CHALMERS | GÖTEBORGS UNIVERSITET

Introduktion till icke-standard analys

Ett bevis av transferprincipen och en explicit konstruktion av Brownsk rörelse

Examensarbete för kandidatexamen i matematik vid Göteborgs universitet

Oscar Anghammar

Mikael Norder

Andreas Petersson

Institutionen för matematiska vetenskaper
Chalmers tekniska högskola
Göteborgs universitet
Göteborg 2013

Introduktion till icke-standard analys

Ett bevis av transferprincipen och en explicit konstruktion av Brownsk rörelse

Examensarbete för kandidatexamen i matematik vid Göteborgs universitet

Mikael Norder

Examensarbete för kandidatexamen i matematik inom matematikprogrammet vid Göteborgs universitet

Oscar Anghammar Andreas Petersson

Handledare: Leif Arkeryd
Examinator: Hjalmar Rosengren

Institutionen för matematiska vetenskaper
Chalmers tekniska högskola
Göteborgs universitet
Göteborg 2013

Sammanfattning

I denna rapport konstrueras inledningsvis hyperreella tal från reella tal. Vi definierar aritmetik för skuggor av hyperreella tal och presenterar en icke-standard teori för serier och konvergens utifrån denna konstruktion. Förfarandet generaliseras sedan genom att skapa en icke-standard modell till matematisk analys utifrån en mängdteoretisk modell i ZF, som ett led i konstruktionen presenteras teori för filter och ultraprodukter. Vi visar sedan transferprincipen och bygger upp internalitetsbegreppet för att studera relationer mellan standard och icke-standard modellerna för att kunna härledda standardresultat från icke-standard resultat. Avslutningsvis tillämpas teorin för att skapa en explicit icke-standard konstruktion av Brownsk rörelse och vi ger ett bevis för att denna konstruktion är ekvivalent med standard formuleringen.

Abstract

In this report we begin by constructing the hyperreals from the real numbers. From these concepts we define arithmetic on shadows of hyperreals and present a non-standard theory for the convergence of numerical series. We generalise the construction method to create a non-standard model of analysis from a set-theoretical model in ZF. In order to accomplish this, a theory of filters and ultraproducts is presented. We then prove the transfer principle for the resulting non-standard structure and make use of the concept in internality to establish deduction procedures between the standard and non-standard results. Finally, the resulting theory is applied to explicitly create a non-standard construction of Brownian motion, which we prove is equivalent to the standard formulation.

Innehåll

Förord

Denna rapport är författad av Oscar Anghammar, Mikael Norder och Andreas Petersson som ett gemensamt kandidatarbete i matematik. Författarna har studerat matematik och logik på Göteborgs Universitet och texten riktar sig till studenter med liknande bakgrund. Rapporten består av en genomgång av grunden till icke-standard analys, en genomgång av vanliga icke-standard tekniker och slutligen en fördjupning inom icke-standard sannolikhetsteori.

De individuella bidragen i rapporten reflekteras i stort av denna indelning. Oscar Anghammar har författat kapitlen 2-4 om de hyperreella talen samt elementära operationer på dessa, Mikael Norder kapitlen 5-9 om icke-standard analys i relation till grundläggande mängd- och modellteori och Andreas Petersson kapitlen 11-13 som rör en tillämpning av icke-standard analys på sannolikhetsteori. Kapitel 1 och kapitel 10 har skrivits gemensamt med Oscar Anghammar som huvudansvarig. Oscar Anghammar och Andreas Petersson har ansvarat för största delen av det redaktionella arbetet att sätta samman rapporten och Mikael Norder har ansvarat för källhanteringen.

En tidslogg och en dagbok har kontinuerligt förts under arbetets gång och dessa beskriver författarnas individuella prestationer i detalj.

Författarna vill uttrycka sin stora tacksamhet till Leif Arkeryd som varit en engagerad handledare för arbetet.

1 Inledning

Detta kapitel utgår framförallt från [Bair J. et. al. 2013].

1.1 Introduktion och historia om icke-standard analys

Inom matematisk analys studerar man teorier om bland annat derivering och integrering av funktioner, som ofta är definerad på de reella talen. Ett fundamentalt verktyg inom analysen är gränsvärdet, som används för att studera hur funktioner uppför sig nära ett visst tal eller oändligheten. Inom standardanalysen använder man sig av den så kallade (ϵ, δ)-defintionen för att definera gränsvärden, men det visar sig att detta inte är det enda sätt man kan konstruera matematisk analys på.

En alternativ metod är att istället utöka mängden av reella tal till en mängd som innehåller infinita och infinitesimala tal, det vill säga oändlig stora och oändligt små (nollskillda) tal. Vi kallar dessa tal hyperreella, och teorin som bygger på dessa kalla vi för *icke-standard analys*.

Termen ”icke-standard analys” skiljer sig från den mindre rigorösa *infinitesimalkalkylen* som bedrevs innan Weierstrass och etablerandet av (ϵ, δ)-defintionen. Det dröjde till 1960-talet innan *Abraham Robinson* konstruerade utifrån formell matematisk logik ett väldefinierat system för infinitesimalkalkyl, och det är detta vi kalla för icke-standard analys. Detta gjorde det möjligt att utifrån denna konstruktion bevisa riktigheten i de informella icke-standard metoder som tidigare använts.

En framstående förespråkare för de tidiga icke-standard analysmetoderna var *Gottfried Leibniz*. För att beskriva hans metoder definerar vi två begrepp, *Bernoulliskt*- och *Arkimediskt* kontinumm.

(i) $\forall \epsilon > 0 \; \exists n \in \mathbb{N} : n\epsilon > 1$ Arkimediskt kontinuum

(ii) $\exists \epsilon > 0 \; \forall n \in \mathbb{N} : \epsilon \leq \frac{1}{n}$ Bernoulliskt kontinuum

Ett talsystem som innehåller hyperreella tal kallar vi alltså för Bernoulliskt, i motsats till det Arkimediska talsystemet som används i standardanalysen.

Vi skall i uppsatsen, givet en motsvarighet till Robinsons konstruktion, visa den så kallade transferprincipen vilken kan användas som härledningsregel mellan ett Arkimediskt och ett Bernoulliskt kontinuum. Transferprincipens ursprung kan vi härleda till Leibniz, som kallade sin motsvarighet till den moderna transferprincipen för ”kontinuitetslagen” (Lex continuitatis). Kontinuitetslagen uttrycker att ”Det som gäller för de ändliga talen också gäller för de oändliga”. Denna lag var ett obevisat antagande — eller ett axiom om man så vill uttrycka det — från Leibniz sida om talen i hans Bernoulliska kontinuum. Detta axiom tillät honom att överföra egenskaper från ett Arkimediska kontinuum till ett Bernoulliskt kontinuum.

$$\{\text{Arkimedisk talkropp}\} \xrightarrow{\text{Kontinuitetslagen}} \{\text{Bernoullisk talkropp}\}$$

Exempelvis gäller enligt kontinuitetslagen den trigonometriska identiteten $\sin^2 x + \cos^2 x = 1$ även i det Bernoulliska kontinuum, just på grund av att den gäller i det Arkimediska. Med andra ord kan alltså x i formeln lika gärna vara en infinitesimal eller ett obegränsat tal.

Analogt formulerade Leibniz ytterligare en lag för att överföra egenskaper från den Bernoulliska till den Arkimediska talkroppen: "Lagen om transcendent homogenitet" ("Lex homogeneorum transcendentalis").

$$\{\text{Bernoullisk talkropp}\} \xrightarrow{\text{Lagen om transcendent homogenitet}} \{\text{Arkimedisk talkropp}\}$$

Enligt denna lag gäller det att man kan bortse från tal som är oändligt små jämfört med andra termer. Exempelvis gäller alltså en form av likhet mellan $a+\epsilon$ och a, där a är ett reellt tal och ϵ en infinitesimal. Generellt vid summation av infinitesimaler av olika ordning kan man bortse från infinitesimalerna av högre ordning. Med andra ord kan vi formulera detta som att två tal i någon mening kan betraktas som lika om de ligger oändligt nära varandra.

Vi kan illustrera användet av lagen om transcendent homogenitet med ett av Leibniz bevis för produktregeln:

$$d(uv) = (u + du)(v + dv) - uv = udv + vdu + dudv \simeq udv + vdu$$

där den sista relationen betecknats $\simeq$[1] för att visa att vi använt lagen om transcendental homogenitet.

Rörande infinitesimalernas ontologiska status — huruvida de *existerar* eller inte — förefaller inte Leibniz ha varit särskilt bekymrad. Hans hållning illustreras av följande citat:

> "Filosofiskt sett, erkänner jag varken magnituder oändligt stora eller oändligt små [...] Jag tar dessa båda som fiktioner, bekväma uttryckssätt anpassade till räknesätten, precis som imaginära rötter inom algebran".

Leibniz var inte den enda som förespråkade icke-standard analys, och kanske är det ingen tillfällighet att olika teorier och tillämplingar av infinitesimaler har dykt upp i historien. Det är möjligt att det på en intiutiv nivå inte upplevs särskilt abstrakt med idén om oändligt stora- och små tal. De infinitesimala talen ligger i själva verket ganska nära till hands när man exempelvis försöker förstå vad en derivata eller integral är för något; i standardanalysen lever Leibniz notationer för integrering och derivering kvar, detta trots att Leibniz ansåg att differentialerna i exempelvis $\frac{dx}{dy}$ skulle betraktas som infinitesimala tal och hela utrycket som en kvot av dessa.

För att ytterligare poängtera de intuitiva fördelarna med icke-standard analys betraktar vi ett par minnesregler som kan användas i standardanalysen. Först har vi kedjeregeln

$$\frac{df}{dy} = \frac{df}{dx}\frac{dx}{dy},$$

[1] Vi kommer i kapitel 3 definera relationen $\simeq$ som "oändligt nära".

där högerledet kan — för att komma ihåg formeln — betraktas som en multiplikation mellan två kvoter av infinitesimaler. Den andra reglen är variabelbyte i en enkelintegral, där förändringen av differentialen är

$$\frac{dx}{dt} = g'(t) \Rightarrow dx = g'(t)dt.$$

Här kan vi tänka att vi multipliceratbåda leden i första utsagan med infinitesimalen dt.

Trots sina intiutiva fördelar är icke-standard analys idag just icke-standard, vilket troligtvis beror på att dess rigorösa grund konstruerades långt efter standaranalysens rigorösa grund.

1.2 Introduktion till icke-standard sannolikhetsteori

Ett område där icke-standard analysen spelat en förhållandevis stor roll är sannolikhetsteori. När man först lär sig sannolikhetsteori så brukar man gå in på så kallade sannolikhetsrum — vi talar om Venndiagram, om händelser som mängder och hur man kan uppfatta sannolikheten av dessa som mängdernas kardinalitet dividerat med sannolikhetsrummets kardinalitet. Men när vi studerar lite mer avancerad sannolikhetsteori släpper vi oftast denna grund för att istället gå in på diskreta och kontinuerliga slumpvariabler, vilket leder till en diskrepans mellan början på sannolikhetskursen och fortsättningen.

Till detta erbjuder icke-standard analys ett alternativ. Vi kan kombinera den tidiga sannolikhetslärans kombinatoriska idéer med en rigorös hantering av kontinuerliga slumpvariabler. Grunden för detta är ett sannolikhetsrum med så kallad hyperändlig kardinalitet — ”storleken” på Ω är ett obegränsat hypernaturligt tal N. Sannolikheten för varje $w \in \Omega$ är definierad som en infinitesimal,

$$P(w \in \Omega) \quad \doteqdot \quad \frac{1}{N}.$$

Detta leder till att väntevärdet för en slumpvariabel på rummet får en intuitiv definition,

$$E(X) \quad \doteqdot \quad \sum_{w \in \Omega} \frac{X(w)}{N}.$$

Allting är definierat som om det underliggande sannolikhetsrummet vore ändligt.

Vi kommer att använda dessa idéer för att definiera så kallad Brownsk rörelse på ett mer explicit sätt än det vanliga. En följd av detta blir att Brownsk rörelse faktiskt existerar, vilket är ett djupt resultat inom sannolikhetsteorin. För att illustrera detta tillvägagångssätt, notera att en slumpvandring av längd 3 kan definieras som

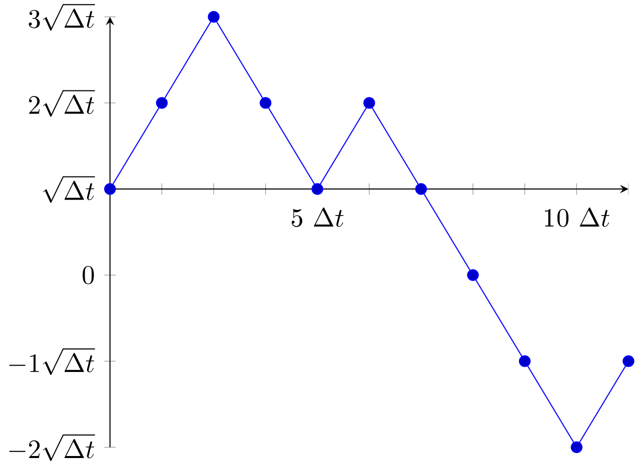
$$X(w,t) \quad \doteqdot \quad \sum_{s=1}^{t} w(s), \qquad t = 1, 2, 3.$$

Ω är här mängden av alla sekvenser av $\{-1, 1\}$ av längd 3;

$$\Omega = \{(-1, -1, 1), (-1, 1, -1), \cdots\} .$$

När vi vill definiera dess kontinuerliga motsvarighet låter vi alltså Ω vara en hyperändlig mängd av sådana här sekvenser med obegränsad längd och stegen som processen tar låter vi vara infinitesimala.

Figur 1: En infinitesimal slumpvandring



Vi kommer att visa hur man kan plocka ut ”standarddelen” av denna slumpvandring och att denna är ekvivalent med den vanlig definitionen av brownsk rörelse — vilket motiverar vårt tillvägagångssätt. För att göra detta kommer vi att behöva en del måtteori, eftersom det används i standardkonstruktionerna av sannolikhetsteori. Om läsarens mål var att lära sig om Brownsk rörelse skulle man naturligtvis kunna hoppa över detta led.

1.3 Syfte

Syftet med arbetet är att undersöka hur grunderna för icke-standard analys ser ut och hur något känt resultat inom sannolikhetsteori formuleras i icke-standard analys.

1.4 Rapportupplägg

Vårt arbete kan delas upp i fyra delar. Den första innehåller kapitel två, tre och fyra som skall fungera som en lätt introduktion till de hyperreella talen. Vi konstruerar här de hyperreella talen i en specifik kontext och använder dessa för att behandla lite grundläggande analys som ligger till grund för de senare delarna i arbetet. Vi undviker tung teori genom att hänvisa till senare kapitel.

Den andra delen innehåller kapitel fem, sex, sju och åtta. Här konstruerar vi en mer allmän grund till de hyperreella talen för att rigoröst bevisa den så kallade ”transferprincipen”. Teorin i dessa kapitel är främst till för beviset till transferprincipen; är man mer intresserad av de senare delarna i arbetet räcker det med att förstå vad transferprincipen är för något.

Den tredje delen innehåller kapitel nio och tio. Här går vi igenom diverse satser som kommer behövas för den sista delen i arbetet.

Den fjärde och sista delen innehåller kapitel elva, tolv och tretton. Här tillämpar vi den icke-standard teori vi tidigare byggt för att konstruera Brownsk rörelse.

2 Konstruktion

Detta kapitel utgår framförallt från [Goldblatt 1998].

2.1 Vilka egenskaper vill vi ha?

De hyperreella talen har egenskaper som skiljer sig från de reella talen. För att få en rigorös grund till icke-standard analysen måste vi veta hur tal med sådana egenskaper konstrueras.

Som tidigare nämnts har positiva infinitesimala tal egenskapen att vara positiva samtidigt som de är mindre än alla positiva reella tal. Infinitesimala tal har även egenskapen att de kan ordnas; exempelvis kan ett infitesimalt tal vara dubbelt så stort som ett annat. För att representera denna egenskap tittar vi på reellvärda talföljder:

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots \tag{1}$$

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \cdots \tag{2}$$

Om vi betraktar gränsvärdet av dessa följder som det *reella* tal som bäst representerar följderna, så ser vi att dessa två matematiska objekt båda har egenskapen att representeras av det reella talet 0. Vi ser också att den första följden är elementvis dubbelt så stor som den andra, så vi betraktar det första matematiska objektet att vara dubbelt så stort som det andra.

Om vi på liknande sätt representerar de reella talen som följder, så är det rimligt att definiera exempelvis 0 och $\frac{1}{2}$ som:

$$0, 0, 0, 0, \cdots$$

och

$$\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \cdots$$

Om vi betraktar talföljderna elementvis kan vi tydligt se hur egenskapen att vara större än 0 representeras hos både (1) och (2). Vi kan även se att alla utom ändligt många element ur de två första talfölderna är mindre än elementen ur en talföljd som representerar ett godtyckligt litet positivt reellt tal.

Det visar sig alltså att reellvärda talföljder på ett intuitivt sätt kan representera de storleksegenskaper vi vill att infinitesimala tal ska ha, och det är just reellvärda talföljder vi kommer använda oss av när vi konstruerar de hyperreella talen.

2.2 Enighetsmängder

Om vi ska betrakta reellvärda följder som tal, måste vi ha något sätt att bestämma om två följder representerar samma tal. Att låta varje unik följd representera ett tal kommer leda till problem med ordningen; jämför exempelvis dessa följder:

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \ldots$$
$$\frac{1}{4}, \frac{1}{2}, \frac{1}{6}, \frac{1}{8}, \ldots$$

Hur ska vi avgöra vilken av dessa två som är störst?

Vad vi i själva verket gör är att vi kommer att betrakta de båda ovanstående följderna som samma infinitesimala tal, eftersom de bara skiljer varandra åt på två tal i följden. Men på hur många positioner måste två följder skilja sig från varandra innan de representerar två olika tal? Vi definierar en *enighetsmängd*:

$$E_{rs} \quad = \quad \{n \in \mathbb{N} : r_n = s_n\}$$

Som alltså mäter hur lika två följder, r_n och s_n, är. Vi säger att r_n och s_n representerar samma tal om E_{rs} är stor, det vill säga om den innehåller tillräckligt många element. För att kunna definiera vad som är ”stort” måste vi först analysera vilka egenskaper vi vill att stora mängder ska ha. Vi vill att:

- $\mathbb{N}$ ska vara stor, så att två likadana följder representerar samma tal.
- $\emptyset$ *inte* ska vara stor, så att två följder som helt saknar gemensamma element representerar olika tal.
- Om r och s representerar samma tal, och om s och t representerar samma tal, så vill vi att r och t representerar samma tal. Eftersom $E_{rs} \cap E_{st} \subseteq E_{rt}$, ger detta kravet att:

  om A och B är stora mängder och $A \cap B \subseteq C$, då är C stor.

- En och endast en av A och $\mathbb{N} \setminus A$ ska vara stor för godtycklig delmängd $A \subseteq \mathbb{N}$.

Det sista vilkoret används för att undvika konflikt när vi definierar ordning mellan de tal som följderna representerar. Låt

$$L_{rs} \quad = \quad \{n \in \mathbb{N} : r_n < s_n\}$$

vara mängden som bestämmer ordningen mellan två följder; om L_{rs} är stor, så är talet representerat av r strängt mindre än talet representerat av s.

Om vi nu har att L_{rs} är stor och $E_{rs} = \emptyset$, så vill vi givetvis inte att $\mathbb{N} \setminus L_{rs} = L_{sr}$ ska kunna vara stor eftersom det skulle leda till en motsägelse. Därför inkluderar vi det sista kravet bland egenskaperna som en stor mängd ska ha.

Vi vet nu vilka egenskaper vi vill att stora mängder ska ha, men det återstår att faktiskt konstruera sådana. För att göra detta använder vi ett så kallat *ultrafilter*, som är en samling delmängder (till mängden man applicerar filtret på) som uppfyller specifika egenskaper. I fortsättningen på detta kapitel kommer vi referera till $\mathfrak{U}$, som betecknar ett fixt icke-principalt ultrafilter på mängden $\mathbb{N}$. Detta filter innehåller precis de delmängder till $\mathbb{N}$ som är stora. Vi behandlar ultrafilter utförligt i kapitel 5.

Tillvägagånsättet att skapa nya tal med hjälp av talföljder — och att betrakta tillräckligt lika följder som samma tal — är ett exemepl på en *ultraprodukt*, som vi introducerar i kapitel 8.1.

2.3 Relationer och operationer på reellvärda följder

En *ekvivalensrelation* på A är en godtycklig relation $\sim$ som uppfyller följande tre villkor:

- Reflexivitet: $a \sim a \qquad \forall a \in A$
- Symmetri: $a \sim b \Rightarrow b \sim a \qquad \forall a, b \in A$
- Transitivitet: $a \sim b \wedge b \sim c \Rightarrow a \sim c \qquad \forall a, b, c \in A$

En *ekvivalensklass* är en delmängd till A som, givet ett element $a \in A$ (kallad representant), består av alla element som är ekvivalenta med a under given ekvivalensrelation.

Låt $\mathbb{R}^{\mathbb{N}}$ vara mängden av alla reellvärda talföljder, där ett elementen betecknas $\langle r_n : n \in \mathbb{N} \rangle$ eller lite förenklat som $\langle r_n \rangle$. Vi definierar en relation $\equiv$ på $\mathbb{R}^{\mathbb{N}}$:

$$\langle r_n \rangle \equiv \langle s_n \rangle \quad \text{omm} \quad E_{rs} \in \mathfrak{U}$$

När denna relation gäller, så säger vi att följderna *överensstämmer på en stor mängd* eller *överensstämmer för nästan alla n*.

Sats 2.1. *$\equiv$ är en ekvivalensrelation på $\mathbb{R}^{\mathbb{N}}$.*

Bevis. Reflexiviteten och symmetrin på $\equiv$ följer direkt av reflexiviteten och symmetrin på $=$ (som är en ekvivalensrelation).

Transiviteten följer av att elementen i $\mathfrak{U}$ har egenskapen att vara ”stora” (se 2.2):

$$\begin{aligned} &\langle r_n \rangle \equiv \langle s_n \rangle \wedge \langle s_n \rangle \equiv \langle t_n \rangle \Rightarrow E_{rs}, E_{st} \in \mathfrak{U} \\ &E_{rs} \cap E_{st} \subseteq E_{rt} \Rightarrow E_{rt} \in \mathfrak{U} \Rightarrow \langle r_n \rangle \equiv \langle t_n \rangle. \end{aligned}$$

$\square$

För att få en lite mer flexibel notation kommer vi beteckna mängden $\{n \in \mathbb{N} : r_n = s_n\}$ med $[\![r = s]\!]$ istället för E_{rs}. Denna notation kan även användas för mängder som mäter hur mycket två följder överensstämmer med andra relationer än "=", exempelvis:

$$
\begin{aligned}
[\![r < s]\!] &= \{n \in \mathbb{N} : r_n < s_n\}, \\
[\![r > s]\!] &= \{n \in \mathbb{N} : r_n > s_n\}, \\
[\![r \leq s]\!] &= \{n \in \mathbb{N} : r_n \leq s_n\},
\end{aligned}
$$

och så vidare.

2.4 Ekvivalensklasser av reellvärda följder

Ekvivalensklassen för en följd $r \in \mathbb{R}^{\mathbb{N}}$ under $\equiv$ kommer betecknas $[r]$, det vill säga:

$$[r] = \left\{s \in \mathbb{R}^{\mathbb{N}} : r \equiv s\right\}.$$

En kvotmängd är en mängd av ekvivalensklasser. Kvotmängden för $\mathbb{R}^{\mathbb{N}}$ under $\equiv$ är:

$$^*\mathbb{R} = \left\{[r] : r \in \mathbb{R}^{\mathbb{N}}\right\}.$$

Det är mängden $^*\mathbb{R}$ som vi kommer referera till som *de hyperreella talen*. Vi vill nu definiera operationer och ordning på ekvivalensklasserna. Vi börjar med att definiera operationerna addition och multiplikation på följder:

$$
\begin{aligned}
r \oplus s &= \langle r_n + s_n : n \in \mathbb{N}\rangle \\
r \odot s &= \langle r_n \cdot s_n : n \in \mathbb{N}\rangle
\end{aligned}
$$

För ekvivalensklasserna definierar vi nu:

$$
\begin{aligned}
[r] + [s] &= [r \oplus s] = [\langle r_n + s_n\rangle], \\
[r] \cdot [s] &= [r \odot s] = [\langle r_n \cdot s_n\rangle],
\end{aligned}
$$

och

$$[r] < [s] \quad \text{omm} \quad [\![r < s]\!] \in \mathfrak{U} \quad \text{omm} \quad \{n \in \mathbb{N} : r_n < s_n\} \in \mathfrak{U}.$$

Dessa notationer är väldefinierade, det vill säga oberoende representanterna.

Om definierar $\mathbf{0} = \langle 0, 0, 0, \cdots \rangle$ och $\mathbf{1} = \langle 1, 1, 1, \cdots \rangle$, så har vi följande viktiga sats som motiverar varför vi kan räkna med hyperreella tal.

Sats 2.2. *Strukturen $\langle {}^*\mathbb{R}, +, \cdot, < \rangle$ är en ordnad kropp med nolla $[\mathbf{0}]$ och etta $[\mathbf{1}]$*

Man kan bevisa detta med den så kallade ”transferprincipen” som introduceras i kapitel 8.4. Man kan även visa det direkt från definitionen av de hyperreella talen, något som vi inte tänker göra här.

Vi kan identifiera ett reellt tal $r \in \mathbb{R}$ med den konstanta följden $\boldsymbol{r} = \langle r, r, r, \cdots \rangle$, vilket gör att vi kan placera den i ${}^*\mathbb{R}$:

$$ {}^*r \quad = \quad [\boldsymbol{r}]. $$

Nästa sats kommer vi inte bevisa, men den följer ganska tydligt från definitionerna. Även denna sats skulle man kunna bevisa med transferprincipen.

Sats 2.3. *Avbildningen $r \mapsto {}^*r$ är en ordningsbevarande ringmonomorfi från $\mathbb{R}$ till ${}^*\mathbb{R}$.*

Vet man inte vad en ringmonomorfi är, så räcker det här att veta att denna sats tillåter oss att identifiera alla reella r med ${}^*r \in {}^*\mathbb{R}$. Detta leder i sin tur till att vi kan betrakta $\mathbb{R}$ som en delkropp till ${}^*\mathbb{R}$.

2.5 Infinitesimala och obegränsade tal

Låt $\epsilon = \langle \frac{1}{n} : n \in \mathbb{N} \rangle$. Vi har att:

$$ [\![\mathbf{0} < \epsilon]\!] \quad = \quad \left\{ n \in \mathbb{N} : 0 < \frac{1}{n} \right\} \quad = \quad \mathbb{N} \in \mathfrak{U}, $$

det vill säga $[\mathbf{0}] < [\epsilon]$ i ${}^*\mathbb{R}$. Men vi har även att om r är ett godtyckligt positivt *reellt* tal, så är:

$$ [\![\epsilon < \boldsymbol{r}]\!] \quad = \quad \left\{ n \in \mathbb{N} : \frac{1}{n} < r \right\} \in \mathfrak{U} $$

Komplementet till mängden ovan — med avseende på $\mathbb{N}$ — är ändlig, och $[\![\epsilon < \boldsymbol{r}]\!]$ tillhör därför $\mathfrak{U}$ eftersom mängder med den egenskapen är stora. Vi har alltså att $[\epsilon] < {}^*r$ i ${}^*\mathbb{R}$ för alla positiva $r \in \mathbb{R}$, och alltså uppfyller $[\epsilon]$ storleksegenskaperna vi vill att en positiv infinitesimal ska ha.

Låt nu $\omega = \langle 1, 2, 3, \cdots \rangle$. Vi har nu för godtyckligt $r \in \mathbb{R}$ att:

$$ [\![\omega > \boldsymbol{r}]\!] \quad = \quad \{ n \in \mathbb{N} : n > r \} \in \mathfrak{U}, $$

där inklusionen i $\mathfrak{U}$ beror på samma argument som ovan. $[\omega]$ uppfyller alltså egenskaperna för att vara obegränsad.

Vi har även att:

$$[\epsilon] \cdot [\omega] \quad = \quad [\mathbf{1}],$$

så $[\omega] = [\epsilon]^{-1}$ och $[\epsilon] = [\omega]^{-1}$.

Med ramverket som vi byggt upp kan vi alltså rigoröst konstruera infinitesimala- och obegränsade tal.

3 Hyperreella tal

Detta kapitel utgår framförallt från [Goldblatt 1998].

3.1 Grundläggande definitioner

Definition 3.1. *Ett hyperreellt tal b är:*

- *Begränsat om $r < b < s$ för något $r,s \in \mathbb{R}$.*
- *Positivt obegränsat om $r < b$ för alla $r \in \mathbb{R}$.*
- *Negativt obegränsat om $b < r$ för alla $r \in \mathbb{R}$.*
- *Obegränsat om den är antingen positivt- eller negativt obegränsat.*
- *Positivt infinitesimalt om $0 < b < r$ för alla positiva $r \in \mathbb{R}$.*
- *Negativt infinitesimalt om $r < b < 0$ för alla negativa $r \in \mathbb{R}$.*
- *Infinitesimalt om den är positivt infinitesimalt, negativt infinitesimalt eller 0.*
- *Väsentligt om den är begränsad men inte infinitesimalt.*

I kapitel 8 konstruerar vi med hjälp av mängdlära samtliga mateamtiska objekt för $^*\mathbb{R}$ på ett mer rigoröst — men mindre intuitivt — sätt. För att ge läsaren en bättre intuition om vad som händer när man *utvidgar* ett objekt från $\mathbb{R}$ till $^*\mathbb{R}$, följer här två specifika definitioner om utvigningar av följder och delmängder till $\mathbb{R}$. Dessa utvidgningar kommer vi att flitigt använda i nästa kapitel.

Definition 3.2. *Utvidningen av godtyckliga delmängder $A \subset \mathbb{R}$ till $^*A$ fås av att för varje $r \in \mathbb{R}^{\mathbb{N}}$ sätta:*

$$[r] \in {}^*A \quad \Leftrightarrow \quad [\![n \in \mathbb{N}]\!] = \{n \in \mathbb{N} : r_n \in A\} \in \mathfrak{U}$$

Definition 3.3. *Utvidningen av talföljden $s : \mathbb{N} \mapsto \mathbb{R}$ till $^*s : {}^*\mathbb{N} \mapsto {}^*\mathbb{R}$ fås av att för varje $n \in \mathbb{R}^{\mathbb{N}}$ där $[n] \in {}^*N$ definiera en annan följd r som*

$$r_k \doteqdot \begin{cases} s_{n_k} & \textit{om } k \in [\![n \in \mathbb{N}]\!] \\ 0 & \textit{om } k \notin [\![n \in \mathbb{N}]\!] \end{cases}$$

och därefter sätta $^*s_n = [r]$.

