f'(x) = g'(h(x))$ po verižnem pravilu.

Definicija 9.2.12. Pravimo, da je $g : J \to J$ SEMIKONJUGIRANA funkciji $f : I \to I$, če obstaja zvezna surjektivna preslikava $h : I \to J$, za katero velja

- $h \circ f = g \circ h$,
- obstaja $m \in \mathbb{N}$, da ima za vsak $x \in J$ praslika $h^{-1}(x)$ največ m elementov.

Opomba. Definicija semikonjugacije lahko variira glede na literaturo.

Opomba. Znova velja $h(O_f(x)) = O_g(h(x))$, vendar pa ta relacija ni več bijektivna.

Primer. Vzemimo $h : [-1, 1] \to [0, 1]$, $h(x) = x^2$ ter funkciji $f : [-1, 1] \to [-1, 1]$, $f(x) = \sqrt{1 - x^2}$ in $g : [0, 1] \to [0, 1]$, $g(x) = 1 - x$. Funkcija f ima eno fiksno točko, vse ostale pa so bodisi predperiodične bodisi del 2-cikla. Podobno ima funkcija g eno fiksno točko, vse ostale točke pa so 2-cikli. Dinamiki torej nista ekvivalentni, funkciji pa sta semikonjugirani. V tem primeru smo sklopili predperiodične in periodične točke.

Trditev 9.2.13. *Naj bo g semikonjugirana f preko h. Če je $x_0 \in I$ periodična za f, je tudi $h(x_0)$ periodična za g, a se perioda ne ohrani nujno.*

Proof. Naj bo $n \in \mathbb{N}$ najmanjše število, da velja $f^n(x_0) = x_0$ in $n \in \mathbb{N}$. Potem velja $g^n(h(x_0)) = g^{n-1}(g(h(x_0))) = g^{n-1}(h(f(x_0))) = \cdots = h(f^n(x_0)) = h(x_0)$. □

Opomba. Da se pokazati, da je nova perioda delitelj stare.

Izrek 9.2.14. *Naj bo g semikonjugirana f preko h in naj bosta I, J kompaktna intervala. Če je f kaotična in je izpolnjen eden od spodnjih treh pogojev, je g kaotična:*

- *h je injektivna,*
- *g je zvezna,*
- *f je zvezna.*

Proof. Točka (C1) sledi iz prejšnje trditve: če vzamemo poljubno $U^{\text{odp}} \subseteq J$, je tudi $h^{-1}(U)$ odprta v I in vsebuje periodično točko x_0, torej je $h(x_0)$ periodična v U. Na enak način dobimo tudi gosto orbito.

Recimo, da je h injektivna. Za (C3) si oglejmo preslikavo $d : I \times I \to \mathbb{R}$, podano s predpisom $d(x, y) = |h(x) - h(y)|$. To je zvezna preslikava iz kompakta v $\mathbb{R}$. Definiramo še

$$\Lambda_\beta = \{(x, y) \in I \times I \mid |x - y| \geq \beta\}.$$

Ker je d zvezna, slika Λ_β v kompakten interval v $\mathbb{R}$, ki ne vsebuje 0. Obstaja $\beta' > 0$, da iz pogoja $|f^n(x) - f^n(y)| > \beta$ sledi $|g^n(h(x)) - g^n(h(y))| > \beta'$. To je občutljivostna konstanta za g.

Če je g zvezna, pogoj (C3) sledi iz pogojev (C1) in (C2). Če pa je f zvezna in g ni, potem se da dokazati, da obstaja $\delta > 0$, za katerega iz $l(V) < \delta$ sledi, da so vse h-praslike dolžine največ β, kjer je $V^{\text{odp}} \subseteq J$ in $l(V)$ njena dolžina. Tedaj je $\frac{\delta}{2}$ iskana občutljivostna konstanta za g. □

Primer. Kompaktnost intervalov je pomembna. Če vzamemo $f(x) = 2x$ na $(0, \infty)$ in $g(x) = x + \log 2$ na $\mathbb{R}$, potem f očitno izpolnjuje pogoj (C3), saj je razširitev. Po drugi strani je očitno, da g ne izpolnjuje (C3), saj je razlika $|g^n(x) - g^n(y)|$ konstantna za vse $n \in \mathbb{N}$. Vseeno pa sta f in g konjugirani preko $h(x) = \log x$.

Primer. Velja $T(D(x)) = T^2(x)$, torej je T semikonjugirana D preko $h = T$. Ker je podvojitvena preslikava kaotična, je tudi T kaotična.

