CHALMERS | GÖTEBORGS UNIVERSITET

Icke-standardanalys

En introduktion och en jämförande studie med tillämpningar inom komplex analys och finansiell matematik.

Kandidatarbete inom civilingenjörsutbildningen vid Chalmers

Lina Berneryd
Victor Ekdahl
Magnus Jedvert
Oskar Paulander

Institutionen för matematiska vetenskaper
Chalmers tekniska högskola
Göteborgs universitet
Göteborg 2012

Icke-standardanalys

En introduktion och en jämförande studie med tillämpningar inom komplex analys och finansiell matematik.

Kandidatarbete i matematik inom civilingenjörsprogrammet Teknisk fysik vid Chalmers

Oskar Paulander

Kandidatarbete i matematik inom civilingenjörsprogrammet Teknisk matematik vid Chalmers

Lina Berneryd Victor Ekdahl Magnus Jedvert

Handledare: Leif Arkeryd
Examinator: Carl-Henrik Fant

Institutionen för matematiska vetenskaper
Chalmers tekniska högskola
Göteborgs universitet
Göteborg 2012

Sammanfattning

Den här rapporten ger en introduktion till icke-standardanalys och den grundläggande teorin som krävs för att nå dit. Icke-standardanalysen bygger på en utvidgning av de reella talen som kallas för de hyperreella talen. Utökningen från de reella talen till de hyperreella talen visas både axiomatiskt och med en konkret ultraproduktkonstruktion. Centrala begrepp som transferprincipen och interna mängder förklaras och bevisas. Två olika tillämpningar av icke-standardanalys tas upp, ett inom komplex analys och ett inom finansiell matematik. Inom den komplexa analysen bevisas Picards stora sats, först med en icke-standardmetod och sedan med en standardmetod. Dessa två tillvägagångssätt jämförs sedan med slutsatsen att icke-standardmetoden är enklare, kortare och mer intuitiv. Inom den finansiella matematiken presenteras den hyperfinita binomialmodellen och det demonstreras hur den kan sägas innehålla prisformeln för både binomialmodellen och Black-Scholes modell samtidigt.

Dessa exempel pekar på att icke-standardanalys är ett kraftfullt verktyg för att såväl lösa problem som att förstå matematiska koncept.

Abstract

Non-standard analysis is a field of mathematics wherein infinitely large and infinitely small (infinitesimal) numbers are used. In this report, this set of numbers, called the hyperreals, are defined using two different approaches. One using an explicit ultraproduct construction and the other with an axiomatic approach. Central concepts such as the transfer principle and internal sets are explained and proven. To emphasize the value of non-standard analysis two applications are given and studied, one in the field of complex analysis and one about financial mathematics. The example given in complex analysis is Picard's big theorem, which is proven using both non-standard and standard techniques. The two proofs are then compared with the conclusion that the non-standard version is shorter, simpler and more intuitive than its standard counterpart. Within the financial mathematics it is demonstrated how the pricing formulae from both the binomial model and Black-Scholes are contained in the hyperfinite binomial model.

In both cases it turns out that the nonstandard approach is a powerful tool for both solving problems as well as understanding mathematical concepts.

Innehåll

Förord

Denna rapport har skrivits för att ge en introduktion till icke-standardanalysen. Författarna har bakgrund inom utbildningarna Teknisk Matematik och Teknisk Fysik, och arbetet är skrivet som en introduktion för studenter med liknande förkunskaper. Vidare så fördjupar sig arbetet i två delar av matematiken, finansiell matematik samt komplex analys, som avspeglar författarnas egna intresseområden.

Författarna vill framföra sitt tack till professor Leif Arkeryd för hans entusiasm och mycket goda handledning, både inom matematiken och praktiska detaljer som litteraturtips och detaljer angående fackspråk.

Här följer en redogörelse för kandidatarbetsgruppens medlemmars individuella bidrag under kandidatarbetet.

- Slutsatser och diskussion - I de två exemplen så har respektive huvudförfattare tillsammans bidragit till diskussionen under arbetets gång och gemensamt kommit fram till slutsatserna.

- Huvudansvarig författare av avsnitt

  1. Abstract och Sammanfattning - Oskar Paulander, Lina Berneryd, Magnus Jedvert, Victor Ekdahl
  2. Förord - Lina Berneryd, Oskar Paulander
  3. Inledning - Magnus Jedvert
  4. Grundläggande Teori - Oskar Paulander
  5. Filter och ultrafilter - Victor Ekdahl
  6. Konstruktion av $\mathbb{R}$ från $\mathbb{Q}$ - Oskar Paulander
  7. Icke-standardutvidgning av $\mathbb{R}$
     (a) Konstruktion av ${}^*\mathbb{R}$ med ultraprodukter - Victor Ekdahl
     (b) Axiomatisk utvidgning av $\mathbb{R}$ - Magnus Jedvert
  8. Transfer - Lina Berneryd
     (a) Integralen blir en summa - Victor Ekdahl
  9. Definitioner av hyperreella begrepp och regler - Oskar Paulander
  10. Komplex analys
      (a) Syfte - Oskar Paulander
      (b) Nödvändig teori inom komplex analys
          i. Topologiska begrepp - Lina Berneryd och Oskar Paulander
          ii. Övertäckningsavbildning - Lina Berneryd och Oskar Paulander
          iii. Konstruktion av övertäckningsavbildningen - Oskar Paulander
      (c) Icke-standardanalys och Picards stora sats
          i. S-kontinuitet - Lina Berneryd
          ii. S-kontinuitet och väsentliga singulariteter - Lina Berneryd
          iii. Ett nytt icke-standardbegrepp; permanens - Oskar Paulander
          iv. Robinsons och Picards stora sats - Lina Berneryd och Oskar Paulander
      (d) Standardanalys och Picards stora sats - Lina Berneryd
      (e) Jämförelse mellan icke-standard och standard Picards stora sats - Lina Berneryd
      (f) Slutsats - Oskar Paulander
  11. Finansiell matematik
      (a) Introduktion till finans - Victor Ekdahl
      (b) Grundläggande stokastiska processer - Victor Ekdahl och Magnus Jedvert
      (c) Matematiska definitioner av de finansiella begreppen - Magnus Jedvert