Vi betecknar mängden av alla obegränsade naturliga tal, det vill säga $^*\mathbb{N} \setminus \mathbb{N}$, som $^*\mathbb{N}_\infty$. På samma sätt betecknar vi mängden av alla obegränsade hyperreella tal som $^*\mathbb{R}_\infty$.

3.2 Halor och skuggor

Definition 3.4.

- *Ett hyperreellt tal b är oändligt nära ett annat hyperreellt tal c om $b - c = \epsilon$ är infinitesimalt. Detta är en ekvivalensrelation på ${}^*\mathbb{R}$ och betecknas $b \simeq c$. Ekvivalensklassen till relationen kallas* halo, *och halon där b ingår betecknas $hal(b)$.*

- *Om $r \in hal(b)$ är ett reellt tal så säger vi att r är en* skugga *eller* standarddel *av b.*

Nästa sats visar en viktig princip om ordningen på de hyperreella talen. Som vi skall se så räcker det med att visa att något element ur $hal(b)$ är mindre än eller lika med något element ur $hal(c)$ för att man skall kunna dra slutsatsen att $b \leq c$.

Sats 3.5. *Om vi för $b, c \in \mathbb{R}$ och $x, y \in {}^*\mathbb{R}$ har att $c \simeq y \leq x \simeq b$, då gäller det att: $b \leq c$.*

Bevis. Vi har $b = x + \epsilon_b$ och $c = y + \epsilon_c$ där $\epsilon_b, \epsilon_c \simeq 0$. Om vi antar, för att få en motsägelse, att $b > c$, så är $b - c$ ett positivt reellt tal. Men:

$$b - c = (x - y) + (\epsilon_b - \epsilon_c)$$

Där den första parentesen är mindre än eller lika med 0 och den andra ett infinitesimalt tal. Summan av dessa två paranter kan således inte vara ett positivt reellt tal och därmed har vi en motsägelse. □

Nästa sats motiverar varför vi kan betrakta skuggan som en funktion $sh(\cdot) : {}^*\mathbb{R} \setminus {}^*\mathbb{R}_\infty \to \mathbb{R}$.

Sats 3.6. *Alla begränsade hyperreella b har exakt en skugga r, som vi betecknar $sh(b)$.*

Bevis. Låt $A = \{r \in \mathbb{R} : r < b\}$. Enligt defintionen av att vara begränsad finns det reella tal r, s s.a. $r < b < s$, alltså har mängden A en majorant s vilket implicerar att det finns ett supremum $c \in \mathbb{R}$ till A (enligt supremumegenskapen hos $\mathbb{R}$). Detta c är den unika skuggan av b.

Vi visar först $b \simeq c$. Tag ett $\epsilon \in \mathbb{R}^+$. Vi har att $c + \epsilon \notin A$, vilket ger $b \leq c + \epsilon$ enligt definitionen av A. Om vi hade att $b \leq c - \epsilon$ skulle inte c vara den minsta majoranten till A, vilket ger $c - \epsilon < b \leq c + \epsilon$ som i sin tur ger $|b - c| \leq \epsilon$. Detta håller för alla reella ϵ vilket ger $b \simeq c$.

Om vi nu antar att $b \simeq c' \in \mathbb{R}$, så får vi $c \simeq c'$ vilket ger $c = c'$ eftersom 0 är det enda reella tal som är infinitesimalt. □

3.3 Aritmetik för skuggor

På grund av sats 3.6 kommer vi i kommande kapitel framförallt använda oss av skugg-notationen. Följande sats ger oss användbar aritmetik för skuggor.

Sats 3.7. *Om b och c är begränsade hyperreella tal och $n \in \mathbb{N}$, då har vi:*

(i) $sh(b \pm c) = sh(b) \pm sh(c)$,

(ii) $b \cdot c = sh(b) \cdot sh(c)$,

(iii) $sh(b/c) = sh(b)/sh(c)$ *om* $sh(c) \neq 0$ *(d.v.s. om x är väsentlig),*

(iv) $sh(b^n) = sh(b)^n$,

(v) $sh(|b|) = |sh(b)|$,

(vi) $sh(\sqrt[n]{b}) = \sqrt[n]{sh(b)}$ *om* $b \geq 0$,

(vii) *om* $b \leq c$ *då är* $sh(b) \leq sh(c)$.

Bevis. (i) till (vi) ovan följer direkt från aritmetiken av de hyperreella talen som i sin tur följer driekt från hur vi konstruerat dem. Exempelvis har vi för additionsfallet i (i), om vi betecknar $b = sh(b) + \epsilon_b$ och $c = sh(c) + \epsilon_c$ där $\epsilon_b, \epsilon_c \simeq 0$:

$$sh(b + c) = sh(sh(b) + sh(c) + \epsilon_b + \epsilon_c) = sh(sh(b) + sh(c)) = sh(b) + sh(c)$$

(vii) följer direkt från sats 3.5. $\square$

4 Konvergens

Detta kapitel utgår framförallt från [Goldblatt 1998].

4.1 Konvergens, begränsning och divergens av följder

Vi använder oss av *transferprincipen* i detta kapitel. En formell beskrivning och bevis av transferprincipen finns i kapitel 8.4. I detta kapitel kommer vi enbart använda oss av transferprincipen som argument för att om ett påstående gäller för $\mathbb{N}$, så kommer det även gälla för ${}^*\mathbb{N}$.

I standard analys säger vi att en följd $\langle s_n : n \in \mathbb{N} \rangle$ är ***konvergent till gränsvärdet*** $L \in \mathbb{R}$ om varje godtyckligt litet intervall runt L innefattar en svans av talföljden (där svans definieras som följden $\langle s_n : n > a \rangle$ för något $a \in \mathbb{N}$). Formellt skriver vi det:

$$(\forall \epsilon \in \mathbb{R}^+)(\exists m_\epsilon \in \mathbb{N})(\forall n \in \mathbb{N})(n > m_\epsilon \Rightarrow |s_n - L| < \epsilon)$$

En naturlig icke-standard formulering av konvergens är — löst beskrivet — att alla tal oändligt långt fram i talföljden är oändligt nära ett och samma gränsvärde (det vill säga de är begränsade och ligger i samma halo). Följande sats motiverar denna formulering:

Sats 4.1. *En reellvärd talföljd $\langle s_n : n \in \mathbb{N} \rangle$ konvergerar till $L \in \mathbb{R}$ om och endast om $s_n \simeq L$ för alla obegränsade n.*

Bevis. Antag först att $\langle s_n : n \in \mathbb{N} \rangle$ konvergerar till L, och välj ett fixt $N \in {}^*\mathbb{N}_\infty$. Vi skall visa att $|s_N - L| < \epsilon$ för alla $\epsilon \in \mathbb{R}^+$, eftersom detta implicerar att $s_N \simeq L$. Givet ett $\epsilon \in \mathbb{R}^+$, så implicerar standard formuleringen av konvergens att det finns ett $m_\epsilon \in \mathbb{N}$ så att standard svansen efter s_{m_ϵ} är inom ϵ av L:

$$(\forall n \in \mathbb{N})(n > m_\epsilon \Rightarrow |s_n - L| < \epsilon)$$

Med transfer så gäller detta påstående även för ***utökade*** svansar (det vill säga $\langle s_n : n > a \rangle$ för något $a \in {}^*\mathbb{N}_\infty$):

$$(\forall n \in {}^*\mathbb{N})(n > m_\epsilon \Rightarrow |s_n - L| < \epsilon) \tag{3}$$

Eftersom N är obegränsad och m_ϵ begränsad, så har vi att $N > m_\epsilon$. Enligt (3) har vi då $|s_N - L| < \epsilon$ som önskat.

För det omvända fallet, anta att $s_n \simeq L$ för alla obegränsade n. För att visa konvergens använder vi att elementen i den utökade svansen är oändligt nära L och sedan applicerar vi transfer.

Välj ett fixt $N \in {}^*\mathbb{N}_\infty$. För alla element n i mängden $\{n \in {}^*\mathbb{N} : n > N\}$ gäller det att de är obegränsade, eftersom N är obegränsad. Alltså gäller det att $|s_n - L| < \epsilon$ givet ett fixt $\epsilon \in \mathbb{R}^+$. Detta visar att följande påstående är sant:

$$(\exists z \in {}^*\mathbb{N})(\forall n \in {}^*\mathbb{N})(n > z \Rightarrow |s_n - L| < \epsilon).$$

Men med transferprincipen får vi även att

$$(\exists z \in \mathbb{N})(\forall n \in \mathbb{N})(n > z \Rightarrow |s_n - L| < \epsilon)$$

är ett sant påstående, vilket i sin tur visar konvergensen eftersom ϵ kan väljas godtyckligt. □

Icke-standard formuleringen av konvergens är minst lika användbar som standard formuleringen. Vi tillämpar icke-standard formuleringen för att bevisa *satsen om monoton konvergens*:

Sats 4.2. *En reellvärd följd $\langle s_n : n \in \mathbb{N} \rangle$ konvergerar i $\mathbb{R}$ om den antingen är:*

(i) uppåt begränsad i $\mathbb{R}$ och icke-minskande: $s_1 \leq s_2 \leq \cdots$

(ii) nedåt begränsad i $\mathbb{R}$ och icke-ökande: $s_1 \geq s_2 \geq \cdots$.

Bevis. Vi kommer enbart visa (i). Låt s_N vara en utökad term. Vi kommer att visa att s_N har en skugga, och att denna skugga är supremum av $\{s_n : n \in \mathbb{N}\}$ i $\mathbb{R}$. Eftersom supremum är unikt, så har vi att alla utökade termer har samma skugga, och med sats 4.1 får vi då att $\langle s_n \rangle$ är konvergent.

Enligt antagandet finns det ett reellt tal b som är en majorant för $\{s_n : n \in \mathbb{N}\}$. Vi har nu att

$$s_1 \leq s_n \leq b, \qquad \forall n \in \mathbb{N} \tag{4}$$

Med transfer gäller (4) även för alla $n \in {}^*\mathbb{N}$, vilket ger att s_N har en skugga L. Detta L är en övre gräns till $\langle s_n \rangle$; eftersom följden är icke-minskande, så har vi med transfer att

$$n \leq m \Rightarrow s_n \leq s_m, \qquad \forall n, m \in {}^*\mathbb{N}.$$

Detta ger oss att $s_n \leq s_N \simeq L$ om $n \in \mathbb{N}$, vilket i sin tur ger att $s_n \leq L$.

För att visa att L är den *minsta* majoranten, betraktar vi godtycklig övre reell gräns r av $\{s_n : n \in \mathbb{N}\}$. Med transfer har vi att $s_n \leq r$ för alla $n \in {}^*\mathbb{N}$, vilket ger att $L \simeq s_N \leq r$ som i sin tur ger att $L \leq r$. □

Här följer ett par satser som visar hur begränsning och divergens av följder kan karaktäriseras med icke-standard analys. Vi tänker inte bevisa dessa.

Sats 4.3. *En reellvärd följd $\langle s_n \rangle$ är begränsad i $\mathbb{R}$ om och endast om alla dess* utökade *termer, det vill säga element i den utökade svansen, är begränsade.*

Mer specifikt är följden övre begränsad om och endast om den inte har några positivt obegränsade utökade termer, och undre begränsad om och endast om den inte har några negativt obegränsade utökade termer.

Sats 4.4. *En reellvärd följd $\langle s_n \rangle$*

- *Divergerar mot ∞ om och endast om alla dess utökade termer är positivt obegränsade.*
- *Divergerar mot $-\infty$ om och endast om alla dess utökade termer är negativt obegränsade.*

4.2 Cauchyföljder och hopningspunkter

Standarddefinitionen av cauchyföljder är alla följder där termerna närmar sig varandra, det vill säga $\lim_{m,n\to\infty} |s_n - s_m| = 0$. Mer formellt skriver vi detta som:

$$(\forall \epsilon \in \mathbb{R}^+)(\exists j \in \mathbb{N})(\forall m, n \in \mathbb{N})(m, n \geq j \Rightarrow |s_m s_n| < \epsilon).$$

Nästa sats är ett ekvivalent icke-standard vilkor för cauchyföljder. Vi lämnar satsen obevisad.

Sats 4.5. *En reellvärd följd $\langle s_n \rangle$ är en cauchyföljd i $\mathbb{R}$ om och endast om alla dess utökade termer är oändligt nära varandra, det vill säga om och endast om $s_m \simeq s_n$ för alla $m, n \in {}^*\mathbb{N}_\infty$*

Ett reellt tal L sägs vara en *hopningspunkt* till den reellvärda följden $\langle s_n : n \in \mathbb{N} \rangle$ om varje öppet intervall runt L inehåller oändligt många termer från följden. Formellt skrivs detta som påståendet:

$$(\forall \epsilon \in \mathbb{R}^+)(\forall m \in \mathbb{N})(\exists n \in \mathbb{N})(n > m \land |s_n - L| < \epsilon). \tag{5}$$

Sats 4.6. *En reellvärd följd $\langle s_n : n \in \mathbb{N} \rangle$ har $L \in \mathbb{R}$ som hopningspunkt om och endast om följden har åtminstone en utökad term oändligt nära L, det vill säga om och endast om $s_N \simeq L$ för något obegränsat N.*

Bevis. Antag att (5) är sann. Låt ϵ vara ett positivt infinitesimalt tal och $m \in {}^*\mathbb{N}_\infty$. Med transfer av (5) får vi att det finns något $n \in {}^*\mathbb{N}$ där $n > m$ (vilket betyder att n är obegränsad) och

$$|s_n - L| < \epsilon \simeq 0.$$

Vilket visar att s_n är en utökad term där $s_n \simeq L$.

Om vi istället antar att det finns ett obegränsat N så att $s_N \simeq L$. Vi tar godtyckligt $\epsilon \in \mathbb{R}^+$ och $m \in \mathbb{N}$, vilket ger oss att $N > m$ och $|s_N - L| < \epsilon$. Detta visar att:

$$(\exists n \in {}^*\mathbb{N})(n > M \wedge |s_n - L| < \epsilon).$$

Med transfer får vi att $|s_n - L| < \epsilon$ för något $n \in \mathbb{N}$ där $n > m$. Detta visar (5) eftersom ϵ och m var godtyckligt valda. □

Denna sats ger ett intuitivt icke-standard bevis till Bolzan-Weierstrass sats: Enligt sats 4.3 är mängden av begränsade tal bland de utökad termerna i en begränsad talföljd icke-tom, och enligt föregående sats har vi att skuggorna av dessa begränsade utökade termer är hopningspunkter. Vi har alltså visat följande:

Sats 4.7 (Bolzano-Weierstrass). *Varje begränsad följd av reella tal har åtminstone en hopningspunkt i $\mathbb{R}$.* [2]

4.3 Serier

En reell oändlig serie $\sum_1^\infty a_i$ är konvergent om och endast om talföljden $s = \langle s_n : n \in \mathbb{N} \rangle$ av delsummor

$$s_n = 1_1 + \cdots + a_n$$

är konvergent. Vi skriver $\sum_1^n a_i$ för s_n, och $\sum_m^n a_i$ för $s_n - s_{m-1}$ när $n \geq m$. Om vi utökar s till en hyperreell talföljd $s = \langle s_n : n \in {}^*\mathbb{N} \rangle$, så får uttrycken $\sum_1^n a_i$ och $\sum_m^n a_i$ betydelse för alla $n, m \in {}^*\mathbb{N}$ och kan därför betraktas som *hyperändliga summor* när n är obegränsad.

Om vi tillämpar vår tidigare resultat om konvergens av utökade talföljder så får vi följande:

[2] Notera att vissa kursböcker formulerar Bolzano-Weierstrass på ett ekvivalent sätt där satsen säger att varje begränsad talföljd har en konvergent delföljd.

- $\lim_{n\to\infty} \sum_1^n a_i = \sum_1^\infty a_i = L \in \mathbb{R}$ omm $\sum_1^n a_i \simeq L$ för alla obegränsade n

- $\sum_1^\infty a_i$ konvergerar i $\mathbb{R}$ omm $\sum_m^n a_i \simeq 0$ för alla obegränsade m, n där $m \leq n$

4.4 Limes supremum och limes infimum

Not 4.1. *Detta delkapitel innehåller teori som inte kommer att användas senare i arbetet, men demonstrerar på ett bra sätt hur icke-standard analys kan tillämpas för att få en intuitiv beskrivning av matematiska objekt.*

Följder som är divergenta men begränsade har liknande egenskaper som de konvergenta följderna. Bolzano-Weiertrass sats garanterar att vi har åtminstone en hopningspunkt för sådana följder, och sats 4.6 ger oss möjlighet att betrakta mängden av dessa hopningspunkter som mängden av alla skuggor av de utökade termerna:

$$C_s = \{sh(s_n) : n \in {}^*\mathbb{N}_\infty\}$$

Där $s = \langle s_n : n \in \mathbb{N}\rangle$ är en begränsad följd.

En reell övre/undre gräns till följden s är även en övre/undre gräns till C_s, enligt proposition 3.5. Detta tillsammans med supremumegenskapen hos $\mathbb{R}$ ger oss att den reella mängden C_s måste ha ett supremum u och infimum l (eftersom s är begränsad). Vi kallar u för *limes supremum* och l för *limes infimum*, och vi använder oss av dessa två beteckningar:

$$u = \limsup_{n\to\infty} s_n = \overline{lim}\, s \quad \text{och} \quad l = \liminf_{n\to\infty} s_n = \underline{lim}\, s.$$

I själva verket är $\limsup_{n\to\infty} s_n$ och $\liminf_{n\to\infty} s_n$ element i C_s, och därmed alltså maximum respektive minimum i mängden.

Här följer några satser om limes supremum och infimum utan bevis:

Sats 4.8. *Ett reellt tal L är lika med $\overline{lim}\, s$ om och endast om*

- *$s_n < L$ eller $s_n \simeq L$ för alla obegränsade n; och*

- *$s_n \simeq L$ för åtminstone ett obegränsat n.*

Motsvarande gäller även för $\underline{lim}\, s$.

Sats 4.9. *En begränsad reellvärd talföljd s konvergerar mot $L \in \mathbb{R}$ om och endast om*

$$\limsup_{n\to\infty} s_n = \liminf_{n\to\infty} s_n = L$$

Sats 4.10. *Om s är en begränsad reellvärd talföljd med limes sumpremum $\overline{lim}$, givet godtyckligt $\epsilon \in \mathbb{R}^+$ har vi att:*

- $s_n < \overline{lim} + \epsilon$ *För alla utom ändligt många $n \in \mathbb{N}$.*
- $\overline{lim} - \epsilon < s_n$ *För oändligt många $n \in \mathbb{N}$.*

5 Ultrafilter

Framställningen i detta kapitel följer Hurd och Loebs 'Introduction to non-standard analysis' och material har även inhämtats från Felix Mendelsons 'Introduction to mathematical logic'. Filterbegreppet, rent allmänt, är emellertid vida spritt.

5.1 Konstruktion av fria ultrafilter

Då vi konstruerade de hyperreella talen i kapitel 2 förutsatte vi att det existerade så kallade 'icke-principala ultrafilter'. Dessa antogs vara mängder som innehöll alla 'stora' enighetsmängder. Dessa mängder var i själva verket oändliga delmängder till $\mathbb{N}$. Detta kapitel har som mål att visa att icke-principala ultrafilter existerar på $\mathbb{N}$ och på så vis färdigställa konstruktionen av de hyperreella talen ${}^*\mathbb{R}$.

Denna konstruktion av ${}^*\mathbb{R}$ är emellertid endast ett specifikt exempel på en allmän kontruktionsmetod som vi i senare kapitel kommer definiera som 'ultraprodukt'. Filterbegreppet kommer då att användas analogt med hur det användes i konstruktionen av de hyperreella talen även för konstruktionen av andra ultraprodukter.

Informellt kan vi använda oss av filter som ett verktyg för att formalisera begreppet 'stor delmängd'. Några av de egenskaper vi förväntar oss att en stor delmängd skall ha presenterades i samband med konstruktionen av de hyperreella talen i kapitel 2, dessa återspeglas nu i definition 5.2 nedan.

Definition 5.1. *Mängden av alla delmängder till en godtycklig mängd I, även känd som potensmängden till I, benämns P(I) $\doteqdot \{A\colon A \subseteq I\}$.*

Definition 5.2. *Ett* **filter** $\mathfrak{F}$ *på* $I \neq \varnothing$ *är en delmängd till P(I), alltså en mängd av delmängder till I, som uppfyller följande:*

(i) $\mathrm{A} \in \mathfrak{F} \wedge \mathrm{B} \in \mathfrak{F} \Rightarrow \mathrm{A} \cap B \in \mathfrak{F}$ — $\mathfrak{F}$ är sluten under snitt.

(ii) $\mathrm{A} \in \mathfrak{F} \wedge \mathrm{A} \subseteq \mathrm{B} \Rightarrow \mathrm{B} \in \mathfrak{F}$ — $\mathfrak{F}$ är sluten under övre delmängder

Ett filter $\mathfrak{F}$ är ett **ultrafilter** om det även uppfyller:

(iii) För alla $X \subseteq \mathrm{I}$ gäller antingen $X \in \mathfrak{F}$ eller $X^c \in \mathfrak{F}$, där alltså X^c betecknar komplementet till X med avseende på I, d.v.s $X^c \doteqdot I \setminus X$

Ett filter som innehåller ändliga delmängder till I kallas principalt. Icke-principala filter kallas även fria. Vidare kallas ett filter $\mathfrak{F}$ skilt från potensmängden till I, alltså $\mathfrak{F} \neq P(I)$, för ett äkta filter. Detta är ekvivalent med att $\emptyset \notin \mathfrak{F}$ vilket är en enkel konsekvens av (ii) eftersom $\emptyset \subseteq A$ för alla $A \subseteq I$.

Definition 5.3. *Vi säger att ett äkta filter $\mathfrak{F}$ på I är maximalt om för alla äkta filter $\mathfrak{G}$ på I sådana att $\mathfrak{F} \subseteq \mathfrak{G}$ gäller att $\mathfrak{F} = \mathfrak{G}$.*

Sats 5.4. *Ett äkta filter $\mathfrak{F}$ på I är ett ultrafilter omm det är maximalt.*

Bevis.

Antag alltså att $\mathfrak{F}$ är ett ultrafilter på I, enligt definition gäller då att för alla $X \subseteq I$ antingen $X \in \mathfrak{F}$ eller $X^c \in \mathfrak{F}$.

Antag vidare, för ett motsägelsebevis, att $\mathfrak{F} \subseteq \mathfrak{G}$ för något äkta filter $\mathfrak{G}$ på I och att det finns ett $X \subseteq I$ sådant att $X \in \mathfrak{G}$ men $X \notin \mathfrak{F}$. Då har vi att $X^c \in \mathfrak{F}$, följaktligen är $X^c \in \mathfrak{G}$ och även $X^c \cap X = \emptyset \in \mathfrak{G}$. Men då är $\mathfrak{G}$ inte ett äkta filter vilket motsäger antagandet att $\mathfrak{G}$ är ett äkta filter, alltså är $\mathfrak{F}$ maximalt.

Antag, för den motsatta implikationen, att $\mathfrak{F}$ är maximalt och att $X \notin \mathfrak{F}$ för något $X \subseteq$ I.

Låt $\mathfrak{G} = \{A \subseteq \mathrm{I} : X \cap F_i \subseteq A,$ för alla $F_i \in \mathfrak{F}\}$. $\mathfrak{G}$ är alltså en utvidgning av $\mathfrak{F}$ som inkluderar X och som enligt definition är sluten under snitt och övre delmängd m.a.p X, det är alltså klart att $\mathfrak{F} \subseteq \mathfrak{G}$ men att $\mathfrak{F} \neq \mathfrak{G}$ då $X \in \mathfrak{G}$. Eftersom vi har antagit att $\mathfrak{F}$ är ett maximalt äkta filter följer att $\mathfrak{G}$ inte kan vara ett äkta filter, det följer att $\emptyset \in \mathfrak{G}$, d.v.s att $X \cap F_i = \emptyset$ för något $F_i \in \mathfrak{F}$. Detta är ekvivalent med $X^c \cap F_i = F_i$, alltså är $F_i \subseteq X^c$ så enligt (ii) är $X^c \in \mathfrak{F}$. Alltså är $\mathfrak{F}$ ett ultrafilter.

$\square$

Vi vill nu visa att alla äkta filter $\mathfrak{F}$ på I är delmängder till något ultrafilter $\mathfrak{U}$. För att visa detta kommer vi behöva urvalsaxiomet i form av Zorns lemma. Detta lemma tar vi som ett axiom om partiella ordningar.

Definition 5.5. *En partiellt ordnad mängd $(X, \leq)$ är ett ordnat par där X är en icke-tom mängd och $\leq$ är en binär relation på X sådan att:*

(i) $\leq$ är reflexiv, d.v.s $x \leq x$, för alla $x \in X$

(ii) $\leq$ är antisymmetrisk, d.v.s $x \leq y$ och $y \leq x$ omm $x = y$

(iii) $\leq$ *är transitiv, d.v.s om* $x \leq y$ *och* $y \leq z$ *så* $x \leq z$

Vi säger att $\leq$ är en partiell ordning på X om (i-iii) gäller och att $\leq$ är en **total** ordning på X om dessutom

(iv) $\forall x, y \in X(x \leq y \vee y \leq x)$ *gäller.*

Definition 5.6. *Om* $(X, \leq)$ *är en partiellt ordnad mängd så kallar vi varje totalt ordnad delmängd* $(K, \leq)$*, där alltså* $K \subseteq X$ *för en* **kedja**.

Vi säger vidare att x är en **övre begränsning** på en delmängd $B \subseteq X$ om $\forall b \in B(b \leq x)$ och att $m \in X$ är **maximalt** om $\forall x \in X(m \leq x \Rightarrow x = m)$.

Zorn's Lemma: *Antag att* $(X, \leq)$ *är en partiellt ordnad mängd. Om varje kedja i* $(X, \leq)$ *har en övre begränsning så har* X *ett* $\leq$*-maximalt element.*

Sats 5.7. Ultrafiltersatsen *För varje äkta filter* $\mathfrak{F}$ *på* I *finns ett ultrafilter* $\mathfrak{U}$ *sådant att* $\mathfrak{F} \subseteq \mathfrak{U}$*.*

Bevis. Vi antar alltså Zorns lemma.

Antag att $\mathfrak{F}$ är ett äkta filter på I. Låt $\mathfrak{F}^{\mathfrak{s}}$ vara mängden av alla äkta filter som innehåller $\mathfrak{F}$ och låt $\leq$ definiera en partiell ordning på $\mathfrak{F}^{\mathfrak{s}}$ enligt $\mathfrak{A} \leq \mathfrak{B}$ omm $A \in \mathfrak{A} \Rightarrow A \in \mathfrak{B}$. (Det är trivialt att verifiera att $\leq$ bildar en partiell ordning på I).

Antag att $\mathfrak{C}^{\mathfrak{s}}$ är en kedja i $\mathfrak{F}^{\mathfrak{s}}$. Låt $\mathfrak{F}^{\mathfrak{t}} = \bigcup_{\mathfrak{C} \in \mathfrak{C}^{\mathfrak{s}}} \mathfrak{C}$, med andra ord låter vi $\mathfrak{F}^{\mathfrak{t}}$ vara en mängd som innehåller elementen från alla mängder i kedjan. Då är $\mathfrak{C} \leq \mathfrak{F}^{\mathfrak{t}}$. Vidare är $\mathfrak{F}^{\mathfrak{t}}$ är ett filter ty:

(i) Antag $A, B \in \mathfrak{F}^{\mathfrak{t}}$. Då har vi att $A \in \mathfrak{C}_1$ och $B \in \mathfrak{C}_2$ för några $\mathfrak{C}_1, \mathfrak{C}_2$ i $\mathfrak{C}^{\mathfrak{s}}$. Eftersom $\mathfrak{C}^{\mathfrak{s}}$ är en kedja ordnad av $\leq$ kan vi anta att $\mathfrak{C}_1 \leq \mathfrak{C}_2$. Då får vi att $A, B \in \mathfrak{C}_2$ enligt definitionen av $\leq$. Eftersom $\mathfrak{C}_2$ är ett filter har vi att $A \cap B \in \mathfrak{C}_2$ och eftersom $\mathfrak{C}_2 \subseteq \mathfrak{F}^{\mathfrak{t}}$ har vi att $A \cap B \in \mathfrak{F}^{\mathfrak{t}}$.

(ii) Antag att $A \in \mathfrak{F}^{\mathfrak{t}}$ och att $A \subseteq B$. P.s.s har vi att $A \in \mathfrak{C}$ för något filter $\mathfrak{C}$. Då är $B \in \mathfrak{C}$ och så följaktligen $B \in \mathfrak{F}^{\mathfrak{t}}$.

Att $\mathfrak{F}^{\mathfrak{t}}$ är ett äkta följer av motsvarande argument, d.v.s om $\emptyset \in \mathfrak{F}_{\mathfrak{t}}$ har vi att $\emptyset \in \mathfrak{C}$ för något $\mathfrak{C}$, o.s.v.

Vi har alltså att varje kedja i $\mathfrak{F}^{\mathfrak{s}}$ av filter som innesluter $\mathfrak{F}$ har en övre begränsning så enligt Zorns lemma finns ett $\leq$-maximalt element. Vi har redan visat (sats 5.4) att ett äkta filter är maximalt omm det är ett ultrafilter så vi är färdiga. $\square$

Vi vill visa att det finns fria ultrafilter tillgängliga för vår konstruktion av de hyperreella talen, d.v.s att det finns ett fritt ultrafilter på $\mathbb{N}$ och mer generellt att det finns fria ultrafilter på alla oändliga mängder vilket kommer att möjliggöra även andra analoga konstruktioner. Vi börjar med att introducera Fréchetfiltret vilket kommer att användas som ett led i att visa detta.

Definition 5.8. Fréchetfiltret: *För en godtycklig mängd I består Fréchetfiltret $\mathfrak{F}^{\mathfrak{co}}$ av komplementen till alla ändliga delmängder till I. Fréchetfiltret kallas även ibland för co-finita filtret.*

$\mathfrak{F}^{\mathfrak{co}} \doteqdot \{X \subseteq I : X^c \text{ är ändlig}\}$.

Sats 5.9. *För oändliga mängder I är Fréchetfiltret ett fritt äkta filter på I.*

Bevis.