Primer (podvojitev argumenta). Naj bo $h : [0, 1) \to S^1$ homeomorfizem, podan z $h(x) = (\cos 2\pi x, \sin 2\pi x)$. Zanima nas, kaj se zgodi s podvojitveno preslikavo, če uporabimo h kot konjugacijo. Trdimo, da je konjugirana preslikavi $f : S^1 \to S^1$, podani z $f(\cos t, \sin t) = (\cos 2t, \sin 2t)$, oziroma $f(z) = z^2$.

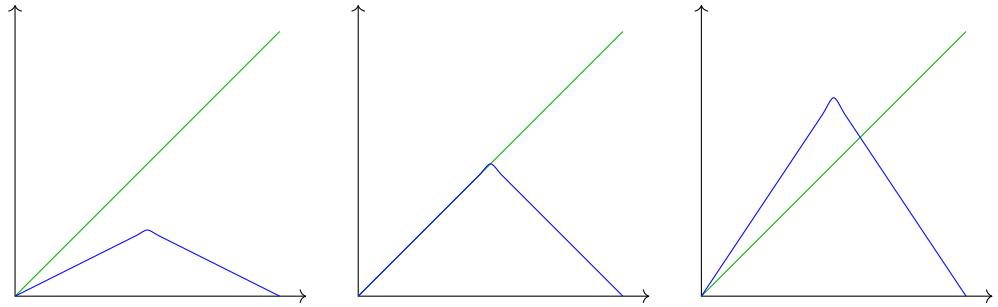
Primer. Opazujemo preslikavo $q_{-2}(x) = x^2 - 2$ na intervalu $[-2, 2]$. Preverimo lahko, da je q_{-2} semikonjugirana s preslikavo iz prejšnjega primera s $h : (\cos t, \sin t) \mapsto 2 \cos t$ oziroma $h(z) = 2 \operatorname{Re} z$. Potem je q_{-2} semikonjugirana tudi D, torej je kaotična.

9.2.4 Bifurkacije

Za motivacijo si oglejmo enoparametrično družino šotorskih preslikav za $c \in [0, 2]$:

$$T_c(x) = c \min(x, 2 - x) = \frac{c}{2} T(x).$$

Vse te preslikave interval $[0, 1]$ slikajo vase, dinamike pa se malce razlikujejo. Če je $c < 1$, je 0 edina fiksna točka, in je privlačna, saj je $T_c(x) < x$ za poljuben $x \in [0, 1]$. Pri $c = 1$ so vse točke v $[0, \frac{1}{2}]$ fiksne, vse ostale točke pa so predperiodične (oz. predfiksne). Če pa je $c > 1$, pa imamo dve fiksni točki, ki sta obe odbojni.



(a) $c = 0.5$ (b) $c = 1$ (c) $c = 1.5$

Figure 9.3: Grafi T_c za različne vrednsoti c

Poglejmo si še, kaj se zgodi z 2-cikli. Pri $c \leq 1$ je situacija praktično enaka kot za fiksne točke, pri $c > 1$ pa se zgodi nekaj bolj zanimivega.

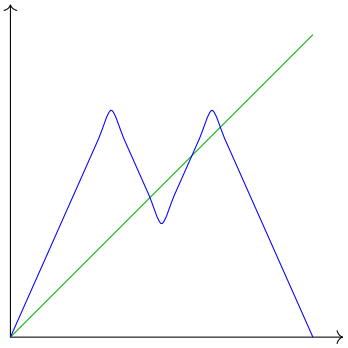


Figure 9.4: Graf $T^2_{1.5}$

V splošnem za to družino velja, da se pri večjih c pojavljajo m-cikli, ki postanejo gosti pri $c = 2$. Pri $c = 1$ se zgodi tudi sprememba, da iz točke $x = \frac{1}{2}$ dobimo ne le odbojno točko, ampak tudi dodaten 2-cikel. Točkam, pri katerih se zgodi "kvalitativna sprememba" dinamike, pravimo BIFURKACIJE.

Mi bomo bifurkacije opazovali na družinah zveznih funkcij $f_c : I \to I$, ki bodo gladko odvisne od parametra $c \in J \subseteq \mathbb{R}$. To pomeni, da za funkcijo $F(x, c) = f_c(x)$ obstajajo vsi parcialni odvodi $\frac{\partial^k}{\partial c^k} F(x, c)$ za $k \in \mathbb{N}$.

Ključna opazka z motivacije je, da so se vse spremembe dinamike zgodile, ko je bil odvod v fiksni (oz. periodični) točki bodisi 1 bodisi ni obstajal. To formalizira spodnji izrek.