(d) Interna och externa mängder - Victor Ekdahl med bidrag [1] av Magnus Jedvert
(e) Mått och sannolikhetsrum - Magnus Jedvert
(f) Hyperfinita binomialmodellen - Magnus Jedvert
(g) Jämförelse med standardmetoder - Magnus Jedvert och Victor Ekdahl

- Redaktionell ansvarsfördelning - Lina Berneryd och Oskar Paulander har ansvarat för den slutgiltiga sammanställningen av rapporten.

En loggbok har förts över de enskilda medverkandes prestationer under arbetet.

[1] Beviset av sats 41

1 Inledning

Icke-standardanalys använder sig av ett utvidgat talsystem, det hyperreella talsystemet, som utöver de vanliga reella talen även innehåller infinitesimala och oändligt stora tal. Detta visar sig vara en kraftfull utökning då man kan bevara alla elementära axiom från standardanalysen så att grundläggande satser sanna inom standardanalysen är sanna även inom icke-standardanalysen och *tvärtom*. Denna förmåga att gå fram och tillbaks mellan de reella och hyperreella talen kallas för transferprincipen. Mer precist är det alla satser inom första ordningens logik som går att överföra. Detta är satser av formen ”för alla tal x...”.

Förmågan att arbeta med infinitesimaler på ett rigoröst sätt är ett kraftfullt verktyg inom både ren och tillämpad matematik, såsom geometri, matematisk analys, sannolikhetslära, finansiell matematik och matematisk fysik.

En omedelbar tillämpning inom matematisk analys är till exempel möjligheten att kunna definiera derivata som en äkta kvot

$$f' = {}^{o}\left(\frac{f(x+\epsilon)-f(x)}{\epsilon}\right)$$

mycket likt det sätt konceptet introducerades och är förhoppningsvis mer intuitivt tilltalande än motsvarande epsilon-delta-gränsvärde.
En stor nytta med icke-standardanalys är möjligheten till kortare och mer intuitiva bevis.

> ”There are good reasons to believe that non-standard analysis, in some version or other, will be the analysis of the future.” [3]
>
> *Kurt Gödel*

1.1 Historisk bakgrund

Historiskt har man ofta använt oändligt små och stora tal i olika resonemang, även om man haft svårt att ge en rigorös definition av vad det är. Arkimedes betraktade till exempel geometriska objekt som summan av oändligt många infinitesimala element. Detta använde han framgångsrikt för att beräkna längder och areor på olika geometriska figurer.

Både Newton och Leibniz använde sig av infinitesimaler när de oberoende utvecklade differential- och integralkalkylen i slutet av 1600-talet, även om deras syn på infinitesimaler skiljde sig något. Ett exempel på hur Leibniz använde sig av infinitesimaler är Leibniz lag för derivata av en produkt. Om $u(x)$ och $v(x)$ är två funktioner så är $(uv)' = u'v + uv'$ eller motsvarande $\mathrm{d}uv = u\,\mathrm{d}v + v\,\mathrm{d}u$. Han visade detta genom att observera

> ”$\mathrm{d}uv$ is the same thing as the difference between two successive uv’s; let one of these be uv, and the other $u + \mathrm{d}u$ into $v + \mathrm{d}v$” [3]

$$\begin{aligned}\mathrm{d}uv &= (u+\mathrm{d}u)(v+\mathrm{d}v) - uv\\ &= u\,\mathrm{d}v + v\,\mathrm{d}u + \mathrm{d}u\,\mathrm{d}v\end{aligned}$$

och sedan konstatera att resultatet följer eftersom

> ”the omission of the quantity $\mathrm{d}u\,\mathrm{d}v$, which is infinitely small in comparison with the rest, for it is supposed that $\mathrm{d}u$ and $\mathrm{d}v$ are infinitely small.” [3]

Newton formulerade sig istället i termer av *fluenter* x, och deras *fluxioner* $\dot{x}$ som

> ”the speeds with which they flow and are increased by their generating motion.” [3]

och ansåg *inte*, till skillnad från Leibniz

> ”Mathematical Quantities as composed of Parts extremaly small, but as generated by a continual motion.” [3]

Historiskt utsattes användandet av infinitesimaler även för mycket kritik. En stor del av kritiken härstammar från att man både gör algebraiska manipulationer med infinitesimalerna och samtidigt sätter dem till noll där det passar. Den mest berömda kritiken kommer från Berkely [14]

> ”And what are these Fluxions? The Velocities of evanescent Increments? And what are these same evanescent Increments? They are neither finite Quantities nor Quantities infinitely small, nor yet nothing. May we not call them the ghosts of departed quantities?”

där han i ”departed” syftar på $\mathrm{d}x = 0$ och ”ghost” $\mathrm{d}x \neq 0$. Han skriver även i [15]

> ”Axiom. No reasoning about things whereof we have no idea. Therefore no reasoning about Infinitesimals.”

När Cantor i slutet av 1800-talet lade den första rigorösa grunden för mängdlära, där man i viss mening kan behandla begreppet oändligheter, trodde han att man skulle kunna bevisa avsaknaden av infinitesimaler med hjälp av detta nya system. Han skriver i ett brev till Weierstrass 1887, som han senare publicerar i [16], ett ’bevis’ för att infinitesimaler är

> ”impossible, i.e., intrinsically inconsistent imaginary things”.

