Lite speciell matematik

Karl-Heinz Fieseler

Uppsala 2015

Contents

1 Mängder

Mängdläran samt oändliga kardinaltal går tillbaka till Georg Cantor (1845 - 1918). Vi påminner om mängdlärans grundbegrepp:

Definition 1.1. En mängd M är en sammanfattning av vissa objekt, dess *element*. Två mängder M, N är lika:

$$M = N$$

om de har samma element. Vi skriver

$$x \in M$$

om objektet x är med i M som element. I synnerhet finns det en mängd, den tomma mängden $\emptyset$, som inte har några element alls.

Den tomma mängden $\emptyset$ kan jämföras med nollan: Det är bekvämt att ha den till sitt förfogande. Tänk dig mängderna som paket: Det finns ju också ett tomt paket.

Example 1.2. 1. $\mathbb{N} := \{0, \underbrace{1, 2, 3, ...}_{\text{Gud}}\}$.

2. Vi ska snart diskutera människornas försök att skapa naturliga tal.

3. $\mathbb{R} := \{x; x \text{ reellt tal}\}$. Här kan vi inte längre rada upp alla element, en sak vi ska visa senare.

Remark 1.3. 1. En mängd kan beskrivas genom ett påstående om dess element: Mängden
$$\{x; P(x)\}$$
har som element alla objekt x sådant att påståendet $P(x)$ är sant.

Exempel:

(a) $\{x; x \in \mathbb{R}, x^2 = 1\} = \{\pm 1\}$,

(b) $\{x; x \in \mathbb{R}, x^2 = -1\} = \emptyset$.

(c) Intervall: Slutna: $[a, b] := \{x \in \mathbb{R}; a \leq x \leq b\}$ och öppna: $(a, b) := \{x \in \mathbb{R}; a < x < b\}$.

2. Låt A, B vara mängder. Vi skriver
$$A \subset B$$
och kallar A en delmängd av B om alla element som tillhör A också tillhör B. Eller symboliskt:
$$x \in A \overset{\text{innebär}}{\Longrightarrow} x \in B.$$

3. Mängddifferens: Låt A, B vara mängder. Deras differens $A \setminus B$ är mängden
$$A \setminus B := \{x; x \in A \text{ och } x \notin B\}.$$
Exempel: $A = [a, b], B = (a + \frac{b-a}{3}, b - \frac{b-a}{3})$, så är
$$A \setminus B = \left[a, a + \frac{b-a}{3}\right] \cup \left[b - \frac{b-a}{3}, b\right].$$

4. Låt $A_n, n \in \mathbb{N}$, vara en följd av mängder. Deras union är mängden
$$\bigcup_{n=0}^{\infty} A_n := \{x; \overbrace{\text{det finns ett}}^{\exists}\ n \in \mathbb{N} : x \in A_n\}.$$
Exempel: Ta $A_n := \left[\frac{1}{n}, 1 - \frac{1}{n}\right], n \in \mathbb{N}_{>0}$. Då har vi
$$\bigcup_{n=1}^{\infty} A_n = (0, 1).$$

5. Låt $A_n, n \in \mathbb{N}$, vara mängder. Deras snitt är mängden

$$\bigcap_{n=0}^{\infty} A_n := \{x; \overbrace{\text{för alla}}^{\forall} \ n \in \mathbb{N} : x \in A_n\}.$$

Exempel: Ta $A_n := [a_n, b_n]$. Om intervallföljden A_n är avstigande, dvs. $A_n \supset A_{n+1}$ för alla $n \in \mathbb{N}$ och längderna konvergerar mot 0, dvs. $b_n - a_n \to 0$, så har vi

$$\bigcap_{n=0}^{\infty} A_n = \{c\}.$$

med det reella talet $c = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

6. **Cantors diskontinuum**: Vi tar

$$T^0 := [0,1], T^1 := \left[1, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right] = T_1^1 \cup T_2^1,$$

får fram i det tredje steget

$$T^2 := \bigcup_{k=1}^{4} T_k^n,$$

där T_1^2 resp. T_3^2 är den första tredjedelen till T_1^1 resp. T_2^1 och T_2^2 resp. T_4^2 den sista tredjedelen. Om vi fortsätter får vi allmänt

$$T^n = \bigcup_{k=1}^{2^n} T_k^n,$$

med delintervall T_k^n av längd $\frac{1}{3^n}$, sådant att T_{k+1}^n ligger höger om T_k^n. Snittet

$$C := \bigcap_{n=0}^{\infty} T^n$$

kallas för Cantors diskontinuum. Punkterna i C motsvarar avstigande intervallföljder $T_{k_n}^n$, dvs. vi har $k_{n+1} = 2k_n$ eller $k_{n+1} = 2k_n - 1$. Motsvarande punkt $x \in C$ definieras genom

$$\bigcap_{n=0}^{\infty} T_{k_n}^n = \{x\},$$

medan en punkt $x \in C$ tillordnas följden $T^n_{k_n}$ med $x \in T^n_{k_n}$. Om vi använder oss av den treadiska utvecklingen, så får vi

$$C = \left\{ x \in \mathbb{R}; x = \sum_{\nu=1}^{\infty} a_\nu 3^{-\nu}, a_n \in \{0, 2\} \right\}.$$

Bevis: "$\subset$": För $x \in C \cap T^n_{k_n}$ sätter vi

$$a_{n+1} := \left\{ \begin{array}{ccl} 0 & , & \text{om } x \in T^{n+1}_{2k_n - 1} \\ 2 & , & \text{om } x \in T^{n+1}_{2k_n} \end{array} \right. .$$

Sedan är $\sum_{\nu=1}^{n+1} a_\nu 3^{-\nu}$ vänstra ändpunkten till intervallet $T^{n+1}_{k_{n+1}} \ni x$ och således $x = \sum_{n=1}^{\infty} a_n 3^{-n}$.

"$\supset$": Om nu $x = \sum_{\nu=1}^{\infty} a_\nu 3^{-\nu}$ är givet, ta k_n, sådant att $\sum_{\nu=1}^{n} a_\nu 3^{-\nu}$ är vänstra ändpunkten till $T^n_{k_n}$. Sedan har vi även $x \in T^n_{k_n}$ och således också $x \in \bigcap_{n=1}^{\infty} T^n_{k_n} \subset C$.

7. Sådana "patologiska" mängder som hans diskontinuuum C har lett Cantor till mängdläran, i synnerhet också till kardinaltal. Den är en källa till många intressanta anmärkningar, t.ex. följande:

   (a) C är en "tunn" delmängd till enhetsintervallet $[0, 1]$: Dess komplement $[0, 1] \setminus C$ har samma längd som $[0, 1]$ pga.

   $$\sum_{\nu=0}^{\infty} \frac{2^\nu}{3^{\nu+1}} = 1,$$

   men det har lika många element som $[0, 1]$.

   (b) Det finns en kontinuerlig icke-avtagande funktion som är deriverbar på $\mathbb{R} \setminus C$ med derivata $f' = 0$, den är icke-avtagande och man har

   $$f(x) = \left\{ \begin{array}{ccl} 0 & , & \text{om } x \leq 0 \\ 1 & , & \text{om } x \geq 1 \end{array} \right. .$$

   Man utvidgar ovanstående funktion stegvis från $\mathbb{R} \setminus T^n$ till $\mathbb{R} \setminus T^{n+1}$. Det tillkommer öppna intervall av längd $\frac{1}{3^{n+1}}$, värdet av f på ett av dem är medelvärdet av dess värden på de två grannintervallen i $\mathbb{R} \setminus T^n$. För $x \in C$ definierar man

   $$f(x) = \lim_{n \to x} f(x_n)$$

   med någon följd $(x_n)_{n \in \mathbb{N}} \subset \mathbb{R} \setminus C$.

8. En avbildning (eller funktion) $f : A \longrightarrow B$ är en regel som tillordnar varje element $x \in A$ ett element $f(x) \in B$.

9. Låt A vara en mängd. Den identiska avvbildningen

$$\mathrm{id}_A : A \longrightarrow A$$

tillordnar $x \in A$ sig själv, dvs. $\mathrm{id}_A(x) = x$. Ibland skriver vi bara id, om det är klart om vilken mängd A det handlar.

10. Avbildningar $f : A \longrightarrow B$ och $g : B \longrightarrow C$ kan sammansättas till avbildningen

$$g \circ f : A \longrightarrow C$$

med $(g \circ f)(x) := g(f(x))$.

11. En avbildning $f : A \longrightarrow B$ kallas *injektiv* om olika element $x, x' \in M$ tillordnas olika värden:

$$x \neq x' \Longrightarrow f(x) \neq f(x').$$

12. En avbildning $f : A \longrightarrow B$ kallas *surjektiv* om varje element $y \in B$ förekommer som ett funktionsvärde: $y = f(x)$ med något $x \in A$, dvs.:

$$\forall y \in B \; \exists x \in A : f(x) = y.$$

13. En avbildning $f : A \longrightarrow B$ kallas bijektiv om den är både injektiv och surjektiv. I så fall finns det en avbildning

$$f^{-1} : B \longrightarrow A$$

med

$$f^{-1} \circ f = \mathrm{id}_A, f \circ f^{-1} = \mathrm{id}_B.$$

Här kommer nu några ytterligare funderingar om mängdbegreppet:

Remark 1.4. 1. Hur kan vi skapa mängder? Om a är något objekt, så kan vi naturligtvis förpacka det i ett paket: Vi får mängden

$$\{a\} := \{x; x = a\},$$

som har precis ett element, objektet a nämligen.

2. Vilka objekt finns det i vårt universum? Mängderna själva är objekt. Då har vi den icke-tomma mängden:

$$\{\emptyset\} \neq \emptyset.$$

Är man nu snål så tar man $\emptyset$ som "urobjekt" och förpacka den på olika upprepade sätt.

3. En liten lek: Naturliga tal i denna värld::

$$0 := \emptyset,$$

samt

$$n + 1 := \{0, 1, ..., n\}.$$

Med andra ord

$$1 = \{\emptyset\}, 2 = \{\emptyset, \{\emptyset\}\}, 3 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, 4 = ...$$

Vi anmärker

$$n \leq m \Longleftrightarrow n \subset m$$

och

$$n < m \Longleftrightarrow n \in m.$$

Uppenbarligen

$$n + 1 = n \cup \{n\}.$$

Sedan definierar vi: $\mathbb{N}$ är den minsta mängden, sådant att

(a) $\emptyset \in \mathbb{N}$,

(b) $n \in \mathbb{N} \Longrightarrow n \cup \{n\} \in \mathbb{N}$.

Frågan är nu, varför det finns någon sådan mängd. Jo, sätt helt enkelt

$$\mathbb{N} := \bigcap_{A \text{ mängd som uppfyller (a) och (b)}} A\ .$$

Vi ska inte använda den här mängdteoretiska tolkningen av naturliga tal, vi har bara nämnt den för rolighetens skull.