Izrek 9.2.15. *Naj bo $f_c : I \to I$ družina zvezno odvedljivih funkcij, ki so gladko odvisne od parametra $c \in J$. Denimo, da za $x_0 \in I$ velja $f_{c_0}(x_0) = x_0$ in $f'_{c_0}(x_0) \neq 1$. Potem obstajata okolici $I' \subseteq I$ in $J' \subseteq J$ ter preslikava $p : J' \to I'$, da je $f_c(p(c)) = p(c)$, za katero je $p(c)$ edina fiksna točka f_c na I'.*

Proof. Uporabimo izrek o implicitni funkciji za $F(x, c) = f_c(x) - x$. Po predpostavki je $F(x_0, c_0) = 0$ in $\partial_x F(x_0, c_0) = f'_{c_0}(x_0) - 1 \neq 0$, torej lahko na majhni okolici c_0 izrazimo $x = p(c)$ ter velja $F(p(c), c) = 0$. □

Posledica 9.2.16. *Privlačne in odbojne točke so stabilne za majhne motnje, prav tako pa tudi točke, v katerih je odvod enak -1.*

Opomba. Izrek podaja le potreben, ne pa tudi zadosten pogoj.

Primer (Tangentna bifurkacija). To je bifurkacija, v kateri iz nevtralne točke dobimo dve fiksni točki, ki jih prej ni bilo. Ena je privlačna, druga odbojna, zato se angleško imenuje tudi *saddle-node bifurcation*. Primer tega je $q_c(x) = x^2 + c$ pri $c = \frac{1}{4}$.

Primer (Potrojitev fiksne točke). Za $f_c(x) = c \arctan(x)$ imamo pri $c \leq 1$ natanko eno fiksno točko pri $x = 0$, ki je privlačna za $c < 1$, pri $c = 1$ pa nevtralna in šibko privlačna. Pri $c > 1$ dobimo tri fiksne točke.

Primer (Podvojitev periode). Naj bo $f_{c_0}(x_0) = x_0$ in $f'_{c_0}(x_0) = -1$. Po izreku se fiksna točka ohrani, za drugi $f^2_{c_0}$ pa velja $(f^2_{c_0})'(x_0) = 1$, torej se lahko zgodi bifurkacija. Če se zgodi, potem iz fiksne točke dobimo eno fiksno točko in 2-cikel (fiksne točke namreč ne moremo izgubiti in ne moremo dobiti). Primer take preslikave je $q_c(x) = x^2 + c$ pri $c = -\frac{3}{4}$.

9.2.5 Simbolična dinamika

Oglejmo si preslikavo $T_3(x) = 3 \min\{x, 1 - x\}$. Funkcija slika le podintervala $[0, \frac{1}{3}]$ in $[\frac{2}{3}, 1]$ slika nazaj v $[0, 1]$. Podobno se za T_3^2 še manjši del intervala slika nazaj v $[0, 1]$. Če nadaljujemo induktivno, ugotovimo, da T_3^n v $[0, 1]$ nese ravno množice $I_n = \frac{1}{3} I_{n-1} \cup (1 - \frac{1}{3} I_{n-1})$. Torej se v $[0, 1]$ za vse $n \in \mathbb{N}$ preslikajo natanko elementi Cantorjeve množice. Radi bi pokazali, da $T_3 : C \to [0, 1]$ deluje kaotično.

Uvedemo prostor indeksov

$$S = \{\bar{s} = (s_1, s_2, \ldots) \mid s_j \in \{0, 1\}\} = \{0, 1\}^\infty$$

in metriko

$$d(\bar{s}, \bar{t}) = \sum_{j=1}^{\infty} \frac{|s_j - t_j|}{2^j}.$$

Vsakemu elementu $x \in C$ priredimo $\bar{s} \in S$ tako, da vzamemo tretjiški zapis x, in zamenjamo dvojke z enicami. Preverimo lahko, da je ta preslikava $h : C \to S$ homeomorfizem. Opazimo, da je $T_3(C) \subseteq C$, torej jo lahko obravnavamo kot preslikavo $C \to C$. Potem lahko h obravnavamo kot konjugacijo

$$\begin{array}{ccc} C & \xrightarrow{T_3} & C \\ \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle h} \\ S & \xrightarrow[f]{} & S \end{array}$$

Na intervalu $[0, \frac{1}{3}]$ preslikava T_3 deluje kot $x \mapsto 3x$, torej $f = h \circ T_3 \circ h^{-1}$ na tem intervalu slika $(0, s_2, s_3, \ldots)$ v $(s_2, s_3, \ldots)$. Podobno, če je $x \in [\frac{2}{3}, 1]$, slikamo $(1, s_2, s_3, \ldots)$ v $(1 - s_2, 1 - s_3, \ldots)$ (modulo 2). Enostavno se lahko prepričamo, da je f kaotična.