En annan del av kritiken mot infinitesimaler är av metafysisk natur, att det är något mystiskt och påhittat med dessa. En mätning av ett instrumentvärde kan till exempel aldrig resultera i en infinitesimal. En av de kraftfullaste kritikerna i denna kategori är Karl Marx som benämner Leibniz och Newton teorier som ”the mystical differential calculus” [17].

I början av 60-talet kom Abraham Robinson på ett sätt att definiera infinitesimaler på ett rigoröst sätt och skriver i [4]

> ”In the fall of 1960 it occured to me that the concepts and methods of contemporary Mathematical Logic are capable of providing a suitable framework for the development of the Differential and Integral Calculus by means of infinitely small and infinitely large numbers.”

och kommenterar Lebniz tidigare arbete

> ”However, neither he [Leibniz] nor his disciples and successors were able to give a rational development leading up to a system of this sort. As a result, the theory of infinitesimals gradually fell into disrepute and was replaced eventually by the classical theory of limits. [...] It is shown in this book that Leibniz’s ideas can be fully vindicated and that they lead to a novel and fruitful approach to classical Analysis and to many other branches of mathematics.”

Det är denna bok som lade grunden för hela den fortsatta utvecklingen inom icke-standard-matematik.

1.2 Syfte och metod

Detta kandidatarbete bestod till stor del i en självstudiekurs i icke-standardanalys. Målsättningen var att först bekanta sig med icke-standardanalysen och sedan använda dessa kunskaper för att belysa fördelarna genom konkreta exempel. Arbetet kan därför delas in i två faser där första fasen bestod av självstudier inom icke-standardanalysen. När förståelse för ämnet byggts upp påbörjades nästa fas: fördjupning i konkreta exempel. Rapporten skrevs löpande under arbetet, i synnerhet den första delen med underliggande teori, såväl icke-standardteori som standardmatematik.

Vid val av fördjupningsområden beaktades en rad kriterier: relevans för matematiken, svårighetsgrad och inte minst eget intresse. Det var även viktigt att exemplena faktiskt lyfte fram fördelarna med icke-standardanalys gentemot standardanalys. Två huvudsakliga områden valdes med detta i beaktande, hyperbinomialmodellen inom finansmatematik, och Picards stora sats inom komplex analys.

2 Grundläggande Teori

2.1 Algebraiska Begrepp

Det finns två huvudsakliga vägar att gå för att konstruera en modell för de hyperreella talen. Ett sätt är att bygga upp teorin axiomatiskt med utgångspunkt i de reella talen. Det andra alternativet är att , på liknande sätt som de reella talen kan definieras ur de rationella, definiera de Hyperrella talen som ekvivalensklasser av talföljder. Ekvivalensrelationen som används kallas ett *filter* eller speciellt *ultrafilter*. Konstruktion av de Hyperreella talen medelst filter är den allmänt mest använda tillika den tolkning som används i denna uppsats (dock kommer även den axiomatiska beskrivningen tas upp).

Detta avsnitt kommer att gå igenom nödvändiga algebraiska begrepp och definitioner som behövs för att beskriva och förstå uppbyggnaden av ett ultrafilter. Läsaren förväntas känna till elementära begrepp såsom "partition av en mängd" och "binär operation".

I slutet av detta avsnitt finns en sammanfattning av avsnittet som belyser de viktigaste slutsatserna och även förklarar varför dessa är intressanta för våra ändamål.

2.1.1 Ekvivalensrelationer och -klasser

Definition 1. Ekvivalensrelation. *En relation $\sim$ på en icke-tom mängd A kallas för en ekvivalensrelation (på A) om $\sim$ uppfyller följande :*

$$a \in A \text{ då } a \sim a \tag{1}$$
$$a,b \in A \wedge a \sim b \Rightarrow b \sim a \tag{2}$$
$$a,b,c \in A \wedge (a \sim b \wedge b \sim c) \Rightarrow a \sim c \tag{3}$$

En ekvivalensrelation sägs vara *reflexiv*(1) , *symmetrisk*(2) och *transistiv*(3) enligt defintion 1.

Definition 2. Ekvivalensklass. *Låt $\sim$ vara en ekvivalensrelation på en mängd A. Låt $a \in A$ och $[a] = \{x \in A : a \sim x\}$. Delmängden $[a]$ av A kallas för as ekvivalensklass (med avseende på $\sim$).*

Sats 1. *Om $\sim$ är en ekvivalensrelation på en mängd S så bildar ekvivalensklasserna som genereras av $\sim$ en partition av S. Omvänt gäller också att givet en partition P och en relation $\sim$ på S enligt $a \sim b$. så är $\sim$ en ekvivalensrelation omm $\exists P_i \subset P$ s.a. $a,b \in P_i$. Resultatet säger att det finns ett naturligt samband mellan alla ekvivalensrelationer på en mängd och dess partitioner.*

Bevis. Låt $\sim$ vara en ekvivalensrelation på S. Om $a \in S$ så tillhör a åtminstone en ekvivalensklass; nämnligen $[a]$. I detta fall är S alltså unionen av alla ekvivalensklasser. Vad som nu återstår att visa är att om två ekvivalensklasser inte är lika så är de disjunkta - eller omvändningen; om två ekvivalensklasser har något gemensamt element så måste de vara lika. Antag nu att $[a] \cap [b] \neq \emptyset$ och låt $c \in [a] \cap [b]$
Tag $x \in [a] \Longrightarrow a \sim c$ och $a \sim x$ och därmed $c \sim a$ och $a \sim x$ varför $c \sim x$ Men vi vet även att $b \sim c$ vilket medför $b \sim x$ d.v.s. $x \in [b]$ Detta visar att $[a] \subseteq [b]$. Men på samma sätt visas det att $[b] \subseteq [a]$, alltså måste måste $[a] = [b]$