(i) Om A^c och B^c är ändliga så är $A^c \cup B^c = (A \cap B)^c$ ändlig, och därför $A \cap B \in \mathfrak{F}^{\mathfrak{co}}$

(ii) $A \subseteq B \Leftrightarrow B^c \subseteq A^c$ så om A^c är ändlig är B^c ändlig. D.v.s om $A \in \mathfrak{F}^{\mathfrak{co}}$ och $A \subseteq B$ så $B \in \mathfrak{F}^{\mathfrak{co}}$

Filtret är äkta: På samma sätt, $\emptyset^c = I$ och I oändlig ger att $\emptyset \notin \mathfrak{F}^{\mathfrak{co}}$ □

Det är dock klart att Fréchetfiltret inte utgör något ultrafilter på $\mathbb{N}$. Exempelvis är varken mängden av alla jämna tal eller dess komplement, mängden av alla udda tal, element i $\mathfrak{F}^{\mathfrak{co}}$ då bägge dessa mängder är oändliga.

Enligt sats 5.7 finns ett ultrafilter $\mathfrak{U}$ sådant att $\mathfrak{U} \supseteq \mathfrak{F}^{\mathfrak{co}}$ så allt vi behöver vissa är att detta ultrafilter är fritt.

Sats 5.10. *Om I är en oändlig mängd så existerar ett fritt ultrafilter på I*

Bevis.

Låt alltså $\mathfrak{U}$ vara ett ultrafilter sådant att $\mathfrak{U} \supseteq \mathfrak{F}^{\mathfrak{co}}$, sådant filter existerar enligt sats 5.7.

Antag att $\mathfrak{U}$ är principalt. Då har $\mathfrak{U}$ någon ändlig delmängd till $I \supset A$ som element. Enligt definition har vi då att $A^c \in \mathfrak{F}^{\mathfrak{co}} \subseteq \mathfrak{U}$. Alltså är både A och A^c element i $\mathfrak{U}$ vilket motsäger att $\mathfrak{U}$ är ett ultrafilter. □

Korollarium 5.11. *Det finns ett fritt ultrafilter $\mathfrak{U}$ på $\mathbb{N}$.*

Följaktligen har vi nu visat allt som behövs för konstruktionen av de hyperreella talen ${}^*\mathbb{R}$ i kapitel 2. Vi kommer, som antytts i inledningen till kapitlet, att använda oss av Ultrafiltersatsen igen i kapitlen 8 och 9 då vi skall konstruera en större icke-standard struktur som inkluderar ${}^*\mathbb{R}$.

Not 5.1. *Användandet av urvalsaxiomet (i form av Zorn's lemma) utgör stundom en källa för kritik av icke-standard analys på grund av urvalsaxiomets icke-konstruktiva natur. Det kan därför vara intressant att veta att det är möjligt att visa ultrafiltersatsen utifrån ett svagare antagende än Zorn's lemma, nämligen 'boolean prime ideal theorem'. Vi redogör för bevisiden bakom detta:*

Filter och ideal är duala koncept. Betrakta återigen definitionen av filter:

(i) $A \in \mathfrak{F} \wedge B \in \mathfrak{F} \Rightarrow A \cap B \in \mathfrak{F}$ $\qquad$ $\mathfrak{F}$ *är sluten under snitt.*

(ii) $A \in \mathfrak{F} \wedge A \subseteq B \Rightarrow B \in \mathfrak{F}$ $\qquad$ $\mathfrak{F}$ *är sluten under övre delmängder*

om vi nu 'vänder på' operationerna i definitionen av filter får vi definitionen av ett ideal:

(i) $A \in \mathfrak{F} \wedge B \in \mathfrak{F} \Rightarrow A \cup B \in \mathfrak{F}$ $\qquad$ $\mathfrak{F}$ *är sluten under unioner.*

(ii) $A \in \mathfrak{F} \wedge A \supseteq B \Rightarrow B \in \mathfrak{F}$ $\qquad$ $\mathfrak{F}$ *är sluten under nedre delmängder*

Givet ett filter $\mathfrak{F}$ på I kan vi alltså associera det med ett ideal $\mathfrak{I}$ på I genom att låta $A \in \mathfrak{I}$ omm $I \setminus A \in \mathfrak{F}$.

'Boolean prime ideal theorem', som alltså är svagare än urvalsaxiomet, säger att varje ideal kan utvidgas till ett maximalt ideal (definitionen av ett maximalt ideal är analogt med definitionen av ett maximalt filter). Specifikt finns alltså ett maximalt ideal $\mathfrak{J}$ sådant att $\mathfrak{J} \subseteq \mathfrak{I}$. Då kan vi helt enkelt konstruera ett ultrafilter $\mathfrak{U}$ sådant att $\mathfrak{U} \supseteq \mathfrak{F}$ genom att låta $\mathfrak{U} \doteqdot \{A : (I \setminus A) \in \mathfrak{J}\}$.

6 Formell logik

Denna (komprimerade) framställning av matematisk logik och elementär modellteori utgår från [Felix Mendelson] och [Christian Bennet]. Materialet är dock, i det närmaste, ett allmänt tankegods och återfinns i väsentligen alla framställningar av matematisk logik och även i flertalet läroböcker i icke-standard analys, om än i något förkortad form. Sanningsdefinitionen som introduceras är Alfred Tarskis.

6.1 Varför vill vi ha formella metoder?

Leibniz kontinuitetslag, såsom den presenterats i kapitel 1.1, är antagandet, eller kanske bättre uttryckt det informella axiomet, att ”obegränsade och infinitesimala tal lyder under samma lagar som vanliga reella tal”. Detta tillämpades i ett flertal bevis i kapitlen 4 och 5, i form av ett axiom som refererades till som transferprincipen.

Vi har i kapitel 2 konstruerat obegränsade och infinitesimala tal i form av de hyperreella talen $^*\mathbb{R}$ - problemet vi står inför då vi vill bevisa Leibniz kontinuitetslag är alltså vad som avses med ”lyder under samma lagar” i utsagan. För att finna en lösning på detta introducerar vi ett formellt språk och likställer lag med ett påstående i detta formella språk. Givet det formella språket är det möjligt att matematiskt definiera begreppen tolkning och sanning. Uttrycker vi då Lebniz lag som att ett uttryck i det formella språket är sant då det tolkas i $\mathbb{R}$ om och endast om samma uttryck tolkas i $^*\mathbb{R}$ får vi möjlighet att matematiskt bevisa denna formulering av Leibniz lag.

Med de formella metoderna är det dock möjligt att visa mer än att transferprincipen gäller mellan $\mathbb{R}$ och $^*\mathbb{R}$ - det är möjligt att bevisa att alla motsvarande konstruktioner uppfyller en generell standard-princip. Detta avser konstruktioner med avseende på sekvenser av mängder av reella tal, sekvenser av funktioner på reella tal, och så vidare, definierade analogt med de sekvenser av reella tal som presenterades i kapitel 2. För att åstadkomma detta skall vi använda oss av en mängdteoretisk modell till matematisk analys - vilken vi kommer att kalla ett matematiskt ”universum” och utifrån denna modell konstruera ett icke-standard universum. Att vi kan bygga upp icke-standard analysen från mängdteorin är intressant i sig då det visar att standard och icke-standard matematik alltså kan stå på samma grundvalar.

Så för att uppnå dessa två ting: att konstruera ett icke-standard universum och att visa att de matematiska objektet däri lyder under transferprincipen börjar vi med att definiera första ordningens logik (FOL). Denna kommer att användas för att formellt definiera vad detta ett matematiskt påstående eller en matematisk lag kan vara samt för att upprätta en mängdteoretisk modell. Utifrån detta kommer vi på ett enkelt sätt kunna jämföra påståenden om reella tal med påståenden om hyperrella tal samt även generellare mellan matematiska objekt i standard respektive icke-standard universa efter dessa introducerats. När detta är klart kommer vi kunna avgöra exakt vad den avsedda tolkningen av ett matematiskt påstående i denna kontext är och utifrån detta bevisa transferprincipen.

6.2 Syntax

Med begreppet syntax avses hur en utsaga, det vill säga för våra ändamål en matematisk formel, får se ut. Detta är i sig helt oberoende utav den avsedda tolkningen av symbolerna i formeln. Formler kommer att bestå av symboler vilka kombineras enligt vissa formationsregler. Nedan presenteras alltså begynnelsevis förekommande symboler i det formella språket samt hur de kombineras för att bilda formler.

Symbolerna i FOL delas in i kategorier av logiska och icke-logiska symboler samt markörer. Enligt praxis har de logiska symbolerna en bestämd tolkning medan tolkningen av de icke-logiska symbolerna är kontextberoende.

Definition 6.1. *Typer av symboler*

Logiska symboler:

$Konnektiv : \neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$

$Kvantifikatorer : \forall, \exists$

$Identitetssymbol : =$

Icke-logiska symboler:

$Individkonstanter : c_0, c_1, ...$

$Funktionssymboler^3 : f_0^n, f_1^n, ..., n \in \mathbb{N}$

$Relationssymboler : R_0^n, R_1^n, ..., n \in \mathbb{N}$

Markörer

$individvariabler : x_0^n, x_1^n, ..., n \in \mathbb{N}$

$parenteser :), ($

Vi kommer att förhålla oss ganska löst till exakt vilka symboler som är tillåtna enligt ovan. Exempelvis kommer även $y, z, ...$ användas som symboler för variabler, ställigheten hos funktioner kommer sällan att göras explicit, et cetera.

[3] n är ställigheten hos f^n respektive R^n. (Med en n-ställig funktion eller relation avses en funktion av n argument respektive en relation mellan n objekt).

Beroende på vilken teori eller vilken axiomatisering av en viss teori som studeras kan de icke-logiska symbolerna i syntaxen variera. Vi definierar att ett **lexikon** $\mathcal{L}$ är en mängd av icke-logiska symboler.

Vad gäller teorin för $\mathbb{R}$ respektive $^*\mathbb{R}$ kan vi föreställa oss att $\mathcal{L}_{\mathbb{R}}$ och $\mathcal{L}_{^*\mathbb{R}}$innehåller symbolerna $+, *, <$ eller liknande beroende på vilket matematiskt problem vi ämnar studera, samt utöver detta, namn på alla individkonstanter (det vill säga reella respektive hyperreella tal) och elementära funktioner definierade på $\mathbb{R}$ respektive $^*\mathbb{R}$. (Vad som kan anses vara ett tillräckligt lexikon för matematisk analys tas upp i kapitel 7 om mängdteori).

Givet ett lexikon $\mathcal{L}$ definieras $\mathcal{L}$-formler induktivt enligt följande:

Definition 6.2. *$\mathcal{L}$-formler*

Individkonstanter, och -variabler samt funktioner av dessa benämns $\mathcal{L}$-termer.

Om t_i benämner en godtycklig L-term är $t_i = t_j$ samt $R(t_1, ..., t_n)$ atomära $\mathcal{L}$-formler, där R är en godtycklig n-ställig relation i $\mathcal{L}$.

Om ϕ_i är en atomär L-formel, $*$ ett konnektiv och Q en kvantifikator så är $\neg\phi, (\phi_i * \phi_j)$ samt $Qx\phi$ $\mathcal{L}$-formler.

Följaktligen är alltså exempelvis

"$\forall\epsilon > 0\ \exists\delta > 0\ \forall x(\mid x - c \mid < \delta \rightarrow \mid f(x) - f(c) \mid < \epsilon)$"

en formel i $\mathcal{L}_{\mathbb{R}}$ enligt ovan, där ϵ, δ är variabelsymboler och c är en individkonstant.

medan Eudoxus-Archimedes princip:

"$\forall x\ \exists n(x < n \wedge n \in \mathbb{N})$"

inte är en formel i $\mathcal{L}_{\mathbb{R}}$ då $\mathbb{N}$ inte finns bland de namngivna konstanterna. Vi kan självklart välja att införa en ny konstantsymbol $\mathbb{N}$ i $\mathcal{L}_{\mathbb{R}}$ så att Eudoxus-Archimedes princip blir en $\mathcal{L}_{\mathbb{R}}$-formel. Notera dock att vi ännu inte specificerat hur $\mathcal{L}_{\mathbb{R}}$-formler skall tolkas och att, följaktligen $\mathbb{N}$ än så länge bara är en symbol intet på något sätt skild från exempelvis konstantsymbolen π eller 7. Hur vi löser detta presenteras i nästa delkapitel där vi presenterar tolkning av logiska och icke-logiska symboler.

En variabelförekomst i en formel ϕ kan vara fri eller bunden. Vi säger att en variabelförekomst är bunden om den ligger inom en kvantifikators räckvidd, en variabelförekomst som inte är bunden benämns fri. Exempelvis är x_1 bunden och x_2 fri i ϕ nedan

$\phi \doteqdot \exists x_1(x_1 < x_2)$

För att poängtera att en $\mathcal{L}$-formel ϕ har $x_1, ..., x_n$ som fria variabler skriver vi ofta, men inte nödvändigtvis, $\phi = \phi(x_1, ..., x_n)$. Detta skulle då innebära:

$\phi = \phi(x_2) = \exists x_1(x_1 < x_2)$ i vårt exempel ovan.

6.3 Semantik; Formella $\mathcal{L}$-strukturer och tolkning

Baserat på ett givet lexikon $\mathcal{L}$, alltså en mängd icke-logiska symboler, vill vi nu definiera formellt begreppen $\mathcal{L}$-struktur och en tolkning T i denna. Grundidèn är egentligen ganska bekant och det nya är kanske främst själva formaliseringen av begreppen. Vi presenterar först denna intuitiva grundidé med ett exempel:

Betrakta följande enkla $\mathcal{L}_\times$-formel för lexikonet $\mathcal{L}_\times = \{\times, =, 1\}$

$\forall x \exists y(x \times y = 1)$

De flesta av oss kommer nog att tolka formeln som utsagan att varje tal har en multiplikativ invers. Detta innebär att vi tolkat den icke-logiska symbolen $\times$ som funktionen multiplikation och att variablerna x och y representerar någon form av tal samt tilldelat symbolen 1 några särskilda egenskaper med avseende på multiplikation. Det är inte svårt att konstatera att huruvida formeln skall betraktas som sann under denna tolkning beror på vilken typ av tal x och y skall tillåtas representera.

(Vi kan även i förbigående notera att formeln strikt taget inte är syntaktiskt korrekt eftersom om $\times$ är en funktionssymbol och vi har definierat att funktionstermen skall vara på formen $\times(x, y)$, så kallad polsk notation).

Det är klart att om formeln tolkas som en utsaga om reella tal skilda från noll eller tillika rationella tal så är den sann enligt det faktum att **det finns** (tolkningen av den logiska symbolen $\exists$) multiplikativa inverser **för alla** (tolkningen av den logiska symbolen $\forall$) nollskilda tal i $\mathbb{R}$ respektive $\mathbb{Q}$. På motsvarande sätt anser vi att formeln är falsk om den tolkas som en utsaga om heltal.

Enligt vårt exempel säger vi att $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{N}$, etc tillsammans med identifikation av symbolen $\times$ med den matematiska operationen multiplikation, tillika symbolen $=$ med den faktiska identitetsrealtionen samt symbolen 1 med med talet 1 är exempel på $\mathcal{L}_\times$-strukturer.

Vi kan redan nu börja föreställa oss att 'lagarna' i kontinuitetslagen kan uttryckas som en mängd formler liknande den ovan och påståendet att 'samma lagar gäller' kan motsvaras av att tolkningen av formler är sanna i en struktur omm de är sanna i en annan. För att matematiskt kunna bevisa resultat om tolkningar behöver vi emmellertid en definition av begreppen tolkning och sanning.

Definition 6.3. *En $\mathcal{L}$-struktur $\mathcal{M}$ består av:*

En mängd $D \neq \varnothing$, domänen till $\mathcal{M}$. Denna mängd beteckanas $Dom(\mathcal{M})$.

En tolkningsfunktion T sådan att:

$T(c) \in D$ *om c är en individkonstant i $\mathcal{L}$*

$T(R^n) \subseteq D^n$ *om R^n är en relationssymbol i $\mathcal{L}$*[4]

$T(f^n)$ *är en funktion från D^n till D*

Normalt används oftast skrivsätten $c^{\mathcal{M}}, (R^n)^{\mathcal{M}}, (f^n)^{\mathcal{M}}$ för tolkningsfunktionen istället för $T(c), T(R^n), T(f^n)$.

Vi låter individvariablerna $x_0, ...$ anta värden ur D vilket svarar mot det informella uttrycket att en $\mathcal{L}$-formel **handlar om** exempelvis reella tal, alltså motsvarande det fall då $Dom(M) = \mathbb{R}$.

Vad vi nu har utgör ett så kallat **extensionellt** semantiskt system, det vill säga att tolkningen av en icke-logisk $\mathcal{L}$-symbol definieras av en **mängd av objekt**, symbolens extension. För individkonstanter ter sig detta helt naturligt, exempelvis låter vi tolkningen av symbolen 'π' vara talet π och vi tillåter oss att skriva $\pi^{\mathcal{M}} = \pi$. Generellt tolkas konstantsymboler som de objekt de namnger. Gällande relationer och tolkningen av dessa kan en ytterligare kommentar vara belysande: exempelvis kommer vi använda i matematiska strukturer en mängd, vi kan namnge den S för 'summa', av ordnade triplar $\langle a, b, c\rangle$ sådana att $a + b = c$, för att modellera operationen $+$. Det vill säga att

$+^{\mathcal{M}} = S = \{\langle a, b, c\rangle : a + b = c\}$ där a, b och c representerar element i $Dom(\mathcal{M})$.

Mängden S är alltså symbolen $+$'s extension under aktuell tolkning.

Nu vill vi kunna tala om att en $\mathcal{L}$-formel ϕ är sann i en $\mathcal{L}$-struktur $\mathcal{M}$. Detta inbegriper utöver vad vi redan har en metod för att tolka även de logiska symbolerna.

Vi kan redan nu definiera tolkningen av konnektiven $\neg, \wedge, \vee, \Rightarrow$ och $\Leftrightarrow$ få ett användbart logiskt resultat utan att behöva definiera tolkningen av kvantifikatorerna $\forall$ och $\exists$.

Definition 6.4. *Tolkning av de logiska konnektiven.*

Antag att ϕ, ψ, ψ_1 och ψ_2 är $\mathcal{L}$-formler.

[4] D^n, den kartesiska produkten av D i n dimensioner, är mängden av alla ordnade n-tupler av element ur D.

$\neg$ *är negation, d.v.s att om* $\phi \doteqdot (\neg\psi)$ *är* ϕ *sann omm* ψ *inte är sann.*

$\wedge$ *är konjunktion, d.v.s att om* $\phi \doteqdot (\psi_1 \wedge \psi_2)$ *är sann omm* ψ_1 *och* ψ_2 *båda är sanna.*

$\vee$ *är disjunktion d.v.s att om* $\phi \doteqdot (\psi_1 \vee \psi_2)$ *är sann omm* ψ_1 *eller* ψ_2 *är sann.*

$\Rightarrow$ *är materiell implikation, d.v.s att om* $\phi \doteqdot (\psi_1 \Rightarrow \psi_2)$ *är sann omm* ψ_1 *är falsk eller* ψ_2 *är sann.*

$\Leftrightarrow$ *är materiell ekvivalens, d.v.s att om* $\phi \doteqdot (\psi_1 \Leftrightarrow \psi_2)$ *är sann omm* ψ_1 *och* ψ_2 *båda är sanna eller båda är falska.*

Sats 6.5. *Under denna tolkning av de logiska konnektiven är* $\vee, \Rightarrow, \Leftrightarrow$ *redundanta symboler.*

Bevis. Vi skall alltså visa att vi med definition 6.4 kan ersätta varje $\mathcal{L}$-formel som innehåller konnektiven $\vee, \Rightarrow, \Leftrightarrow$ med en ekvivalent formel som endast innehåller konnektiven $\neg$ och $\wedge$.

$\vee$: Vi har enligt definition att $\psi_1 \vee \psi_2$ är sann omm ψ_1 eller ψ_2 är sann. Följande är logiskt ekvivalent:

inte både ψ_1 och ψ_2 är falska

inte både $\neg\psi_1$ och $\neg\psi_2$ är sanna

$\neg(\neg\psi_1 \wedge \neg\psi_2)$ är sann, alltså är $\vee$ en redundant symbol då den kan ersättas av en kombination av $\neg$ och $\vee$.

$\Rightarrow$: På samma sätt har vi att $\psi_1 \Rightarrow \psi_2$ är sann är logiskt ekvivalent med att $(\neg\psi_1 \vee \psi_2)$ är sann.

$\Leftrightarrow$: Återigen på samma sätt: Enligt tolkningen av konnektiven är $\psi_1 \Leftrightarrow \psi_2$ ekvivalent med $(\psi_1 \Rightarrow \psi_2 \wedge \psi_2 \Rightarrow \psi_1)$.

□

Vi hade kunnat begränsa konnektiven till $\neg$ *och* $\vee$ *med motsvarande redundans bevis, etc.*

För att kunna hantera kvantifikatorer och fria variabler introducerar vi först begreppet **satisfiering** ur vilket en fullständig definition av sanning för för $\mathcal{L}$-strukturer naturligt följer.

Låt S vara mängden av alla sekvenser $s = (s_0, s_1, ...)$ av element ur $Dom(\mathcal{M})$

För en given sekvens s definierar vi nu en funktion s^* som tillskriver varje term t i $\mathcal{L}$ ett element $s^*(t)$ i $dom(\mathcal{M})$

- Om t är en variabel x_i låter vi $s^*(t) = s_i$
- Om t är en konstant c_i låter vi $s^*(t) = c_i^{\mathcal{M}}$ alltså tolkningen av c_i i $\mathcal{M}$
- För funktioner f_i^n och termer $t_0, ..., t_n$ låter vi

  $s^*(f_i^n(t_0, ..., t_n)) = (f_i^n)^M(s^*(t_0), ..., s^*(t_n))$

Definition 6.6. *Satisfierbarhet*

- *Om ϕ är en atomär formel $P^n(t_0, ..., t_n)$ vars tolkning är relationen $(P^n)^{\mathcal{M}}$ så satisfierar s ϕ omm n-tupeln $(s^*(t_0), ..., s^*(t_n)) \in (P^n)^{\mathcal{M}}$*
- *s satisfierar $\neg\phi$ omm s inte satisfierar ϕ*
- *s satisfierar $\phi \Rightarrow \psi$ omm s inte satisfierar ϕ eller s satisfierar ψ*
- *s satisfierar $\forall x_i \phi$ omm för alla element c i $dom(M)$ sekvensen $s = (s_0, s_1, ..., s_{(i-1)}, c, ...)$ satisfierar ϕ (x_i kan förekomma som fri variabelförekomst i ϕ eller inte).*
- *s satisfierar $\exists x_i \phi$ omm s satisfierar $\neg\forall x_i \neg\phi$.*

Definition 6.7. *Sanning*

- *En formel ϕ är sann för tolkningen $\mathcal{M}$ (skrivs $\mathcal{M} \models \phi$) omm varje sekvens $s \in S$ satisfierar ϕ*
- *ϕ är falsk omm ingen sekvens s satisfierar ϕ*
- *Vi säger att $\mathcal{M}$ är en* **modell** *till en mängd $\mathcal{L}$-formler Γ om alla varje L-formel i Γ är sann i $\mathcal{M}$.*

Utifrån detta behöver en formel ϕ, tolkad i en viss struktur, varken vara sann eller falsk. Antag exempelvis att $\phi = \phi(x)$ är formeln:

$x < 1$

ϕ är då en atomär $\mathcal{L}_{\mathbb{R}}$-formel i enligt definition 6.2 och satisfieras av en sekvens $s = (s_0), s_1, ...)$ om och endast om $s_0 < 1$. Så om vi tolkar ϕ på talmängden $\mathbb{R}$ finns både sekvenser som satisfierar ϕ och sekvenser som inte satisfierar ϕ, alltså är ϕ varken sann eller falsk.

7 Mängdteoretiska modeller

Detta kapitel bygger på material om mängdteorin ZF om vilken det finns mycket tillgängligt material. Vi har använt oss av [Hurd och Loeb] och [Robert Goldblatt].

Vi vill nu konstruera en struktur $V(\mathbb{R})$ för lämpligt lexikon $\mathcal{L}$ som är en modell till matematisk analys i vidare bemärkelse — ett matematiskt ”universum”. Alla satser från matematisk analys skall gå att uttrycka som $\mathcal{L}$-formler och ha en sann tolkning i $V(\mathbb{R})$.

Modellen definieras som en mängdteoretisk struktur vars domän är en så kallad superstruktur vilket vi nu skall definiera.

7.1 Superstrukturer

Låt $\mathrm{P}(X)$ vara mängden av alla delmängder till X.

Definition 7.1. *Superstruktur*

Låt $V_0(\mathbb{R}) = \mathbb{R}$

$V_{n+1}(\mathbb{R}) = V_n(\mathbb{R}) \cup P(V_n(\mathbb{R}))$,

Låt sedan $V(\mathbb{R}) = \bigcup_{n=0}^{\infty} V_n(\mathbb{R})$

Vi säger då att $V(\mathbb{R})$ är en superstruktur över $\mathbb{R}$.

Definition 7.2. *Rank*

Rank(a) *är det minsta* n *för vilket* $a \in V_n(\mathbb{R})$.

Vi väljer att kalla elementen i superstrukturen för **entiteter** i ett matematiskt universum. Vi inför detta begrepp för att dra associationerna lite bort från den gängse bilden av vad en mängd eller ett element är då vi skall fortsätta argumentera för att en funktion eller i princip vilket matematiskt objekt som helst kan tolkas som en mängd på ett tillfredsställande sätt.

Det är klart att alla entiteter i $V(\mathbb{R})$ av $Rank > 0$ är mängder. Vi kallar de enklaste entiteterna, de av rank 0, alltså elementen i $\mathbb{R}$, för individer och inför som konvention att dessa saknar element.

Definition 7.3. *Individerna i $V(\mathbb{R})$ är urelement, d.v.s att de är icke-tomma men saknar element från $V(\mathbb{R})$. Detta uttrycks av följande axiom:*

$\forall x \in V_0(\mathbb{R})(x \neq \emptyset \wedge \forall y \in V(\mathbb{R})\ (y \notin x))$

Detta är i någon mening godtyckligt, vi skulle även kunna valt att definiera att reella tal **är** ekvivalensklasser av Cauchysekvenser om det passat vårat syfte bättre, i sådant fall hade vi definierat $V_0(\mathbb{Q}) = \mathbb{Q}$ och bildat en superstruktur över $\mathbb{Q}$ helt analogt.

7.2 Lexikonet för en superstruktur

Låt $\mathcal{L}_{V(\mathbb{R})} = \{\in, =, [\text{konstantsymboler för alla entiteter}]\}$. Den avsedda tolkningen för symbolen '$\in$' är relationen 'att vara element i'. Mellan individer är likhet '$=$' en primitiv relation[5], medan vi definierar likhet mängdteoretiskt (alltså i termer av $\in$) för övriga entiteter:

Definition 7.4. *Två mängder är lika omm de har samma element, d.v.s:*

$X = Y \doteqdot \forall x \in V(\mathbb{R})(\ x \in X \Leftrightarrow\ x \in Y)$

Vi ska argumentera för att $\mathcal{L}_{V(\mathbb{R})}$ är ett lämpligt lexikon för matematisk analys i strukturen $V(\mathbb{R})$. Att lexikonet är relativt begränsat m.a.p antalet predikats-/relations- symboler gör att vi får ett enklare arbete att bevisa satser om språket ifråga då detta begränsar hur en atomär $\mathcal{L}_{V(\mathbb{R})}$-formel kan se ut. Å andra sidan måste vi visa att vi kan definiera övriga funktioner och relationer utifrån $=$ och $\in$.

Vi vill att ett kvantifierat påstående $\forall x \in V(\mathbb{R})$ skall tolkas på samtliga entiteter i $V(\mathbb{R})$, så om en entitet är element i en mängd skall den även vara element i domänen. Uttryckt i symboler vill vi alltså att $a \in A \in V(\mathbb{R}) \Rightarrow a \in V(\mathbb{R})$ skall gälla. Om detta gäller kallar vi mängden A **transitiv**. Vi definierar vidare att en struktur X är **starkt transitiv** i om det för varje mängd $A \in X$ finns en transitiv mängd $B \in X$ sådan att $A \subseteq B \subseteq X$.

Enligt konstruktion är alla $A \in V(\mathbb{R})$ transitiva och därmed är alltså $V(\mathbb{R})$ starkt transitiv: Antag att a och A är entiteter sådana att $a \in A \in V(\mathbb{R})$. Då är $A \in V_{(k+1)}(\mathbb{R})$ för något k. Enligt 7.1 har vi att:

[5]Med primitiv relation avses att relationen inte definieras utifrån 'enklare' begrepp, relationen är grundläggande. Vi säger alltså i paragrafen att '$\in$' och '$=$' mellan individer är primitiva relationer.

$V_{k+1}(\mathbb{R}) = V_k(\mathbb{R}) \cup \mathrm{P}(V_k(\mathbb{R}))$

så antingen är $A \in V_k(\mathbb{R})$ eller $A \in \mathrm{P}(V_k(\mathbb{R}))$

$a \in A$ implicerar $\{a\} \subseteq A$ så i bägge fallen är $\{a\} \in \mathrm{P}(V_k(\mathbb{R}))$

7.3 Funktioner och relationer i $V(\mathbb{R})$

Vi har redan tidigare definierat att vi tolkar funktioner och relationer som mängder, enligt det elementära exemplet i 6.3 lät vi S (summa) vara en mängd av ordnade triplar $\langle a, b, c\rangle$ sådana att $a + b = c$.

Generellt gäller samma resonemang för n-ställiga funktioner, att de kan uttryckas som en mängd av ordnade $(n+1)$-tupler $\langle a_1, ..., a_n, f(a_1, ..., a_n)\rangle$ - funktioner representeras som en typ av relationer (se 6.3)

Definition 7.5. *En entitet f är en funktion $A \longrightarrow B$ om:*

(i) $\langle a, b\rangle$ *implicerar* $a \in A$ *och* $b \in B$

(ii) $\langle a, b\rangle \in f$ *och* $\langle a, c\rangle \in f$ *implicerar* $b = c$

(iii) För alla $a \in A$ *finns* $b \in B$ *sådant att* $\langle a, b\rangle \in f$

Vi kommer inte strikt visa allt som behövs för att garantera att, som utlovat, alla matematiska entiteter finns som mängder i superstrukturen $V(\mathbb{R})$. För att visa detta måste vi visa att superstrukturen är sluten under kartesiska produkter, par, et cetera. Ett sådant bevis anses ändå inte särskilt belysande[6]. Definitionen av relationer i $V(\mathbb{R})$ (som nu alltså även inkluderar funktioner) tarvar emellertid en teori om ordnade n-tupler:

Definition 7.6. *Det ordnade talparet* $\langle a, b\rangle$ *definieras som mängden* $\{\{a\}, \{a, b\}\}$.