4. Till sist blir vi lite galna och sammanfattar alla mängder i "allmängden"

$$A := \{M; M \text{mängd}\}.$$

Då gäller i synnerhet $A \in A$. Fast det verkar lite besynnerligt, måste det också vara farligt? Jo, i alla fall blir det riktigt farligt, om man tittar på följande delmängd $B \subset A$, som har som element alla "harmlösa mängder":

$$B := \{M \in A; M \notin M\}.$$

Anta nu $B \notin B$. Då får vi $B \in B$, medan $B \in B$ leder till $B \notin B$, med andra ord, det blir motsägelse! (**Russellska antinomin**, Bertrand Russell (1872 - 1970)). Problemet här är självreferensen: Man har inte än skapat mängden och frågar redan då vilka egenskaper den har. Men denna tidsaspekt är ju ingenting som spelar någon roll i matematiken. I stället måste man för att undvika sådana motsägelser komma fram till några begränsningar i skapandet av mängder.

5. **ZFC-axiomsystemet** är det allmänt accepterade regelsystemet för beteendet och skapandet av mängder (Ernst **Z**ermelo (1871 - 1953), Adolf Abraham Halevi **F**raenkel (1891 - 1965), axiom of **C**hoice), det handlar om ett mängduniversum, där alla objekt själva är mängder.

6. För ZFC-mängderna gäller alltid $M \notin M$: **Funderingsaxiomet** (som ingår i ZFC) säger att varje mängd A har ett element $B \in A$, sådant att $B \cap A = \emptyset$. Funderingsaxiomet förbjuder nämligen att det finns oändliga följder $(B_n)_{n \in \mathbb{N}}$ av mängder B_n med

$$B_n \ni B_{n+1}$$

för alla $n \in \mathbb{N}$. Om så är fallet tittar vi på mängden

$$A := \{B_n, n \in \mathbb{N}\}$$

och ser att $B_{n+1} \in B_n \cap A$ för alla $n \in \mathbb{N}$. I synnerhet förbjuds alltså $M \in M$.

Här gör vi oss inte besväret att diskutera alla ZFC-axiom, utan förblir naiva i förhoppningen att det inte går snett.

2 Kardinaltal

Målet med det första avsnittet är att berätta lite grann om oändliga *kardinaltal*. Ändliga kardinaltal är helt enkelt naturliga tal, som används för att räkna antalet element i en ändlig mängd. Men begreppet kan generaliseras så det passar för oändliga mängder också. I själva verket säger vi inte explicit, vad ett oändligt kardinaltal är för någonting, utan bara hur de hanteras. Men så ska vi också göra med reella tal i femte avsnittet!

Definition 2.1. Vi säger att två mängder M, N har samma *kardinalitet* om det finns en bijektion

$$f : M \longrightarrow N.$$

I så fall skriver vi också $M \sim N$.

Remark 2.2. Relationen $\sim$ har följande egenskaper: Den är

1. *reflexiv*: $M \sim M$. Ta $f = \mathrm{id}_M$
2. *symmetrisk*: $M \sim N$ om och endast om $N \sim M$. Om $f : M \longrightarrow N$ är bijektiv, så är den inversa avbildningen $f^{-1} : N \longrightarrow M$ också bijektiv.
3. *transitiv*: $M \sim N, N \sim O$ innebär $M \sim O$. Om $f : M \longrightarrow N$ och $g : N \longrightarrow O$ är bijektiva, blir sammansättningen $g \circ f : M \longrightarrow O$ också en bijektion.

Definition 2.3. Låt M vara en mängd. Kardinaltalet $|M|$ är ett mått för storleken av mängden M, vi har

$$|M| = n \in \mathbb{N}$$

för en ändlig mängd M med n element. Annars förväntar vi oss av våra kardinaltal

$$|M| = |N| \text{ om och endast om } M \sim N.$$

Kardinaltalens storlek jämförs så här: Vi skriver

$$|M| \leq |N|,$$

om och endast om det finns en injektiv avbildning $\iota : M \hookrightarrow N$. (Krokpilen $\hookrightarrow$ används ibland för att betona att en avbildning är injektiv.)

Remark 2.4. Anta att injektionen $\iota : M \longrightarrow N$ inte är surjektiv. För ändliga mängder innebär detta $|M| < |N|$. För oändliga mängder är det annorlunda: Ta $M = \mathbb{N} = N$ och $\iota : \mathbb{N} \hookrightarrow \mathbb{N}, x \mapsto x + 1$.

Proposition 2.5. *Låt $f : M \twoheadrightarrow N$ vara en surjektiv avbildning. Sedan har vi $|M| \geq |N|$.*

Proof. Plocka från varje *fiber*

$$F_y := \{x \in M; f(x) = y\},\ y \in N,$$

ett element och samla dem ihop i en mängd $L \subset M$. Då är $f|_L : L \longrightarrow N$ en bijektion och således $|N| = |L| \leq |M|$. □

Att det går bra med framplockningen krävs det i urvalsaxiomet:

Urvalsaxiom: Givna icke-tomma parvis disjunkta mängder $A_i, i \in I$, dvs. $A_i \cap A_j = \emptyset$ for $i, j \in I, j \neq i$, finns det en mängd B med

$$|B \cap A_i| = 1$$

för alla $i \in I$.

Definition 2.6. En mängd M kallas *räknebar* om M är ändlig eller

$$|M| = \omega := |\mathbb{N}|.$$

Här används notationen ":=" för att betona att det inte handlar om ett påstående, utan att vänstra ledet, i det här fallet ω, införs som ett nytt namn på högra ledet, de naturliga talens kardinalitet.

Remark 2.7. En oändlig mängd M är räknebar om dess element kan radas upp i en följd:

$$M = \{a_0, a_1, a_2, ...\}.$$

Nämligen om $f : \mathbb{N} \xrightarrow{\cong} M$ är bijektiv, så sätt

$$a_n := f(n).$$

Å andra sidan ger en följd $(a_n)_{n \in \mathbb{N}}$ upphov till avbildningen

$$f : \mathbb{N} \longrightarrow M, n \mapsto a_n.$$

Example 2.8. 1. $\mathbb{N}$ är räknebar.

2. $\mathbb{Z}$ är räknebar:

$$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, 4, ...\}.$$

3. $\mathbb{N}^2$ är räknebar:

$$\mathbb{N}^2 = \{(0,0), (1,0), (0,1), (0,2), (1,1), (2,0), (3,0), (2,1), (1,2), ...\}.$$

Rita en skiss genom att förbinda successiva punkter i planets första kvadrant!

Proposition 2.9. *Vi har $\omega \leq \aleph$ för alla oändliga kardinaltal $\aleph$.*

Proof. Vi visar att varje oändlig mängd M har en oändlig räknebar delmängd. Induktivt skapar vi en följd $(a_n)_{n\in\mathbb{N}} \subset M$ med parvis olika element och sätter $L := \{a_n; n \in \mathbb{N}\}$. Ta något som helst element $a_0 \in M$. Om $a_0, ..., a_n$ är hittade, plocka $a_{n+1} \in M \setminus \{a_0, ..., a_n\} \neq \emptyset$. □

Nu är det dags för nya mängder:

Definition 2.10. Låt M, B vara mängder.

1. Pontensmängden $\mathcal{P}(M)$ är

$$\mathcal{P}(M) := \{A; A \subset M\},$$

mängden av alla delmängder till M.

2. Med

$$M^B := \{f; f : B \longrightarrow M\}$$

betecknas mängden av alla avbildningar $f : B \longrightarrow M$.

Example 2.11. 1. $\mathcal{P}(\emptyset) = \{\emptyset\}$.

2. $\mathcal{P}(\{0\}) = \{\emptyset, \{0\}\}$.

3. $\mathcal{P}(\{0,1\}) = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$.

4. Om M är ändlig, så har vi

$$|\{0,1\}^M| = 2^{|M|}.$$

5. Mängden $M^{\{1,...,n\}}$ kan uppfattas som mängden M^n av alla n-tipplar $(a_1, ..., a_n)$ med komponenter i M.

6. Mängden $M^{\mathbb{N}}$ kan uppfattas som mängden av alla följder $(a_n)_{n\in\mathbb{N}}$ med element $a_n \in M$.

Definition 2.12. För ett kardinaltal $|M|$ definierarar vi

$$2^{|M|} := |\{0,1\}^M|.$$

Proposition 2.13. *Vi har*

$$|\mathcal{P}(M)| = 2^{|M|}.$$

Proof. Vi skapar en bijektion

$$\Phi : \{0,1\}^M \longrightarrow \mathcal{P}(M).$$

Nämligen

$$\Phi(f) := E_f := \{x \in M; f(x) = 1\}.$$

Den tillhörande inversa avbildningen är

$$\Psi : \mathcal{P}(M) \longrightarrow \{0,1\}^M,$$

med $\Psi(A) := \chi_A$, där $\chi_A : M \longrightarrow \{0,1\}$ är den *karaktäristiska funktionen* till delmängden $A \subset M$, dvs.

$$\chi_A(x) := \begin{cases} 1 & , \quad \text{if } x \in A \\ 0 & , \quad \text{if } x \notin A \end{cases} \quad .$$

□

Proposition 2.14. *Vi har*

$$2^\omega > \omega.$$

Proof. Pga. Prop.2.9 räcker det att visa att $\{0,1\}^{\mathbb{N}}$ intev är räknebar. Anta

$$\{0,1\}^{\mathbb{N}} = \{f_0, f_1, f_2, ...\}$$

med funktionerna $f_n : \mathbb{N} \longrightarrow \{0,1\}$. Sedan är funktionen $g : \mathbb{N} \longrightarrow \{0,1\}$ med

$$g(n) := 1 - f_n(n)$$

inte med i vår uppräkning! Motsägelese! □

Olikheten $\aleph < 2^{\aleph}$ för $\aleph = n, n \in \mathbb{N}$, och $\aleph = \omega$ kan generaliseras:

Proposition 2.15. *För alla kardinaltal $\aleph$ har vi*

$$\aleph < 2^{\aleph}.$$

Proof. För en mängd M måste vi visa

$$|M| < |\mathcal{P}(M)|.$$

Vi har $|M| \leq |\mathcal{P}(M)|$, eftersom

$$M \longrightarrow \mathcal{P}(M), x \mapsto \{x\},$$

är en injektiv avbildning. Anta att det finns en bijektion

$$F : M \longrightarrow \mathcal{P}(M).$$

Vi använder oss av samma argument som i Russellska antionomin: Titta på mängden

$$A := \{x \in M; x \notin F(x)\}.$$

Men F är ju en bijektion, således

$$A = F(y)$$

med något $y \in M$. Nu frågar vi om $y \in A$ eller ej:

1. Om $y \in A = F(y)$, så hittar vi $y \notin A$ enligt definition av mängden A, så det går alltså inte.

2. Om $y \notin A = F(y)$, så hittar vi $y \in A$, så det går inte heller.

Med andra ord, vårt antagande, existensen av en bijektiv avbildning $F : M \longrightarrow \mathcal{P}(M)$, var fel. $\square$