Podobno bi lahko obravnavali preslikave, ki na vsakem koraku razdelijo interval na $N > 2$ podintervalov. Tedaj definiramo h tako, da slika x v $(a_1, a_2, \ldots) \in S$, kjer je a_i enak zaporedni številki intervala, v katerem je i-ta komponirana slika x.

9.2.6 Izrek Šarkovskega

Izrek 9.2.17. *Naj bo $I \subseteq \mathbb{R}$ kompakten interval in $f : I \to I$ zvezna. Če obstaja 3-cikel za f, obstajajo tudi n-cikli za vse ostale $n \in \mathbb{N}$.*

Dokaz je osnovan na dveh preprostih dejstvih za kompaktna intervala I, J in zvezno f:

- Če je $I \subseteq f(J)$, obstaja interval $J_0 \subseteq J$, da je $f(J_0) = I$.
- Če je $J \subseteq f(J)$, obstaja fiksna točka v J.

Naj bodo $a < b < c$ točke v 3-ciklu. Brez škode za splošnost naj velja $f(a) = b$, $f(b) = c$ in $f(c) = a$ (sicer imamo simetrični primer). Začnemo z $J_0 = [b, c]$. Vidimo, da $f(J_0)$ pokriva $[a, c]$. Potem obstaja interval $J_1 \subseteq J_0$, da je $f(J_1) = J_0$. Potem $f(J_1)$ pokriva J_0, in nadaljujemo induktivno. Tako dobimo zaporedje intervalov

$$[b, c] = J_0 \supseteq J_1 \supseteq J_2 \supseteq \cdots$$

za katere velja $f(J_k) = J_{k-1}$, oziroma $f^n(J_n) = J_0 \supseteq J_n$.

Naj bo $n \in \mathbb{N}$. Ker je $f^{n-1}(J_{n-2}) = f(J_0) \supseteq [a, c]$, obstaja podinterval $I_{n-1} \subseteq J_{n-2}$, za katerega je $f^{n-1}(I_{n-1}) = [a, b]$. Če to slikamo z f, dobimo $f^n(I_{n-1}) \supseteq J_0$. Torej obstaja $I_n \subseteq I_{n-1}$, da je $f(I_n) = J_0$. To pa pomeni, da je $I_n \subseteq f^n(I_n)$ in imamo fiksno točko x preslikave f^n. Vemo, da je $f^k(x) \in J_0$ za $1 \leq k \leq n - 2$, ter $f^{n-1}(x) \in [a, b)$ (lahko preverimo, da $x \neq b$). Torej mora imeti x periodo n, in ne manjše periode.

Primer. Pri družini šotorskih preslikav $T_c(x)$ se 3-cikel prvič pojavi pri $c = \phi$. Za vse večje c imamo dva 3-cikla, in cikle vseh ostalih dolžin.

Primer. Zveznost je potreben pogoj: za funkcijo $f(x) = 1 - \frac{1}{x}$ velja $f^3(x) = x$.

Imamo še splošnejši izrek. Oglejmo si naslednjo ureditev naravnih števil:

$$\begin{aligned}
&3 \triangleright 5 \triangleright 7 \triangleright 9 \triangleright \cdots \\
\triangleright\ &2\cdot 3 \triangleright 2\cdot 5 \triangleright 2\cdot 7 \triangleright 2\cdot 9 \triangleright \cdots \\
\triangleright\ &4\cdot 3 \triangleright 4\cdot 5 \triangleright 4\cdot 7 \triangleright 4\cdot 9 \triangleright \cdots \\
\triangleright\ &8\cdot 3 \triangleright 8\cdot 5 \triangleright 8\cdot 7 \triangleright 8\cdot 9 \triangleright \cdots \\
\triangleright\ &\cdots \\
\triangleright\ &\cdots \triangleright 8 \triangleright 4 \triangleright 2 \triangleright 1
\end{aligned}$$

Izrek 9.2.18 (Šarkovski)**.** *Naj bo $I \subseteq \mathbb{R}$ kompakten interval in $f : I \to I$ zvezna. Če obstaja p-cikel za f, obstajajo tudi vsi q-cikli za števila $p \triangleright q$.*

Izrek 9.2.19 (komplement Šarkovskega)**.** *Za vsaka $q \triangleright p$ obstaja funkcija s p-ciklom in brez q-cikla.*

Posledica 9.2.20. *Če ima f le končno mnogo periodičnih točk, so vse njihove periode dolžine potence 2.*

9.2.7 Realna dinamika v višjih dimenzijah

Obravnavane pojme lahko uporabimo tudi za funkcije $F : U \subseteq \mathbb{R}^m \to U$ za $m \geq 2$.