Sats 7.7. $\langle a, b\rangle = \langle c, d\rangle$ *omm* $a = c$ *och* $b = d$

Bevis. Följande är ekvivalent:

$\langle a, b\rangle = \langle c, d\rangle$

[6] Dylika bevis återfinns vanligtvis i introduktionsböcker till mängdteori, exempelvis Patrick Suppes 'Axiomatic Set Theory'. Påståendet att dessa bevis inte 'är särskilt belysande' motiveras av att bevisen i stort bygger på axiom som exempelvis postulerar att givet två mängder finns en mängd som innehåller elementen från bägge dessa mängder (Unionsaxiomet).

$\{\{a\}, \{a,b\}\} = \{\{c\}, \{c,d\}\}$

$(\{a\} = \{c\}$ och $\{a,b\} = \{c,d\})$ eller $(\{a\} = \{c,d\}$ och $\{a,b\} = \{c\})$

$(a = c$ och $b = d)$ eller $(a = b = c = d)$

$a = c$ och $b = d$ □

Definition 7.8. *Ordnade n-tupler* $\langle x_1, ..., x_n \rangle$ *definieras induktivt som ordnade par:*

$\langle x_1, ..., x_n \rangle = \langle x_1, ..., x_{n-1} \rangle, x_n \rangle$

Det följer enkelt av 7.7 ovan och induktion att definitionen av $\langle x_1, ..., x_n \rangle$ har de egenskaper vi kräver. Tag som exempel en ordnad trippel $\langle a, b, c \rangle = \langle \langle a, b \rangle, c \rangle$.

Enligt satsen har vi då att $\langle a, b, c \rangle = \langle d, e, f \rangle$

omm $\langle \langle a, b \rangle, c \rangle = \langle \langle d, e \rangle, f \rangle$

omm $\langle a, b \rangle = \langle d, e \rangle$ och $c = f$

omm $a = b$ och $b = e$ och $c = f$.

Sats 7.9. *Antag att* a *och* b *är entiteter av Rank* $\leq k$*. Då är det ordnade paret* $\langle a, b \rangle$ *en entitet av Rank* $k + 2$*.*

Bevis. Enligt konstruktionen av $V(\mathbb{R})$ kan vi anta att $a, b \in V_k(\mathbb{R})$.

Detta är ekvivalent med att $\{a, b\} \subseteq V_k(\mathbb{R})$ så $\{a, b\} \in V_{(k+1)}(\mathbb{R})$ då vi definierat att alla delmängder till $V_k(\mathbb{R})$ är element i $V_{(k+1)}(\mathbb{R})$.

Klart att både $\{a\}$ och $\{a, b\}$ är delmängder till $\{a, b\}$ och följaktligen har vi att det ordnade paret $(a, b) = \{\{a\}, \{a, b\}\} \in V_{(k+2)}(\mathbb{R})$

□

Specifikt har vi då att om a, b är individer, alltså av *Rank* 0, så är det ordnade paret en entitet av *Rank* 2 och en binär relation på individer, som vi identifierar med en mängd av ordnade par, blir följaktligen en entitet av *Rank* 3.

Sannolikhetsmått på $\mathbb{R}$, som skall diskuteras i senare kapitel, är funktioner från en mängd av delmängder av $\mathbb{R}$, som vi kallar $\mathcal{A}$, till det reella intervallet $(0,1)$. Delmängder till $\mathbb{R}$ ligger i $V_1(\mathbb{R})$, alltså både reella intervall och elementen i $\mathcal{A}$. följaktligen ligger mängder av delmängder, och däribland $\mathcal{A}$ i $V_2(\mathbb{R})$. Alltså är $\mathcal{A}$ av *Rank* 2. Funktionen $f : \mathcal{A} \to \mathbb{R}$ är alltså en mängd av ordnade par med element i $\mathcal{A} \cup \mathbb{R}$. Rang $(0,1) \leq$ rang $(\mathcal{A})$ har vi redan. *Rank* $(\mathcal{A}) = 2$ så funktionens *Rank* kommer, enligt ovan, att vara *Rank* $(\mathcal{A}) + 2 = 4$.

7.4 $\mathcal{L}_{V(\mathbb{R})}$-formler

Som redan antytts begränsar vi alltså formellt relationssymbolerna i $\mathcal{L}_{V(\mathbb{R})}$ till $=$ och $\in$. Dessutom låter vi lexikonet innehålla konstantsymboler för alla entiteter i $V(\mathbb{R})$.

Enligt 6.2 har vi att atomära $\mathcal{L}_{V(\mathbb{R})}$ formler formas av termer och relationssymbolerna i lexikonet, d.v.s $=, \in$. Därav följer att alla atomära $\mathcal{L}_{V(\mathbb{R})}$-formler är på formen:

$s = t$ och $s \in t$,

Där alltså s, t antingen är variabler eller konstanstsymboler som betecknar en godtycklig entitet i $V(\mathbb{R})$.

Logiska konnektiv hanteras som för $\mathcal{L}$-formler.

Vi inför en begränsning på hur kvantifikatorer får användas för att forma $\mathcal{L}_{V(\mathbb{R})}$-formler: då en kvantifikator binder en variabel x måste variabeln vara på formen

$x \in t$

För en godtycklig $\mathcal{L}_{\mathbb{R}}$-formel ϕ är alltså

$\forall x \in t(\phi)$ och $\exists x \in t(\phi)$ $\mathcal{L}_{\mathbb{R}}$-formler,

medan varken $\forall x(\phi)$ eller $\exists x(\phi)$ är det.

Vi har alltså inte tillåtit några funktionssymboler i $\mathcal{L}_{V(\mathbb{R})}$ och heller inte sedvanliga relationssymboler som exempelvis $\leq$. Enligt tidigare diskussion är detta inte en begränsning för uttryckskraften i språket då vi istället infört namn på alla entiteter. Detta förenklar beviset för transferprincipen som vi leder upp till men tenderar att göra språket mer svårhanterligt för läsare. Av denna anledning kommer vi, oftare än inte, att använda sedvanlig notation för funktioner och använda oss av symboler som $\leq, \times$ och $\int_a^b$ som om de var element i $\mathcal{L}_{V(\mathbb{R})}$. Den springande punkten är då att de resulterande formlerna **kan** *omvandlas till $\mathcal{L}_{V(\mathbb{R})}$-formler i strikt mening.*

Vi kan illustrera detta med hjälp av ett tidigare exempel: Addition kunde definieras som en mängd S *av ordnade triplar* $\langle a, b, c \rangle$ *sådana att* $a + b = c$*. Denna mängd* S *är nu en entitet i* $V(\mathbb{R})$*. För att uttrycka påståendet att* $a_0 + b_0 = c_0$ *strikt i* $\mathcal{L}_{V(\mathbb{R})}$ *har vi alltså formeln* $\langle a_0, b_0, c_0 \rangle \in S$ *men vi kommer alltså att godta även uttrycket* $a_0 + b_0 = c_0$ *för läsbarhetens skull då vi anser det klart att alla nödvändiga entiteter existerar i* $V(\mathbb{R})$*.*

8 Icke-standard ramverk

Detta kapitel bygger på material från följande källor: [Hurd och Loeb], [Stroyan och Luxenburg] och [Jerome Keisler].

Givet två superstrukturer $V(X)$ och $V(Y)$ som båda tolkar $\in$ och $=$, sådana att $V(X) \subseteq V(Y)$ samt en funktion $* : V(X) \to V(Y)$ som bevarar den matematiska strukturen hos $V(X)$ säger vi att trippeln $\langle V(X), V(Y), * \rangle$ bildar ett **icke-standard ramverk**.

Att funktionen $* : V(X) \to V(Y)$ bevarar matematisk struktur syftar till att $\in, =$ och $Rank$ relationerna bevaras. Denna ide om strukturbevarande axiomatiseras i begreppet monomorfi:

Definition 8.1. *En injektion $* : V(X) \to V(Y)$ kallas en* **monomorfi** *om*

(i) ${}^*(\emptyset) = \emptyset$

(ii) $a \in X \Longrightarrow {}^*a \in {}^*X$ *samt* ${}^*n = n, \forall n \in \mathbb{N}$

(iii) $a \in V_{n+1}(X) \setminus V_n(X) \Longrightarrow {}^*a \in V_{n+1}({}^*X) \setminus V_n({}^*X)$

(iv) *Om* $b \in {}^*V_n(X), n \geq 1$ *och* $a \in b$ *så* $a \in {}^*V_{n-1}(X)$

(v) Transferprincipen: *Om ϕ är en $\mathcal{L}_{V(X)}$-formel utan fria variabelförekomster så är tolkningen av ϕ sann i $V(X)$ omm tolkningen av ϕ är sann i $V({}^*X)$*

$$\begin{tikzcd} & \mathcal{L}_{V(X)} \arrow[dl, "T"'] \arrow[dr, "{}^*T"] & \\ V(X) \arrow[rr, "*"] & & V(Y) \end{tikzcd}$$

Tolkningen ${}^*T$ av $\mathcal{L}_{V(X)}$-formler i superstrukturen $V({}^*X)$ som introduceras i diagrammet låter konstantsymboler $c \in \mathcal{L}_{V(X)}$ tolkas som entiteten ${}^*c \in V({}^*X)$ som definieras av $*$-avbildningen mellan superstrukturerna.

(Ekvivalent används även notationen ${}^*\phi$ för $*$-transformen av ϕ, resultatet av att ersätta varje konstantsymbol c i ϕ med konstantsymbolen ${}^*c$, alltså en rent syntaktisk operation. Tolkningen av ${}^*\phi$ blir då, analogt med tolkningen av ϕ i $V(X)$, på de entiteter som namnges i $V({}^*X)$).

Från definition 8.1 är det klart att exempelvis $\langle V(\mathbb{R}), V(\mathbb{R}), *\rangle$ om $*$ är identitetsfunktionen, bildar ett icke-standard ramverk. Vi vill konstruera $V(Y)$ större än $V(X)$ i den meningen att $* : V(X) \to V(Y)$ inte kan vara en surjektion på $V(Y)$, det vill säga att $V(X) \subset V(Y)$. De entiteter $b \in V(Y)$ sådana att $b \neq {}^*a, \forall a \in V(X)$ kommer kallas för **icke-standard entiteter**. $V(Y)$ kommer då kallas för en **utvidgning** av $V(X)$.

Vi skall nu, utifrån superstrukturen för matematisk analys $V(\mathbb{R})$, vårt "standard-universum", konstruera en ytterligare superstruktur $V({}^*\mathbb{R})$ som vi kallar för "icke-standard universum" och funktionen $* : V(\mathbb{R}) \to V({}^*\mathbb{R})$ på så vis att $\langle V(\mathbb{R}), V({}^*\mathbb{R}), *\rangle$ bildar ett icke-standard ramverk. I kommande kapitel kommer vi även att visa att denna konstruktion av $V({}^*\mathbb{R})$ kan resultera i en utvidgning av $V(\mathbb{R})$ - icke-standard universum är alltså inte entydigt bestämt utan kommer att bero på valet av ultrafilter[7].

8.1 Ultraprodukter

Att konstruera ultraprodukter är en generell algebraisk metod att konstruera nya modeller till en teori givet en eller flera modeller till denna teori. Vi redan sett exempel på detta i konstruktionen av de hyperreella talen ${}^*\mathbb{R}$ som en ultraprodukt av identiska kopior av de reella talen $\mathbb{R}$. Vi betraktade alltså då begynnelsevis mängden av alla sekvenser av reella tal $r_1, r_2, ...$, d.v.s element i $\mathbb{R}^{\mathbb{N}}$.

Sedan definierade vi en tolkning av identitetsrelationen "=", där $r = s$ var sann i $\mathbb{R}^{\mathbb{N}}$ om r_i och s_i var lika nästan överallt, vilket formaliserades av filterbegreppet. Vi visade att denna tolkning av $=$ i $\mathbb{R}^{\mathbb{N}}$ gav upphov till ekvivalensklasser och vi definierade ${}^*\mathbb{R}$ som mängden av dessa ekvivalensklasser.

Vi definierar nu analogt ultraprodukter allmänt på entiteter i $V(\mathbb{R})$.

Låt $\mathfrak{U}$ vara ett ultrafilter på en indexmängd I. För en godtycklig entitet $S \in V(\mathbb{R})$ inför vi symbolen $\prod S$ för mängden av alla funktioner $a : I \to S$. På denna funktionsmängd definierar vi $\mathfrak{U}$-ekvivalens $=_{\mathfrak{U}}$.

Definition 8.2. $\mathfrak{U}$-*ekvivalens*

$a =_{\mathfrak{U}} b$ *omm* $\{i : a(i) = b(i)\} \in \mathfrak{U}$

Nu definierar $=_{\mathfrak{U}}$ en ekvivalensrelation på $\prod S$ och ekvivalensklassen a modulo $\mathfrak{U}$ definieras av $\{b : a =_{\mathfrak{U}} b\}$.

Bevis. Beviset är, i allt väsentligt, det samma som för sats 2.1 för ekvivalensklasser i $\mathbb{R}^{\mathbb{N}}$

Reflexivitet: Att $a =_{\mathfrak{U}} a$ är klart då $I \in \mathfrak{U}$ enligt definition.

[7] Teorin för ultrafilter behandlades i kapitel 5.

Symmetri: Följer direkt från symmetri hos $=$.

Transitivitet: Observera först att $\{\ i : a(i) = b(i)\} \cap \{\ i : b(i) = c(i)\} \subseteq \{\ i : a(i) = c(i)\}$. Enligt definition har vi att $\mathfrak{U}$ är sluten under både snitt och delmängder varför om vi antar att $\{\ i : a(i) = b(i)\} \in \mathfrak{U}$ och $\{\ i : b(i) = c(i)\} \in \mathfrak{U}$ så har vi $\{\ i : a(i) = b(i)\} \cap \{\ i : b(i) = c(i)\} \in \mathfrak{U}$ samt $\{\ i : a(i) = c(i)\}$. $\square$

Definition 8.3. *Vi definierar ultraprodukten av S som mängden av alla* **ekvivalensklasser** *i $\prod S$ och inför symbolen $\prod_{\mathfrak{U}} S$ för denna mängd. Ekvivalensklassen för $a \in \prod S$ skrivs $[a] \in \prod_{\mathfrak{U}} S$.*

8.2 Superstrukturen över en ultraprodukt

Vår avsikt är alltså att konstruera ett icke-standard ramverk $\langle V(\mathbb{R}), V(^*\mathbb{R}), *\rangle$ där $V(^*\mathbb{R})$ är superstrukturen över ultraprodukten $\prod_{\mathfrak{U}} \mathbb{R}$.

Definition 8.4. *Superstrukturen $V(^*\mathbb{R})$*

$V_0(^*\mathbb{R}) = \prod_{\mathfrak{U}} \mathbb{R}$

$V_{(n+1)}(^*\mathbb{R}) = V_n(^*\mathbb{R}) \cup P(V_n(^*\mathbb{R}))$

Låt sedan $V(^*\mathbb{R}) = \bigcup_{n=0}^{\infty} V_n(^*\mathbb{R})$

Lexikonet $\mathcal{L}_{V(^*\mathbb{R})}$ och formationsreglerna för $\mathcal{L}_{V(^*\mathbb{R})}$-formler definieras analogt med motsvarande definitioner för superstrukturen $V(\mathbb{R})$. Vi låter alltså även $V(^*\mathbb{R})$ innehålla namn på alla entiteter i $V(^*\mathbb{R})$, så de enda syntaktiska objekt som kan skilja en $\mathcal{L}_{V(^*\mathbb{R})}$-formel från en $\mathcal{L}_{V(\mathbb{R})}$-formel är namn på icke-standard entiteter.

8.3 Definition av $*$-avbildningen

För att återknyta till motsvarande situation i kapitel 2, konstruktionen av de hyperreella talen, kan $*$-avbildningen mellan reella och hyperreella tal definieras direkt genom att låta ett reellt tal r avbildas på ekvivalensklassen för den konstanta sekvensen av talet r, d.v.s

$* : r \to [r],\ \forall r \in \mathbb{R}$

För att $*$ skall vara en strukturbevarande monomorfi mellan superstrukturer krävs dock en mer komplicerad definition av $*$. Vi definierar $*$ som sammansättningen av två funktioner e och M, där e definieras enligt samma ide som $*$-avbildningen mellan reella och hyperreella talen. Målmängden till e kommer vara en tredje struktur, vilken vi kommer kalla den "begränsade ultraprodukten" - alltså inte superstrukturen $V(^*\mathbb{R})$ i definition 8.4.

Definition 8.5. *Den begränsade ultraprodukten över $V(\mathbb{R})$.*

Låt $V_{(-1)}(\mathbb{R}) \doteqdot \emptyset$

$\prod_{\mathfrak{U}}^0 V(\mathbb{R}) \doteqdot \bigcup_{n=0}^{\infty}(\prod_{\mathfrak{U}}(V_n(\mathbb{R}) \setminus V_{n-1}(\mathbb{R})))$

Den begränsade ultraprodukten är alltså en struktur i vilken elementen är 'sekvenser' av entiteter av rang n. För $\langle a_1, a_2, ...\rangle = [a] \in \prod_{\mathfrak{U}}^0 V(\mathbb{R})$ gäller alltså att $a_i \in (V_n(X)/V_{(n-1)}(X))$, för alla i och något fixt n då alla $(V_n(X)/V_{(n-1)}(X))$ är disjunkta, d.v.s att vi kan säga att $Rank([a]) = Rank(a_i)$.

Definition 8.6. *Avbildningen* $e : V(\mathbb{R}) \to \prod_{\mathfrak{U}}^0 V(\mathbb{R})$ *definieras av*

$e(a) = [\bar{a}]$, *där* $\bar{a}(i) = a$, *för alla* $i \in I$

I den begränsade ultraprodukten har vi alltså konstruerat, analogt med konstruktionen av de hyperreella talen i kapitel 2, icke-standard utvidgningar genom att betrakta sekvenser av entiteter av en fix godtycklig rang. Funktionen e associerar entiteterna i standard universum med de konstanta sekvensernas ekvivalensklasser i den begränsade ultraprodukten.

Vi har redan definierat en tolkning av '=' i ultraprodukter (def 8.2) och därmed även i den begränsade ultraprodukten. För att kunna tolka $\mathcal{L}_{V(\mathbb{R})}$-formler behövs även en tolkning av '$\in$'.

Definition 8.7. *Elementrelationen modulo* $\mathfrak{U}$

För $[a]$, $[b] \in \prod_{\mathfrak{U}}^0 V(\mathbb{R})$ *låter vi* $[a] \in_{\mathfrak{U}} [b]$ *omm* $\{i \in I : a_i \in b_i\} \in \mathfrak{U}$.

Bevis. $\in_{\mathfrak{U}}$ är väldefinierad:

Antag att $[a] \in_{\mathfrak{U}} [b]$ och även att $a =_{\mathfrak{U}} c$ och $b =_{\mathfrak{U}} d$.

Vi har alltså enligt antagandet att:

$\{i : a_i \in b_i\} \in \mathfrak{U}$, $\{i : a_i = c_i\} \in \mathfrak{U}$ samt $\{i : b_i = d_i\} \in \mathfrak{U}$

Filter-egenskaperna ger att snittet av dessa mängder är ett element i $\mathfrak{U}$ vilket är ekvivalent med att $\{i : a_i = c_i \wedge b_i = d_i \wedge a_i \in b_i\} \in \mathfrak{U}$.

Klart att denna mängd är $\{i : c_i \in d_i\}$, d.v.s $[c] \in_{\mathfrak{U}} [d]$

$\square$

Nu har vi allt som behövs för att definiera M (och därmed $*$). På så vis kommer vi att hamna i en superstruktur igen och kan använda våra vanliga definitioner av $=$ och $\in$.

Definition 8.8. $M : \prod_{\mathfrak{U}}^{0} V(\mathbb{R}) \to V({}^*\mathbb{R})$ *definieras induktivt på Rank i den begränsade ultraprodukten:*

Givet $b \in \prod_{\mathfrak{U}} [V_n(\mathbb{R}) \setminus V_{n-1}(\mathbb{R})]$ *låter vi*

$M([b]) = [b]$, *för* $n = 0$

$M([b]) = \Big\{ \ M([a]) : [a] \in \prod_{\mathfrak{U}} \bigcup_{k=0}^{n-1} [V_k(\mathbb{R}) \setminus V_{k-1}(\mathbb{R})] \, och [a] \in_{\mathfrak{U}} [b] \Big\}$,

för $[b] \in \prod_{\mathfrak{U}} [V_n(\mathbb{R}) \setminus V_{n-1}(\mathbb{R})]$ *där* $n \geq 1$.

$$
\begin{tikzcd}
 & \mathcal{L}_{V(\mathbb{R})} \arrow[ddl, "T"'] \arrow[d, "T_{\mathfrak{U}}"] \arrow[ddr, "{}^*T"] & \\
 & \prod V(\mathbb{R}) \arrow[dr, "M"] & \\
V(\mathbb{R}) \arrow[ur, "e"] \arrow[rr, "*"] & & V({}^*\mathbb{R})
\end{tikzcd}
$$

8.4 Bevis av transferprincipen

Vi har definierat transferprincipen som punkten **(v)** i definition 8.1 av monomorfi, alltså:

Transferprincipen:

Om ϕ är en $\mathcal{L}_{V(\mathbb{R})}$-formel utan fria variabelförekomster så:

$V(\mathbb{R}) \models \phi$ omm $V(^*\mathbb{R}) \models \phi$

Detta kommer att följa som ett korollarium av *Łoś*[8]*s sats.*

Sats 8.9. Łośs sats:

Om $\phi(x_1, ..., x_n)$ *är en formel i* $\mathcal{L}_{V(\mathbb{R})}$ *med* $x_1, ..., x_n$ *som enda fria variabler och* $[a_1], ..., [a_n] \in \prod_{\mathfrak{U}}^0 V(\mathbb{R})$ *så gäller:*

$V(^*\mathbb{R}) \models \phi(M([a_1]), ..., M([a_n]))$ *omm* $\{i \in I : V(\mathbb{R}) \models \phi(a_1(i), ..., a_n(i))\} \in \mathfrak{U}$.

Bevis. Beviset följer definition 6.6 av begreppet sanning och använder induktion över antalet logiska symboler n förekommande i ϕ. Vi definierar då n som formelns komplexitet.

Basfall: Om $n = 0$ är ϕ en atomär $\mathcal{L}_{V(\mathbb{R})}$-formel. Enligt stycke 7.4 är ϕ då på endera av formerna $t_1 = t_2$ eller $t_1 \in t_2$ där t_1, t_2 är antingen konstant- eller variabelsymboler.

- Antag ϕ är $\mathcal{L}_{V(\mathbb{R})}$-formeln $t_1 = t_2$. Då är utvärderingen av ϕ under någon sekvens s av element ur $V(^*\mathbb{R})$ på formen $M([a_1]) = M([a_2])$ där $M([a_1])$ och $M([a_2])$ antingen är första respektive andra komponent i sekvensen s om t_1 respektive t_2 är varibelsymboler, eller tolkningen av t_1 respektive t_2 om de är konstantsymboler.

  Motsvarande utvärdering i $\prod_{\mathfrak{U}}^0 V(\mathbb{R})$ är på formen $\{i \in I : a_1(i) = a_2(i)\} \in \mathfrak{U}$ där $a_1(i)$ och $a_2(i)$ är de i:te komponenterna av $[a_1]$ respektive $[a_2]$.

  Vidare kan vi anta att $[a_1], [a_2] \in \prod V_0(\mathbb{R})$ vilket ger att $M([a]) = [a]$, annars är likheten mängdteoretisk (d.v.s definierad i termer av $\in$).

  Enligt definition är $\{i \in I : a_1(i) = a_2(i)\} \in \mathfrak{U}$ omm $a_1 =_{\mathfrak{U}} a_2$ omm $[a_1] = [a_2]$.

- Antag att ϕ är formeln $x \in y$.

  På samma sätt får vi uttrycken $\{i \in I : a_1(i) \in a_2(i)\} \in \mathfrak{U}$ och $M([a_1]) \in M([a_2])$. Antag att $\{i \in I : a_1(i) \in a_2(i)\} \in \mathfrak{U}$ är sann, enligt definition gäller då $[a_1] \in_{\mathfrak{U}} [a_2]$ och $M([a_1]) \in M([a_2])$ om M bevarar *Rank*, men det är fallet enligt konstruktion (Se 8.5).

Induktionsfall: Antag sant för alla formler ψ med komplexitet $< n$. Enligt 6.5 räcker det att visa tre fall: $\phi = \neg\psi$, $\phi = \psi_1 \wedge \psi_2$ samt $\phi = (\exists x \in c)\psi$

- $\phi = \neg\psi$. Då är följande ekvivalent:

[8] Uttalas "wash". Tydligen betyder det även "älg" på polska.

$V(^*\mathbb{R}) \models \phi(M([a_1]), ..., M([a_n]))$

$V(^*\mathbb{R}) \models \neg\psi(M([a_1]), ..., M([a_n]))$

$V(^*\mathbb{R}) \not\models \psi(M([a_1]), ..., M([a_n]))$ enligt sanningsdefinitionen

$\{i \in I : V(\mathbb{R}) \models \psi(a_1(i), ..., a_n(i))$ är sann$\} \notin \mathfrak{U}$, enligt hypotesen

$\{i \in I : V(\mathbb{R}) \models \neg\psi(a_1(i), ..., a_n(i))$ är sann$\} \in \mathfrak{U}$ då $\mathfrak{U}$ är ett ultrafilter.

- $\phi = \psi_1 \wedge \psi_2$. Då är följande ekvivalent:

  $V(^*\mathbb{R}) \models \phi(M([a_1]), ..., M([a_n]))$

  $V(^*\mathbb{R}) \models (\psi_1 \wedge \psi_2)(M([a_1]), ..., M([a_n]))$

  $V(^*\mathbb{R}) \models \psi_1(M([a_1]), ..., M([a_n]))$ och $V(^*\mathbb{R}) \models \psi_2(M([a_1]), ..., M([a_n]))$

  enligt sanningsdefinitionen

  $\{i \in I : V(\mathbb{R}) \models \psi_1(a_1(i), ..., a_n(i))\} \in \mathfrak{U}$ och

  $\{i \in I : V(\mathbb{R}) \models \psi_2(a_1(i), ..., a_n(i))\} \in \mathfrak{U}$ enligt hypotesen

  $\{i \in I : V(\mathbb{R}) \models \psi_1(a_1(i), ..., a_n(i))\} \cap \{i \in I : V(\mathbb{R}) \models \psi_2(a_1(i), ..., a_n(i))\} \in \mathfrak{U}$

  då $\mathfrak{U}$ är ett filter och alltså slutet under snitt.

  $\{i \in I : \psi_1 \wedge \psi_2(a_1(i), ..., a_n(i))$ är sann$\} \in \mathfrak{U}$

- $\phi = (\exists x \in c)\psi$.

  Vi visar först ena implikationen:

  Antag att $V(^*\mathbb{R}) \models (\exists x \in c)\phi(M([a_1]), ..., M([a_n]), x))$

  Enligt sanningsdefinitionen finns då $M([a]) \in V(^*\mathbb{R})$ sådant att

  $V(^*\mathbb{R}) \models M([a]) \in c \wedge \phi(M([a_1]), ..., M([a_n]), M([a]))$, för något $c \in V(^*\mathbb{R})$.

  Detta motsvarar fallet $\phi = \psi_1 \wedge \psi_2$ ovan och enligt detta får vi att

$\{i \in I : V(\mathbb{R}) \models a(i) \in c \wedge \phi(a_1(i), ..., a_n(i), a(i))\} \in \mathfrak{U}$.

Klart att denna mängd är en delmängd till $\{i \in I : V(\mathbb{R}) \models \exists x \in c \wedge \phi(a_1(i), ..., a_n(i), x)\} = U$. Ultrafiltret $\mathfrak{U}$ slutet under övre delmängd ger då att $U \in \mathfrak{U}$.

Antag å andra sidan att $\{i \in I : V(\mathbb{R}) \models \exists x \in c \wedge \phi(a_1(i), ..., a_n(i), x)\} \in \mathfrak{U}$.

Då finns alltså en entitet $a \in V(\mathbb{R})$ som vi med urvalsaxiomet kan välja som vittne. Detta a satisfierar ψ och $V(\mathbb{R}) \models \phi$. Vidare kan vi anta att $a \in c$ för något c. Enligt konstruktion är ekvivalensklassen för sekvensen $a_i = a$, alltså $[\bar{a}] \in \prod_{\mathfrak{U}}^0 V(\mathbb{R})$.

Vi har alltså att $V({}^*\mathbb{R}) \models M([\bar{a}]) \in {}^*c \wedge \phi(M([a_1]), ..., M([a_n]), M([a]))$

□

Korollarium 8.10. *Bevis av egenskap (v) Transferprincipen*

Låt ϕ vara en $\mathcal{L}_{V(\mathbb{R})}$-formel som saknar fria varibelförekomster. Då har vi enligt Łośs sats att ${}^*\phi$ är sann i $V({}^*\mathbb{R})$ omm $\{i \in I : \phi$ är sann i $V(\mathbb{R})\} \in \mathfrak{U}$. Så att ϕ saknar fria variabler ger att $\{i \in I : \phi$ är sann i $V(\mathbb{R})\}$ antingen är hela I eller $\emptyset$. Enligt definitionen av ultrafilter har vi att $I \in \mathfrak{U}$ och $\emptyset \notin \mathfrak{U}$ alltså är ${}^*\phi$ sann i $V({}^*\mathbb{R})$ omm ϕ är sann i $V(\mathbb{R})$.

Vi vill nu visa att $*$ som sammansättningen av avbildningarna e och M är en monomorfi mellan $V(\mathbb{R})$ och $V({}^*\mathbb{R})$ enligt 8.1. Då vi redan visat Transferprincipen kommer vi använda oss av den i beviset.

Lemma 8.11. ${}^*(A \setminus B) = {}^*A \setminus {}^*B$ *för* $A, B \in V(\mathbb{R})$

Bevis.

Vi kan karakterisera mängden $(A \setminus B) \in V(\mathbb{R})$ i $\mathcal{L}_{V(\mathbb{R})}$ med följande formel:

$x \in (A \setminus B) \Leftrightarrow (x \in A \wedge x \notin B)$

Föjdaktligen har vi enligt transferprincipen att

$x \in {}^*(A \setminus B) \Leftrightarrow (x \in {}^*A \wedge x \notin {}^*B)$

karakteriserar mängden ${}^*(A \setminus B) \in V({}^*\mathbb{R})$

$\square$

Sats 8.12. $* : V(\mathbb{R}) \to V({}^*\mathbb{R})$ *är en monomorfi*

Bevis.