Det återstår att visa omvändningen. Låt vara som ovan. Om $a \in S$ gäller $a \sim a$ ty $\exists P_i \subset P$ s.a. $a \in P_i$. Att visa att en delmängd är symmteriskt är trivialt ; om en mängd innehåller a och b så innehåller den b och a. Slutligen skall transistiviteten visas. Antag $a \sim b$ och $b \sim c$. Detta ger att $\exists P_i,P_j \subset P$ s.a. $a,b \in P_i$ och $b,c \in P_j$. Men $b \in P_i \cap P_j \Rightarrow P_i \cap P_j \neq \emptyset \Rightarrow P_i = P_j$, ty P är en partition. Då både a och c tillhör denna mängd följer det att $a \sim c$ - transistiviteten är nu bevisad.

$\square$

2.1.2 Algebraiska Strukturer

Definition 3. Grupp. *En grupp $(G,\oplus)$ (skrivs oftast bara som G) utgörs av en mängd G samt en operation $\oplus$ och uppfyller följande ;*

G är Sluten : $a \oplus b \in G$ *om* $a,b \in G$

Associativ : $(a \oplus b) \oplus c = a \oplus (b \oplus c)$

Neutralt element : $\exists e$ *s.a.* $e \oplus a = a \oplus e = a$, $e,a \in G$

Inverser : $\exists a^{-1}$ $\forall a \in G$ *s.a.* $a \oplus a^{-1} = a^{-1} \oplus a = e$

anm: Det går att visa att det endast existerar ett neutralt element samt att varje element a har en entydig invers a^{-1}.

Definition 4. Abelsk Grupp. *En grupp kallas abelsk om operationen* $\oplus$ *är kommutativ på mängden, dvs :* $a \oplus b = b \oplus a, a,b \in G$.

Inför följande definition förtydligas att additiv- respektive multiplikativ notation tillämpas. De additiva inverserna skrivs $-a$ och operationen $a \cdot b$ skrivs kort ab. Det bör noteras att operationerna $+$ och $\cdot$ inte nödvändigtvis behöver vara vanlig addition eller multiplikation.

Definition 5. Ring. *En ring är en algebraisk struktur bestående av en mängd* R *samt två binära operationer* $+$ *och* $\cdot$ *och som uppfyller att :*

$(R, +)$ *är en abelsk grupp*

$a(bc) = (ab)c$; $\cdot$ *är associativ*

$a(b + c) = ab + ac$ *och* $(a + b)c = ac + bc$; $\dot{}$ *är distributiv m.a.p.* $+$.

2.2 Filter och ultrafilter

Ett filter är en delmängd av en partiellt ordnad mängd som uppfyller vissa villkor. Vid konstruktionen av de hyperreella talen så kommer det att finnas behov av att jämföra oändliga följder utan att kräva likhet mellan alla termer. Man kan då istället ”välja ut” de termer som man vill jämföra genom att använda ett filter.

Definition 6. Filter. *Ett filter* $\mathcal{F}$ *på* S *är den delmängd av* $(P(S), \leq)$ *som uppfyller villkoren*

I: *Om* $A, B \in \mathcal{F}$ *så gäller att* $A \cap B \in \mathcal{F}$.

II: *Om* $A \in \mathcal{F}$ *och* $A \subseteq B \subseteq S$ *så gäller att* $B \in \mathcal{F}$.

III: $\emptyset \notin \mathcal{F}$ *och* $S \in \mathcal{F}$

$P(S)$ *är potensmängden av* S *det vill säga* $P(S) = \{A : A \subseteq S\}$ *(mängden av delmängder). Ett filter uppfyller kriteriet* $\emptyset \notin \mathcal{F}$ *när det är skilt från* $P(S)$ *och kallas då äkta.*

Ett enkelt exempel på ett filter på en ändlig mängd är ett filter där basen utgörs av ett element i mängden.

Exempel 1. *Låt* $\mathcal{F}$ *vara ett filter på* $S = \{1, 2, 3\}$ *och låt filtrets bas vara* $\{2\}$. *Då gäller att*

$$P(S) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

Från filtrets bas $\{2\}$ *får vi då med hjälp av definitionen av filter att*

$$\mathcal{F} = \{\{2\}, \{1,2\}, \{2, 3\}, \{1, 2, 3\}\}$$

Notera att $\mathcal{F} \subseteq P(S)$.

Eftersom en oändlig följd kan definieras som en funktion $f : \mathbb{N} \to S$ för någon mängd S (där S innehåller termerna) så är det särskilt intressant för oss att diskutera filter på $\mathbb{N}$. Ett specifikt filter som är relevant är Fréchet-filtret på $\mathbb{N}$.

Exempel 2. *Fréchet-filtret på* $\mathbb{N}$ *ges av*

$$\mathcal{F}_0 = \{A \subseteq \mathbb{N} \mid \mathbb{N} \setminus A \text{ är ändligt } \}$$

En delmängd $A \subseteq \mathbb{N}$ *tillhör alltså* $\mathcal{F}_0$ *om och endast om komplementet till* A *i* $P(\mathbb{N})$ *är ändligt.*

Ett filter som inte kan utvidgas kallas ultrafilter.