Corollary 2.16. *Det finns oändligt många oändliga kardinaltal.*

Vi har definierat relationen "$\leq$" för kardinaltal och hoppas att den har samma egenskaper som den har när man jämför naturliga tal (= ändliga kardinaltal). Här är ett första steg åt det hållet:

Proposition 2.17. *Låt M, N vara mängder. Olikheterna*

$$|M| \leq |N| \leq |M|$$

innebär

$$|M| = |N|.$$

Proof. Let $\iota : M \hookrightarrow N$ and $j : N \hookrightarrow M$ vara injektioner. Med $A := j(i(M)), B := j(N) \setminus A, C := M \setminus j(N)$ får vi en disjunkt uppdelning

$$M = A \cup B \cup C.$$

Och så använder vi nedanstående lemma med $f = j \circ i$. $\square$

Lemma 2.18. *Anta att mängden M är en disjunkt union $M = A \cup B \cup C$ och att det finns en bijektion*

$$f : M \longrightarrow A.$$

Sedan finns det också en bijektiv avbildning

$$h : M \longrightarrow A \cup B.$$

Proof. Vi rekommenderar läsaren att rita en skiss med

$$A = [-1, 2], B = (2, 3], C = (3, 4]$$

samt

$$f(x) = \begin{cases} \frac{x}{2}, & , \quad \text{if } x \geq 0 \\ x, & , \quad \text{if } x < 0 \end{cases} ,$$

för att hänga med i följande argumentation:

Vi har en avstigande följd

$$A_0 = A \supset A_1 := f(A) \supset A_2 := f^2(A) \supset ....$$

och med $B_n := f^n(B), C_n := f^n(C)$ får vi den disjunkta uppdelningen

$$A_n = A_{n+1} \cup B_{n+1} \cup C_{n+1}.$$

Vi tar nu

$$D := \bigcap_{n=0}^{\infty} f^n(A)$$

och får en oändlig disjunkt uppdelning

$$M = D \cup \bigcup_{n=0}^{\infty} (B_n \cup C_n).$$

Vi definierar nu bijektionen

$$h : M \longrightarrow A \cup B = D \cup \bigcup_{n=0}^{\infty} (B_n \cup C_{n+1})$$

genom

$$h|_D := \mathrm{id}_D, h|_{B_n} := \mathrm{id}_{B_n},$$

medan

$$h|_{C_n} := f|_{C_n}.$$

$\square$

Corollary 2.19. *Vi har*

$$|\mathbb{R}| = 2^\omega.$$

Proof.

1. "$2^\omega \leq |\mathbb{R}|$": Avbildningen

   $$\{0,1\}^{\mathbb{N}} \longrightarrow \mathbb{R}, (a_n)_{n\in\mathbb{N}} \mapsto \sum_{n=0}^{\infty} \frac{a_n}{10^n},$$

   är injektiv.

2. "$2^\omega \geq |\mathbb{R}|$": Avbildningen

   $$\{0,1\}^{\mathbb{N}} \longrightarrow [0,2], (a_n)_{n\in\mathbb{N}} \mapsto \sum_{n=0}^{\infty} \frac{a_n}{2^n},$$

   är surjektiv. Således

   $$2^\omega \geq |[0,2]| \geq |(0,2)| = |\mathbb{R}|,$$

   eftersom

   $$(0,2) \longrightarrow \mathbb{R}, x \mapsto \frac{1-x}{x(x-2)}$$

   är en bijektion.

$\square$

Andra steget i vårt program blir:

Proposition 2.20. *Kardinaltalen till två mängder M, N kan jämföras: Det gäller*

$$|M| \leq |N| \text{ eller } |M| \geq |N|.$$

Problemet man har är att man inte har riktigt koll på vilka avbildningar som finns mellan två mängder. Knepet är att man inte betraktar alla avbildningar, utan inför en "ordningsstruktur" på våra mängder och sedan kräver att avbildningarna bevarar ordningstrukurerna. Vi behöver några nya begrepp:

Definition 2.21. En **linjär ordningsralation** på en mängd M är en relation "$\leq$", som är

1. *reflexiv*: För alla $x \in M$ har vi $x \leq x$,
2. *antisymmetrisk*: För alla $x, y \in M$ medför olikheterna $x \leq y \leq x$ likhet: $x = y$,
3. *transitiv* : För alla $x, y, z \in M$ medför olikheterna $x \leq y \leq z$ olikheten $x \leq z$
4. *linjär* (eller total): Två godtyckliga element $x, y \in M$ kan jämföras: Vi har $x \leq y$ eller $y \leq x$.

En **välordning** på M är en linjär ordning, sådant att varje icke-tom delmängd $M_0 \subset M$ har ett första element $a \in M_0$, dvs. sådant att $a \leq x$ gäller för alla $x \in M_0$.

Example 2.22. 1. $\mathbb{N}$ med den naturliga ordningsrelationen är välordnad.

2. $\mathbb{R}$ med den naturliga ordningsrelationen är **inte** välordnad. Själva reella linjen har inte något första element, men inte heller har det öppna enhetsintervallet $(0, 1) \subset \mathbb{R}$ det - vi har ju $0 \notin (0, 1)$.
3. $\mathbb{N}^2$ med den lexikografiska ordningen är välordnad: Den lexikografiska ordningen fungerar så här: Vi skriver

   $$(a, b) \leq (c, d),$$

   om $a < c$ eller $a = c$ samt $b \leq d$.

Definition 2.23. Låt M vara välordnad. Givet ett element $a \in M$ sådant att

$$M_{>a} := \{x \in M; x > a\} \neq \emptyset$$

definierar vi dess omedelbara efterträdare $\sigma(a) \in M$ (successor) som det första elementet i $M_{>a}$, dvs.

$$M_{>a} = M_{\geq \sigma(a)}.$$

Remark 2.24. Ett element $a \in M$ har inte nödvändigtvis en omedelbar företrädare, dvs. ett element $b \in M$, sådant att $a = \sigma(b)$, även om det inte är M's första element. T.ex. ta $M = \mathbb{N}^2$ med den lexikografiska ordningen och $a = (1, 0)$. Däremot gäller $\sigma(a) = (1, 1)$.

Definition 2.25. En delmängd $M_0 \subset M$ till en välordnad mängd kallas ett *initialsegment* omm $x \leq b \in M_0 \Longrightarrow x \in M_0$, dvs. med ett element $b \in M_0$ ligger också alla dess företrädare i M_0.

Remark 2.26. För ett initialsegment $M_0 \subset M$ till en välordnad mängd gäller:

1. antingen $M_0 = M$
2. eller
$$M_0 = M_{<a} := \{x \in M; x < a\}$$
med något $a \in M$, nämligen det första elementet i $M \setminus M_0$.

Theorem 2.27 (Välordningssats). *På varje mängd M finns det en välordning.*

Vi ska först visa Prop.3.10: Vi behöver:

Lemma 2.28. *Låt $M_0 \subset M, N_0 \subset N$ vara initialsegment till de välordnade mängderna M, N. Anta vi har ordningsbevarande abildningar f, g som i diagrammet*

$$\begin{array}{ccc} M & \stackrel{f}{\hookrightarrow} & N \\ \cup & & \cup \\ M_0 & \stackrel{g}{\longrightarrow} & N_0 \end{array},$$

där g är bijektiv och f injektiv med $f(M) \subset N$ ett initialsegment. Sedan har vi $g = f|_{M_0}$.

Proof. För att börja med så har vi

$$f(a) = b = g(a),$$

där a resp. b är första elementet i M resp. N. Ytterligare

$$f(M_{<c}) = N_{<f(c)}$$

samt

$$g(M_{<c}) = N_{<g(c)}$$

för $c \in M_0$. Om inte $f = g$ på M_0 ta $c \in M_0$ som det första elementet, där f, g inte överensstämmer. Sedan har vi

$$N_{<g(c)} = g(M_{<c}) = f(M_{<c}) = N_{<f(c)}$$

och således $g(c) = f(c)$. Motsägelse! □

Prop.3.10 följer nu från välordningssatsen och:

Lemma 2.29. *Låt M, N vara välordnade mängder. Sedan finns det*

1. *ett initialsegment $M_0 \subset M$ tillsammans med en ordningsbevarande bijektion $M_0 \longrightarrow N$*

2. *eller ett initialsegment $N_0 \subset N$ tillsammans med en ordningsbevarande bijektion $M \longrightarrow N_0$.*

Proof. Vi kan anta att både M och N är välordnade. Vi tittar på par (M_1, N_1), där $M_1 \subset M$ och $N_1 \subset N$ är initialsegment och det finns en (entydig) ordningsbevarande bijektion $f_1 : M_1 \longrightarrow N_1$. Anta nu att (M_2, N_2) är ett sådant par till. Vi får anta $M_1 \subset M_2$. Enligt Lemma 2.28 har vi sedan $f_2|_{M_1} = f_1$. Således kan vi ta unionen över alla M_1 och N_1 och får ett maximalt par $(M_{\max}, N_{\max})$ tillsammans med $f_{\max} : M_{\max} \longrightarrow N_{\max}$. Pga. maximaliteten måste nu $M_{\max} = M$ eller $N_{\max} = N$. □

Bevis av välordningssatsen. Vi vill skapa en välordning på M genom successiv framplockning. Den styrs av en funktion som i nedanstående Proposition:

Proposition 2.30. *Given en funktion*

$$\gamma : \mathcal{P}(M) \setminus \{M\} \longrightarrow M$$

sådant att $\gamma(A) \in M \setminus A$ för alla $A \subsetneqq M$, finns det en entydig välordning på mängden M, som uppfyller

$$\gamma(M_{<a}) = a$$

för alla $a \in M$, i synnerhet gäller alltså

$$\sigma(a) = \gamma(M_{\leq a}).$$

(Man har ju $M_{\leq a} = M_{<\sigma(a)}$.)

Funktionen γ finns pga.

Lemma 2.31. *Given en mängd M finns det en funktion*

$$\gamma : \mathcal{P}(M) \setminus \{M\} \longrightarrow M$$

sådant att $\gamma(A) \in M \setminus A$ för alla $A \subsetneqq M$.

Proof. Skriv

$$\mathcal{P}(M) \setminus \{\emptyset\} = \{A_i; i \in I\},$$

sådant att $A_i \neq A_j$ för $i \neq j$. Mängderna

$$A_i \times \{i\}, i \in I,$$

är då disjunkta och urvalsaxiomet ger oss en mängd B, som innehåller precis ett element (a_i, i) ur varje mängd $A_i \times \{i\}$. Definiera

$$\varphi : \mathcal{P}(M) \setminus \{\emptyset\} \longrightarrow M, A_i \mapsto a_i,$$

och sätt

$$\gamma(A) := \varphi(M \setminus A).$$

□

Vi tittar på potentiella initialsegment:

Definition 2.32. Vi kallar $K \subset M$ en γ-kedja omm K har en välordning $\leq$ med $\gamma(K_{<a}) = a$ för $a \in K$.