(i) ${}^*(\emptyset) = \emptyset$

Enligt lemmat har vi: ${}^*\emptyset = {}^*(A \setminus A) = {}^*A \setminus {}^*A = \emptyset$

(ii) $a \in X \implies {}^*a \in {}^*X$ samt ${}^*n = n, \forall n \in \mathbb{N}$

Antag $a \in X$.

Följande är ekvivalent:

$a \in X$

$\{a\} \subseteq X$

$\{a\} \setminus X = \emptyset$

Enligt lemmat har vi då att:

${}^*\emptyset = {}^*(\{a\} \setminus X) = {}^*\{a\} \setminus {}^*X = \emptyset.$

Enligt ekvivalenserna ovan är detta ekvivalent med att ${}^*a \in {}^*X$

Att ${}^*n = n, \forall n \in \mathbb{N}$ följer helt enkelt av att vi definierat ${}^*n = [n]$ för individer och att både n och $[n]$ betecknar talet n[9].

(iii) $a \in V_{n+1}(\mathbb{R}) \setminus V_n(\mathbb{R}) \implies {}^*a \in V_{n+1}({}^*\mathbb{R}) \setminus V_n({}^*\mathbb{R})$

Antag $a \in (V_{n+1}(\mathbb{R}) \setminus V_n(\mathbb{R}))$.

[9]Likheten i uttrycket ${}^*n = n, \forall n \in \mathbb{N}$ är alltså inte relationen '=' i varken $\mathcal{L}_{V(\mathbb{R})}$ eller $\mathcal{L}_{V({}^*\mathbb{R})}$ utan istället en likhet 'utifrån' sett, som påstår att de båda entiteterna i respektive struktur utgör namn på samma naturliga tal n, talet n anses då existera utanför dessa strukturer. Detta alltså i motsats till det matematiska uttrycket $n = (n_0, n_1, ...)$ vars tolkning i både $V(\mathbb{R})$ och $V({}^*\mathbb{R})$ naturligtvis är falsk.

Enligt transferprincipen är antagandet ekvivalent med:

$$^*a \in {}^*(V_{n+1}(\mathbb{R}) \setminus V_n(\mathbb{R})) = {}^*V_{(n+1)}(\mathbb{R})/{}^*V_n(\mathbb{R})$$

Påstår nu att ${}^*V_k(\mathbb{R}) \subseteq V_k({}^*\mathbb{R})$. Även detta följer av transfer, betrakta $\mathcal{L}_{V(\mathbb{R})}$-formeln ϕ:

$$\phi \doteqdot \forall x \in P(X)(\forall y \in x)(y \in X)$$

ϕ uttrycker alltså en egenskap hos potensmängder, nämligen att om x är element i potensmängden till X så är x en delmängd till X, det vill säga att alla element y till x är även element i X. Detta är uppenbarligen sant, alltså har vi att tolkningen av ϕ i $V({}^*)$ är sann, med andra ord gäller:

$$\forall x \in {}^*P(X)(\forall y \in x)(y \in {}^*X)$$

D.v.s att för $x \in {}^*P(X)$ gäller att $x \subseteq {}^*X$, vilket är ekvivalent med att $x \in P({}^*X)$ så vårt ursprungliga antagande implicerar ${}^*a \in V_{n+1}({}^*X) \setminus V_n({}^*X)$.

(iv) Om $b \in {}^*V_n(X), n \geq 1$ och $a \in b$ så $a \in {}^*V_{n-1}(X)$

Omedelbart från definitionen av M.

$\square$

8.5 Exempel på tillämpning av transferprincipen

Bevis av sats 2.2: $< {}^*\mathbb{R}, +, \times \leq >$ är en ordnad kropp med etta och nolla

Bevis.

Bevisidén är helt enkelt att kroppsaxiomen för $\mathbb{R}$ kan uttryckas i $\mathcal{L}_{V(\mathbb{R})}$. Vi väljer att uttrycka dessa med sedvanlig notation för addition och multiplikation istället för med de enda relationssymbolerna $=$ och $\in$ i $\mathcal{L}_{V(\mathbb{R})}$.

Associativitet: $\forall x, y, z \in \mathbb{R}((x + y) + z = x + (y + z) \wedge (x \times y) \times z = x \times (y \times z))$

Kommutativitet: $\forall x, y \in \mathbb{R}((x + y = y + x) \wedge (xy = yx))$

Distributivitet: $\forall x, y, z \in \mathbb{R}(x(y + z) = xy + xz)$

Nolla: $\forall x \in \mathbb{R}(x + 0 = x)$, där 0 är en individkonstant i $Dom(\mathbb{R})$

$\forall x \in \mathbb{R}(x \times 0 = 0)$

Etta: $\forall x \in \mathbb{R}(x \times 1 = x)$, där 1 är en individkonstant i $Dom(\mathbb{R})$

Additiv invers: $\forall x \in \mathbb{R} \exists y(x + y = 0)$

Multiplikativ invers: $\forall x \in (\mathbb{R} \setminus 0) \exists y(xy = 1)$

Ordningen $\leq$ reflexiv: $\forall x, y \in \mathbb{R}(x \leq y \lor y \leq x)$

Ordningen $\leq$ antisymmetrisk: $\forall x, y \in \mathbb{R}(x \leq y \land y \leq x \Longrightarrow x = y)$

Ordningen $\leq$ transitiv: $\forall x, y, z \in \mathbb{R}(x \leq y \land y \leq z \Longrightarrow x \leq z)$

Ordningen $\leq$ är total: $\forall x, y \in \mathbb{R}(x \leq y \land y \leq x)$

Vi vet att $< \mathbb{R}, +, \times \leq >$ är en ordnad kropp med etta och nolla, alltså att tolkningen av kroppsaxiomen i $\mathbb{R}$ är sanna så enligt transferprincipen är även tolkningen av kroppsaxiomen i ${}^*\mathbb{R}$ alla sanna och därför är $< {}^*\mathbb{R}, +, \times \leq >$ är en ordnad kropp med etta och nolla.

$\square$

9 Utvidgningar

Detta kapitel följer [Hurd och Loeb].

9.1 Standard och icke-standard entiteter

Vi vill nu genomföra motsvarande argument som i kapitel 2.5, alltså visa att $V({}^*\mathbb{R})$ innehåller icke-standard entiteter. Detta kommer även i nuvarande kontext av superstrukturer bero på vilket ultrafilter som använts i konstruktionen.

Definition 9.1. *$b \in V({}^*\mathbb{R})$ kallas en standardentitet om $b = {}^*a$ för något $a \in V(\mathbb{R})$*

Vi kallar mängden av alla standardentiteter i ${}^*A$ för bilden av A (under *-transformen) och inför skrivsättet:

Definition 9.2. ${}^{im}A = \{{}^*a : a \in A\}$

Sats 9.3. *Om $A \in V(\mathbb{R})$ har ändlig kardinalitet så är ${}^{im}A = {}^*A$*

Bevis.

Om A är ändlig kan vi anta att $A = \{a_1, a_2, ..., a_n\}$, för något $n \in \mathbb{N}$

Låt ϕ vara $\mathcal{L}_{V(\mathbb{R})}$-formeln $\forall x \in A(x = a_1 \vee x = a_2 \vee ... \vee x = a_n) \wedge (a_1 \in A \wedge a_2 \in A, ..., \wedge a_n \in A)$

ϕ är uppenbarligen sann enligt antagandet och karakteriserar entiteten A. Enligt transferprincipen är tolkningen av ϕ i $V({}^*\mathbb{R})$ sann. Vi har alltså

$\forall x \in {}^*A(x = {}^*a_1 \vee x = {}^*a_2 \vee ... \vee x = {}^*a_n) \wedge ({}^*a_1 \in {}^*A \wedge {}^*a_2 \in {}^*A, ..., \wedge {}^*a_n \in {}^*A)$

Alltså att ${}^*A = \{{}^*a_1, {}^*a_2, ..., {}^*a_n\} = {}^{im}A$ $\square$

Om nu $V({}^*\mathbb{R})$ inte innehåller några icke-standard entiteter så har vi följdaktligen att $V({}^*\mathbb{R})$ helt enkelt är bilden av $V(\mathbb{R})$ och allt $*$ gör är att byta namn på entiteterna. Detta är exakt vad som inträffar ifall $V({}^*\mathbb{R})$ konstrueras med ett principalt ultrafilter.

För att visa att, givet ett lämpligt ultrafilter, vi kan konstruera ett icke-standard universum som innehåller icke-standard entiteter kommer vi införa begreppet ”utvidgning”. I en utvidgning kommer sats 9.3 gälla **endast om** entiteten A har ändlig kardinalitet.

9.2 Hyperändliga mängder

Vår definition av utvidgning kommer att bygga på tolkningen av ändlighet i $V(^*\mathbb{R})$. För att genomföra detta behöver vi ett $\mathcal{L}_{V(\mathbb{R})}$-uttryck som karakteriserar de ändliga mängderna i $V(\mathbb{R})$.

Normalt, d.v.s när vi bedriver matematik i $V(\mathbb{R})$, säger vi att en mängd A har ändlig kardinalitet ifall det existerar en bijektion mellan A och något initialsegment till $\mathbb{N}$. Vi skall nu se att analogin till detta i $V(^*\mathbb{R})$ helt enkelt mosvaras av en bijektion till ett initialsegment till $^*\mathbb{N}$, vilket är en uppenbar konsekvens av transferprincipen.

För att detta inte helt enkelt skall bli ett namnbyte enligt ovan kan vi tillsvidare anta att $^*\mathbb{N} \supsetneq \mathbb{N}$ utifrån våra intuitioner från kapitel 2. Detta kommer emellertid följa, i vår formella kontext, som ett enkelt korrolarium av sats 9.11 nedan.

Definition 9.4.

Antag att A är en entitet i $V(\mathbb{R})$ av rank strikt större än noll. Låt då $P_F(A)$ vara mängden av alla ändliga delmängder till A. Vi kallar då $^*P_F(A)$ i $V(^*\mathbb{R})$ för mängden av alla **hyperändliga** *delmängder till $^*A$.*

Sats 9.5. *Låt J vara ett initialsegment till $^*\mathbb{N}$, $J = \{n \in {^*\mathbb{N}} : n \leq j\}$ för något $j \in {^*\mathbb{N}}$.*

Antag att $B \in {^*V_k(\mathbb{R})}$ är en hyperändlig mängd. Då finns en bijektion $f : J \longrightarrow B$ i $^*V_{(k+4)}(\mathbb{R})$.

Bevis.

Enligt antagandet har vi alltså att $B \in {^*P_F(A)}$ för något $A \in V_n(\mathbb{R})$, $n \geq 1$.

Betrakta följande $\mathcal{L}_{V(\mathbb{R})}$-formler (uppdelade i segment för att öka läsbarheten):

$\forall B \in P_F(V_n(X)) \ \exists j \in \mathbb{N} \ \exists f \in V_{n+4}(X))$

Om B är en ändlig entitet så finns ett naturligt tal j och en funktion f

$(\forall x, y \in \mathbb{N} : x, y \leq j(x \neq y \Rightarrow f(x) \neq f(y))$

f är injektiv på en domän av naturliga tal mindre eller lika med j

$(\forall b \in B \; \exists x \leq j \in \mathbb{N} : f(x) = b)$

f är surjektiv på B

Låt ϕ vara sammansättningen av dessa formler, ϕ uttrycker då helt enkelt att för ändliga B gäller den definition som introducerats ovan nämligen att B är ändlig om det finns en bijektion på något initialsegment $\{1, 2, 3, ..., j\}$ av $\mathbb{N}$. ϕ är alltså sann då formeln uttrycker vår definition av ändlighet i $V(\mathbb{R})$. Följdaktligen gäller alltså enligt transferprincipen att tolkningen av ϕ i $V(^*\mathbb{R})$ är sann. Denna tolkning är att en mängd $^*B$ är hyperändlig om motsvarande bijektion existerar på något initialsegment till $^*\mathbb{N}$.

□

Enligt vårt antagande att $^*\mathbb{N} \setminus \mathbb{N}$ är icketom innebär detta att vi i någon mening utvidgat begreppet ändlighet från begränsade till obegränsade tal genom $*$-transformen.

Som en följd av satsen ovan definierar vi för hyperändliga mängder den **interna kardinaliteten**[10] som j.

Karakteriseringen av ändliga mängder som de mängder som har en bijektion på något initialsegment på $\mathbb{N}$ i $V(\mathbb{R})$ är inte den enda möljliga. Definitionen är bevisbart ekvivalent med både Dirichlets lådprincip och Tarskis definition av ändlighet:

- Dirichlets lådprincip säger att en mängd är ändlig omm varje injektion från en mängd på sig själv även är en surjektion.

  Enligt transferprincipen får vi då på samma sätt att en mängd är hyperändlig omm varje injektion från mängden på sig själv även är en surjektion.

- Tarskis definition av ändlighet säger att en mängd F är ändlig omm varje icke-tom delmängd till $P(F)$ har ett $\subseteq$-minimalt element.

  På samma sätt för vi då enligt transferprincipen att F är hyperändlig omm varje icketom delmängd till $^*P(F)$ har ett $\subseteq$-minimalt element.

[10] Bijektionen i sats 9.5 är en så kallad 'intern' funktion, definitionen av interna entiteter kommer presenteras i kapitel 10. Begreppet internalitet är dock endast intressant i kontext av en utvidgning.

9.3 Utvidgningar

Definition 9.6. *Superstrukturen $V(^*\mathbb{R})$ är en* **utvidgning** *av superstrukturen $V(\mathbb{R})$ med avseende på en monomorfi $*$ om:*

För alla entiteter $A \in V(\mathbb{R})$ av Rank > 0 finns en entitet $B \in {}^*P_F(A)$ sådan att $^*a \in B$ för alla $a \in A$, d.v.s att en hyperändlig mängd B innehåller bilden av godtycklig entitet A under $*$-avbildningen.

Om ultraproduktkonstruktionen $V(^*\mathbb{R})$ är en utvidgning av $V(\mathbb{R})$ beror, som redan antytts, på valet av ultrafilter. Detta är analogt med att $\omega \in {}^*\mathbb{R}$ i konstruktionen av de hypereella talen i kapitel 2.5 berodde på att ultrafiltret innehöll alla cofinita mängder av naturliga tal. Vi konstruerar nu ett filter vilket gör $V(^*\mathbb{R})$ till en utvidgning av $V(\mathbb{R})$ enligt definitionen ovan.

Definition 9.7.

Låt J vara mängden av alla icketomma ändliga entiteter i $V(\mathbb{R})$ av Rank > 0. Så om $a \in J$ gäller enligt definition att det finns en entitet i $b \in V(\mathbb{R})$ av Rank > 0 där $a \in (P_F(b)/\emptyset)$. För $a \in J$ definierar vi $J_a = \{b \in J : a \subseteq b\}$, J_a är alltså mängden av ändliga entiteter som har a som delmängd.

Lemma 9.8. *Mängden $\mathfrak{F} = \{A \subseteq J : \exists a \in J$ sådant att $J_a \subseteq A\}$ är ett fritt äkta filter på J.*

Bevis.

$\mathfrak{F}$ är sluten under snitt:

Antag att $A_1, A_2 \in \mathfrak{F}$. Enligt definition har vi då att det finns $a_1, a_2 \in J$ sådana att $A_1 \supseteq J_{a_1}$ och $A_2 \supseteq J_{a_2}$. Följdaktligen har vi att $A_1 \cap A_2 \subseteq J_{a_1} \cap J_{a_2}$.

Men $J_{a_1} \cap J_{a_2} = \{b \in J : a_1 \subseteq b \wedge a_2 \subseteq b\} = \{b \in J : a_1 \cup a_2 \subseteq b\} = J_{a_1 \cup a_2}$. Om a_1 och a_2 bägge är ändliga och icketomma är även a_1 och a_2 det. Följdaktligen är $A_1 \cap A_2 \in \mathfrak{F}$

$\mathfrak{F}$ är sluten under övre delmängd:

Antag att $A_1 \in \mathfrak{F}$ och att $A_1 \subseteq B \in J$. Enligt definition har vi $\mathfrak{F} = \{A \subseteq J : \exists a \in J$ sådant att $J_a \subseteq A\}$. Vi har $J_a \subseteq A_1$ och $A_1 \subseteq B$ så klart att $J_a \subseteq B$ varför $B \in \mathfrak{F}$.

$\mathfrak{F}$ är icke-principalt:

Antag $a \in J$. Givet ett $b \in J$ sådant att $a \cap b = \emptyset$ gäller då enligt definition av J_b att $a \notin J_b$ eftersom $b \not\subseteq a$. Då är det klart att $J_b \subseteq (J/\{a\}) \subseteq J$ varför $(J/\{a\}) \in \mathfrak{F}$. Det är alltså klart att om $\mathfrak{F}$ innehåller någon ändlig entitet $a \in J$ så innehåller $\mathfrak{F}$ även komplementet till denna entiet och därmed, som redan visats ovan, även snittet av dessa, d.v.s $\emptyset$. Detta strider direkt mot definitionen av $\mathfrak{F}$ eftersom vi definierat att $a \in (P_F(b)/\emptyset)$, d.v.s att alla entiteter i $a \in (P_F(b)/\emptyset)$ är icke-tomma.

$\square$

Att $\mathfrak{F}$ är icke-principalt visar att ett ultrafilter som innehåller $\mathfrak{F}$ kommer att innehålla komplementet till alla ändliga entiteter analogt med Frechetfiltret i kapitel 5. Enligt ultrafiltersatsen garanteras existensen av ett sådant ultrafilter $\mathfrak{U} \supseteq \mathfrak{F}$ och vi kan visa:

Sats 9.9. *Om $V(^*\mathbb{R})$ konstruerats med ett ultrafilter $\mathfrak{U} \supseteq \mathfrak{F}$ enligt lemmat så är $V(^*\mathbb{R})$ en utvidgning av $V(\mathbb{R})$.*

Bevis.

Antag att A är en mängd i $V(\mathbb{R})$ av $Rank > 0$.

Vi definierar en funktion $\Gamma : J \to P_F(A)$ där $\Gamma_a = a \cap A$. Då är $\Gamma \in \prod P_F(A)$ och följaktligen kan vi definiera $B = M([\Gamma])$. Då är $B \in {}^*P_F(A)$. Om $x \in A$ så är självklart $J_x = \{a \in J : x \in a\} \subseteq \{a \in J : x \in a \cap A\} = C$ och följaktligen har vi att $\{a \in J : x \in a \cap A\} \in \mathfrak{U} \supseteq \mathfrak{F}$ enligt definitionen av $\mathfrak{F} = \{C \subseteq J : \exists x \in J \text{ sådant att } J_x \subseteq C\}$. Enligt vår gängse notation har vi alltså att $[\bar{x}] \in_{\mathfrak{U}} [\Gamma]$ vilket enligt LOs sats är ekvivalent med ${}^*x \in B$. $\square$

Som utlovat ska vi visas att vår definition av utvidgning implicerar existensen av icke-standard entiteter. För att göra detta införs ytterligare ett logiskt begrepp:

Definition 9.10. *En två-ställig relation R kallas* **ändligt satisfierbar** *på $A \subseteq Dom(R)$ om det för varje ändlig mängd $\{x_1, x_2, ..., x_n\} \in A$ finns ett y i målmängden till R sådant att $\langle x_i, y\rangle \in R$ för alla $i \leq n$. Vi säger att R är ändligt satisfierbar om den är ändligt satisfierbar på hela $Dom(R)$.*

Sats 9.11. *Om $V(^*\mathbb{R})$ är en utvidgning av $V(\mathbb{R})$ så finns för varje ändligt satisfierbar relation $R \in V(\mathbb{R})$ ett element ett element b i $Range(^*R)$ sådant att $\langle {}^*x, b\rangle \in {}^*R$ för alla $x \in Dom(R)$.*

Bevis.

Låt alltså R vara en godtycklig ändligt satisfierbar relation. Enligt definition 9.6 finns $B \in {}^*P_F(Dom(R))$ sådant att för alla $x \in Dom(R)$ gäller ${}^*x \in B$. Följande $\mathcal{L}_{V(\mathbb{R})}$-formel uttrycker att R är en ändligt satisfierbar relation i $V(\mathbb{R})$ och är alltså sann enligt antagande:

$\forall w \in P_F(Dom(R))\exists y \in Range(R)\forall x \in w(\langle x, y\rangle \in R)$

Enligt transferprincipen är tolkningen av formeln i $V({}^*\mathbb{R})$ sann, vilket visar satsen. $\square$

Korollarium 9.12. *${}^*\mathbb{N} \setminus \mathbb{N}$ är icke-tom.*

Bevis. Det är klart att $<$ är en ändligt satisfierbar relation på $\mathbb{N}$ eftersom varje ändlig delmängd till $\mathbb{N}$ har ett största element. För att följa terminologin kallar vi relationen $<$ för R. Enligt sats 9.11 har vi då att det finns $b \in {}^*R$ sådant att ${}^*x < b$ för alla $x \in Dom(R) = \mathbb{N}$. D.v.s att b är ett obegränsat naturligt tal och heller inte på formen $b = {}^*a$ för något $a \in V(\mathbb{R})$. $\square$

På entiteter av oändlig kardinaltitet är negationen av identitetsrelationen $\neq$ ändligt satisfierbar. Av detta följer att $*$-avbildningen av varje sådan mängd innehåller icke-standard entiteter. Vi har:

Sats 9.13. *Antag att $A \in V(\mathbb{R})$ har oändlig kardinalitet. Då innehåller ${}^*A$ icke-standard entiteter.*

Bevis.

Relationen $\neq$ är ändligt satisfierbar på A eftersom vi för varje ändlig delmängd $A_i \subseteq A$ kan välja ett $a \in (A/A_i)$. Klart att detta a då uppfyller $a \neq a_i$ för alla $a_i \in A_i$. Enligt sats 9.11 finns då $b \in {}^*A$ sådant att $b \neq {}^*a$ för alla $a \in A$, d.v.s att A innehåller icke-standard entiteter. $\square$

Abraham Robinsons ursprungliga konstruktion av icke-standard analys utgick från satsen om ändlig satisfierbarhet framför vår definition av utvidgning utifrån begreppet hyperändlighet. I själva verket är satsen vi just visat ekvivalent med vår definition av utvidgning vilket följer enkelt genom att vår definition av utvidgning är kan uttryckas av en $\mathcal{L}_{V(\mathbb{R})}$-formel och att relationen $R = \{\langle x, y\rangle : x \in y \wedge \text{y är en ändlig delmängd till A}\}$ är ändligt satisfierbar.

Bevis.

Antag alltså att det för varje ändligt satisfierbar relation R finns ett element b i $Range({}^*R)$ sådant att $\langle {}^*x, b\rangle \in {}^*R$ för alla $x \in Dom(R)$. Låt R vara relationen att vara vara element i en ändlig delmängd till någon mängd A enligt ovan. Denna relation är alltså ändligt satisfierbar. Enligt antagandet (alltså sats 9.11) finns då en hyperändlig delmängd $b \subseteq {}^*A$ sådan att ${}^*x \in b$ för alla $x \in A$, alltså var definition av utvidgning. $\square$

10 Interna mängder, permanens och mättnad

I detta kapitel utgår vi från [Goldblatt] och [Hurd och Loeb].

Vi definierar en entitet i $V({}^*\mathbb{R})$ som *intern* om den är ett element i någon standard mängd:

$$\text{a är intern om och endast om } a \in {}^*A \text{ för något } A \in V(\mathbb{R})$$

Notera att alla standardentiteter är interna, eftersom vi har att ${}^*a \in {}^*\{a\}$. En $\mathcal{L}_{V({}^*\mathbb{R})}$-formel Φ sägs vara intern om dess konstanter är interna.

Element i $V({}^*\mathbb{R})$ som inte är interna kallar vi för *externa*. Interna mängder har många trevliga egenskaper vilka vi ska utforska i senare kapitel. Vi börjar med att genom en rad satser undersöka vilka mängder som är interna och vilka som inte är det.

10.1 Interna mängder

Sats 10.1. *Mängden av alla interna element i $V({}^*\mathbb{R})$ är ${}^*V(\mathbb{R}) = \bigcup_{n=0}^{\infty} {}^*V_n(\mathbb{R})$.*

Bevis. Om $b \in {}^*V(\mathbb{R})$, så är $b \in {}^*V_n(\mathbb{R})$ för något $n \in \mathbb{N}$ vilket ger att b är intern eftersom ${}^*V_n(\mathbb{R})$ är standard. Detta visar att alla element i ${}^*V(\mathbb{R})$ är interna. Om b är intern så har vi att $b \in {}^*a$ där $a \in V_{n+1}(\mathbb{R}) \setminus V_n(\mathbb{R})$ för något heltal $n \geq 1$, så $a \subseteq V_n(X)$ vilket ger $b \in {}^*a \subseteq {}^*V_n(\mathbb{R})$. Detta visar att alla interna element i $V({}^*\mathbb{R})$ finns i ${}^*V(\mathbb{R})$. $\square$

Sats 10.2. *Varje intern mängd är ett element i en standardmängd som är transitiv vilket medför att ${}^*V(\mathbb{R})$ är starkt transitiv.*

Bevis. Låt A vara intern och säg att $A \in {}^*B$ för något $B \in V(\mathbb{R})$. Eftersom $V(\mathbb{R})$ är starkt transitiv så finns det ett transitivt $T \in V(\mathbb{R})$ med $B \subseteq T$. Men eftersom $*$-avbildningen bevarar transitivitet (vilket ses genom ett enkelt transferargument) är även ${}^*T$ transitivt. Dessutom har vi att ${}^*B \subseteq T$ vilket medför att $A \in {}^*T$ vilket bevisar första delen av satsen. Eftersom ${}^*T$ är transitiv och varje medlem av ${}^*T$ är intern så har vi också att $A \subseteq {}^*T \subset {}^*V(\mathbb{R})$ vilket bevisar andra delen av satsen. $\square$

Vi får också följande viktiga korollarium.

Korollarium 10.3. *Varje medlem av en intern mängd är intern.*

Följande sats knyter ihop interna mängder med $\mathcal{L}_{V(\mathbb{R})}$-formler på ett sätt som kommer visa sig vara intressant när vi ska *bilda* mängder som vi vill ska vara interna.

Sats 10.4 (**Interna Definitionsprincipen**). *Låt $\Phi(x)$ vara en intern $\mathcal{L}_{V(^*\mathbb{R})}$-formel där x är den enda fria variabeln, och låt A vara en intern mängd. Då är $\{x \in A : \Phi(x) \text{ är sann }\}$ intern.*

Bevis. Låt $c_1, \cdots, c_n$ vara konstanterna $\Phi(x)$ och beteckna $\Phi(x) = \Phi(c_1, \cdots, c_n, x)$. Låt nu $A, c_1, \cdots, c_n, x \in {}^*V_k(\mathbb{R})$ för något $k \in N$. Då är formeln

$$(\forall x_1, \cdots, x_n, y \in V_k(\mathbb{R}))(\exists z \in V_{k+1}(\mathbb{R}))(\forall x \in V_k(\mathbb{R}))[x \in z \Leftrightarrow [x \in y \wedge \Phi(x_1, \cdots, x_n, x)]]$$

i $\mathcal{L}_{V(\mathbb{R})}$ sann i $V(\mathbb{R})$. Enligt transferprincipen är då även tolkningen i $V(^*\mathbb{R})$ sann, enligt detta är $\{x \in A : \Phi(x) \text{ är sann}\} \in {}^*V_{k+1}(\mathbb{R})$ (genom att välja $y = A, x_1 = c_1$ o.s.v.). $\square$

Sats 10.5. *Om A och B är interna mängder, så är även följande det: $A \cap B$, $A \cup B$, $A \setminus B$ och $A \times B$.*

Bevis. Vi bevisar först operationen $\cap$. Betrakta följande $\mathcal{L}_{V(\mathbb{R})}$-formel (som säger att $V(\mathbb{R})$ är sluten under snitt):

$$(\forall W, Y \in V_{n+1}(\mathbb{R}))(\exists Z \in V_{n+1}(\mathbb{R}))(\forall x \in V_n(\mathbb{R}))[x \in Z \Leftrightarrow x \in W \wedge x \in Y] \tag{6}$$

(6) är sann eftersom $V(\mathbb{R})$ är sluten under snitt och därmed även tolkningen av formeln i $V(^*\mathbb{R})$ sann enligt transferprincipen. Enligt denna tolkning säger formeln att om A och B är två entiteter i ${}^*V_{n+1}(\mathbb{R})$ så finns en entitet C i ${}^*V_{n+1}(\mathbb{R})$ som har exakt samma element från ${}^*V_n(\mathbb{R})$ som $A \cap B$. Enligt definition 8.1 (iv) gäller att alla element i A, B och C är från ${}^*V_n(\mathbb{R})$, vilket ger $C = A \cap B$.

För $A \cup B$ använder man samma argument fast med $\vee$ istället för $\wedge$ i (6). För $A \setminus B$ så byter vi ut $x \in Y$ mot $x \notin Y$. Fallet $A \times B$ blir lite annorlunda, men inses lätt om man i (6) betraktar $Z \in V_{n+2}(X)$ och $\langle x, y \rangle$ som $\{\{x\}, \{x, y\}\} \in V_{n+2}(X)$. $\square$

I enlighet med 7.5 så kan vi uppfatta funktioner som mängder och därför går det ihop att prata om funktioner som interna och externa.

Sats 10.6. *För en intern funktion f gäller att:*

(i) Om A är en intern delmängd till f:s domän så är $f(A)$ intern.

(ii) Om B är en intern delmängd till bilden av f så är $f^{-1}(B)$ intern.

Bevis.

Enligt korollarium 10.3 har vi att element i interna mängder är interna. Eftersom vi definierat funktioner i superstrukturen som mängder av ordnade par av entiteter gäller alltså att om $f : C \to D$ är intern så är alla dessa ordnade par $\langle a, b \rangle$ interna. Enligt definition har vi att $\langle a, b \rangle = \{\{a\}, \{a, b\}\}$ så på samma sätt gäller då komponentvis att $\langle a, b \rangle$ intern ger att elementen i paren är interna.

Låt $C = \{c_1, c_2, ...\}$ och $D = \{d_1, d_2, ...\}$ för lämpliga indexmängder I och J. Vi har alltså att om $f : C \to D$ är intern så är alla c_i och d_j interna, dvs att det finns x_i och y_j sådana att $c_i \in {}^*x_i$ och $d_j \in {}^*y_j$ för alla i, j i indexmängderna.

Låt nu $X = \bigcup_{i \in I} x_i$ och $Y = \bigcup_{j \in J} y_j$. Enligt transferprincipen gäller $x_i \in X \Rightarrow {}^*x_i \in {}^*X$, vi har alltså att $C \in {}^*X$ och $D \in {}^*Y$, alltså är både $Dom(f)$ och $Range(f)$ interna.