Definition 7. Ultrafilter. *Ett filter på* S *är ett ultrafilter* $\mathcal{U}$ *om det gäller att*

I: $\mathcal{U} \subseteq \mathcal{F} \Rightarrow \mathcal{U} = \mathcal{F}$

eller ekvivalent

II: För alla $A \subseteq S$ *gäller att antingen* $A \in \mathcal{U}$ *eller* $A^c \in \mathcal{U}$ *där* $A^c = S \setminus A$.

Ett ultrafilter är principialt om det finns ett element $s \in S$ *så att*

$$\mathcal{U} = \{A \subseteq S \mid s \in A\}$$

Det räcker att kontrollera ett av villkoren i definition 7, eftersom de är ekvivalenta. Filtret i exempel 3 är ett ultrafilter. Det är också principialt eftersom $\{2\} \subseteq S$. Fréchet-filtret i exempel 2 är inte ett ultrafilter och inte heller principialt.

2.2.1 Zorns Lemma

När de hyperreella talen konstrueras med ultraprodukter så används ett ultrafilter. Det kommer att visa sig att ultrafilter som innehåller Fréchet-filtret är särskilt användbara. För att bevisa existensen av lämpliga ultrafilter så används man Zorns lemma, som kan härledas ur urvalsaxiomet. Det görs som nedan men mer utförligt av i Goldblatts bok [3].

Definition 8. Zorns Lemma *Låt* $(P, \leq)$ *vara en partiellt ordnad mängd. Om varje totalt ordnad delmängd i* P *har en övre begränsning i* P *så innehåller* P *ett maximalt element.*

Med Zorns lemma kan man bevisa att det existerar ett ultrafilter $\mathcal{U}$ på $\mathbb{N}$ så att Fréchet-filtret $\mathcal{F}_0$ på $\mathbb{N}$ är en delmängd av $\mathcal{U}$. Med det som antagande kan man bevisa följande.

Sats 2. *Det finns ett icke-principialt ultrafilter på* $\mathbb{N}$.

Bevis. Enligt antagandet så har vi att $\mathcal{F}_0 \subseteq \mathcal{U}$ där $\mathcal{U}$ är ett ultrafilter på $\mathbb{N}$. För varje $n \in \mathbb{N}$ gäller att $\mathbb{N} \setminus \{n\} \in \mathcal{F}_0 \subseteq \mathcal{U}$. Eftersom inte både en delmängd och delmängdens komplement kan tillhöra $\mathcal{U}$ medför det att $\{n\} \notin \mathcal{U}$, vilket innebär att $\mathcal{U}$ inte är principialt. $\mathcal{U}$ är alltså icke-principialt. □

Sats 2 är viktig för att konstruktionen skall gå att genomföra.

2.3 Sammanfattning

I detta kapitel har en rad begrepp definierats och en viktig sats bevisats. De flesta definitionerna är relativt abstrakta och har långtgående och intressanta tillämpningar. Anledningen till att de presenterats här är dock för ett speciellt syfte ; de utgör verktyg för att konstruera de Hyperreella talen. Viktiga lärdomar att ha med sig från det här kapitlet är att *filterna är ekvivalensrelationer* [2] och att de därför *bildar de en partition* av mängden av alla reella talföljder (vilka kommer att användas som utgånspunkt för konstruktion av de Hyperreella talen - detta diskuteras närmare i senare kapitel) enligt sats

Satsen visar sig alltså vara viktig eftersom detta garanterar att ekvivalensklasserna är disjunkta och därmed blir de Hyperreella talen väldefinierade. Varför filtrerna ser ut som de gör kommer även det att förtydligas senare, men i korta drag beror det på att en för svag ekvivalensrelation "filtrerar inte bort" tillräckligt många talföljder. Vidare så kan en alltför stark ekvivalensrelation filtrera bort alldeles för många tal - och resultat blir bara de reella talen. Det visar sig att ultrafiltret är precis den ekvivalensrelation som behövs för att definiera den mängd av tal som önskats; de Hyperreella talen ${}^*\mathbb{R}$.

[2] Filtrerna är ekvivalensrelationer på mängden av alla reella talföljder för att vara exakt.

3 Konstruktion av $\mathbb{R}$ från $\mathbb{Q}$

I detta kapitel kommer ett sätt att definiera de reella talen med utgångspunkt i de rationella presenteras. Vägen från Peanos Axiom till $\mathbb{Q}$ utelämnas här, men den intresserade kan fördjupa sig i denna genom att läsa exempelvis "Standard and Nonstandard Analysis" av R.F. Hoskins, som även utgör en grund för detta avsnitt [11].

Varför just utvidgningen från $\mathbb{Q}$ till $\mathbb{R}$ tas med i denna rapport beror på två saker; dels är det betydligt svårare att utvidga $\mathbb{Q}$ till $\mathbb{R}$ än exempelvis $\mathbb{R}$ till $\mathbb{C}$. Vidare så är det intressant för våra ändamål då många tekniker som används för att definiera $\mathbb{R}$ på detta sätt kan modifieras och användas för att få fram de hyperreella talen från $\mathbb{R}$. Alltså är detta avsnitt inte bara av intresseväckande natur utan det syftar även till att introducera vissa frågeställningar och tankesätt som är av stor vikt i sökandet efter de hyperreella talen.

De reella talen går att definiera från de rationella med hjälp av Dedekindsnitt, men i detta kapitel kommer istället ett tillvägagångssätt med Cauchy-följder att tillämpas. Utgående från följder av rationella tal skall de reella skapas - tanken är att låta varje reellt tal motsvara en följd av rationella tal. En viktig frågeställning är dock "när är två följder lika?" det vill säga "vilken ekvivalensrelation skall användas?". Utmaningen ligger alltså i att skräddarsy en ekvivalensrelation som sorterar ut de reella talen från de rationella följderna, och samtidigt har de önskade egenskaperna som $\mathbb{R}$ ska ha. Låt oss inledningsvis titta närmare på följder av rationella tal.