Example 2.33. Om M har oändligt många element, är

$$K = \{a_n; n \in \mathbb{N}\},$$

där $a_0 := \gamma(\emptyset), a_{n+1} = \gamma(\{a_0, ..., a_n\})$, en γ-kedja.

Man visar nu följande:

Lemma 2.34. *Om K, L är γ-kedjor, så gäller antigen $K = L$ eller $K = L_{<b}$ med något $b \in L$ eller $L = K_{<a}$ med något $a \in K$, och ordningsrelationerna överensstämmer.*

Sedan är

$$M_0 := \bigcup_{K\ \gamma-\text{kedja}} K$$

en maximal γ-kedja. Men då måste vi redan ha $M_0 = M$, eftersom annars vore ju $M_1 := M_0 \cup \{\gamma(M_0)\}$ en större γ-kedja. □

Bevis till lemmat 2.34. Vi tittar på alla initialsegment $K_0 \subset K$, som samtidigt är initialsegment $L_0 \subset L$, dvs. $K_0 = L_0$, och sådant att ordningsrelationerna på K_0 och L_0 överensstämmer. Deras union U har då samma egenskap. Vi måste utesluta att $U = K_{<a} = L_{<b}$. Men i detta fall vore ju $a = \gamma(U) = b$ och

$$V := U \cup \{\gamma(U)\}$$

ett större initialsegment till både K och L, motsägelse. □

Men inte bara kan varje mängd välordnas, utan vi har även som följdsats:

Proposition 2.35. *Kardinaltalen är välordnade.*

Remark 2.36. Helheten av alla kardinaltal utgör inte någon mängd, utan bara en *klass*: Klasser behöver inte leva upp till alla ZFC-axiom.

Proof. Låt K vara en icke-tom mängd av kardinaltal. Plocka fram $\aleph = |M| \in K$. Om $\aleph' \geq \aleph$ för alla $\aleph' \in K$ behöver vi inte visa någonting.

Annars tar vi fram en välordning "$\leq$" på M. Vi tittar på den icke-tomma delmängden

$$M_0 := \{a \in M; |M_{<a}| \in K\}.$$

Om $b \in M_0$ är dess första element, så är $|M_{<b}|$ det första kardinaltalet i K. □

Kardinaltalens välordning kan nu förtydligas:

$$0 < 1 < 2 < 3 < .... < \aleph_0 = \omega < \aleph_1 < ....$$

där $\aleph_{i+1} := \sigma(\aleph_i)$ är den omedelbara efterträdaren till $\aleph_i$. Man kan nu fråga sig, var 2^ω finns i den här följden:

$$\aleph_1 \stackrel{?}{=} 2^\omega,$$

kontinuumshypotesen, eller om kanske t.o.m. gäller

$$\sigma(\aleph) \stackrel{?}{=} 2^\aleph$$

för varje oändligt kardinaltal $\aleph$, den generaliserade kontinuumshypotesen.
Lite historik:

1. Första Hilbertska problemet, 1900: Gäller kontinuumshypotesen eller ej?

   Kontinuumshypotesen: Om $A \subset \mathbb{R}$ är oändlig, så finns det

   (a) antingen en bijektion $\mathbb{N} \longrightarrow A$

   (b) eller en bijektion $\mathbb{R} \longrightarrow A$.

   ("Det finns inga mellanmängder i $\mathbb{R}$".)

2. Kurt Gödel, 1938: KH går ihop med ZFC.

3. Paul J. Cohen: 1965: $2^\omega > \sigma(\omega)$ strider inte mot ZFC. (Fieldsmedalj 1966, ≤ 40).

Några funderingar:

1. $2^n \gg n$ för $n \in \mathbb{N}$. Varför skulle då gälla $2^\omega = \sigma(\omega)$?

2. Sannolikt finns det förskräckligt många delmängder till $\mathbb{R}$, men med våra verktyg förbiser man dem.

3. Gödel skapar ett snålt mängdsuniversum (konstruerbara mängder).

4. Cohen kan inte peka på en explicit mellanmängd, utan han skapar en sofistikerad modell för ZFC inkl. $2^\omega > \sigma(\omega)$ med ett försvagat sanningsbegrepp. Men det räcker för att visa att $2^\omega > \sigma(\omega)$ inte strider mot ZFC.

5. ZFC är inte ett fullständigt axiomsystem, man kan lägga till KH eller Icke-KH och det stannar motsägelsefritt, om ZFC var motsägelsefritt.

6. Finns det näraliggande kompletteringar till ZFC? Hittills har ingen hittat några sådana.

3 Reella tal

Första analyskursen bygger på ett intuitivt förståelse av reella tal. I själva verket skulle man väl svara på frågan vad ett reellt tal egentligen är för någonting så här: Det är ett ändligt eller oändligt decimalbråk med tecken. Nästa sak att göra är då att härleda alla räkneregler för reella tal man är van vid! Men pga. att decimalutvecklingen inte är helt entydigt:

$$1 = 0,99999....,$$

trasslar man sig snabbt in i ett ogenomtrångligt snår av tekniska problem.

Då kan man försöka att kringgå problemen genom att formulera sina förväntningar som axiom: Det är det första vi gör i det här avsnittet. Sedan konstrueras en modell som uppfyller våra axiom: Det man gör är att man identifierar ett positivt reellt tal med mängden av alla mindre positiva rationella tal.

Låt oss börja med axiomen: Reella talen kommer med två räkneoperationer:

1. Den första kallas för additionen och skrivs så här

$$\mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}, (x, y) \mapsto x + y,$$

2. den andra för multiplikationen

$$\mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}, (x, y) \mapsto x \cdot y = xy,$$

3. och ett "positivitetsområde" $\mathbb{R}^+ \subset \mathbb{R}$.

Axiomen handlar nu om deras egenskaper.
Följande vill vi ha:

1. de kommutativa lagarna

$$x + y = y + x, xy = yx,$$

2. de associativa lagarna

$$(x + y) + z = x + (y + z),\ (xy)z = x(yz),$$

3. den distributiva lagen
$$x(y+z) = xy + xz,$$
4. det finns ett element $0 \in \mathbb{R}$, sådant att
$$x + 0 = 0$$
för alla $x \in \mathbb{R}$,
5. det finns ett element $1 \in \mathbb{R} \setminus \{0\}$, sådant att
$$1 \cdot x = x$$
för alla $x \in \mathbb{R}$,
6. för varje $x \in \mathbb{R}$ finns det ett element $-x \in \mathbb{R}$ med
$$x + (-x) = 0,$$
7. för varje $x \in \mathbb{R} \setminus \{0\}$ finns det ett element $x^{-1} \in \mathbb{R}$ med
$$x \cdot x^{-1} = 1,$$
8. vi har en disjunkt uppdelning
$$\mathbb{R} = \mathbb{R}^+ \cup \{0\} \cup (-\mathbb{R}^+),$$
9. om $x, y \in \mathbb{R}^+$, så gäller också $x + y, xy \in \mathbb{R}^+$,

Remark 3.1. En mängd K med en addition och en multiplikation som uppfyller de första sju axiomen kallas en *kropp*, om axiom 8 och 9 tillkommer så pratar man om en *ordnad kropp*.

Example 3.2. En kropp för snåla: Ta $K := \{0, 1\}$ med additionstabellen

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad ,$$

och multiplikationstabellen

$$\begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array} \qquad .$$

Lägg märke till att K inte är en ordnad kropp: Om så vore fallet, hade vi nödvändigtvis $K^+ = \{1\}$ och sedan $0 = 1 + 1 \in K^+$, motsägelse!

Innan vi presenterar det sista axiomet samlar vi ihop några räkneregler:

1. Elementen $0, 1$ är entydiga och kallas för nollan resp. ettan. Om nämligen $\tilde{0}$ resp. $\tilde{1}$ skulle ha samma egenskap, fick vi

$$\tilde{0} = \tilde{0} + 0 = 0 + \tilde{0} = 0$$

och

$$\tilde{1} = \tilde{1} \cdot 1 = 1.$$

2. Elementet $-a \in \mathbb{R}$ är entydigt bestämt: Om b är ett element med samma egenskap, så har vi

$$b = 0 + b = (a + (-a)) + b = (a + b) + (-a) = 0 + (-a) = -a.$$

3. Nollan uppfyller

$$0 \cdot a = 0 = a \cdot 0$$

för alla $a \in \mathbb{R}$: Vi har $a = 1 \cdot a = (1 + 0)a = a + 0 \cdot a$ och adderar $-a$ till båda led.

4. Vi har $(-1)a = -a$ pga.

$$0 = (1 + (-1)) \cdot a = a + (-1)a,$$

i synnerhet med $a = -1$ blir det

$$(-1)^2 = (-1) \cdot (-1) = -(-1) = 1.$$

5. Elementet x^{-1} i axiom 7 är entydigt bestämt. Vi skriver också $\frac{1}{x} := x^{-1}$.

6. Vi har $1 \in \mathbb{R}^+$. Annars hade vi $-1 \in \mathbb{R}^+$ och får då $1 = (-1)^2 \in \mathbb{R}^+$, motsägelse.

7. Om vi definierar $x \leq y$ genom $y - x \in \mathbb{R}^+ \cup \{0\}$, blir $\leq$ en linjär ordningsrelation på $\mathbb{R}$, som uppfyller de sedvanliga räknereglerna. Vi visar $0 < x < y \Longrightarrow x^2 < y^2$. Nämligen:

$$x < y \Longleftrightarrow y - x \in \mathbb{R}^+ \Longrightarrow (y - x)(y + x) \in \mathbb{R}^+$$

$$\Longleftrightarrow y^2 - x^2 \in \mathbb{R}^+ \Longleftrightarrow x^2 < y^2.$$

Det sista axiomet, fullständighetsaxiomet, gör att "allt" är entydigt bestämt: Vi behöver:

Definition 3.3. En mängd $M \subset \mathbb{R}$ kallas begränsad uppåt, om det finns ett tal $a \in \mathbb{R}$ med $M \leq a$, dvs. $x \leq a$ gäller för alla $x \in M$. Talet a kallas då en övre gräns till M.

Fullständighetsaxiom: En uppåt begränsad mängd $M \subset \mathbb{R}$ har en minsta övre gräns $b \in \mathbb{R}$, dvs.

1. $M \leq b$ och
2. $M \leq a \Longrightarrow b \leq a$.

Vi skriver också

$$b = \sup M,$$

och kallar talet b för mängden M's *supremum*.

Example 3.4. Mängden

$$M := \{x \in \mathbb{Q}; x^2 \leq 2\}$$

har inte någon minsta övre gräns i $\mathbb{Q}$.