Antag nu, för att visa (ii), att B är en intern delmängd till $Range(f)$. Vi skall använda Interna definitionsprincipen. Betrakta formeln:

$\phi \doteqdot \exists b \in B : \langle a, b \rangle \in f$

Enligt antagandet är då ϕ en intern formel så enligt interna definitionsprincipen är följande mängd intern:

$\{a \in A : \Phi(b) \text{ är sann}\} = f^{-1}(B)$

(i) visas analogt.

$\square$

Nästföljande tre satser kommer att leda fram till flera intressanta bevistekniker i kommande delkapitel.

Lemma 10.7. *Om $a \in V(\mathbb{R}) \setminus \mathbb{R}$ då är ${}^*\mathcal{P}(a)$ mängden av alla interna entiteterna i $\mathcal{P}({}^*a)$*

Bevis. Betrakta följande påstående i $V(\mathbb{R})$ som är sann för $n \geq 1$:

$$(\forall x \in V_n(\mathbb{R}))[(\forall y \in x)[y \in a] \Leftrightarrow x \in \mathcal{P}(a)] \tag{7}$$

[d.v.s. för alla $x \in V_n(\mathbb{R})$, x är en delmängd av a om och endast om $x \in \mathcal{P}$]. Påståendets transfer säger att för alla $x \in {}^*V_n(\mathbb{R})$ så gäller det att x är en delmängd av ${}^*a$ om och endast om $x \in {}^*\mathcal{P}(a)$. Från sats 10.1 har vi att om x är en intern mängd i $V({}^*\mathbb{R})$, d.v.s. $x \in {}^*V(\mathbb{R})$, då är $x \in {}^*V_n(\mathbb{R})$ för något $n \in \mathbb{N}$ och vi kan alltså tillämpa transfern av påståendet (7) på x.

Vi har alltså att x är en delmängd av ${}^*a$ om och endast om den är ett element i ${}^*\mathcal{P}(a)$, vilket ger ${}^*V(\mathbb{R}) \cap \mathcal{P}({}^*a) = {}^*V(\mathbb{R}) \cap {}^*\mathcal{P}(a) = {}^*\mathcal{P}(a)$ (där ${}^*V(\mathbb{R}) \cap \mathcal{P}({}^*a)$ är mängden av alla interna entiteter i $\mathcal{P}({}^*a)$). ☐

Lemma 10.8. *Varje icke-tom intern delmängd till ${}^*\mathbb{N}$ har ett minsta element.*

Bevis. Välordningsprincipen för $\mathbb{N}$ uttrycker att varje delmängd till $\mathbb{N}$ har ett minsta element och kan enkelt uttryckas i $L_{V(X)}$:

$$\phi \quad := \quad (\forall x \in \mathcal{P}(\mathbb{N}) \setminus \emptyset)\,[\,(\exists y \in x)\,(\forall z \in x)\,[y \leq z]\,].$$

Med transfer får vi att ${}^*\phi$ är sann då de naturliga talen är välordnade. Vi har enligt lemma 10.7 att alla interna delmängderna till ${}^*\mathbb{N}$ är exakt ${}^*\mathcal{P}(\mathbb{N})$, vilket tillsammans med ${}^*\phi$ visar satsen. ☐

Satsen kan enkelt generaliseras till ${}^*\mathbb{Z}$, där icketomma uppåt begränsade mängder har ett största element, och nedåt begränsade ett minsta element. När vi nu visat några exempel på interna mängder kan det vara intressant att se på två *externa* mängder.

Sats 10.9. *I en utvidgning $V({}^*\mathbb{R})$ av $V(\mathbb{R})$ är mängderna $\mathbb{N}$ och ${}^*\mathbb{N}_\infty(X)$ externa.*

Bevis. Antag att ${}^*\mathbb{N}_\infty \in \mathcal{P}({}^*\mathbb{N})$ är intern. Enligt lemma 10.8 finns det då ett minsta $b \in {}^*\mathbb{N}_\infty$, men detta är en motsägelse eftersom vi enligt konstruktionen av de hyperreella talen har att även $b - 1 \in {}^*\mathbb{N}_\infty$. Alltså måste ${}^*\mathbb{N}_\infty$ vara extern.

Eftersom ${}^*\mathbb{N}_\infty = {}^*\mathbb{N} \setminus \mathbb{N}$ så har vi att om $\mathbb{N}$ skulle vara intern, så skulle även ${}^*\mathbb{N}_\infty$ vara det enligt sats 10.5. Alltså måste även $\mathbb{N}$ vara extern. ☐

Dessa tre satser ger alltså upphov till nya bevistekniker vilka vi nu ska gå igenom men vi betonar att de är avhängiga på att $V({}^*\mathbb{R})$ är en utvidning! Om inte 9.12 gällde skulle inte 10.9 gälla och 10.8 skulle vara ointressant.

10.2 Permanens och Överspillning

Inom icke-standard analys använder man sig genomgående av bevistekniker som alla grundar sig på att ”interna mängder inte är externa”. Påståendet är visserligen en tautologi, men i en utvidgning finns det flera fall där det inte är självklart hur man ska utnyttja detta faktum. Nedanstående satser klargör.

Sats 10.10 (Permanensprincipen). *Låt $\Phi(x)$ vara en intern formel i $\mathcal{L}_{V({}^*\mathbb{R})}$ med x som den enda fria variablen.*

(i) Om $\Phi(x)$ är sann för varje $x \in \mathbb{N}$, så finns det ett $k \in {}^*\mathbb{N}_\infty$ så att $\Phi(b)$ är sann för varje $b \leq k$ i ${}^*\mathbb{N}$.

(ii) Om $\Phi(x)$ är sann för varje $x \in {}^*\mathbb{N}_\infty$, så finns det ett $k \in \mathbb{N}$ så att $\Phi(b)$ är sann för varje $b \geq k$ i ${}^*\mathbb{N}$.

(iii) Om $\Phi(x)$ är sann för varje infinitesimalt x, så finns det ett reellt $r > 0$ så att $\Phi(b)$ är sann för alla b där $|b| \leq r$ i ${}^*\mathbb{R}$

Bevis. Samtliga tre bevis följer från sats 10.4 (interna definitionsprincipen) och lemma 10.8 (satsen om minsta element för intern delmängd till ${}^*\mathbb{N}$).

(i) Låt $A = \{x \in {}^*\mathbb{N} : \neg\Phi(x)$ är sann i $V({}^*\mathbb{R})\}$. Då är A intern enligt sats 10.4 och $A \subseteq {}^*\mathbb{N}_\infty$ enligt förutsättningen. Om $A = \emptyset$ är vi färdiga, annars har A ett minsta element m enligt lemma 10.8 och vi kan sätta $k = m - 1$.

(ii) Om $\Phi(x)$ Givet den interna mängden A definierad i beviset till (i), så är $A \subseteq \mathbb{N}$ och har en övre begränsning, vilket — enligt en lite modiferad version av lemma 10.8 — ger att mängden har ett största element s och vi kan sätta $k = s + 1$.

(iii) Låt $A = \left\{x \in {}^*\mathbb{N} \setminus \emptyset : \Phi(x) \text{ är sann för alla } y \text{ där } |y| \leq \frac{1}{x}\right\}$ och använd (ii) på $\Phi'(x) := [x \in A]$.

□

Korollarium 10.11 (Överspillningsprincipen). *Låt A vara en intern delmängd till ${}^*\mathbb{R}$.*

(i) Om A innehåller alla begränsade naturliga tal, så innehåller A ett obegränsat naturligt tal.

(ii) Om A innehåller alla obegränsade naturliga tal, så innehåller A ett begränsat naturligt tal.

(iii) Om A innehåller alla positivt infinitesimala tal, så innehåller A ett positivt reellt tal.

Bevis. Påståendena fås direkt från permanensprincipen genom att sätta $\Phi(x) = [x \in A]$. □

Sats 10.12 (Robinsons sekventiella lemma). *Låt $\langle s_n : n \in {}^*\mathbb{N}\rangle$ vara en intern ${}^*\mathbb{R}$-värd följd så att $s_n \simeq 0$ för varje $n \in \mathbb{N}$. Då finns ett obegränsat naturligt tal ω så att $s_n \simeq 0$ för alla naturliga tal $n \leq \omega$.*

Bevis. Använd 10.10(i) med $\Phi(n)$ som den interna formeln $[|ns_n| \leq 1]$ för att få ett $\omega \in {}^*\mathbb{N}_\infty$ så att $|s_n| \leq \frac{1}{n}$ om $n \leq \omega$. Då är $s_n \simeq 0$ om $n \in {}^*\mathbb{N}_\infty$ och $n \leq \omega$ (eftersom $\frac{1}{n}$ är infinitesimalt för sådana n), och alltså är $s_n \simeq 0$ för alla $n \leq \omega$. □

10.3 Uppräknelig mättnad

Den kanske mest användbara egenskapen hos interna mängder är att snittet av en minskande följd av interna icke-tomma mängder också är icke-tom. Detta ger upphov till fenomenet *uppräknelig mättnad* vilket vi förklarar nedan.

Sats 10.13 (Uppräknelig mättnad)**.** *Snittet av en minskande följd*

$$X_1 \supseteq X_2 \supseteq \cdots \supseteq X_k \supseteq \cdots$$

av icketomma interna mängder är alltid icketom, d.v.s.:

$$\bigcap_{k\in\mathbb{N}} X_k \neq \emptyset. \tag{8}$$

Bevis. I beviset kommer vi betrakta interna mängder som följder av reella mängder, vi argumenterar alltså delvis i den begränsade ultraprodukten $\prod_{\mathfrak{U}}^0 V(\mathbb{R})$, riktigheten i detta garanteras av LOs sats.

För varje fixt $k \in \mathbb{N}$, låt $X^k = [A_n^k]$[11] så att $\langle A_n^k : n \in \mathbb{N}\rangle$ är en representant i den begränsade ultraprodukten för den interna mängden X^k av delmängder till $^*\mathbb{R}$ Enligt LOs sats har vi då alltså att $\left\{n \in \mathbb{N} : A_n^k \neq \emptyset\right\} \in \mathfrak{U}$ och $\left\{n \in \mathbb{N} : A_n^k \supseteq A_n^{k+1}\right\} \in \mathfrak{U}$ på grund av motsvarande egenskaper hos X_k .

Eftersom $\mathfrak{U}$ är sluten under ändliga snitt så har vi att för varje $k \in \mathbb{N}$ att:

$$J^k = \left\{n \in \mathbb{N} : A_n^1 \supseteq \cdots \supseteq A_n^k \neq \emptyset\right\} \in \mathfrak{U}.$$

Notera att J_i är ”diagonaliseringsmängder” som uppfyller $J^1 \supseteq J^2 \supseteq \cdots$.

Vi vill konstruera ett hyperreellt tal $[s_n]$ så att $[s_n] \in X^k$ för alla k, d.v.s.

$$\left\{n \in \mathbb{N} : s_n \in A_n^k\right\} \in \mathfrak{U} \quad \forall k \in \mathbb{N}.$$

Eftersom $\{n \in \mathbb{N} : k \leq n\} \in \mathfrak{U}$ för varje fixt $k \in \mathbb{N}$ räcker det att visa att s_n uppfyller

$$\{n \in \mathbb{N} : k \leq n\} \cap J^k \subseteq \left\{n \in \mathbb{N} : s_n \in A_n^k\right\}. \tag{9}$$

För $n \in J^1$ låt

$$k_n = \max\left\{k : k \leq n \wedge n \in J^k\right\}. \tag{10}$$

[11] Detta k är samma k som indexet i (8), och alltså inte en exponent. Skrivsättet används p.g.a att vi är tvungna att använda dubbel indexering.

Vi har då att $n \in J^{k_n}$, och för detta n väljer vi ett $s_n \in A_n^{k_n}$ vilket — enligt hur vi definierat J^{k_n} — ger

$$s_n \in A_n^1 \cap \cdots \cap A_n^{k_n}. \tag{11}$$

För att visa (9) kan vi tillåta s_n att vara godtycklig för $n \notin J^1$. Vi har nu att om $k \leq n$ och $n \in J^k$, så är $k \leq k_n$ enligt (10) och vi har $s_n \in A_n^k$ enligt (11). $\square$

Denna egenskap kallas för uppräknelig mättnad eftersom det faktum att snittet är icke-tomt implicerar att $V(^*X)$ är ”mättat” på mängder. Faktum är att detta är ekvivalent med att vi befinner oss i en utvidgning och som kuriosa nämns att man även kan skapa strukturer som är mer mättade än bara uppräkneligt. Från uppräknelig mättnad följer en mängd olika resultat av vilka vi ska visa två.

Korollarium 10.14. *Om $\{X_n : n \in \mathbb{N}\}$ är en familj av interna mängder och om $X := \bigcup_{n\in\mathbb{N}} X_n$ är intern, då finns det ett $k \in \mathbb{N}$ så att*

$$X = \bigcup_{n\leq k} X_n$$

Bevis. Den högra inklusionen är uppenbar. För den vänstra har vi att om $X \not\subseteq \bigcup_{n\leq k} X_n$ för alla k, och alltså

$$\bigcap_{n\leq k} (X \setminus X_n) = X \setminus (\bigcup_{n\leq k}) \neq \emptyset,$$

så är

$$Y^1 \supseteq Y^2 \supseteq \cdots \supseteq Y^k = \left\{ \bigcap_{n\leq k} (X \setminus X_n) : n \in \mathbb{N} \right\}$$

en minskande följd av icketomma interna mängder. Enligt sats (10.13) har vi då att

$$\bigcap_{k\in\mathbb{N}} Y^k \neq \emptyset.$$

Men vi har även att

$$\bigcap_{k\in\mathbb{N}} Y^k = X \setminus \bigcup_{k\in\mathbb{N}} \bigcup_{n\leq k} X_n = X \setminus X,$$

vilekt är en motsägelse. $\square$

Korollarium 10.15. *Varje följd av interna mängder $< A_n : n \in \mathbb{N} >$ kan utvidgas till en intern följd $< A'_n : n \in {^*\mathbb{N}} >$ sådan att $A'_n = A_n$ för alla $n \in \mathbb{N}$.*

Bevis. Låt $B_m = \{$alla interna följder $< C_n : n \in {}^*\mathbb{N} >$ sådana att $C_n = A_n$ för $n \leq m\}$. Då är B_m en avtagande följd av icke-tomma mängder som dessutom är interna på grund av interna definitionsprincipen. Således ger 10.13 att det finns en följd $A'_n \in \bigcap_{m \in \mathbb{N}} B_m$ med $A'_n = A_n$ för alla $n \in \mathbb{N}$. □

11 Loebmåttet

I detta kapitel har vi utgått från [Cutland] och [Cutland et. al.].

Vi ska nu använda de idéer vi utvecklat till ett lite mer konkret exempel - vi ska konstruera brownsk rörelse som skuggan av en hyperändlig slumpvandring. Men för att göra detta behöver vi en hel del måtteori. Detta är precis vad det låter som - en matematisk teori som studerar vilka operationer som är lämpliga för att storleksbestämma mängder. För ändliga mängder duger det att bara räkna elementen men för uppräkneliga eller överuppräkneliga mängder fungerar inte detta. Istället introduceras så kallade måttrum bestämda på så kallade algebror, vilka är slutna under unioner av uppräkneligt många av dess element.

En samling av interna mängder däremot - något som vi ofta vill arbeta med på grund av dess trevliga egenskaper - är som regel inte sluten under uppräkneliga unioner. Peter Loeb upptäckte 1973 att detta faktum kunde ge upphov till ett nytt sätt att konstruera standard måttrum utifrån icke-standard storheter, något vi kommer utforska i detta kapitel.

11.1 En intern algebra

För att förklara vad Loeb-måttet är för något behöver vi först och främst lite elementär måtteori. Vi börjar med begreppet *algebra*. Poängen med detta begrepp är att ur en samling av delmängder utesluta de som ger upphov till ”konstiga” egenskaper. För att göra detta visar det sig att vi behöver bra slutenhetsegenskaper, vilka vi nu ska formulera.

Definition 11.1. *En* ***algebra*** *är en samling delmängder $\mathcal{A}$ till en mängd Ω sådan att*

- $\Omega \in \mathcal{A}$
- $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$
- $A \in \mathcal{A} \Rightarrow A^C \in \mathcal{A}$

om vi dessutom har att

- $(\forall n \in \mathbb{N})(A_n \in \mathcal{A}) \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$

så kallas $\mathcal{A}$ för en σ**-algebra**.

För den senare egenskapen säger vi att $\mathcal{A}$ är *sluten under uppräkneliga unioner*. Några uppenbara följder av definitionen är:

- $\emptyset \in \mathcal{A}$
- $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$
- $A, B \in \mathcal{A} \Rightarrow A \backslash B \in \mathcal{A}$
- $A, B \in \mathcal{A} \Rightarrow A \triangle B$[12]$\in \mathcal{A}$
- $(\forall n \in \mathbb{N})(A_n \in \mathcal{A}) \Rightarrow \bigcap\limits_{n \in \mathbb{N}} A_n \in \mathcal{A}$

Om nu $\mathcal{A}$ ligger i vårt kända universum $V(\mathbb{R})$ så säger transferprincipen (se tidigare kapitel) att ${}^*\mathcal{A}$ också är en algebra men i $V({}^*\mathbb{R})$. Däremot kan vi inte tillämpa transferprincipen på defintionen av en *σ-algebra* eftersom den innehåller en referens till $\mathbb{N}$ - vi skulle då få att algebran blir sluten under ”hyperuppräkneliga” unioner istället för uppräkneliga. Faktum är att ${}^*\mathcal{A}$ sällan en σ-algebra även om $\mathcal{A}$ är det. För att se detta, låt $\langle A_n : n \in \mathbb{N}\rangle$ vara en följd av mängder i ${}^*\mathcal{A}$ med union A. Enligt definition så är varje A_n intern och om $A \in {}^*\mathcal{A}$ måste A också vara intern. Enligt korollarium 10.14 har vi då att $A = \bigcup\limits_{n \leq k} A_n$ för något $k \in \mathbb{N}$, så om A är en genuint uppräknelig union av $\langle A_n : n \in \mathbb{N}\rangle$ (det vill säga, A kan inte skrivas som en union av ändligt många mängder från följden) kan den inte ligga i $\mathcal{A}$. För varje intern algebra (eftersom alla element i en intern mängd är interna) får vi alltså följande sats:

Sats 11.2. *För varje intern algebra av mängder $\mathcal{A}$ där $A_n \in \mathcal{A}$ för alla $n \in \mathbb{N}$ så gäller att*

$$\bigcup A_n \in \mathcal{A} \iff \bigcup A_n = \bigcup_{n \leq k} A_n$$

för något $k \in \mathbb{N}$.

11.2 Mått

När vi väl har en algebra så vill vi gärna uppskatta storleken på de mängder som ingår däri. Beroende på vilken algebra man har och vilka egenskaper man vill undersöka kan man behöva olika sorters funktioner som ”mäter” storleken på mängderna. Dessa behöver dock alltid ha vissa egenskaper för att man överhuvudtaget ska kunna prata om att de ”mäter” storlekar. Nedan ska vi införa begreppet *mått* explicit och redogöra för dessa egenskaper. Eftersom vår tillämpning främst kommer använda sig av så kallade *sannolikhetsmått*, då vi ska modellera en slumpmässig

[12]Den symmetriska differensen mellan A och B är helt enkelt deras union minus deras snitt. Kort sagt; $A \triangle B \doteqdot (A \backslash B) \cup (B \backslash A)$.

process, är det just dessa vi definierar. Dessa har egenskapen att deras målmängd är det reella intervallet $[0,1]$. För mått generellt kan dock hela $\mathbb{R}$ vara målmängd.

Definition 11.3. *Låt $\mathcal{A}$ vara en algebra av delmängder till mängden Ω. En funktion $\mu : \mathcal{A} \to [0,1]$ kallas för ett* ***ändligt additivt sannolikhetsmått på*** *Ω om*

- $\mu(\emptyset) = 0$
- $\mu(\Omega) = 1$
- $\mu(A \cup B) = \mu(A) + \mu(B)$ *för disjunkta $A, B \in \mathcal{A}$*

Om vi dessutom har att

- *För varje följd $\langle A_n : n \in \mathbb{N} \rangle$ av element i $\mathcal{A}$ som är parvis disjunkta och vars union $A \in \mathcal{A}$ så gäller att $\mu(A) = \sum_{n \in \mathbb{N}} \mu(A_n)$*

så sägs μ istället vara ***uppräkneligt additivt****.*

Notera att μ:s målmängd är det *reella* intervallet $[0,1]$ även om algebran som den är definierad på skulle vara icke-standard. Om $\mathcal{A}$ är en σ-algebra på mängden Ω och μ ett uppräkneligt additivt sannolikhetsmått så säger vi att $(\Omega, \mathcal{A}, \mu)$ är ett *sannolikhetsrum*[13].

Senare ska vi se att Ω kommer uppfattas som en mängd av så kallade ”utfall” – vi hänvisar hädanefter till Ω som just ett *utfallsrum*. Vi tänker oss att vi genomför ett *sannolikhetsexperiment* när vi tar ett $w \in \Omega$ och måttet μ mäter sannolikheten för detta w. Alla element i $\mathcal{A}$ kallas nu för *händelser* - de är mängder som samlar ihop ett antal element i Ω och $\mu(A)$ mäter sannolikheten för att just dessa ska hända. För att konkretisera; tänk dig $\Omega \doteqdot \{1,2,3,4,5,6\}$ som utfallsrummet för ett tärningsslag. Vi kan kalla $A \doteqdot \{2,4,6\}$ för händelsen ”tärningen visar ett jämnt antal prickar” och rent intuitivt borde vi ha att $\mu(A) = \frac{1}{2}$.

Härnäst visas några användbara resultat för ändligt additiva sannolikhetsmått.

Sats 11.4. *Låt $\mathcal{A}$ vara en algebra av delmängder till mängden Ω och låt μ vara ett ändligt additivt sannolikhetsmått på $\mathcal{A}$. Då gäller att:*

(i) $\mu(A_1 \cup ... \cup A_n) = \mu(A_1) + ... + \mu(A_n)$ *för alla parvis disjunkta $A_i \in \mathcal{A}$*

[13] Vi betonar att detta inte är en operation utan en notation, ett enkelt sätt att prata om en mängd, en algebra och ett mått samtidigt.

(ii) om $A, B \in \mathcal{A}$ så $\mu(A) = \mu(A \backslash B) + \mu(A \cap B)$

(iii) om $A, B \in \mathcal{A}$ och $A \subseteq B$ så $\mu(A) \leq \mu(B)$

(iv) $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$

Bevis. (i) Antag att $A_1, ..., A_n$ är parvis disjunkta element i $\mathcal{A}$. Eftersom μ är ändligt additivt gäller att $\mu(A_1 \cup ... \cup A_n) = \mu(A_1) + \mu(A_2 \cup ... \cup A_n)$ och resten följer av induktion.

(ii) Eftersom $A = (A \backslash B) \cup (A \cap B)$ och $(A \backslash B) \cap (A \cap B) = \emptyset$ så ger den ändliga additiviteten att $\mu(A) = \mu(A \backslash B) + \mu(A \cap B)$.

(iii) Vi åberopar (ii) för att se att

$$\mu(B) = \mu(B \backslash A) + \mu(A \cap B) = \mu(B \backslash A) + \mu(A) \Rightarrow \mu(B) - \mu(A) \geq 0$$

(iv) Eftersom $A \cup B = (A \backslash B) \cup (B)$ och $(A \backslash B) \cap (B) = \emptyset$ så ger den ändliga additiviteten att $\mu(A \cup B) = \mu(A \backslash B) + \mu(B)$ vilket ger $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.

$\square$

Efter denna genomgång av elementär mätteori ska vi nu återgå till vårt icke-standard universum.

11.3 Loebmåttet

Låt μ vara en intern funktion definierad såsom i definition 11.3 men med målmängden ${}^*[0,1]$ istället för $[0,1]$. Låt $\mathcal{A}$ vara en intern algebra av delmängder till den interna mängden Ω som är μ:s definitionsmängd. Vi skriver ${}^s\mu(x)$ för $sh(\mu(x))$. sh-funktionen bevarar nu ändlig additivitet (och därmed också uppräknelig additivitet eftersom vi är på en intern algebra) så vi kallar strukturen $(\Omega, \mathcal{A}, {}^s\mu(x))$ för ett *standard ändligt additivt sannolikhetsmåttrum* även om det tekniskt sett inte är ett måttrum då en intern algebra inte är sluten under uppräkneliga unioner.

Däremot ska vi nu presentera en metod som utvidgar $(\Omega, \mathcal{A}, {}^s\mu(x))$ till ett ”riktigt” måttrum. Det kommer nämligen visa sig att ett internt ”måttrum” har många egenskaper som gör det lätt att förstå rent intuitivt. De kan under rätt omständigheter bete sig på nästan samma sätt som ändliga strukturer, vilket gör de väldigt lätta att hantera. Men vi behöver ett sätt att göra om dem till ”riktiga” måttrum för att de kunskaper om dessa som finns på plats i standardmatematiken ska gälla även våra konstruktioner.

Definition 11.5. *Låt $B \subseteq \Omega$ (vi har inte nödvändigtvis att $B \in \mathcal{A}$). Vi säger att B är en* ***Loeb-nollmängd*** *om det för varje reellt $\epsilon > 0$ existerar en mängd $A \in \mathcal{A}$ sådant att $B \subseteq A$ och $\mu(A) < \epsilon$.*

Uppenbarligen så är varje delmängd till en Loeb-nollmängd också en Loeb-nollmängd. De som är bekanta med måtteori ser att detta kommer leda till ett så kallat *fullständigt* mått. Unionen av två Loeb-nollmängder ses också lätt vara en Loeb-nollmängd och vi ska se nedan att detta även gäller uppräkneliga unioner.

Lemma 11.6. *Låt $\langle A_n : n \in \mathbb{N}\rangle$ vara en följd växande*[14] *mängder i $\mathcal{A}$ och låt $B = \bigcup_{n\in\mathbb{N}} A_n$. Då finns det en mängd $A \in \mathcal{A}$ sådan att*

(i) $B \subseteq A$

(ii) ${}^s\mu(A) = \lim_{n\to\infty} {}^s\mu(A_n)$

(iii) $A \backslash B$ *är en Loeb-nollmängd*

Om $\langle A_n : n \in \mathbb{N}\rangle$ istället är en parvis disjunkt samling mängder i $\mathcal{A}$ existerar A och B på samma sätt men istället för (ii) gäller ${}^s\mu(A) = \sum_{n\in\mathbb{N}} {}^s\mu(A_n)$.

Bevis. Låt $\alpha = \lim_{n\to\infty} {}^s\mu(A_n)$. Detta gränsvärde existerar eftersom följden är uppåt begränsad och växande. För varje $n \in \mathbb{N}$ har vi att

$$\mu(A_n) \leq {}^s\mu(A_n) + \frac{1}{n} \leq \alpha + \frac{1}{n}$$

Enligt 10.15 kan vi ta en utvidgad växande intern följd av mängder i $\mathcal{A}$, $\langle A_n : n \in {}^*\mathbb{N}\rangle$. 10.11 ger oss existensen av ett obegränsat $N \in {}^*\mathbb{N}$ sådant att

$$\mu(A_N) \leq \alpha + \frac{1}{N}$$

Låt nu $A = A_N$. Transferprincipen ger oss då att $A_n \in A$ för alla begränsade n varför (i) håller. Vidare har vi att $\mu(A_n) \leq \mu(A)$ för varje ändligt n vilket medför att ${}^s\mu(A_n) \leq {}^s\mu(A) \leq \alpha$ varför ${}^s\mu(A) = \alpha$ vilket visar (ii). Slutligen har vi att

$${}^s\mu(A \backslash A_n) = {}^s\mu(A_n) - {}^s\mu(A) \to 0.$$

Detta tillsammans med det faktum att $(A\backslash B) \subseteq (A\backslash A_n)$ gör att $A\backslash B$ blir en Loeb-nollmängd.

[14] En växande följd mängder är en följd $\langle A_n : n \in \mathbb{N}\rangle$ sådan att $A_1 \subseteq A_2 \subseteq A_3 \subseteq ...$

För att visa det sista påståendet, bilda följden $\langle C_n : n \in \mathbb{N}\rangle$ genom att sätta $C_1 = A_1$ och $C_n = \bigcup_{k=1}^{n} A_k$. Då är C_n en växande följd i $\mathcal{A}$ och vi kan tillämpa (i), (ii) och (iii) på denna så att vi med $\bigcup\limits_{n\in\mathbb{N}} A_n = \bigcup\limits_{n\in\mathbb{N}} C_n \subseteq A$ får:

$$^{s}\mu(A) = \lim_{n\to\infty} {}^{s}\mu(C_n) = \lim_{n\to\infty} {}^{s}\mu(\bigcup_{k=1}^{n} A_k) = \lim_{n\to\infty} \sum_{k=1}^{n} {}^{s}\mu(A_k) = \sum_{n\in\mathbb{N}} {}^{s}\mu(A_n).$$

□

Definition 11.7.

(i) Låt $B \subseteq \Omega$. B sägs nu vara ***Loebmätbar*** *om det finns en mängd $A \in \mathcal{A}$ sådan att $A \triangle B$*[15] *är en Loeb-nollmängd.*

(ii) Samlingen av alla Loebmätbara mängder betecknas $L(\mathcal{A})$.

(iii) För $B \in L(\mathcal{A})$ definierar vi ***Loebmåttet*** *$\mu_L(B) \doteqdot {}^{s}\mu(A)$ för vilket som helst $A \in \mathcal{A}$ där $A \triangle B$ är en Loeb-nollmängd.*

Innan vi ska visa vårt slutgiltiga resultat för Loeb-mått, notera följande.

- Relationen "$A \triangle B$ är en Loeb-nollmängd"är transitiv eftersom för alla mängder A, B, C gäller att $A \triangle C \subseteq (A \triangle B) \cup (B \triangle C)$. Detta gör bland annat att vårt mått blir väldefinierat.
- Om $B \in L(\mathcal{A})$ och $B \triangle A$ är en Loeb-nollmängd så är $A \in L(\mathcal{A})$ eftersom om $C \triangle B$ är en Loeb-nollmängd så är även $C \triangle A$ det enligt noten ovan, för $C \in \mathcal{A}$.
- Om B är en Loeb-nollmängd så är $\mu_L(B) = 0$ eftersom vi i definitionen av Loebmätbarhet ovan kan ta $A = \emptyset$ så att B blir Loebmätbar och $\mu_L(B) = {}^{s}\mu(A) = {}^{s}\mu(\emptyset) = 0$. Ett liknande argument ger att $\mu_L(B) = 0$ och $B \in L(\mathcal{A}) \Rightarrow B$ är en Loeb-nollmängd.