3.1 Rationella talföljder

En talföljd kan ses som en funktion $f : \mathbb{N} \to M$, där M i fallet av en rationell talföljd är Q. Varje naturligt tal n avbildas alltså på ett rationellt tal q. Ofta skrivs *termerna* i talföljden med index n för att representera ett visst tal i följden. Exempelvis skrivs $f(4)$ som q_4. En hel följd kan följaktligen enkelt skrivas som: $(r_n)_{1\leq n<\inf} \equiv r_1,r_2,r_3,\dots,r_n,\dots$ Notera att funktionen inte behöver vara injektiv så termerna behöver nödvändigtvis inte vara olika; den reella följden $(\pi, \pi, \dots)$ är ett bra exempel på en talföljd som definieras av den konstanta avbildningen $f : \mathbb{N} \to \mathbb{R}$ där $f(n) \equiv r_n = \pi$, $\forall n \in \mathbb{N}$.

Begrepp som är viktiga att känna till för fortsättningen av detta kapitlet är konvergens, uppåt/nedåt begränsad, supremum och infimum. Läsaren förväntas bekant med dessa så närmare definitioner utelämnas, föreligger det ett behov av repetition så rekommenderas "Analys i en variabel" av A. Persson och L-C. Böiers eller Hoskins som nämndes i början av avsnittet.

3.1.1 Cauchyföljder

Ett mer stringent krav för att en talföljd skall konvergera än att den är begränsad är att den uppfyller olikheten:

$$|q_n - q_m| = |(q_n - q) + (q - q_m)| \leq |(q_n - q)| + |(q - q_m)|. \tag{4}$$

En talföljd som uppfyller detta sägs vara en *Cauchyföljd* eller en *fundamental* följd. Vidare säger ekvation (4) att q_m och q_n kommer godtyckligt nära varandra då m och n växer oberoende (det vill säga att differensen i vänsterledet går mot noll) om följden konvergerar.

Varje konvergent rationell talföljd är en Cauchyföljd, dock måste inte varje rationell Cauchyföljd konvergera mot ett rationellt tal. Betrakta till exempel följden

$$d_{n+1} = (d_n + 2/d_n)/2 \text{ , för } n \geq 1 \; d_1 = 1 \tag{5}$$

Då $n = 1, 2, 3, 4, \dots$ antar följden värden $1, 3/2, 17/2, 557/408(1.4142..), \dots$. Den uppmärksamma läsaren kanske noterar att denna följd alltså konvergerar mot $\sqrt{2}$, vilket som bekant inte är ett rationellt tal. Att följden faktiskt konvergerar mot just $\sqrt{2}$ går att argumentera för; antag att följden faktiskt konvergerar mot ett tal x. Om så är fallet så går det att approximera x godtyckligt väl genom att välja ett tillräckligt stort n alltså: $d_{n+1} \cong d_n \cong x, \;\; \Rightarrow 2d_n \cong (d_n + 2/d_n)$ eller $2d_n \cong (d_n^2 + 2)/d_n$, alltså erhålls $d_n^2 \cong 2$ vilket gör att vi kan sluta oss till att $x^2 = 2$.

3.2 Konstruktion av $\mathbb{R}$ med Cauchyföljder

I föregående avsnitt visade det sig att varje konvergent rationell talföljd är en Cauchyföljd, men att den inte nödvändigtvis behöver konvergera mot ett rationellt tal. Ofta betraktas de reella talen som punkter på tallinjen - varje punkt på linjen kan beskrivas av ett reellt tal. En Cauchyföljd f_n lever även den på tallinjen och när n växer, säg är större än något k, så kommer delföljden $f_n, n \geq k$ att inrymmas på ett allt mindre intervall på linjen. Rent intuitivt känns det som att termerna till slut kommer att hopa sig kring en punkt och att det finns ett tal tillskrivet denna punkt. Enligt tidigare kapitel räcker inte $\mathbb{Q}$ till för att beskriva alla punkter på tallinjen och detta faktum skvallrar om det som vi hela tiden haft i bakhuvudet; det behövs ett större talsystem för att ge en komplett beskrivning av tallinjen.

För att åstadkomma en utvidgning måste vi på något sätt göra vårt talsystem större. När detta väl är avklarat skall det skräddarsys för att endast innehålla precis det som vill uppnås. Låt till att börja med varje rationellt tal q motsvaras av en (oändlig) konstant följd ($q_n = q \ \forall n \in \mathbb{N}$). Dessa följder är medlemmar i den mycket större mängden av alla rationella följder, konvergenta eller ej. Denna följd betecknas Q^N, vilket skall symbolisera att en rationell talföljd kan ses som en avbildning $\mathbb{N} \to \mathbb{Q}$. När väl addition och multiplikation har definierats strikt har alltså ett nytt talsystem erhållits, som garanterat är större än det ursprungliga (det innehåller ju redan $\mathbb{Q}$). De rationella talen motsvars i Q^N av konstanta följder, följaktligen är de nya talen i Q^N de icke-konstanta följderna (som ju utgör en mycket större del av Q^N).