Remark 3.5. Det finns ett motstycke till den minsta övre gränsen, den största lägre gränsen eller infimum, som existerar för varje delmängd $M \subset \mathbb{R}$ som har en lägre gräns -M är begränsad nedåt. Genom att gå över till $-M \subset \mathbb{R}$ kan vi beskriva den:

$$\inf M = -\sup(-M).$$

Example 3.6. Egenskaper av $\mathbb{R}$:

1. Vi har $\mathbb{N} \subset \mathbb{R}$. Ett naturligt tal $n \in \mathbb{N}$ identifieras med den n-faldiga summan av $1 \in \mathbb{R}$. Vi har $n \neq 0$ pga. $n \in \mathbb{R}^+$.
2. $\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N})$.
3. $\mathbb{Q} \subset \mathbb{R}$. Rationella tal är reella tal pq^{-1} med $p, q \in \mathbb{Z}, q \neq 0$.

4. Det arkimediska ”axiomet” (ett klassiskt namn, fast här är det inte något axiom, utan en följdsats.): För alla $x \in \mathbb{R}$ finns det $n \in \mathbb{N}$ med $n \geq x$. Annars är ju talet x en övre gräns för $\mathbb{N}$, och så finns det en minsta övre gräns $b \geq \mathbb{N}$. Eftersom $b-1$ inte är en övre gräns, hittar vi $m \in \mathbb{N}$ med $m > b-1$, således $m+1 > b$. Motsägelse!

5. Om $x > 0$, så finnes det $n \in \mathbb{N}$ med $\frac{1}{n} < x$.

6. En kvadratrot ur 2: Ta
$$M := \{x \in \mathbb{R}_{>0}; x^2 \leq 2\}.$$
Sedan:
   (a) $1 \in M$, således $M \neq \emptyset$.
   (b) $M \leq 2$, eftersom $x > 2 \Longrightarrow x^2 > 4 > 2$, alltså $x \notin M$.

   Vi påstår nu: $b^2 = 2$ för $b := \sup M$. Vi måste visa
   (a) $b^2 \leq 2$
   (b) $2 \leq b^2$.

   (a) För $x = b + \frac{1}{n} \notin M$ har vi
   $$2 < x^2 = b^2 + \frac{1}{n}\left(2b + \frac{1}{n}\right) < b^2 + \frac{2b+1}{n},$$
   alltså
   $$2 - b^2 < \frac{2b+1}{n}$$
   för alla $n \in \mathbb{N}$ resp.
   $$2 - b^2 \leq 0.$$
   (b) Å andra sidan ta $x = b - \frac{1}{n}$. Det finns $y \in M$ med $x < y < b$, Således
   $$2 \geq y^2 \geq x^2 = b^2 - \frac{1}{n}\left(2b - \frac{1}{n}\right) > b^2 - \frac{2b}{n},$$
   alltså
   $$2 - b^2 > -\frac{2b}{n}$$
   för alla $n \in \mathbb{N}$ resp.
   $$2 - b^2 \geq 0.$$

7. Låt $n \in \mathbb{N}_{>0}$. För varje $a > 0$ finns det ett entydigt bestämt tal $b > 0$ med $a = b^n$. Beviset fungerar som i fallet $n = 2$. Använd

$$(y^n - x^n) = (y - x)(y^{n-1} + y^{n-2}x + ... + yx^{n-2} + x^{n-1}).$$

Example 3.7. Låt

$$\mathbb{R}(X) := \left\{ f(X) = \frac{a_k X^k + ...a_1 X + a_0}{b_\ell X^\ell + ... + b_1 X + b_0}; a_i, b_j \in \mathbb{R}, b_\ell \neq 0 \right\}$$

vara kroppen av alla "rationella funktionsuttryck" med positivitetsområdet

$$\mathbb{R}(X)^+ := \left\{ f(X) = \frac{a_k X^k + ...}{b_\ell X^\ell + ...}; a_i, b_j \in \mathbb{R}, a_k, b_\ell > 0 \right\}.$$

Varje $f(X)$ definierar en funktion $f : \mathbb{R} \backslash P_f \longrightarrow \mathbb{R}$, där P_f är polställemängden till $f(X)$, och

$$f \in \mathbb{R}(X)^+ \Longleftrightarrow f(x) > 0 \text{ för } x \gg 0$$

Sedan har vi $X > n$ för alla $n \in \mathbb{N}$, dvs. det arkimediska axiomet gäller inte!

Definition 3.8. 1. En följd $(a_n)_{n \in \mathbb{N}}$ kallas begränsad, om det finns $c, d \in \mathbb{R}$, sådant att $c \leq a_n \leq d$ gäller för alla $n \in \mathbb{N}$.

2. En delföljd till $(a_n)_{n \in \mathbb{N}}$ är en följd $(a_n(k))_{k \in \mathbb{N}}$, där $k \mapsto n(k)$ är strängt växande.

Theorem 3.9 (Sats av Bolzano-Weierstraß)**.** *En begränsad följd $(a_n)_{n \in \mathbb{N}}$ av reella tal har en konvergent delföljd $(a_{n(k)})_{k \in \mathbb{N}}$.*

Proof. Mängden

$$M := \{a \in \mathbb{R}; a \leq a_n \text{ för oändligt många } n \in \mathbb{N}\}$$

är icke-tom pga. $c \in M$ och har d som en övre gräns, ta $b := \sup M$. Vi bestämmer $n(k)$ induktivt:

1. Ta $n(0) = 0$.

2. Anta $n(k)$ är hittad. Sedan finns det $a \in M$ med

$$b - \frac{1}{k+1} \leq a \leq b,$$

således ett index $n > n(k)$ med

$$b - \frac{1}{k+1} \leq a_n \leq b + \frac{1}{k+1},$$

eftersom $b + \frac{1}{k+1} \leq a_n$ bara gäller för ändligt många $n \in \mathbb{N}$. Sätt $n(k+1) := n$.

Uppenbarligen gäller

$$b - \frac{1}{k} \leq a_{n(k)} \leq b + \frac{1}{k},$$

för alla $k \geq 1$, således:

$$\lim_{k\to\infty} a_{n(k)} = b.$$

□

Vårt axiomsystem för reella tal är fullständigt:

Theorem 3.10. *Låt K vara en ordnad kropp (axiom 1-9) som uppfyller det arkimediska axiomet. Sedan finns det en injektion*

$$\iota : K \hookrightarrow \mathbb{R},$$

som bevarar addition och multiplikation

$$\iota(x+y) = \iota(x) + \iota(y),\ \iota(xy) = \iota(x) \cdot \iota(y)$$

samt positivitetsområdet, dvs.

$$\iota(K^+) \subset \mathbb{R}^+.$$

Om fullständighetsaxiomet gäller för K, är ι också surjektiv.

Example 3.11. Vi tittar på delmängder $K \subset \mathbb{R}$, sådant att $0, 1 \in K$ samt $x, y \in K \Longrightarrow x+y, xy, -x, x^{-1} \in K$ och tar $\iota :=$ inklusionen $K \hookrightarrow \mathbb{R}$. T.ex.

1. $K = \mathbb{Q} \overset{\iota}{\hookrightarrow} \mathbb{R}$,
2. $K = \mathbb{Q} + \mathbb{Q}\sqrt{2} \overset{\iota}{\hookrightarrow} \mathbb{R}$.

Proof. En skiss: Vi har

1. $K^+ \supset \mathbb{N}_{>0} \subset \mathbb{R}^+$

2. och $K \supset \mathbb{Q} \subset \mathbb{R}$.

För godtyckligt $x \in K$ sätter vi

$$\iota(x) := \begin{cases} x & , \quad \text{om } x \in \mathbb{Q} \\ \sup \mathbb{Q}_{\leq x} \quad \in \mathbb{R} & , \quad \text{om } x \in K \setminus \mathbb{Q} \end{cases} .$$

Vi måste visa att ι är "väldefinierat":

1. $\mathbb{Q}_{\leq x} \neq \emptyset$: Ta $m \in \mathbb{N}$ med $-x \leq m$ resp. $-m \leq x$, således

$$-m \in \mathbb{Q}_{\leq x}.$$

2. $\mathbb{Q}_{\leq x} \subset \mathbb{R}$ är begränsad uppåt. Ta $n \in \mathbb{N} \subset K$ med $x \leq n$, således

$$\mathbb{Q}_{\leq x} \leq n \in \mathbb{R}.$$

Till sist: Om K uppfyller fullständighetsaxiomet och $y \in \mathbb{R}$, så gäller $y = \iota(x)$ med

$$x := \sup \mathbb{Q}_{\leq y} \in K.$$

$\square$

En modell för mängden $\mathbb{R}$. Vi tar inspirationen från beviset till Th.3.10 och definierar positiva reella tal som lämpliga delmängder till $\mathbb{Q}^+$.

Definition 3.12. Ett positivt reellt tal är en icke-tom delmängd

$$X \subsetneq \mathbb{Q}^+,$$

som är

1. fullständig nedåt: Med $x \in X$ tillhör alla $y \leq x$ mängden X,
2. och öppen uppåt: För alla $x \in X$ finns det $y \in X$ ovanför x.

Vi betecknar med

$$\mathbb{R}^+ := \{X \in \mathcal{P}(\mathbb{Q}^+); X \text{ positivt reellt tal}\}$$

mängden av alla positiva reella tal.

Definition 3.13. 1. Summan och produkten av två reella tal är elementvisa

$$X + Y = \{x + y; x \in X, y \in Y\},\ XY = \{xy; x \in X, y \in Y\}.$$

2. Vi definierar $X \leq Y$ genom $X \subset Y$.

Theorem 3.14. *1. De kommutativa och associativa lagarna och den distributiva lagen gäller för additionen och multiplikationen i $\mathbb{R}^+$,*

2. det finns ett element $1 \in \mathbb{R}^+$, sådant att

$$1 \cdot x = x$$

för alla $x \in \mathbb{R}^+$, nämligen $\mathbb{Q}^+_{<1}$,

3. för varje $x \in \mathbb{R}^+$ finns det ett element $x^{-1} \in \mathbb{R}^+$ med

$$x \cdot x^{-1} = 1,$$

4. och fullständighetsaxiomet gäller.

Proof. En skiss:

1. Kommutativa och associativa lagarna och distributiva lagen gäller eftersom de gäller i $\mathbb{Q}$,
2. Gjort.
3. Övning!
4. Lägsta övre gränsen (supremum) till en icke-tom begränsad mängd $M \subset \mathbb{R}^+$ har formen

$$\sup M = \bigcup_{X \in M} X \quad \in \mathbb{R}^+,$$

eftersom unionen inte är hela $\mathbb{Q}^+$: Mängden M har ju en övre gräns $Y \in \mathbb{R}^+$, dvs.

$$X \subset Y, \quad \forall\ X \in M.$$

□

Hur kommer man nu fram till hela $\mathbb{R}$? För läsbarhetens skull återgår vi till notationen med små bokstäver: $x, y, ..$ betecknar positiva reella tal. Vi tar

$$\mathbb{R} := (\mathbb{R}^+ \times \mathbb{R}^+)/\sim$$

med ekvivalensrelationen

$$(x, y) \sim (s, t) :\Longleftrightarrow x + t = y + s.$$

För att förstå denna ekvivalensrelation är det bäst att tänka sig första komponenten x som tillgånger och andra komponenten y som skulder.