Slutligen får vi våra två huvudresultat i detta kapitel:

Sats 11.8. *$L(\mathcal{A})$ är en σ-algebra på Ω.*

Bevis. Tag $A \in \mathcal{A}$. Då är $A \triangle A = \emptyset \subseteq \emptyset$ och $\mu(\emptyset) = 0$. Detta visar att $\mathcal{A} \subseteq L(\mathcal{A}$ och speciellt att $\Omega \in L(\mathcal{A})$.

[15] $A \triangle B$ kan ses som ett mått på skillnaden mellan mängderna A och B. En mängd är alltså Loebmätbar om den är så lik en mängd i $\mathcal{A}$ att dessa överlappar nästan helt.

Tag ett $B \in L(\mathcal{A})$. Då finns ett $A \in \mathcal{A}$ sådant att $B \triangle A$ är en Loeb-nollmängd. Dessutom $A^c \in \mathcal{A}$ och $B^c \triangle A^c = B \triangle A$ så $B^c \in L(\mathcal{A})$.

Låt $\langle B_n : n \in \mathbb{N} \rangle$ vara en följd av mängder där varje $B_n \in L(\mathcal{A})$. Vi ska visa att $\bigcup_{n \in \mathbb{N}} B_n \in L(\mathcal{A})$. Från definitonen av $L(\mathcal{A})$ ser vi att det existerar en följd $\langle A_n : n \in \mathbb{N} \rangle$ med $A_n \in \mathcal{A}$ för varje $n \in \mathbb{N}$ sådan att $B_n \triangle A_n$ är en Loeb-nollmängd. Med hjälp av 11.6 ser vi lätt att $\bigcup_{n \in \mathbb{N}} A_n \in L(\mathcal{A})$ och enligt noterna ovan räcker det då att visa att $(\bigcup_{n \in \mathbb{N}} B_n) \triangle (\bigcup_{n \in \mathbb{N}} A_n)$ är en Loeb-nollmängd.

Tag nu ett reellt $\epsilon > 0$. För varje $B_n \triangle A_n$ kan vi enligt definitionen av en Loeb-nollmängd välja ett C_n sådant att $B_n \triangle A_n \subseteq C_n$ och $\mu(C_n) < \frac{\epsilon}{3^n}$ vilket medför att ${}^s\mu(C_n) \leq \frac{\epsilon}{3^n}$. Det är lätt att se att

$$(\bigcup_{n \in \mathbb{N}} B_n) \triangle (\bigcup_{n \in \mathbb{N}} A_n) \subseteq \bigcup_{n \in \mathbb{N}} (B_n \triangle A_n) \subseteq \bigcup_{n \in \mathbb{N}} C_n.$$

Låt oss nu definiera en växande följd på följande sätt; sätt $D_1 = C_1$ och $D_n = \bigcup_{k \leq n} C_k$. Notera att $\bigcup_{n \in \mathbb{N}} C_n = \bigcup_{n \in \mathbb{N}} D_n$. Då $D_n \in \mathcal{A}$ så kan vi använda lemma 11.6 för att se att det finns en mängd $D \in \mathcal{A}$ sådan att $\bigcup_{n \in \mathbb{N}} D_n \subseteq D$ och

$${}^s\mu(D) = \lim_{n \to \infty} {}^s\mu(D_n) = \lim_{n \to \infty} {}^s\mu(\bigcup_{k \leq n} C_k) \leq \lim_{n \to \infty} \sum_{k \leq n} \frac{\epsilon}{3^k}$$

vilket medför att $\mu(D) < \epsilon$. Således är $(\bigcup_{n \in \mathbb{N}} B_n) \triangle (\bigcup_{n \in \mathbb{N}} A_n)$ en Loeb-nollmängd och vi har visat att $L(\mathcal{A})$ är en σ-algebra.

Sats 11.9. *μ_L är ett uppräkneligt additivt mått på $L(\mathcal{A})$.*

Att $\mu_L(\emptyset) = 0$ följer direkt av att $\emptyset$ är en Loeb-nollmängd. Ändlig additivitet är också lätt att visa ty om $A, B \in L(\mathcal{A})$ och $C, D \in \mathcal{A}$ är sådana att $A \triangle C$ och $B \triangle D$ är Loeb-nollmängder så är också $(A \cup B) \triangle (C \cup B)$ en Loeb-nollmängd eftersom

$$(A \cup B) \triangle (C \cup D) \subseteq (A \triangle C) \cup (B \triangle D).$$

Således är

$$\mu_L(A \cup B) = {}^s\mu(C \cup D) = {}^s\mu(C) + {}^s\mu(D) = \mu_L(A) + \mu_L(B).$$

Låt oss nu visa den uppräkneliga additiviteten. Tag en parvis disjunkt följd $\langle B_n : n \in \mathbb{N} \rangle$ med $B_n \in L(\mathcal{A})$ och låt $\langle A_n : n \in \mathbb{N} \rangle$ vara en följd som i beviset ovan. Låt oss dessutom anta att följden A_n också är parvis disjunkt. Då får vi att $\mu_L(\bigcup_{n\in\mathbb{N}} B_n) = \mu_L(\bigcup_{n\in\mathbb{N}} A_n)$ enligt föregående bevis och enligt lemma 11.6 existerar $A \in \mathcal{A}$ så att

$$\mu_L(\bigcup_{n\in\mathbb{N}} A_n) = 0 + \mu_L(\bigcup_{n\in\mathbb{N}} A_n) = \mu_L(A \backslash \bigcup_{n\in\mathbb{N}} A_n) + \mu_L(\bigcup_{n\in\mathbb{N}} A_n) = \mu_L(A)$$

enligt den ändliga additiviteten och

$$\mu_L(A) = {}^s\mu(A) = \sum_{n\in\mathbb{N}} {}^s\mu(A_n) = \sum_{n\in\mathbb{N}} {}^s\mu(B_n).$$

Om nu $\langle A_n : n \in \mathbb{N}\rangle$ inte är parvis disjunkt så bildar vi en ny disjunkt följd genom att låta $A_1' = A_1$ och $A_n' = A_n \backslash \bigcup_{k\langle n} A_k$. Notera därefter att $B_n \triangle A_n' \subseteq \bigcup_{k\leq n} B_k \triangle A_k$ och att $\bigcup_{n\in\mathbb{N}} A_n' = \bigcup_{n\in\mathbb{N}} A_n$ och upprepa argumentet ovan.

□

För detta mått gäller nu ytterligare en egenskap. Denna ger en än god intuitiv bild av μ_L och kommer att komma till användning senare.

Sats 11.10. *För varje $B \in L(\mathcal{A})$ och varje positivt $\epsilon \in \mathbb{R}$ finns det mängder $B^-, B^+ \in \mathcal{A}$ sådana att*

(i) $B^- \subseteq B \subseteq B^+$ *och*

(ii) $\mu_L(B^+) - \epsilon \leq \mu_L(B) \leq \mu_L(B^-) + \epsilon$

Bevis. Tag ett reellt $\epsilon > 0$. Enligt definitionerna av Loebmätbarhet och Loebnollmängd så tar vi ett $A \in \mathcal{A}$ sådant att $A \triangle B \subseteq C \in \mathcal{A}$ med $\mu(C) < \epsilon$. Då gäller att

$$\mu(A) \leq \mu(A \backslash C \cup C) = \mu(A \backslash C) + \mu(C) \leq \mu(A \backslash C) + \epsilon$$

och

$$\mu(A \cup C) \leq \mu(A) + \mu(C) \leq \mu(A) + \epsilon$$

Eftersom vi också har att

$$A \backslash C \subseteq A \backslash (A \triangle B) \subseteq B$$

och

$$B \subseteq A \cup (A \triangle B) \subseteq A \cup C$$

så följer satsen om vi sätter $B^- = A \backslash C$ och $B^+ = A \cup C$.

$\square$

11.4 Ytterligare måtteori

För att kunna jämföra vår konstruktion av Loebmåttet med etablerade mått på standardsidan ger vi här en mycket kort introduktion till Lebesguemåttet och Borelalgebran. I vår definition av mått ovan antog vi att dessas målmängd alltid var intervallet $[0,1]$. Här ska vi dock göra en mycket kort exkursion och hantera mått som kan anta alla reella värden. Egenskaperna för dessa är väsentligen likadana, för en mer detaljerad diskussion se [Folland].

Definition 11.11. *Den minsta σ-algebran på $\mathbb{R}$ som innehåller alla öppna intervall kallas för* ***Borelalgebran*** *och betecknas $\mathcal{B}$. Dess element kallas för* ***Borelmängder***.

Definition 11.12. *Det uppräkneligt additiva mått λ som är definierat på alla Borelmängder och för vilket gäller att $\lambda((a,b)) = b - a$ för alla öppna intervall (a,b) kallas för* ***Lebesguemåttet***.

Detta mått är unikt och skapar en ny algebra på följande sätt.

Definition 11.13. *En mängd $B \subseteq \mathbb{R}$ sägs vara* ***Lebesguemätbar*** *om det för varje positivt reellt ϵ existerar en sluten mängd $C \subseteq B$ och en öppen mängd $D \supseteq B$ sådana att $\lambda(D \backslash C) < \epsilon$. Samlingen av dessa mängder är också den en σ-algebra kallad* ***Lebesguealgebran*** *och betecknas $\mathcal{L}$.*

Inte bara mängder utan också funktioner kan sägas vara mätbara:

Definition 11.14. *En funktion $f : A \to \mathbb{R}$ med sägs vara* ***Lebesguemätbar*** *om det för varje Borelmängd B gäller att $f^{-1}(B) \in \mathcal{L}$ där A är en Lebesguemätbar mängd med $A \subseteq \mathbb{R}$. En funktion $f : \Omega \to \mathbb{R}$ sägs vara* ***Loebmätbar*** *om det för varje Borelmängd B gäller att $f^{-1}(B) \in L(\mathcal{A})$ där Ω är intern med en intern algebra $\mathcal{A}$.*

Not 11.1. *Läsaren kan säkert komma att undra varför vi inte talar om $\mathcal{A}$-mätbara funktioner - vi ska nämligen behandla funktioner från $(\Omega, \mathcal{A}, \mu)$ till ${}^*\mathbb{R}$ där $\mathcal{A}$ är en intern algebra (och således inte en σ-algebra). Anledningen till detta är att så länge funktionen är intern är dess inversa bild av en intern mängd intern enligt 10.6. Givet en σ-algebra till ${}^*\mathbb{R}$ som är "tillräckligt rik" på interna mängder så är alltså alla interna fuktioner på en algebra $\mathcal{A}$ som består av* alla *interna delmängder till en intern mängd Ω $\mathcal{A}$-mätbara. Vi visar inte detta här utan hänvisar till exempelvis* [Albeverio et. al.].

Vi säger också att trippeln $(\mathbb{R}, \mathcal{L}, \lambda)$ är ett *Lebesguemåttrum.* Dessa begrepp ger upphov till den så kallade Lebesgueintegralen, vilken infördes efter att man upptäckt flera funktioner som inte lät sig integreras på normalt "Riemann-vis" men som "borde" vara lika med noll. Lebesgueintegralen klarar av fler sorters funktioner och den kan även generaliseras till alla måttrum. Vi definierar den på följande vis:

Definition 11.15. *Låt f vara en Loeb- eller Lebesguemätbar funktion. Då definieras* ***Lebesgueintegralen*** *med avseende på måttet m som*

$$\int f dm = \lim_{n \to \infty} \sum_{k=-n^2}^{n^2} \frac{k}{n} m(f^{-1}(I)).$$

*där I är intervallet $(\frac{k-1}{n}, \frac{k}{n})$. Vi kallar f för (loeb- eller lebesgue-)**integrerbar** om gränsvärdet för $\int |f| dm$ existerar.*

Lebesgueintegralen delar många egenskaper med Riemannintegralen. Bland annat är den en linjär funktion och överensstämmer med Riemannintegralen på alla funktioner som är Riemannintegrerbara. Vi avslutar med att infoga en mycket användbar sats som gäller godtyckliga måttrum (där integralen är Lebesgueintegralen):

Sats 11.16 (Chebyshevs olikhet)**.** *Låt f vara en mätbar reellvärd funktion på måttrummet $(\Omega, \mathcal{A}, m)$ och g en reellvärd mätbar växande och icke-negativ funktion på målmängden till f. Då gäller att*[16]

$$m(\{w \in \Omega : f(w) \geq t\}) \leq \frac{1}{g(t)} \int_\Omega g(f) dm$$

[16]Detta är en generalisering av den ”vanliga” olikheten, $P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$ där X är en slumpvariabel med standardavvikelsen σ och väntevärdet μ

12 Hyperändliga sannolikhetsrum

I följande kapitel kommer vi framförallt att utgå från [Albeverio et. al.], [Anderson] och [Loeb]. Flera satser och observationer kommer att lämnas utan bevis, men för den intresserade är de likartat formulerade inklusive bevis i litteraturen.

Vi kommer i detta kapitel att introducera så kallade stokastiska processer. Detta begrepp är något som de flesta som läst grundläggande sannolikhetsteori är bekanta med – det används oftast när man på något sätt vill modellera en slumpmässig händelse som rör sig över ett tidsintervall.

Fördelen med icke-standardanalys, som vi ska utforska, är att ”rummet” som de slumpmässiga händelserna sker i kan byggas upp på ett sätt så att det i mångt och mycket beter sig som ändliga mängder – vi använder de hyperändliga mängder vi definierat i avsnitt 9.2. Samtidigt har vi visat hur ett sådant rum kan omvandlas till ett Loebmåttrum med allt vad vi vill ha av måtteoretiska egenskaper.

I nästa avsnitt skall vi definiera precis vad vi menar med ett hyperändligt sannolikhetsrum för att därefter gå vidare med stokastiska processer. Vårt arbete kommer till slut leda fram till ett konkret exempel i nästa kapitel.

12.1 Hyperändliga sannolikhetsrum och tidslinjer

En viktig struktur får vi genom att kombinera våra idéer om måttrum och hyperändlighet. Givet en hyperändlig mängd Ω och en tillhörande intern algebra $\mathcal{A}$ så definierar vi det hyperändliga räknemåttet P genom att för varje intern mängd $A \subseteq \Omega$ sätta $P(A) = \frac{|A|}{|\Omega|} = \sum\limits_{w \in A} \frac{1}{|\Omega|}$ där $|\cdot|$ är den interna kardinaliteten hos mängder. Existens och väldefinierbarhet följer med transfer från det ändliga till det hyperändliga fallet. Man kan likna denna konstruktion med att räkna antalet element i A och jämföra detta med antalet element i Ω - kanske det mest intuitiva sättet att hantera storleksbegreppet.

Definition 12.1. *Låt Ω vara en hyperändlig (och således intern per definition) mängd med tillhörande intern algebra $\mathcal{A}$ och låt P vara det hyperändliga räknemåttet på $\mathcal{A}$ med $P(\Omega) = 1$. Då säger vi att trippeln $(\Omega, \mathcal{A}, P)$ är ett **hyperändligt sannolikhetsrum**.*

Ett viktigt exempel på ett hyperändligt sannolikhetsrum får vi genom följande konstruktion. Välj ett $N \in {}^*\mathbb{N}_\infty$, sätt $\triangle t \doteqdot \frac{1}{N}$ och låt

$$T \doteqdot \{0, \triangle t, 2\triangle t, ..., N\triangle t = 1\}.$$

Detta blir så att säga en "hyperändlig motsvarighet" till det reella intervallet $[0,1]$ som innehåller alla rationella tal givet att $N = \eta!$ för något $\eta \in {}^*\mathbb{N}_\infty$ (detta visas lätt genom ett transferargument). Faktum är att $sh : T \to [0,1]$ är surjektiv - inget irrationellt tal ingår i T men för varje irrationellt tal $r \in [0,1]$ finns ett unikt $t \in T$ sådant att $t < r < t + \triangle t$ (men detta visas inte här, se [Albeverio]). Denna konstruktion kallar vi för *den hyperändliga tidslinjen*.

Till hyperändliga sannolikhetsrum $(\Omega, \mathcal{A}, P)$ kan vi som i avsnitt 11.3 associera ett Loebmåttrum. Vi gör om det hyperändliga sannolikhetsrummet til ett standard ändligt additivt måttrum och sedan till ett Loebmåttrum genom att ta skuggan av P och sedan upprepa processen i 11.3:

$$(\Omega, \mathcal{A}, P) \to (\Omega, \mathcal{A}, {}^sP) \to (\Omega, L(\mathcal{A}), P_L).$$

Naturligtvis kan detta också göras med T istället för Ω så att $(T, \mathcal{A}, P)$ associeras med $(T, L(\mathcal{A}), P_L)$.

12.2 Stokastiska processer och lyftningar

Med dessa begrepp på plats är det dags att introducera två sannolikhetstermer.

Definition 12.2. *Vi börjar med två så kallade stokastiska processer.*

(i) För något standard sannolikhetsrum $(S, \mathcal{B}, \mu)$ definierar vi en ***standard stokastisk process*** *x som en avbildning $x : \Omega \times [0,1] \to \mathbb{R}$.*

(ii) För den hyperändliga tallinjen T och något hyperändligt sannolikhetsrum $(\Omega, \mathcal{A}, P)$ definierar vi en ***hyperändlig stokastisk process*** *som en* intern *avbildning $X : \Omega \times T \to {}^*\mathbb{R}$.*

Vårt mål med det här avsnittet är att från en hyperändlig stokastisk process få en standard stokastisk process genom att på något sätt "plocka ut standarddelen" ur den hyperändliga stokastiska processen. Dessa stokastiska processer är funktioner av två variabler så vi skall alltså "approximera hyperändligt i två led" för att uppnå vårt resultat. Vi ska titta på hur mätbarheten bevaras mellan de olika rummen men även hur vi kan relatera olika integraler med varandra.

Vi ska visa nästan alla steg vi gör i "Ω-led" men inte några steg i "T-led", detta på grund av det faktum att T och $[0,1]$ har olika mått förknippade med sig så vi måste på något sätt översätta mellan Loebmåttet och Lebesguemåttet. De bevis som krävs för detta går utanför ramen för detta arbete.

Not 12.1. *När vi vill ta skuggan av ett hyperreellt objekt A nedan kommer vi i likhet med 11.3 skriva* $^sA \doteqdot sh(A)$

Vi går vidare med att undersöka när skuggan av hyperändliga stokastiska processer är mätbara i Ω-led och T-led. Vi börjar med att definiera så kallade *lyftningar*.

Definition 12.3.

(i) Låt $f : \Omega \to \mathbb{R}$ *och* $F : \Omega \to {}^*\mathbb{R}$. *F kallas för en* ***lyftning*** *av f om F är en intern funktion och om* $\{w \in \Omega : {}^sF(w) \neq f(w)\}$ *är en Loeb-nollmängd*[17].

(ii) Låt $f : [0,1] \to \mathbb{R}$ *och* $F : T \to {}^*\mathbb{R}$. *F kallas för en* ***lyftning*** *av f om F är en intern funktion och om* $\{t \in T : {}^sF(t) \neq f({}^st)\}$ *är en Loeb-nollmängd.*

Vi får nu följande två resultat varav det andra lämnas obevisat (se tidigare diskussion).

Sats 12.4.

(i) Låt $(\Omega, \mathcal{A}, P)$ *vara ett hyperändligt sannolikhetsrum. och* $(\Omega, L(\mathcal{A}), P_L)$ *dess associerade Loebmåttrum. En funktion* $f : \Omega \to \mathbb{R}$ *är Loebmätbar om och endast om f har en lyftning* $F : \Omega \to {}^*\mathbb{R}$.

(ii) Låt $([0,1], \mathcal{B}, \mu)$ *vara Lebesgue-måttrummet och* $(T, L(\mathcal{A}), P_L)$ *Loebrummet associerat med den hyperändliga tidslinjen T. En funktion* $f : [0,1] \to \mathbb{R}$ *är Lebesgue-mätbar om och endast om den har en lyftning* $F : T \to {}^*\mathbb{R}$.

Bevis. $\Leftarrow$)

Låt F vara en lyftning av $f : \Omega \to \mathbb{R}$. Tag ett standard öppet intervall B med radien $\frac{1}{n}$, $n \in \mathbb{N}$, kring en godtycklig punkt $r \in \mathbb{R}$, det vill säga $B_{1/n}(r) = \{r' \in \mathbb{R} : |r - r'| < \frac{1}{n}\}$. Vi ska visa att $f^{-1}(B_{1/n}(r))$ är en Loebmätbar mängd. Låt nu $U = \{w \in \Omega : {}^sF(w) = f(w)\}$. Observera att

$$P_L(U^c) = P_L(\{w \in \Omega : {}^sF(w) \neq f(w)\}) = 0$$

enligt definitionen av en lyftning. Tag så ett $w \in U$. Då gäller att

$$f(w) \in B_{1/n}(r) \iff |r - {}^sF(w)| < \frac{1}{n} \iff {}^s|r - F(w)| < \frac{1}{n}$$
$$\iff |r - F(w)| \leq \frac{1}{n} - \frac{1}{m}$$

[17] Vi säger att ${}^sF(w) = f(w)$ *nästan överallt.*

för något $m \in \mathbb{N}$. Den första ekvivalensen följer per definition, den andra och tredje genom räkneregler för skuggor, sats 3.3. Det viktigaste är att vi ersätter en extern egenskap med en intern egenskap!

För nu ger sats 10.4 att $\{w \in \Omega : |r - F(w)| \leq \frac{1}{n} - \frac{1}{m}\} \in \mathcal{A}$ och i likhet med beviset till sats 11.8 så har vi att

$$\bigcup_{m \in \mathbb{N}} \{w \in \Omega : |r - F(w)| \leq \frac{1}{n} - \frac{1}{m}\} \in L(\mathcal{A}) \Rightarrow U \cap f^{-1}(B_{1/n}(r)) \in L(\mathcal{A}).$$

Eftersom

$$U \cap f^{-1}(B_{1/n}(r)) \triangle f^{-1}(B_{1/n}) = f^{-1}(B_{1/n}) \cap U^c \subseteq U^c$$

är en Loeb-nollmängd (då $P_L(U^c) = 0$) ger diskussionen efter Definition 11.7 att $f^{-1}(B_{1/n}) \in L(A)$ varför f är Loebmätbar.

$\Rightarrow$)

Låt nu $N_1, N_2 ...$ vara en uppräknelig öppen topologisk bas[18] till $\mathbb{R}$. Låt $B_n \doteqdot f^{-1}(N_n) \in L(\mathcal{A})$ enligt antagande. Enligt 11.10 kan vi hitta $A_{n,m} \in \mathcal{A}$ sådana att $P_L(B_n) \leq P(A_{n,m}) + \frac{1}{m}$ och $A_{n,m} \subseteq A_{n,m+1} \subseteq B_n$. Då blir $B_n \backslash \bigcup_m A_{n,m}$ en Loeb-nollmängd och eftersom dessa är slutna under uppräkneliga unioner är $C \doteqdot \bigcup_n (B_n \backslash \bigcup_m A_{n,m})$ också en Loeb-nollmängd. Vi ska nu hitta en intern funktion F sådan att ${}^sF(w) = f(w)$ överallt utom på C.

För varje $n \in \mathbb{N}$ låt $\mathcal{P}_{n,m} \doteqdot \{$alla interna funktioner $F : \Omega \to {}^*\mathbb{R} : F(A_{k,l}) \subseteq {}^*N_k$ för alla $k \leq n, l \leq m\}$. Varje $\mathcal{P}_{n,m}$ är icke-tom och Interna Definitionsprincipen ger att den är intern vilket innebär att det existerar ett $F \in \bigcap \mathcal{P}_{n,m}$ då snittet av en följd avtagande icke-tomma interna mängder är icke-tomt. Tag nu ett $w \notin C$ och sedan ett n sådant att $f(w) \in N_n$. Då har vi att $w \in B_n$ varför det finns ett k sådant att $w \in A_{n,k}$ eftersom vi lät $w \notin C$. Alltså är $F(w) \in {}^*N_n$ på grund av definitionen av $\mathcal{P}_{n,m}$ och $F(w) \in \bigcap\{{}^*N_n : f(w) \in N_n\}$ eftersom argumentet kan upprepas för varje n. Således ligger $F(w)$ i (den hyperreella versionen av) *alla* reella intervall kring $f(w)$ och därmed drar vi slutsatsen att ${}^sF(w) = f(w)$ överallt utom på C.

$\square$

Not 12.2. *Beviset av den andra delen i satsen ovan baseras på del 1 och det faktum att alla Lebesguemätbara mängder i* $[0, 1]$ *har Loebmätbara motsvarigheter i* T *och de har samma mått! På ett sätt kan vi alltså säga att* Lebesguemåttet på $[0, 1]$ är ekvivalent med (Loebversionen av) räknemåttet på T - *ett mycket intressant faktum vars bevis tyvärr går utanför detta arbete.*

[18]Detta är en term inom topologin som inte närmare definieras här. Det som vi behöver veta är att det är en samling mängder som "täcker" $\mathbb{R}$ och att varje öppet intervall kan skrivas som en union av ett godtyckligt antal element i basen. Att en uppräknelig sådan bas existerar är ett känt resultat från topologin.

12.3 Hyperändlig integrationsteori

När vi nu är klara med hur vi ska hantera mätbarhet för stokastiska processer så behöver vi, för att kunna använda våra processer till något, en teori kring integration. Denna visar sig vara särskilt elegant i vår hyperändliga värld.

Definition 12.5. *Låt $(\Omega, \mathcal{A}, P)$ vara ett hyperändligt sannolikhetsrum där P är räknemåttet och $\mathcal{A}$ algebran av alla interna mängder. Låt $F : \Omega \to {}^*\mathbb{R}$ vara en intern funktion. Väntevärdet, $E(F)$, av F[19] definieras som $E(F) = \sum\limits_{w \in \Omega} \frac{F(w)}{|\Omega|}$*

Den hyperändliga summan $\sum\limits_{w \in \Omega} \frac{F(w)}{|\Omega|}$ existerar och är väldefinierad genom ett transferargument, vilket gör vår integrationsteori särskilt enkel. De som är bekanta med sannolikhetsteori tycker säkert att vår definition av väntevärdet är underligt eftersom de flesta bara sett det definierat för så kallade diskreta och kontinuerliga slumpvariabler. Faktum är dock att vår definition ligger närmare den allmänna definitionen av väntevärdet. De som inte är bekanta med sannolikhetsteori kan tänka sig väntevärdet av en funktion från ett utfallsrum till den reella linjen som det ”genomsnittliga” värdet av funktionen vid många sannolikhetsexperiment.

Vi behöver nu ett sätt att "översätta"vårt väntevärde med den vanliga Lebesgue-integralen. Vi börjar med följande definition.

Definition 12.6. *Låt $(\Omega, \mathcal{A}, P)$ vara ett hyperändligt sannolikhetsrum där P är räknemåttet och $\mathcal{A}$ algebran av alla interna mängder låt och $F : \Omega \to {}^*\mathbb{R}$ en intern funktion. F sägs vara S-**integrerbar** om*

(i) $E(|F|)$ *är ett begränsat hyperreellt tal och*

(ii) *för $A \in \mathcal{A}$ gäller att* $P(A) \simeq 0 \implies \sum\limits_{w \in A} \frac{|F(w)|}{|\Omega|} \simeq 0$[20]

Följande sats knyter samman tankarna om S-integrerbarhet och vårt väntevärde. Som vanligt avstår vi från att bevisa den andra delen.

Sats 12.7.

(i) Låt $(\Omega, \mathcal{A}, P)$ vara ett hyperändligt sannolikhetsrum. och $(\Omega, L(\mathcal{A}), P_L)$ dess associerade Loebmåttrum. En funktion $f : \Omega \to \mathbb{R}$ är Loeb-integrerbar om och endast om f har en S-integrerbar lyftning $F : \Omega \to {}^*\mathbb{R}$. I detta fall gäller att $E(F) \simeq \int\limits_{\Omega} f(w) dP_L$

[19]Här kan man tänka sig F som en *slumpvariabel*, något vi ska beskriva närmare i nästa kapitel.
[20]Notera att operationen $|\cdot|$ har två olika betydelser i uttrycket, som absolutbelopp av $F(w)$ samt som intern kardinalitet av Ω.

(ii) Låt $([0,1],\mathcal{L},\mu)$ *vara Lebesgue-måttrummet och* $(T,\mathcal{A},P)$ *den hyperändliga tidslinjen. En funktion* $f:[0,1]\to\mathbb{R}$ *är Lebesgue-integrerbar om och endast om den har en* S*-integrerbar lyftning* $F:T\to {}^*\mathbb{R}$*. I detta fall har vi att* $E(F)\simeq\int_0^1 f(r)d\mu$*.*

För att bevisa detta behöver vi gå igenom en serie lemman. Vi börjar med att införa så kallade ändliga funktioner.

Definition 12.8. *Vi säger att en intern funktion* F *är* ***ändlig*** *om den är begränsad av något naturligt tal, det vill säga* $F:\Omega\to {}^*[-n,n]$ *för något* $n\in\mathbb{N}$*.*

Dessa ändliga funktioner har följande viktiga egenskap, vilken formuleras som ett lemma.

Lemma 12.9. *Låt* F *vara en ändlig funktion. Då gäller att* $E(F)\simeq\int\limits_{\Omega} {}^sF(w)dP_L$*.*

Beviset av detta är inte alltför svårt men ganska långt, varför vi inte redovisar det här. Beviset bygger på observationen att integralen respektive väntevärdet av en konstant på en delmängd till Ω bara är konstanten gånger P_L- respektive P-måttet av delmängden. Därefter kan vi approximera väntevärdet ovanifrån och loebintegralen underifrån med hjälp av *enkla funktioner*[21] och vi kan visa att dessa approximationer ligger godtyckligt nära varandra varför resultatet följer. För en mer detaljerad beskrivning, se exempelvis [Loeb]. Vi får efter detta följande karakteristik för S-integrerbara funktioner.