Primer (Pekova preslikava). Imamo naslednjo preslikavo $F : [0,1)\times[0,1) \to [0,1)\times[0,1)$:

$$F(x,y) = (2x - \lfloor 2x \rfloor, \frac{1}{2}y + \frac{1}{2}\lfloor 2x \rfloor).$$

Oglejmo si to preslikavo v dvojiškem zapisu (brez neskončnega zaporedja enic). Pišimo $x = 0.x_1x_2x_3\ldots$ in $y = 0.y_1y_2y_3\ldots$, kar se transformira v

$$F(x,y) = (0.x_2x_3x_4\ldots, 0.x_1y_1y_2y_3\ldots).$$

To lahko predstavimo tudi kot operator zamika v zaporedjih z indeksi, ki gredo proti ∞ in $-\infty$:

$$\ldots x_3x_2x_1|y_1y_2y_3\ldots \mapsto \ldots x_4x_3x_2|x_1y_1y_2\ldots$$

Iz tega ni težko dokazati, da je F kaotična.

Za $x < \frac{1}{2}$ velja $F(x,y) = (2x, y/2)$. V y smeri je preslikava torej privlačna, v x smeri pa odbojna. Izhodišče ni stabilna točka in ni niti šibko privlačna niti šibko odbojna točka. Takim točkam pravimo SEDLA. ⊡

Primer. Definiramo preslikavo

$$F(x,y) = \frac{1}{12}\begin{bmatrix}5 & 3\\ 3 & 5\end{bmatrix}\cdot\begin{bmatrix}x\\ y\end{bmatrix}.$$

Dinamiko lahko analiziramo v lastnih koordinatah (tj. koordinate, ki jih razpenjata lastna vektorja matrike), potem pa jo transformiramo v običajne koordinate. V tem primeru je izhodišče asimptotsko stabilna fiksna točka.

Primer. Za preslikavo

$$F(x,y) = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$$

imamo kompleksni lastni vrednosti. Če pogledamo lastna vektorja, vidimo, da se povsod v spirali oddaljujemo od izhodišča.

Primer. Imejmo Fibonaccijevo zaporedje $a_{n+2} = a_{n+1} + a_n$. Za $x_n = a_{n-1}$ in $y_n = a_n$ lahko to prevedemo na sitem

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_n \\ y_n \end{bmatrix}.$$

Obe lastni vrednosti matrike sta večji od 1 po absolutni vrednosti.

Trditev 9.2.21. *Naj bo $A \in \mathbb{R}^{n\times n}$ obrnljiva matrika, katere lastne vrednosti so manjše od 1 po absolutni vrednosti. Potem je izhodišče asimptotsko stabilna fiksna točka preslikave $F(x) = Ax$. Podobno, če so vse lastne vrednosti večje od 1 po absolutni vrednosti, je izhodišče šibko odbojna točka.*

Opomba. Primere z $\lambda = 0$ smo izvzeli, ker lahko pri takih primerih zmanjšaš dimenzijo.

Izkaže se, da linearna teorija pove veliko tudi o lokalnih lastnostih nelinearnih sistemov v okolici fiksnih točk.

Definicija 9.2.22. Naj bo $F : U \subseteq \mathbb{R}^m \to U$ razreda $\mathcal{C}^1$. Točka $x_0 \in U$ je HIPERBOLIČNA, če za vse lastne vrednosti Jacobijeve matrike v tej točki velja $0 < |\lambda| \neq 1$.

Izrek 9.2.23 (Hartman-Grobman)**.** *Naj bo $F : U \subseteq \mathbb{R}^m \to U$ razreda $\mathcal{C}^1$ in $x_0 \in U$ njena hiperbolična točka. Potem obstaja homeomorfizem φ, ki neko okolico izhodišča slika v okolico x_0, ter za katerega velja $\varphi \circ DF(x_0) = F \circ \varphi$.*

Opomba. Blizu točke x_0 se do zvezne preslikave natančno preslikava F obnaša kot linearni sistem za $A = DF(x_0)$.