Definition 3.15. Ett reellt tal är en ekvivalensklass $[x, y]$, där $x, y \in \mathbb{R}^+$. Addition och multiplikation definieras genom

$$[x, y] + [s, t] := [x + s, y + t]$$

samt

$$[x, y] \cdot [s, t] := [xs + yt, xt + ys].$$

Positivitetsområdet består av alla $[x, y]$ med $x > y$.

4 Algebrans fundamentalsats

I föregående avsnitt har vi sett att varje reellt tal $a \in \mathbb{R}_{>0}$ har en (entydig) positiv n-te rot $b \in \mathbb{R}_{>0}$, dvs. sådant att $b^n = a$. Nu letar vi lite mer allmänt efter reella lösningar till ekvationen

$$x^n - a = 0,\ \ a \in \mathbb{R}.$$

och konstaterar:

1. För $a = 0$ är $x = 0$ den enda lösningen.
2. För udda n finns det precis en lösning.
3. För jämnt n och $a > 0$ finns två lösningar.
4. För jämnt n och $a < 0$ finns inga lösningar.

Det enklaste fallet där lösningar saknas är

$$x^2 + 1 = 0.$$

Eftersom det visade sig vara bekvämt, har man hittat på ett nytt tal i, som skulle uppfylla $i^2 = -1$ och sedan kommit genom addition och multiplikation med reella tal fram till komplexa tal $z = x + iy, x, y \in \mathbb{R}$, som kan adderas och multipliceras med varandra, sådant att axiomen 1 - 7 är uppfyllda: De utgör *komplexa talkroppen* $\mathbb{C}$. Och så har nu varje ekvation $x^n - a = 0$ en lösning i $\mathbb{C}$.

Men måste man vara fortsatt påhittig när det nu i stället gäller komplexa ekvationer

$$z^n - a = 0, \ a \in \mathbb{C},$$

och förstora $\mathbb{C}$?

För $n = 2$ kan man utan större problem hitta komplexa lösningar, medan för $n > 2$ hjälper en geometrisk tolkning av komplexa tal: De är ju faktiskt bara par $(x, y) \in \mathbb{R}^2$ av reella tal, som kan adderas

$$(x, y) + (s, t) := (x + s, y + t)$$

och multipliceras

$$(x, y) \cdot (s, t) := (xs - yt, xt + ys).$$

Och om man nu använder sig av polärkoordinater, så ser man att det även finns n-te rötter till ett givet komplext tal

$$a = r(\cos(\varphi) + i \sin(\varphi)),$$

nämligen talen

$$b_\nu := \sqrt[n]{r}(\cos\left(\frac{1}{n}(\varphi + 2\pi\nu)\right) + i \sin\left(\frac{1}{n}(\varphi + 2\pi\nu)\right), \nu = 0, ..., n-1.$$

Så det var mycket tillfredställande! Slutligen blir man lite djärvare och vill veta hur det står till med *polynomekvationer*

$$z^n + a_{n-1}z^{n-1} + ... + a_1 z + a_0 = 0,$$

där koefficienterna $a_{n-1}, ..., a_0 \in \mathbb{C}$ är komplexa tal. För $n \leq 4$ finns det lösningsformlar som innehåller förutom de vanliga räkneoperationerna rötter

- så då finns det alltid någon komplex lösning. Men för $n \geq 5$ finns inte längre sådana formlar.

Låt oss först titta på motsvarande problem för reella polynomekvationer:

$$g(x) = x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 = 0,$$

där koefficienterna $a_{n-1}, ..., a_0 \in \mathbb{R}$ är reella tal. Om n är udda, så har vi $g(x) < 0$ för $x \ll 0$ och $g(x) > 0$ för $x \gg 0$, eftersom för stora $x \in \mathbb{R}$ är x^n den dominerande termen:

$$g(x) = x^n \left(1 + \frac{a_{n-1}}{x} + ... + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}\right),$$

där andra faktorn går mot 1 för $x \to \infty$. Då finns det mitt emellan ett nollställe till funktionen $g(x)$. Det är enkelt att skriva upp ett nollställe

$$x_0 := \sup\{x \in \mathbb{R}; g(x) < 0\}$$

är ett nollställe, där man utnyttjar att funktionen $g(x)$ är kontinuerlig. Man kan också göra så här:

Theorem 4.1 (Algebrans fundamentalsats). *Varje polynomekvation*

$$f(z) = z^n + a_{n-1}z^{n-1} + ... + a_1 z + a_0 = 0,$$

av grad $n > 0$ med komplexa koefficienter $a_{n-1}, ..., a_0 \in \mathbb{C}$ har ett komplext nollställe.

Bevis: Beviset görs i två steg:

1. Funktionen $f : \mathbb{C} \longrightarrow \mathbb{C}$ har ett minimalställe $z_0 \in \mathbb{C}$, dvs. sådant att

$$|f(z)| \geq |f(z_0)|$$

gäller för alla $z \in \mathbb{C}$.

2. Minimalställen för en polynomfunktion är redan nollställen. I själva verket visar man motsatsen: En punkt $z_0 \in \mathbb{C}$ med $f(z_0) \neq 0$ är aldrig ett minimalställe för f: Nära till z_0 finns alltid punkter z med

$$|f(z)| < |f(z_0)|.$$

Detta är uppenbarligen fel för $\mathbb{R}$ i stället för $\mathbb{C}$ (tänk på $g(x) = x^2 + 1$ med $x_0 = 0$). Skillnaden beror på att på reella linjen en punkt bara kan approximeras ovan- eller nedifrån, medan i det komplexa planet finns oändligt många linjer längs vilka man kan närma sig en given punkt.

1.) **Existens av minimalställen:** Mängden

$$M := \{|f(z)|; z \in \mathbb{C}\} \subset \mathbb{R}$$

har 0 som lägre gräns. Om

$$b := \inf M \geq 0$$

är dess största lägre gräns (infimum) finns det en följd $(z_\mu)_{\mu \geq 1} \subset \mathbb{C}$ av komplexa tal med

$$\lim_{\mu \to \infty} |f(z_\mu)| = b.$$

Skriv

$$f(z) = z^n \left(1 + \sum_{\nu=0}^{n-1} \frac{a_\nu}{z^{n-\nu}} \right) = z^n K(z).$$

Klammeruttrycket $K(z)$ uppfyller

$$\lim_{z \to \infty} K(z) = 1,$$

eftersom varje ickekonstant term i summan går mot 0 för $z \to \infty$. Således

$$\lim_{z \to \infty} f(z) = \lim_{z \to \infty} z^n = \infty.$$

I synnerhet är följden (z_μ) begränsad, eftersom $(f(z_\mu))$ är begränsad. Enligt Th.3.9 (för komplexa talföljder) finns det en delföljd som konvergerar mot något $z_0 \in \mathbb{C}$. Vi får ersätta den ursprungliga följden genom den där delföljden resp. anta $\lim_{\mu \to \infty} z_\mu = z_0$ och får

$$f(z_0) = f(\lim_{\mu \to \infty} z_\mu) = \lim_{\mu \to \infty} f(z_\mu).$$

resp.

$$|f(z_0)| = |f(\lim_{\mu \to \infty} z_\mu)| = \lim_{\mu \to \infty} |f(z_\mu)| = b,$$

dvs. $z_0 \in \mathbb{C}$ är ett minimalställe.

2.) **Ett icke-nollställe z_0 är aldrig ett minimalställe**: Vi får anta $z_0 = 0$ - om inte det är fallet ersätt $f(z)$ med $f(z + z_0)$. Vi skriver

$$f(z) = a_0 + a_m z^m + z^{m+1} g(z)$$

med $a_0, a_m \neq 0$ och

$$g(z) = a_n z^{n-m-1} + \sum_{\nu=m+1}^{n-1} a_\nu z^{\nu-m-1}$$

för $n > m$, medan

$$g(z) \equiv 0$$

för $n = m$. Ta $b \in \mathbb{C}$ med

$$b^m = -\frac{a_0}{a_m}.$$

och undersök funktionen $f(z)$ på halvstrålen

$$\mathbb{R}_{\geq 0} \cdot b = \{tb, t \geq 0\}.$$

Det blir så här:

$$f(tb) = a_0(1 - t^m + t^{m+1}h(t))$$

med

$$h(t) := \frac{b^{m+1}}{a_0} g(tb).$$

Eftersom ju $f(0) = a_0 \neq 0$, räcker det att visa

$$|1 - t^m + t^{m+1}h(t)| < 1$$

för tillräckligt litet $t > 0$. Med

$$M := \max_{0 \leq t \leq 1} |h(t)|$$

och triangelolikheten får vi uppskattningen:

$$|1 - t^m + t^{m+1}h(t)| \leq |1 - t^m| + t^{m+1}M$$

$$= 1 - t^m + t^{m+1}M = 1 - t^m(1 - tM) < 1,$$

om $0 < t < \min(1, \frac{1}{M})$. Ty för $0 \leq t \leq 1$ har vi ju $|1 - t^m| = 1 - t^m$ och för $t < \frac{1}{M}$ blir det $1 - tM > 0$. $\square$

Vad som ligger bakom ovanstående argument kan uttryckas på ett enkelt sätt. Först behöver vi begreppet av en öppen mängd:

Definition 4.2. 1. Mängden

$$D_r(a) := \{z \in \mathbb{C}; |z - a| < r\}$$

kallas cirkelskivan kring $a \in \mathbb{C}$ med radien $r > 0$.

2. En delmängd $W \subset \mathbb{C}$ kallas öppen om för varje $a \in W$ finns det $\varepsilon = \varepsilon(a) > 0$ med

$$D_\varepsilon(a) \subset W.$$

Example 4.3. Varje cirkelskiva $D_r(b)$ är en öppen mängd: För $a \in D_r(b)$ ta $\varepsilon(a) = r - |a|$.

Theorem 4.4. *Låt $f : \mathbb{C} \longrightarrow \mathbb{C}$ vara en icke-konstant polynomfunktion, $W \subset \mathbb{C}$ en öppen mängd. Sedan är också dess bild*

$$f(W) := \{f(z); z \in W\} \subset \mathbb{C}$$

öppen.