Lemma 12.10. *En funktion* $F:\Omega\to {}^*\mathbb{R}$ *är* S*-integrerbar om och endast om det existerar en följd av ändliga funktioner* $\langle F_n : n\in\mathbb{N}\rangle$ *sådana att* ${}^sE(|F-F_n|)\to 0$ *när* $n\to\infty$

Bevis. $\Rightarrow$)

Antag att $F:\Omega\to {}^*\mathbb{R}$ är S-integrerbar. Vi ska nu definiera en följd av ändliga funktioner på följande sätt: för varje $n\in {}^s\mathbb{N}$ definierar vi

$$F_n(w)\doteqdot\begin{cases}F(w)\text{ if }|F(w)|\le n\\ n\text{ if }F(w)>n\\ -n\text{ if }F(w)<-n\end{cases}$$

Då är F_n en ändlig funktion om $n\in\mathbb{N}$. För $m\in {}^s\mathbb{N}_\infty$ så har vi att

$$P(\{w\in\Omega:|F(w)|>m\})\le\frac{1}{m}E(F)\simeq 0$$

[21] Enkla funktioner är mätbara linjärkombinationer av funktioner som är konstanta på givna delmängder till Ω

på grund av Chebyshevs olikhet och $E(|F|) < \infty$. Vi har också att

$$E(|F - F_m|) \leq \sum_{w \in \Omega : F(w) > m} \frac{|F(w)|}{|\Omega|} \simeq 0$$

på grund av $P(\{w \in \Omega : |F(w)| > m\}) \simeq 0$, S-integrabiliteten och icke-standard versionen av konvergens.

$\Leftarrow$)

Antag nu att vi har en följd såsom den formuleras i satsen och anta vidare att $sup_w |F_n| < n$ (annars kan vi bara justera vårt index). ${}^sE(|F - F_N|) \simeq 0$ för $N \in {}^s\mathbb{N}_\infty$ ger att $\sum_{F(w)>N} \frac{|F|}{|\Omega|} \simeq 0$ för obegränsade N varför ${}^sE(|F|) = sh(\sum_{F(w) \leq N} \frac{|F|}{|\Omega|}) < \infty$.

Låt $\epsilon > 0$ vara ett godtyckligt reellt tal. Välj nu ett $n \in \mathbb{N}$ sådant att ${}^sE(|F - F_n|) < \frac{\epsilon}{2}$. Tag därefter ett $A \in \mathcal{A}$ med $P(A) < \frac{\epsilon}{2n}$. Då har vi att

$$\sum_A \frac{|F|}{|\Omega|} \leq \sum_A \frac{|F_n|}{|\Omega|} + \sum_A \frac{|F - F_n|}{|\Omega|} < \epsilon.$$

Således medför $P(A) \simeq 0$ att $\sum_{w \in A} \frac{|F(w)|}{|\Omega|} \simeq 0$.

$\square$

Lemma 12.11. *Antag att $F : \Omega \to {}^*\mathbb{R}$ är S-integrerbar. Då är ${}^sF$ Loebintegrerbar och ${}^sE(F) = \int_\Omega {}^sF dP_L$*

Bevis. Vi använder Lemma 12.10 för att välja en följd av ändliga funktioner F_n sådan att ${}^sE(|F - F_n|) \to 0$ när $n \to \infty$. Varje ${}^sF_n$ är Loebintegrerbar. För varje $\epsilon \in \mathbb{R}_+$ medför detta att $P_L(\{w \in \Omega : |{}^sF - {}^sF_n| \geq \epsilon\}) \to 0$ när $n \to \infty$. Detta säger vi innebär att ${}^sF_n \to {}^sF$ i mått.
Dessutom medför Lemma 12.9 att

$$\int_\Omega |{}^sF_m - {}^sF_n| dP_L = {}^sE(|F_m - F_n|) \leq {}^sE(|F_n - F|) + {}^sE(|F_m - F|) \to 0$$

när m och n $\to \infty$. Detta betyder att ${}^sF_n$ är en Cauchyföljd i $L^1(\Omega, L(\mathcal{A}), P_L)$[22]. Då säger en sats från standardanalysen[23] att ${}^sF \in L^1(\Omega, L(\mathcal{A}), P_L)$ varför ${}^sF$ är Loebintegrerbar och $\int_\Omega {}^sF dP_L = \lim_{n\to\infty} \int_\Omega {}^sF_n dP_L = \lim_{n\to\infty} {}^sE(F_n) = {}^sE(F)$. $\square$

[22]Ett funktionsrum bestående av alla funktioner på loebmåttsrummet vars absolutbelopps integral är begränsad.
[23]Riesz-Fischers sats ger att L_p-rummen är fullständiga.

Lemma 12.12. *Låt $f : \Omega \to \mathbb{R}$ vara en Loeb-integrerbar funktion. Då har f en S-integrerbar lyftning $F : \Omega \to {}^*\mathbb{R}$*

Bevis. Vi börjar med att låta

$$f_n(w) = \begin{cases} n \text{ om } f(w) > n \\ f(w) \text{ om } |f(w)| \leq n \\ -n \text{ om } f(w) < -n \end{cases}$$

för $n \in \mathbb{N}$

Den så kallade Dominerade konvergenssatsen[24] ger nu att $\int_\Omega |f - f_n| dP_L \to 0$ när $n \to \infty$. Eftersom $f_n(w)$ är ändlig och Loebmätbar så ger oss Sats 12.4 ändliga lyftningar F_n av f_n. Lemma 12.9 gör att

$$^sE(|F_n - F_m|) = \int_\Omega |f_n - f_m| dP_L \to 0$$

när m och n $\to \infty$. Vi kan då med hjälp av 10.11 välja ett $N \in {}^s\mathbb{N}_\infty$ sådant att $^sE(|F_n - F_N|) \to 0$ när $n \to \infty$. Låt $F \doteqdot F_N$. Vi ser med hjälp av Lemma 12.10 direkt att F är S-integrerbar så Lemma 12.11 ger att ^sF är Loebintegrerbar och att

$$\begin{aligned} \int_\Omega |^sF - f| dP_L &\leq \int_\Omega |^sF - {}^sF_n| dP_L + \int_\Omega |^sF_n - f| dP_L \\ &= {}^sE(|F - F_n|) + \int_\Omega |f_n - f| dP_L \to 0 \end{aligned}$$

när $n \to \infty$. Eftersom vänsterledet är oberoende av n så är $\int_\Omega |^sF - f| dP_L = 0$. Således är F en S-integrerbar lyftning av f. □

Slutligen bevisar vi nu första delen av Sats 12.7.

Bevis. $\Leftarrow$ följer direkt av Lemma 12.11. För $\Rightarrow$, tag enligt Lemma 12.12 en S-integrerbar lyftning, F, av f. Då ger Lemma 12.11 att ^sF är Loebintegrerbar och

$$^sE(F) = \int_\Omega {}^sF dP_L = \int_\Omega f dP_L$$

eftersom Lebesgueintegralen är noll på nollmängder. □

[24] se exempelvis [Folland].

Det som återstår är att visa en sats mycket lik Sats 12.7 men som utgår från en intern hyperreellvärd funktion F istället för en reellvärd funktion f.

Sats 12.13. *Låt $(\Omega, \mathcal{A}, P)$ vara ett hyperändligt sannolikhetsrum. och $(\Omega, L(\mathcal{A}), P_L)$ dess associerade Loebmåttrum. Låt $F : \Omega \to {}^*\mathbb{R}$ vara en intern och icke-negativ funktion. Då är F S-integrerbar om och endast om ${}^sF$ är Loeb-integrerbar och ${}^sE(F) = \int\limits_\Omega {}^sF dP_L$*

Bevis. $\Rightarrow$ följer direkt av Lemma 12.11. För $\Leftarrow$, tag med hjälp av Lemma 12.12 en lyftning G till ${}^sF$. Eftersom G är en lyftning gäller att

$$P_L(\{w : {}^sG(w) \neq {}^sF(w)\} = 0 \Rightarrow P_L(\{w : |G(w) - F(w)| > \frac{1}{n}\}) = 0$$

för alla $n \in \mathbb{N}$ varför

$$P(\{w : |G(w) - F(w)| > \frac{1}{n}\}) \simeq 0.$$

Detta gör att den interna mängden

$$I \doteqdot \{n : P(\{w : |G(w) - F(w)| > \frac{1}{n}\}) < \frac{1}{n}\}$$

innehåller alla $n \in \mathbb{N}$ varför den enligt 10.11 måste innehålla något $N \in {}^s\mathbb{N}_\infty$. Vi låter nu $B \doteqdot \{w : |G(w) - F(w)| > \frac{1}{N}\}$ och noterar att $P(B) \simeq 0$ samt att ${}^sF \neq {}^sG$ enbart på Loebnollmängden B. På grund av att G är S-integrerbar får vi att

$$\infty > {}^sE(G) = \int\limits_\Omega {}^sG dP_L = \int\limits_{\Omega \backslash B} {}^sG dP_L = \int\limits_{\Omega \backslash B} {}^sF dP_L = \int\limits_\Omega {}^sF dP_L = {}^sE(F)$$

Vi ska nu visa den andra förutsättningen i definitionen av S-integrerbarhet. Tag så en mängd $A \in \mathcal{A}$ med $P(A) \simeq 0$. Eftersom $P(A \backslash B) \simeq 0$ och G är S-integrerbar så har vi att

$$\sum_{A \backslash B} \frac{F(w)}{\Omega} = \sum_{A \backslash B} \frac{G(w)}{\Omega} \simeq 0.$$

Dessutom så ger ${}^sE(F) = {}^sE(G)$ att

$$E(F) \simeq E(G) \simeq \sum_{\Omega \backslash B} \frac{G(w)}{\Omega} \simeq \sum_{\Omega \backslash B} \frac{F(w)}{\Omega}$$

varför

$$\sum_B \frac{F(w)}{\Omega} = E(F) - \sum_{\Omega \backslash B} \frac{F(w)}{\Omega} \simeq 0.$$

$\square$

Nu lämnar vi måtteorin därhän och går vidare med ett konkret exempel där vi använder det ramverk för transfer av icke-standard stokastiska processer till standard stokastiska processer som vi utvecklat här.

13 Brownsk rörelse

I detta kapitel, där vi främst följer [Albeverio et. al.] kommer vi att konstruera så kallad Brownsk rörelse genom skuggan av en så kallad *hyperändlig slumpvandring*. Brownsk rörelse är en stokastisk process som kan användas för att modellera flera olika fenomen inom exempelvis fysik, ekonomi. Den är uppkallad efter botanikern Robert Brown som 1827 upptäckte hur pollenkorns rörelse i vatten under hans mikroskop tycktes helt slumpmässig. Idag är det fysiska fenomenet bakom Brownsk rörelse väl förstått men den underliggande matematiska modellen är fortfarande intressant och användbar. En känd tillämpning utanför fysiken är Black-Scholes ekvation, som används för att prissätta optioner på en aktiekurs vars rörelse är baserad på en Brownsk rörelse.

I standardteori brukar Brownsk rörelse implicit definieras som en stokastisk process som har vissa mer eller mindre abstrakta egenskaper. Denna definition är inte särskilt intuitiv och ett bättre alternativ skulle kunna vara att börja med den välkända diskreta processen slumpvandringen och nå Brownsk rörelse på hyperändlig väg. Efter detta vill vi gärna visa ekvivalensen med den vanliga definitionen. Det är detta som ska ske i detta kapitel med hjälp av det ramverk som vi skapade i det förra. Vi börjar med att definiera vår hyperändliga stokastiska process precist.

Definition 13.1. *Låt T vara den hyperändliga tidslinjen och låt $\Omega = \{-1, +1\}^T$ vilken definieras som den interna mängden av alla möjliga följder av -1:or och 1:or av längd $|T| \in {}^*\mathbb{N}_\infty$. Då definieras* ***hyperändlig slumpvandring*** *$B : \Omega \times T \to {}^*\mathbb{R}$ som $B(w,t) = \sum_0^t w(s)\sqrt{\triangle t}$ där $w(s)$ är den s:e posten i $w \in \Omega$. Vi låter även $B(w,0) \doteqdot 0$.*

Man kan se detta som att ”partikeln” B rör sig en distans $\sqrt{\triangle t}$ till vänster eller till höger med sannolikheten $\frac{1}{2}$. Vi påpekar också att vi använder konventionen

$$\sum_0^t X(w,s) = X(w,u) + ... + X(w,t-\triangle t) \tag{12}$$

så termen $X(w,t)$ inkluderas ej i summan. Notera att i vår definition innan antar s indexen $0, \triangle t, 2\triangle t, ..., 1$.

Not 13.1. *För att motivera varför hoppstorleken för processen är just $\sqrt{\triangle t}$, låt oss ersätta detta tal med $\triangle x$ i definitionen. Vi får:*

$$E(B(t)^2) = E((\sum_0^t w(s)\triangle x)^2) = E(\sum_{s,r<t} w(s)w(r)\triangle x^2) =$$
$$\triangle x^2 E(\sum_{r\neq s} w(r)w(s) + \triangle x^2 E(\sum_{s<t} w(s)^2) =$$
$$\triangle x^2 \frac{t}{\triangle t}$$

eftersom $w(r)w(s) = \pm 1$ med sannolikheten $\frac{1}{2}$ och $w(s)^2 = 1$. Så vill vi att variansen av processen ska vara begränsad måste hoppstorleken vara av storleksordning $\sqrt{\triangle t}$.

Innan vi går in hur Brownsk rörelse normalt definieras måste vi beskriva vad så kallade oberoende slumpvariabler är för något.

13.1 Tre definitioner av oberoende slumpvariabler

De som är bekanta med sannolikhetsteori känner kanske igen sig i denna definition.

Definition 13.2. *Givet ett hyperändligt sannolikhetsrum $(\Omega, \mathcal{A}, P)$ samt dess associerade Loebrum $(\Omega, L(\mathcal{A}), P_L)$ så sägs funktionen $X : \Omega \to {}^*\mathbb{R}$ vara en slumpvariabel om den är intern och $Y : \Omega \to \mathbb{R}$ vara en slumpvariabel om den är Loebmätbar.*

Vi ska nu definiera tre olika versioner av oberoende för våra två sannolikhetsrum.

Definition 13.3. *Låt $(\Omega, \mathcal{A}, P)$ vara ett hyperändligt sannolikhetsrum och $(\Omega, L(\mathcal{A}), P_L)$ dess associerade Loebrum.*

*i) En samling av slumpvariabler $\{X_i\}_{i \in I}$ i Loebrummet sägs vara **oberoende** om vi för varje ändlig delmängd $\{X_1, ..., X_n\}$, $n \in \mathbb{N}$ och varje n-tipel $(\alpha_1, ..., \alpha_n) \in \mathbb{R}$ har att*

$$P_L(\{w \in \Omega | X_1(w) < \alpha_1, ..., X_n(w) < \alpha_n\}) = \prod_{k=1}^{n} P_L(\{w \in \Omega | X_k(w) < \alpha_k).$$

ii) En samling av interna *slumpvariabler $\{X_i\}_{i \in I}$ det hyperändliga sannolikhetsrummet sägs vara **hyperoberoende** om vi för varje hyperändlig delmängd $\{X_1, ..., X_n\}$, $n \in {}^*\mathbb{N}$ och varje* intern *n-tipel $(\alpha_1, ..., \alpha_n) \in {}^*\mathbb{R}$ har att*

$$P(\{w \in \Omega | X_1(w) < \alpha_1, ..., X_n(w) < \alpha_n\}) = \prod_{k=1}^{n} P(\{w \in \Omega | X_k(w) < \alpha_k).$$

*iii) Samma samling i det hyperändliga sannolikhetsrummet sägs vara S-**oberoende** om vi för varje ändlig delmängd $\{X_1, ..., X_n\}$, $n \in \mathbb{N}$ och varje n-tipel $(\alpha_1, ..., \alpha_n) \in \mathbb{R}$ har att*

$$P(\{w \in \Omega | X_1(w) < \alpha_1, ..., X_n(w) < \alpha_n\}) \simeq \prod_{k=1}^{n} P(\{w \in \Omega | X_k(w) < \alpha_k).$$

Nästa lemma knyter ihop två av oberoende-begreppen.

Lemma 13.4. *Låt $\{X_i\}_{i\in I}$ vara S-oberoende på $(\Omega, \mathcal{A}, P)$. Då är $\{{}^sX_i\}_{i\in I}$ oberoende på dess associerade Loebrum.*

Bevis. Låt $m \in \mathbb{N}$ och $(\alpha_1, ..., \alpha_m) \in \mathbb{R}^m$. Då har vi att

$$\begin{aligned}
&P_L(\{w \in \Omega : {}^sX_{i_1}(w) < \alpha_1, ..., {}^sX_{i_m}(w) < \alpha_m\}) \\
&= \lim_{n\to\infty} {}^sP(\{w \in \Omega : X_{i_1}(w) < \alpha_1 - \frac{1}{n}, ..., X_{i_m}(w) < \alpha_m - \frac{1}{n}\}) \\
&= \lim_{n\to\infty} sh(\prod_{k=1}^{m} P(\{w \in \Omega : X_{i_k}(w) < \alpha_k - \frac{1}{n}\})) \\
&= \prod_{k=1}^{m} \lim_{n\to\infty} {}^sP(\{w \in \Omega : X_{i_k}(w) < \alpha_k - \frac{1}{n}\}) \\
&= \prod_{k=1}^{m} P_L(\{w \in \Omega : {}^sX_{i_k}(w) < \alpha_k\})
\end{aligned}$$

Vi har här använt S-oberoende i andra ekvivalensen. Resterande delar förljer av räknereglerna för skuggor samt Lemma 11.6. Kommutativiteten för gränsvärdet följer om det betraktas på icke-standardvis. □

Den kanske viktigaste satsen för hyperoberoende slumpvariabler är den så kallade centrala gränsvärdessatsen som av de flesta läsare säkert känns igen från standard sannolikhetsteori. Satsen säger att fördelningen[25] av genomsnittet av en följd av likafördelade oberoende slumpvariabler med väntevärde 0 och varians 1 går mot normalfördelningen.

Sats 13.5 (Centrala gränsvärdessatsen)**.** *Låt $\langle X_n : n \in {}^*\mathbb{N}\rangle$ vara en intern följd av hyperoberoende slumpvariablar på $(\Omega, \mathcal{A}, P)$ med en gemensam standard fördelningsfunktion F och med väntevärdet 0 och variansen 1. Då gäller, för varje $m \in {}^*\mathbb{N}_\infty$ och varje $\alpha \in {}^*\mathbb{R}$, att*

$$P(\{w \in \Omega : \frac{1}{\sqrt{m}} \sum_{k=1}^{m} X_k(w) \le \alpha\}) \simeq {}^s\Psi(\alpha)$$

där

$$\Psi(\alpha) \doteqdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{\frac{-x^2}{2}} dx$$

är den så kallade normalfördelningen.

Vi bevisar inte detta exakt, utan nöjer oss med en skiss. Eftersom

[25] Fördelningsfunktionen är den reellvärda funktion som ger sannolikheten för att en slumpvariabel ska anta ett värde $x \in \mathbb{R}$ eller mindre. I det här sammanhanget kan man se den som funktionen som räknar alla $w \in \Omega$ för vilka $X(w) \le x$.

fördelningsfunktionen ovan är standard kan vi se att skuggan av den måste vara fördelningsfunktionen för skuggorna av slumpvariablerna. Vi kan visa att dessa skuggor har väntevärde 0 och varians 1[26] och vi kan använda den vanliga centrala gränsvärdessatsen för att uppnå "samma" resultat för skuggorna. Därefter använder vi transfer från $V(^*\mathbb{R})$ till $V(\mathbb{R})$ för att få vårt resultat.

13.2 Den hyperändliga slumpvandringen och Brownsk rörelse

Nu är det dags att visa vårt slutgiltiga resultat. Vi börjar med en formell definition av Brownsk rörelse. Det finns flera ekvivalenta sätt att göra detta men vi utgår från följande ganska vanliga definition:

Definition 13.6. *En stokastisk process $\beta(w,t) : \Omega \times [0,1]$ på något Loebrum $(\Omega, L(\mathcal{A}), P_L)$ sägs vara en Brownsk rörelse om följande är uppfyllt:*

(i) $\beta(w,0) = 0$.

(ii) $\beta(\cdot,t)$ *är en slumpvariabel (och alltså Loebmätbar) för alla* $t \in [0,1]$.

(iii) För $s < t$ så är $\beta(w,t) - \beta(w,s)$ normalfördelad med väntevärden 0 och variansen $t - s$.

(iv) Om

$$s_1 < t_1 \leq s_2 < t_2 \leq ... \leq s_n < t_n \ i\ [0,1]$$

så är

$$\{\beta(w,t_1) - \beta(w,s_1), ..., \beta(w,t_n) - \beta(w,s_n)\}$$

en mängd oberoende slumpvariabler.

(v) $\{w \in \Omega : \beta(w,\cdot)$ *är* inte *en kontinuerlig funktion av* t $\}$ *är en Loeb-nollmängd*

Därefter följer vårt huvudsakliga resultat i detta kapitel:

Sats 13.7. *Låt B vara den hyperändliga slumpvandringen på $(\Omega, \mathcal{A}, P)$ och $b(w,t) \doteqdot sh(B(w,t'))$ där t' är punkten i T omedelbart till höger om t. Då är b Brownsk rörelse på $(\Omega, L(\mathcal{A}), P_L)$.*

Vi börjar med beviset av de tre första delarna i definitionen.

[26] Här är det lämpligt att använda den generella definitionen av väntevärde $E(X) \doteqdot \int_\Omega X(w)dP_L$ och motsvarande definition av variansen. Detta nämns eftersom vissa läsare kanske enbart är bekanta med definitionerna för diskreta och kontinuerliga slumpvariabler.

Bevis. (i) Med vår konvention (12) får vi att $B(w, \triangle t) = 0$ och resultatet följer direkt.

(ii) Följer även det direkt från definitionen och vår diskussion i föregående kapitel.

(iii) Vi ska använda vår version av Centrala gränsvärdessatsen men för att göra detta behöver vi återigen gå från en extern egenskap i P_L till en intern egenskap i P:

$$\begin{aligned}
&P_L(\{w \in \Omega : b(w, {}^s t) - b(w, {}^s s) \leq \alpha\}) \\
&= P_L(\{w \in \Omega : {}^s B(w, t) - {}^s B(w, s) \leq \alpha\}) \\
&= P_L(\{w \in \Omega : {}^s (\sum_{k=s}^{t} w_k \sqrt{\triangle t}) \leq \alpha\}) \\
&= \lim_{n \to \infty} {}^s P(\{w \in \Omega : \sum_{k=s}^{t} w_k \sqrt{\triangle t} \leq \alpha + \frac{1}{n}\}) \\
&= \lim_{n \to \infty} {}^s P(\{w \in \Omega : \frac{1}{\sqrt{m}} \sum_{k=1}^{m} w_{i_k} \leq \frac{\alpha + 1/n}{\sqrt{t-s}}\})
\end{aligned}$$

Här har vi låtit $t - s = m\triangle t$ och i_k numrera slumpvariablerna "mellan" s och t. Nu följer resultatet av Sats 13.5.

(iv) Vi visar detta med en skiss när $n = 2$. Det generella beviset följer samma princip. Om vi kan visa att $\{B(w, t_1) - B(w, s_1), B(w, t_2) - B(w, s_2)\}$ är hyperoberoende så följer resultatet från Lemma 13.4 eftersom hyperoberoende $\Rightarrow$ S-oberoende. Summorna har inga gemensamma index på grund av vår konvention (12). Den delföljd till $w \in \Omega$ som avgör hur stor den första summan är har alltså ingen inverkan på den andra och vice versa. Detta innebär att "antalet" $w \in \Omega$ som uppfyller $B(w, t_1) - B(w, s_1) \leq \alpha_1$ och $B(w, t_2) - B(w, s_2) \leq \alpha_2$ är precis så många som bara uppfyller $B(w, t_1) - B(w, s_1) \leq \alpha_1$ multiplicerat med *andelen* $w \in \Omega$ som bara uppfyller $B(w, t_2) - B(w, s_2) \leq \alpha_2$. Men detta innebär ju att

$$\begin{aligned}
&P(\{w \in \Omega : B(w, t_1) - B(w, s_1) \leq \alpha_1, B(w, t_2) - B(w, s_2) \leq \alpha_2\}) \\
&= P(\{w \in \Omega : B(w, t_1) - B(w, s_1) \leq \alpha_1\}) \\
&\cdot P(\{w \in \Omega : B(w, t_2) - B(w, s_2) \leq \alpha_2\})
\end{aligned}$$

vilket är vad vi ville visa.

□

För nästa del behöver vi först observera några saker. Låt oss införa notationen $\triangle B(s) = B(s + \triangle t) - B(s)$. Detta är alltså steget som B tar mellan s och $s + \triangle t$ och har värdet $\pm\sqrt{\triangle t}$. Således är $(\triangle B(s))^2 = \triangle t$. Vi noterar vidare att

$$B(t)^2 = \sum_{s=0}^{t} (B(s) + \triangle B(s))^2 - B(s)^2.$$

För att se detta, notera att den första termen i summan lägger till kvadraten för det aktuella indexet medan den senare termen drar ifrån kvadraten för det förra indexet. Nu får vi alltså

$$E(B(t)^2) = E(\sum_{s=0}^{t} (B(s) + \triangle B(s))^2 - B(s)^2)$$
$$= \sum_{s=0}^{t} E(2B(s)\triangle B(s) + \triangle t) = \sum_{s=0}^{t} \triangle t = t$$

där vi noterat att $E(2B(s)\triangle B(s)) = 0$ eftersom $\triangle B(s)$ antar ± 1 med sannolikhet $\frac{1}{2}$ oavsett vad $B(s)$ är.

Från detta får vi att:

$$E(B(t)^4) = E(\sum_{s=0}^{t} ((B(s) + \triangle B(s))^4 - B(s)^4))$$
$$= \sum_{s=0}^{t} E(4B(s)^3 \triangle B(s) + 6B(s)^2 B(s)^2 + 4B(s)\triangle B(s)^3 + \triangle B(s)^4)$$
$$= \sum_{s=0}^{t} E(6B(s)^2 \triangle t + (\triangle t)^2)$$

av samma anledning som $E(2B(s)\triangle B(s)) = 0$. Vidare är

$$\sum_{s=0}^{t} E(6B(s)^2 \triangle t + (\triangle t)^2) = \sum_{s=0}^{t} 6s\triangle t + (\triangle t)^2 = 3t(t - \triangle t) + t\triangle t \leq 3t^2$$

där vi använt den slutna formeln för aritmetiska summor[27]. Samma process ger oss att

$$E((B(T) - B(s))^4) \leq 3(t-s)^2 \tag{13}$$

Nu är vi redo för beviset av (v).

[27] $\sum_0^t s\triangle t = t(t - \triangle t)/2$, se vår konvention (12).

Bevis. Låt oss för varje par av heltal m, n definiera mängden

$$\Omega_{m,n} \doteqdot \{w \in \Omega : (\exists i < n)(\exists s \in T \cap I)(|B(w,s) - B(w, \frac{i}{n})| \geq \frac{1}{m})\}$$

där $I = (\frac{i}{n}, \frac{i+1}{n}]$. Vi har att $b(w, \cdot)$ är diskontinuerlig om $w \in \bigcup_m \bigcap_n \Omega_{m,n}$ (detta lämnas obevisat) så vi ska alltså visa att $P_L(\bigcup_m \bigcap_n \Omega_{m,n}) = 0$. För detta räcker det med att visa att ${}^sP(\Omega_{m,n}) \to 0$ när $n \to \infty$ för alla $m \in \mathbb{N}$ eftersom Loeb-nollmängder är slutna under uppräkneliga unioner och snitt. Notera också att

$$P(\Omega_{m,n}) \leq \sum_{i<n} P(\{w \in \Omega : (\exists s \in T \cap I)(|B(w,s) - B(w, \frac{i}{n})| \geq \frac{1}{m})\})$$

på grund av P:s måttegenskaper. Detta tal är dessutom mindre än

$$2\sum_{i<n} P(\{w \in \Omega : |B(w,s) - B(w, \frac{i}{n})| \geq \frac{1}{m})\})$$

på grund av följande argument: Antag att $|B(w,(i+1)/n) - B(w, i/n)| < 1/m$ men att det finns ett $s \in I$ sådant att $|B(w,s) - B(w, i/n)| > 1/m$. Tag det minsta s:et och definiera *reflektionen* av w :

$$w'(t) \doteqdot \begin{cases} w(t) \text{ om } t < s \\ -w(t) \text{ om } t \geq s \end{cases}$$

Då gäller ju att

$$\left|B(w',(i+1)/n) - B(w', i/n)\right| \geq 1/m$$

så för varje

$$w \in \{w \in \Omega : (\exists s \in T \cap I)(|B(w,s) - B(w, \frac{i}{n})| \geq \frac{1}{m})\}$$

finns det ett

$$w' \in \{w \in \Omega : |B(w,s) - B(w, \frac{i}{n})| \geq \frac{1}{m})\}$$

varför olikheten gäller. Nu får vi alltså, med hjälp av Chebyshevs olikhet, att

$$P(\Omega_{m,n}) \leq 2\sum_{i<n} P(\{w \in \Omega : |B(w,s) - B(w, \frac{i}{n})| \geq \frac{1}{m})\})$$

$$\leq 2\sum_{i<n} m^4 E(|B(w,s) - B(w, \frac{i}{n})|^4) \leq 6m^4 \sum_{i<n} \frac{1}{n^2} = \frac{6m^4}{n} \to 0$$

när $n \to \infty$ där vi använt identiteten (13). Detta räcker alltså för att visa satsen! $\square$

14 Källförteckning

Albeverio S. et. al. (1986) *Non-standard methods in stochastic analysis and mathematical physics*, London: Academic press.

Anderson R.M. (1976) *A non-standard representation for Brownian motion and Itô integration*, Israel Journal of mathematics, vol. 25.

Bair J. et. al. *Is mathematical history written by the victors?* http://u.cs.biu.ac.il/ katzmik/abc.pdf (2013-05-14).

Bennett C. (1994) *Grundläggande Modellteori*, Göteborg: Institutionen för filosofi GU.

Cutland N. et al (1995) *Loeb measure theory: Developments in nonstandard mathematics*, ss 151-177. Guildford: Biddles Ltd.

Cutland N. (2000) *Loeb Measures in Practice: Recent Advances*, Berlin: Springer-Verlag.

Folland G.B. (1999) *Real analysis: Modern Techniques and Their Applications*, New York: John Wiley & Sons.

Goldblatt R. (1998) *Lectures on the Hyperreals: an introduction to nonstandard analysis*, New York: Springer-Verlag.

Hurd A.E. och Loeb P.A. (1985) *An Introduction to Nonstandard Real Analysis*, Orlando: Academic Press.

Keisler J. (2010) *The ultrafilter construction: Ultrafilters across mathematics*, red. Bergelson et al., ss 163-179. Providence: American mathematical society.

Loeb P.A. (1975) *Conversion from nonstandard to standard measure sequences and applications in probability theory*, Transactions of the American Mathematical Society, vol.211, ss. 113-122.

Mendelson F. (2001) *Introduction to mathematical logic*, Boca Raton: Chapman and Hall/CRC.

Stroyan K.D. och Luxemburg W.A.J. (1976) *Introduction to the theory of infinitesimals*, New York: Academic Press.