3.2.1 Definition av Q^Ns aritmetik

Låt addition av två element $\bar{r}, \bar{s} \in R^N$ definieras som termvis addition i följderna. Definiera multiplikation på samma sätt;

$$\bar{r} \oplus \bar{s} = (r_1 + s_1, r_2 + s_2, ..., r_k + s_k, ...) \tag{6}$$

$$\bar{r} \otimes \bar{s} = (r_1 s_1, r_2 s_2, ..., r_k s_k, ...) \tag{7}$$

Då addition och multiplikation definieras på detta sätt så är de både associativa och kommutativa - vidare så kommer de fungera precis som $+$ och $\cdot$ för $\mathbb{Q}$. Med det faktum att de rationella talen motsvaras av konstanta följder i åtanke definieras de neutral elementen för addition respektive multiplikation naturligt av:

$$\bar{0} = (0, 0, 0, ...) \tag{8}$$

$$\bar{1} = (1, 1, 1, ...) \tag{9}$$

3.3 Första ansats: svag ekvivalens

Det krävs ingen djupgående analys för att konstatera att Q^N är en alldeles för stor mängd för att vara användbar. Tag exempelvis följden $(0, 0, 1, 0, ..., 0, ...)$; det är uppenbart att denna följd inte är konstant 0 men onekligen så är följden väldigt lik den konstanta följden $< 0 >$. Nu är det dags att införa en fiffig ekvivalensrelation för att gruppera följder i ekvivalensklasser så att följder som är "tillräckligt lika" definieras som lika - det vill säga som medlemmar av samma ekvivalensklass och därmed motsvarar samma reella tal.

Som ett första försök att reducera Q^N till $\mathbb{R}$ införs följande ekvivalensrelation:

Definition 9. Svag Ekvivalens *Två oändliga följder (r_n) och (s_n) sägs vara ekvivalenta i svag mening om och endast om mängden $n \in \mathbb{N} : x_n = y_n$ är en kofinit delmängd av $\mathbb{N}$. Kofinit innebär är att mängdens komplement (med avseende på $\mathbb{N}$) är begränsat (finit - det vill säga att två följder bara skiljer sig på ett begränsat antal ställen).*

Ett exempel på två följder som är ekvivalenta i svag mening är $(r_n) = (1, 2, 3, 3, 3, 3,..)$ och $(s_n) = (9, 8, 7, 3, 3, 3, ...)$ då de endast skiljer sig åt på med ett begränsat antal termer (här de tre första). Vi säger att två talföljder som är ekvivalenta i svag bemärkelse är *lika nästan överallt.*

Uppenbarligen är mängden av alla ekvivalensklasser $[\bar{r}]$ på $\mathbb{Q}^{\mathbb{N}}$ mindre än $\mathbb{Q}^{\mathbb{N}}$, men är det tillräckligt och framförallt ger det en konsistent modell av de reella talen? Efter en kort granskning av den nya mängden, låt oss använda Hoskins notation och kalla den för ${}^{\#}\mathbb{Q}$, visar det sig att svaret är nej. I algebraiska termer visar det sig att ${}^{\#}\mathbb{Q}$ förvisso är en ring, men inte en kropp. Vidare har ${}^{\#}\mathbb{Q}$ nolldelare vilket inte är önskvärt. En explicit förklaring av denna problemställning följer nu.

Låt addition och multiplikation på ${}^{\#}\mathbb{Q}$ definieras av:

$$[\bar{x}] + [\bar{y}] = [\bar{z}] \iff n \in \mathbb{N} : z_n = x_n + y_n \text{ kofinit delmängd av } \mathbb{N} \tag{10}$$

$$[\bar{x}] \cdot [\bar{y}] = [\bar{z}] \iff n \in \mathbb{N} : z_n = x_n \cdot y_n \text{ kofinit delmängd av } \mathbb{N} \tag{11}$$

Betrakta de två följderna $r = (1, 0, 1, 0, ...)$ och $s = (0, 1, 0, 1, 0, ...)$. Ingen av dessa följder är ekvivalent med nollföljden (ty de skiljer sig i varannan term det vill säga uppräkneligt oändligt många gånger) icke desto mindre är den punktvisa produkten mellan de båda lika med nollföljden. Om följderna då betraktas som tal i den mängd som vi vill skall vara $\mathbb{R}$ så visar det sig att två "tal" skiljda från 0, har en produkt som är 0 - detta är inte sant för de reella talen och därför är inte ${}^{\#}\mathbb{Q}$ med dessa definitioner av addition och multiplikation en motsvarighet till $\mathbb{R}$. Vad som bör noteras är dock att mängden ${}^{\#}\mathbb{Q}$ har ett antal intressanta egenskaper vilka kan vara av intresse för andra ändamål, dessvärre ligger det utanför ramarna för det här arbetet.

Ytterligare problem med ${}^{\#}\mathbb{Q}$ är hur talen skall ordnas; hur skall följderna r och s från ovan jämföras? För att råda bot på problemen som tagits upp i detta kapitel (och vissa andra som inte nämnts här) krävs en annan typ av ekvivalensrelation - denna presenteras i nästa avsnitt.

3.4 Definition av de reella talen med Cauchy-Följder

Det visade sig att svag ekvivalens var ett otillräckligt villkor för att ta fram $\mathbb{R}$ från mängden ${}^{\#}\mathbb{Q}$. Inte nog med att nolldelare existerar, det finns inte heller något element $[(r_n)]^2 = [(2)]$ i ${}^{\#}\mathbb{Q}$. Nästa steg i sökandet på en definition av de reella talen blir att istället för att betrakta alla rationella följder istället inskränka sig till Cauchyföljderna i $\mathbb{Q}^{\mathbb{N}}$. Det är nu dags att presentera ekvivalensrelationen som är precis den som behövs; mycket stark ekvivalens.

Definition 10. Den mycket starka ekvivalensrelationen
Två rationella Cauchyföljder sägs vara ekvivalenta i väldigt stark bemärkelse (eller i Cauchybemärkelse) om $(\forall k \in \mathbb{N})(\exists n_0 \in \mathbb{N})$ *sådant att*

$$|r_n - s_m| < 1/k \quad \forall m \geq n_0, n \geq n_0 \tag{12}$$

Är detta uppfyllt skrivs det $(r_n) \sim (s_m)$.