Corollary 4.5. *Låt $f : \mathbb{C} \longrightarrow \mathbb{C}$ vara en icke-konstant polynomfunktion, och $z_0 \in \mathbb{C}$ med $f(z_0) \neq 0$. Sedan finns det godtyckligt nära till z_0 punkter z med*

$$|f(z)| < |f(z_0)|.$$

Proof. Ta $r = |f(z_0)| > 0$ och $\varepsilon > 0$ med $D_\varepsilon(f(z_0)) \subset f(\mathbb{C})$. Sedan har vi

$$f(\mathbb{C}) \cap D_r(0) \supset D_\varepsilon(f(z_0)) \cap D_r(0) \neq \emptyset.$$

□

Proof. En skiss: Given $z_0 \in \mathbb{C}$ måste vi hitta $\varepsilon(f(z_0))$. Igen antar vi $z_0 = 0$ och skriver

$$f(z) = a_0 + a_m \psi(z)$$

med ett polynom $\psi(z)$. Vi letar efter ett funktionsuttryck

$$\varphi(z) = z + c_1 z^2 + ...$$

med

$$\varphi(z)^m = \psi(z) = z^m + b_1 z^{m+1} + ... + b_r z^{m+r}.$$

Anta för ögonblicket att vi har lyckats. Eftersom den dominerande termen i $\varphi(z)$ för argument nära 0 är z, så hittar vi $\delta > 0$ och en öppen mängd $U \ni 0$, sådant att

$$f|_U : U \longrightarrow D_\delta(0)$$

är bijektiv. (Mängden $U \subset \mathbb{C}$ kan tänkas som det inre till en enkelt sluten kurva kring origo.) Eftersom

$$f(z) = a_0 + a_m \varphi(z)^m,$$

kan restriktionen $f|_U$ faktoriseras:

$$U \xrightarrow{\varphi(z)} D_\delta(0) \xrightarrow{z^m} D_{\delta^m}(0) \xrightarrow{z+a_0} D_{\delta^m}(a_0),$$

Eftersom den mellersta pilen är surjektiv och de andra två bijektiva kan vi ta $\varepsilon(f(0)) = \delta^m$. För att hitta $\varphi(z)$ skriver vi

$$\psi(z) = z^m(1 + b_1 z + ... + b_r z^r)$$

och letar efter

$$\varphi(z) = z(1 + c_1 z + ...).$$

Ekvationen

$$(1 + c_1 z + ...)^m = 1 + b_1 z + ... + b_r z^r$$

leder till $mc_1 = b_1$ och induktiva formlar

$$b_k = mc_k + f_k(c_1, ..., c_{k-1})$$

med $b_k = 0$ för $k > r$. Fast $\psi(z)$ är ett polynom, får vi vanligtvis oändligt många $c_k \neq 0$, funktionsuttrycket $\varphi(z)$ blir alltså ett oändligt polynom (potensserie), som lyckligtvis konvergerar för z tillräckligt nära origo. $\square$

I komplexa analysen studerar man funktioner som $\varphi(z)$, dvs. funktioner som lokalt kan beskrivas genom en potensserie.

Definition 4.6. En hel funktion är en funktion

$$f : \mathbb{C} \longrightarrow \mathbb{C}, z \mapsto f(z) := \sum_{\nu=0}^{\infty} a_\nu z^\nu.$$

Example 4.7. Komplexa exponentialfunktionen

$$\exp(z) = \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!}.$$

Den utvidgar reella exponentialfunktionen

$$\exp(x) = e^x, \ x \in \mathbb{R}$$

samt

$$\exp(z+w) = \exp(z)\exp(w).$$

Således

$$\exp(x+iy) = \exp(x)\exp(iy) = e^x \cdot \sum_{\nu=0}^{\infty} \frac{(iy)^\nu}{\nu!}$$

$$= e^x \left(\sum_{k=0}^{\infty} (-1)^k \frac{y^{2k}}{(2k)!} + i \sum_{\nu=0}^{\infty} (-1)^k \frac{y^{2k+1}}{(2k+1)} \right)$$

$$= e^x(\cos(y) + i\sin(y)).$$

Remark 4.8. Algebrans fundamentalsats kan också formuleras så här:

$$f(\mathbb{C}) = \mathbb{C}$$

för ett icke konstant polynom. Däremot gäller

$$\exp(\mathbb{C}) = \mathbb{C} \setminus \{0\}.$$

Theorem 4.9 (Lilla sats av Picard)**.** *För en hel funktion* $f : \mathbb{C} \longrightarrow \mathbb{C}$ *har vi*

$$f(\mathbb{C}) = \mathbb{C}$$

eller

$$f(\mathbb{C}) = \mathbb{C} \setminus \{a\}$$

med något komplext tal $a \in \mathbb{C}$.

5 Lorentztransformationer

Här använder vi våra LA-kunskaper för att härleda Lorentztransformationerna på ett enkelt sätt.

Definition 5.1. Ett *koordinatsystem* (eller en *bas*) i planet $\mathbb{R}^2$ är ett par $K = (\mathbf{v}_1, \mathbf{v}_2) \in \mathbb{R}^2$ av linjärt oberoende vektorer. Koordinaterna av en vektor $\mathbf{u} \in \mathbb{R}^2$ m.a.p. K är talen λ_1, λ_2, sådant att

$$\mathbf{u} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2.$$

Standardkoordinatsystemet består av enhetsvektorerna

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Remark 5.2. Koordinattransformationen från standardkoordinatsystemet till ett allmänt koordinatsystem fungerar så här: Låt $A = (\mathbf{v}_1, \mathbf{v}_2) \in \mathbb{R}^{2,2}$ vara matrisen med $\mathbf{v}_1, \mathbf{v}_2$ som kolonnvektorer. Om $\mathbf{u} \in \mathbb{R}^2$ har de kartesiska koordinaterna x_1, x_2, dvs.

$$\mathbf{u} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2,$$

så har vi uppenbarligen

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

resp.

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = A^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Definition 5.3. Ett koordinatsystem $K = (\mathbf{v}_1, \mathbf{v}_2)$ kallas *ortogonalt* resp. en *ortonormalbas* om

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = 0, \ ||\mathbf{v}_i|| = 1 \text{ for } i = 1, 2.$$

För ortogonala koordinatsystem är det enkelt att beräkna en vektor $\mathbf{u}$'s koordinater: Vi har

$$\mathbf{u} = (\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u} \cdot \mathbf{v}_2)\mathbf{v}_2,$$

eftersom VL och HL har samma skalärprodukter med basvektorerna $\mathbf{v}_1, \mathbf{v}_2$.

Vi ska nu diskutera vissa icke-ortogonala koordinatsystem i planet $\mathbb{R}^2$. En punkt $\begin{pmatrix} x \\ t \end{pmatrix} \in \mathbb{R}^2$ tolkas som en händelse i en endimensionell värld med rumskoordinat x och tidskoordinat t. En partikel som rör sig med konstant hastighet q öster- eller västerut representeras genom en linje $x = \pm qt + x_0$, en foton genom en linje $x = \pm t + x_0$, dvs. ljusets hastighet har normerats till $c = 1$.

Vi vill bestämma koordinatsystemet $\tilde{K}$ som tillhör en observatör som rör sig med konstant hastighet q österut och var vid tid $t = 0$ hos standard-observatören med koordinatsystemet $K = (\mathbf{e}_1, \mathbf{e}_2)$. I Newtonska mekaniken tillordnar denna observatör händelsen (x, t) koordinaterna

$$\tilde{x} = x - qt, \ \tilde{t} = t$$

M.a.o.

$$\tilde{K} = (\mathbf{e}_1, \mathbf{e}_2 + q\,\mathbf{e}_1)$$

pga.

$$\begin{pmatrix} x \\ t \end{pmatrix} = \tilde{x} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \tilde{t} \begin{pmatrix} q \\ 1 \end{pmatrix}.$$

Vi vill nu veta hur hans koordinatsystem $\tilde{K} = (\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2)$ ser ut i en relativistisk värld, där fotonerna har samma hastighet $c = 1$ för alla observatörer. För det första rör han sig inte med avseende på sitt eget koordinatsystem $\tilde{K}$, detta innebär att

$$\tilde{\mathbf{e}}_2 = k \begin{pmatrix} q \\ 1 \end{pmatrix}$$

med någon skalningsfaktor $k > 0$. Eftersom fotonerna har hastighet 1 måste

$$\tilde{\mathbf{e}}_1 + \tilde{\mathbf{e}}_2 = \begin{pmatrix} \lambda \\ \lambda \end{pmatrix} \in \mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Det enklaste vore väl att välja

$$\tilde{\mathbf{e}}_1 = k \begin{pmatrix} 1 - q \\ 0 \end{pmatrix}, \text{ sådant att } \lambda = k.$$

Men eftersom K och $\tilde{K}$ har samma rättigheter, måste följande relativitets-princip uppfyllas:

Relativitetsprincip: Om

$$\tilde{\mathbf{e}}_i = A(q)\mathbf{e}_i, i = 1, 2$$

med en (2,2)-matris $A(q)$, så måste också

$$\mathbf{e}_i = A(-q)\tilde{\mathbf{e}}_i, i = 1, 2,$$

eftersom $\tilde{K}$ ju anser att K rör sig med hastighet q västerut (eller med hastighet $-q$ österut).

I synnerhet

$$A(q)A(-q) = I_2.$$

Men med ovanstående ansats får vi

$$A(q) = k \begin{pmatrix} 1-q & q \\ 0 & 1 \end{pmatrix}$$

och

$$A(q)A(-q) = k^2 \begin{pmatrix} 1-q^2 & q^2 \\ 0 & 1 \end{pmatrix} \neq I_2.$$

Så det verkar inte vara bra. Vi måste alltså diskutera den allmänna situationen

$$A(q) = \begin{pmatrix} \lambda - kq & kq \\ \lambda - k & k \end{pmatrix}$$

och antar för enkelhetens skull att $k = k(q)$ är en jämn funktion av q. Låt oss nu skriva

$$A(-q) = \begin{pmatrix} \tilde{\lambda} + kq & -kq \\ \tilde{\lambda} - k & k \end{pmatrix}.$$

Sedan vill vi att

$$A(q)A(-q) = \begin{pmatrix} \lambda - kq & kq \\ \lambda - k & k \end{pmatrix} \begin{pmatrix} \tilde{\lambda} + kq & -kq \\ \tilde{\lambda} - k & k \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

i synnerhet

$$(\lambda - kq)(-kq) + k^2 q = 0$$

resp.

$$\lambda = k(1+q).$$

Å andra sidan

$$1 = (\lambda - k)(-kq) + k^2 = k^2 - k^2q^2,$$

dvs.

$$k = \frac{1}{\sqrt{1-q^2}}.$$

Således

$$\tilde{\mathbf{e}}_1 = k \begin{pmatrix} 1 \\ q \end{pmatrix}.$$

Vi lämnar det åt läsaren att kolla att våra förväntningar nu blir uppfyllda.

Men lägg märke till att vi måste betala för det: Händelser som är samtidiga m.a.p. K är det inte längre m.a.p. $\tilde{K}$ och tvärtom. Linjerna $t = \text{const}$ är ju horisontella, medan linjerna $\tilde{t} = \text{const}$ är sneda för $q \neq 0$.

Sammanfattningsvis har vi alltså

Proposition 5.4. *En observatör som rör sig med konstant hastighet q österut har koordinatsystemet*

$$\tilde{K} = \left(\tilde{\mathbf{e}}_1 = \frac{1}{\sqrt{1-q^2}} \begin{pmatrix} 1 \\ q \end{pmatrix}, \tilde{\mathbf{e}}_2 = \frac{1}{\sqrt{1-q^2}} \begin{pmatrix} q \\ 1 \end{pmatrix} \right).$$

Remark 5.5. Som i den "tidslösa" euklidiska geometrin i $\mathbb{R}^2$ finns det också i den relativistiska världen en utvecklingsformel, bara man ersätter skalärprodukten

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$$

med

$$g(\mathbf{u}, \mathbf{v}) := \mathbf{u}^T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{v} = xy - st$$

för $\mathbf{u} = \begin{pmatrix} x \\ s \end{pmatrix}, \mathbf{v} = \begin{pmatrix} y \\ t \end{pmatrix}$. Nämligen ovanstående koordinatsystemen $\tilde{K}$ blir då g-ortogonala, dvs. $g(\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2) = 0 = g(\tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_1), g(\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_1) = 1, g(\tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_2) = -1$, och vi kan utveckla

$$\mathbf{u} = g(\mathbf{u}, \tilde{\mathbf{e}}_1)\tilde{\mathbf{e}}_1 - g(\mathbf{u}, \tilde{\mathbf{e}}_2)\tilde{\mathbf{e}}_2$$

och får *Lorentztransformationerna*

$$\tilde{x} = g(x\mathbf{e}_1 + t\mathbf{e}_2, \tilde{\mathbf{e}}_1), \tilde{t} = -g(x\mathbf{e}_1 + t\mathbf{e}_2, \tilde{\mathbf{e}}_2).$$

Vi anmärker att vektorerna $\mathbf{u} \in \mathbb{R}^2$ med $g(\mathbf{u}, \mathbf{u}) = 1$ resp. $g(\mathbf{u}, \mathbf{u}) = -1$ ligger på hyperblerna $x^2 - t^2 = \pm 1$.

På ett liknande sätt som ortogonala koordinatsystem i $\mathbb{R}^2$ kan realiseras som kolonnvektorer till rotationsmatriserna

$$R(\vartheta) = \begin{pmatrix} \cos(\vartheta) & -\sin(\vartheta) \\ \sin(\vartheta) & \cos(\vartheta) \end{pmatrix}, \vartheta \in \mathbb{R},$$

kan vi också göra det med Lorentzkoordinatsystemen: Vi inför de hyperbolisk-trigonometriska funktionerna

$$\cosh(\tau) = \frac{1}{2}(e^{\tau} + e^{-\tau}), \sinh(\tau) = \frac{1}{2}(e^{\tau} - e^{-\tau}),\ \tanh(\tau) := \frac{\sinh(\tau)}{\cosh(\tau)}$$

och anmärker att

$$\cosh^2(\tau) - \sinh^2(\tau) = 1.$$

Lorentzmatriserna har nu följande form:

$$L(\tau) = \begin{pmatrix} \cosh(\tau) & \sinh(\tau) \\ \sinh(\tau) & \cosh(\tau) \end{pmatrix}, \tau \in \mathbb{R},$$

med tillhörande hastighet

$$q = \tanh(\tau) = \tan(\vartheta),$$

där ϑ är vinkeln mellan $\mathbf{e}_1$ och $\begin{pmatrix} \cosh(\tau) \\ \sinh(\tau) \end{pmatrix}$.

Medan rotationer bevarar skalärprodukten, bevarar Lorentzmatriserna g-produkten

$$g(L(\tau)\mathbf{u}, L(\tau)\mathbf{v}) = g(\mathbf{u}, \mathbf{v}).$$

I själva verket gäller även här

$$L(\tau + \sigma) = L(\tau)L(\sigma),$$

men medan "banan" $\{R(\vartheta)\mathbf{u}, \vartheta \in \mathbb{R}\}$ till någon vektor $\mathbf{u} \in \mathbb{R}^2 \setminus \{0\}$ är en cirkel 0, utgör familjen $\{L(\tau)\mathbf{u}, \tau \in \mathbb{R}\}$ för $\mathbf{u} \notin \mathbb{R} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$ en hyperbel.

Example 5.6. Låt oss säga, någon reser vid tiden $t = 0$ ut från 0 med hastighet q österut och vänder om vid tiden t_0 för att komma hem vid $t = 2t_0$. Hur mycket tid har gått för honom under hela resan?

Vi anmärker först att tiden mellan två händelser $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ m.a.p. koordinatsystemet $\tilde{K}$ är $-g(\mathbf{u} - \mathbf{v}, \tilde{\mathbf{e}}_2)$.

På utresan har resenären koordinatsystemet

$$\tilde{K} = \left(\tilde{\mathbf{e}}_1 = \frac{1}{\sqrt{1-q^2}} \begin{pmatrix} 1 \\ q \end{pmatrix}, \tilde{\mathbf{e}}_2 = \frac{1}{\sqrt{1-q^2}} \begin{pmatrix} q \\ 1 \end{pmatrix} \right),$$

på hemresan

$$\hat{K} = \left(\hat{\mathbf{e}}_1 = \frac{1}{\sqrt{1-q^2}} \begin{pmatrix} 1 \\ -q \end{pmatrix}, \hat{\mathbf{e}}_2 = \frac{1}{\sqrt{1-q^2}} \begin{pmatrix} -q \\ 1 \end{pmatrix} \right).$$

Således upplever han själv det hela så här: Tiden för utresan är

$$-g(qt_0\mathbf{e}_1 + t\mathbf{e}_2, \tilde{\mathbf{e}}_2) = -kg(qt_0\mathbf{e}_1 + t_0\mathbf{e}_2, q\mathbf{e}_1 + \mathbf{e}_2)$$
$$-k(q^2 - 1)t_0 = \sqrt{1-q^2}\, t_0.$$

Samma resonemang gäller för hemresan, sådant att tiden mellan utfärden och hemkomsten för resenären bara är $2\sqrt{1-q^2}\, t_0$, medan det är $2t_0$ för den orörliga observatören.

6 Zorns lemma

Vi slutar med ett resultat som används ofta i matematiken när man sysslar med oändliga mängder och vars bevis bygger på liknande argument som välordningssatsen.

Definition 6.1.

1. En **partiell ordning** på en mängd M är en relation "$\leq$", som är reflexiv, antisymmetrisk och transitiv.
2. Ett element $a \in M$ kallas maximalt m.a.p. "$\leq$", om det inte finns något element $x \in M$ med $x > a$.

Example 6.2. Låt elementen i M vara (vissa) delmängder till en given mängd U, dvs.

$$M \subset \mathcal{P}(U).$$

Ordningsrelationen $A \leq B$ definieras genom $A \subset B$.

Remark 6.3.

1. Två element x, y i en partiell ordnad mängd M kan i allmänheten **inte** jämföras med varandra, dvs. det kan hända att varken $x \leq y$ eller $y \leq x$ gäller.

2. Ett maximalt element $a \in M$ behöver inte uppfylla $a \geq x$ för alla $x \in M$.

3. Låt M vara en mängd med en partiell ordning $\leq$. Vi säger att en delmängd $T \subset M$ är linjärt ordnad om två gottyckliga element $x, y \in T$ kan jämföras: Vi har $x \leq y$ eller $y \leq x$.

Theorem 6.4. (Zorns lemma) *(Max August Zorn, 1906-1993): Låt mängden M vara utrustad med den partiella ordningen "$\leq$". Om det för alla m.a.p. $\leq$ linjärt ordnade delmängder $T \subset M$ finns en övre gräns $b \in M$, dvs. sådant att $t \leq b$ för alla $t \in T$ (eller i förkortad notation $T \leq b$), så finns det maximala element $a \in M$. (OBS: Vi kräver inte $b \in T$!)*

Example 6.5. Om $M \subset \mathcal{P}(U)$ som i Ex.6.2 och $T \subset M \subset \mathcal{P}(U)$ är en linjärt ordnad delmängd, så är unionen

$$B := \bigcup_{A \in T} A$$

av alla mängder $A \in T$ en kandidat för en övre gräns. Bara man måste kolla $B \in M$.

Här är en tillämpning i linjära algebran:

Theorem 6.6. *Varje vektorrum V har en bas.*

Proof. Vi gör precis som i Ex.6.5. Ta $M \subset \mathcal{P}(V)$ så här: En delmängd $A \subset V$ tillhör M omm dess element är linjärt oberoende. Sedan kan övre gränsen till en linjärt ordnad delmängd $T \subset M$ hittas som unionen C av alla $A \in T$. Den tillhör M: Om $\mathbf{v}_1, ..., \mathbf{v}_r \in C$, låt oss säga $\mathbf{v}_i \in A_i \in T$, så får vi, eftersom T är linjärt ordnad, omnumrera sådant att $A_1 \subset A_2 \subset ... \subset A_r$. Således är vektorerna $\mathbf{v}_1, ..., \mathbf{v}_r \in A_r$ linjärt oberoende pga. $A_r \in M$.

Slutligen: Om $B \in M$ är en maximal mängd av linjärt oberoende vektorer, så är B en bas: Vi måste visa att varje vektor $\mathbf{v} \in V$ är en entydig linjärkombination av vektorer i B. Entydigheten har vi pga. $B \in M$.

Så ta någon vektor $\mathbf{v} \in V$. Om $\mathbf{v} \in B$, är det klart. Annars gäller $B \cup \{\mathbf{v}\} \notin M$ – mängden $B \in M$ är ju maximal – och således finns det en icke-trivial relation

$$0 = \lambda \mathbf{v} + \lambda_1 \mathbf{v}_1 + ... + \lambda_r \mathbf{v}_r$$

med $\mathbf{v}_1, ..., \mathbf{v}_r \in B$ och $\lambda, \lambda_1, ..., \lambda_r \in \mathbb{R}$. Men $\lambda \neq 0$, eftersom vektorerna $\mathbf{v}_1, ..., \mathbf{v}_r$ är linjärt oberoende, m.a.o.

$$\mathbf{v} = \sum_{i=1}^{r} (-\lambda^{-1} \lambda_i) \mathbf{v}_i.$$

$\square$

Proof of Th.6.4. Vi antar, att det inte finns några maximala element. Låt $\mathrm{Lin}(M) \subset \mathcal{P}(M)$ bestå av alla m.a.p. $\leq$ linjärt ordnade delmängder $T \subset M$. Vi väljer en funktion

$$\gamma : \mathrm{Lin}(M) \longrightarrow M,$$

som tillordnar varje T en äkta övre gräns $\gamma(T) > T$. Det finns ju i alla fall en övre gräns $b \geq T$ och eftersom b inte är maximalt, finns det $\gamma(T) > b$.

Igen tittar vi på "γ-kedjor": En delmängd $K \subset M$ kallas γ-kedja, om $K \in \mathrm{Lin}(M)$ är t.o.m. välordnad m.a.p. $\leq$ och

$$a = \gamma(K_{<a})$$

gäller för alla $a \in K$. Mängden av sådana γ-kedjor är linjärt ordnad m.a.p. $\subset$, således är union av alla γ-kedjor igen en γ-kedja, en maximal sådan. Men $K' = K \cup \{\gamma(K)\}$ är en ännu större γ-kedja: Motsägelse! $\square$