Om (r_n) är en rationell sekvens som konvergerar mot ett rationellt tal r och $(r_n) \sim (s_m)$ så följer det direkt ur olikheten

$$|r_n - s_m| \leq |r - r_n| + |r_n - s_m| \tag{13}$$

att även (s_m) konvergerar mot r. Alltså kan Cauchyföljderna av rationella tal delas in i ekvivalensklasser; om två följder är ekvivalenta i väldigt stark bemärkelse så är de medlemmar i samma ekvivalensklass. Varje ekvivalensklass definierar ett reellt tal, nämligen det tal som serierna i ekvivalensklassen konvergerar mot. Vidare om serien $\mathbf{r_s}$ konvergerar mot ett rationellt tal r så säger vi att ekvivalensklassen definierar det *det rationella reella talet* r. Följaktligen är sekvenserna $a_n = (2n+1)/n$, $b_n = (2^n - 1)/2^{n-1}$ och $c_n = 2$ alla medlemmar av samma ekvivalensklass ; nämligen den som motsvarar det reella rationella talet 2. På samma sätt definieras $\sqrt{2}$ av varje Cauchyföljder som konvergerar mot $\sqrt{2}$, exempelvis $d_{n+1} = (d_n + 2/d_n)/2$.

3.4.1 De reella talens egenskaper

Nu när de "nya" talen är definierade så återstår det att definiera aritmetiken för dessa. Detta utgör ingen större utmaning; låt oss använda vanlig rationell aritmetik punktvis i Cauchyföljderna som utgör ekvivalensklasserna. Om a_n och b_n är är Cauchyföljder så kommer även $(a_n + b_n)$ och $a_n \cdot b_n$ att vara det. Det gäller även att om $a_n \sim a'_n$ och $b_n \sim b'_n$ så kommer även $(a_n + b_n) \sim (a'_n + b'_n)$ vara Cauchyföljder. Så om två reella tal $\mathbf{a}$ och $\mathbf{b}$ motsvaras av de rationella sekvenserna a_n och b_n så definieras naturligt summan och produkten av dessa som $\mathbf{a} + \mathbf{b} = (a_n + b_n)$ respektive $\mathbf{a} \cdot \mathbf{b} = (a_n \cdot b_n)$.

Ett problem med ${}^{\#}\mathbb{Q}$ som faktiskt inte togs upp i avsnitt 3.4 är att ${}^{\#}Q$, till skillnad från $\mathbb{R}$ inte har någon supremumegenskap; en delmängd har en övre begränsning men inte att det existerar *ett minsta tal* X s.a. $x \leq X$, $\forall x \in X$ det vill säga en minsta övre begränsning. Detta kallas som bekant för *supremum* av mängden. Egenskapen att det alltid går, givet två tal $x < y$, att hitta ett naturligt tal n sådant att $y < xn$ kallas för Arkimedisk egenskap. Detta enkla resultat leder till en del intressant egenskaper. Resultatet säger att till vilket rationellt tal som helst (eller reellt för den delen) så går det att hitta ett naturligt tal n som är större; $\mathbb{N}$ är en obegränsad delmängd av såväl $\mathbb{Q}$ som $\mathbb{R}$. Vidare bevisar detta att det inte finns några infinitesimalt små tal i exempelvis $\mathbb{Q}$, ty om det hade funnits ett sådant tal $0 < \epsilon < 1/n, \forall n \in \mathbb{N}$ så skulle talet $1/\epsilon = r$ vara en övre begränsning till $\mathbb{N}$. Som tidigare nämnts i detta stycke innehåller alltså inte heller $\mathbb{R}$ några infinitesimalt små tal. Icke desto mindre så skiljer sig $\mathbb{R}$ från $\mathbb{Q}$ i en väldigt viktig bemärkelse; nämligen supremumegenskapen - och det är denna egenskap som gör att "hålen" som $\mathbb{Q}$ lämnar på tallinjen kan fyllas ut (det vill säga att de irrationella talen kan beskrivas).

Ett snabbt motexempel visar att $\mathbb{Q}$ inte har någon supremumegenskap; betrakta mängden $\{r \in \mathbb{Q} : r^2 < 2\}$. Att mängden är uppåt begränsad, av exempelvis 100 eller $\frac{3}{2}$, är förvisso sant men det finns ingen minsta övre begränsning i $\mathbb{Q}$. Däremot hade motsvarande mängd av reella tal begränsats uppåt av $\sqrt{2}$.

Att visa att $\mathbb{R}$ faktiskt innehar en supremumegenskap går att visa, men tar ett tag. Då det ligger i periferin till denna uppsats så kommer inte ett helt komplett bevis att presenteras; dock följer här en överskådlig skiss över hur ett sådant skulle genomföras;

Sats 3. $\mathbb{R}$ *har en supremumegenskap.*

Bevis. (Skiss). Låt $A \subset \mathbb{R}$, $A \neq \emptyset$ och låt $a,b \in \mathbb{Q}$. Antag vidare att:

- $\exists x \in A$ sådant att $a < x$
- $b > x \quad \forall x \in A$

Eftersom $\mathbb{R}$ har en Arkimedisk egenskap och varje $(b - a)$ är ett positivt rationellt tal så gäller det att: $\exists m$ för varje n så att $(b - a) < m \cdot (1/n)$, där både m och n är naturliga tal. Detta medför att $b < a + \frac{m}{n} = M_n$. Alltså begränsas A uppåt av mängden

$$\{m \in \mathbb{N} : a + \frac{m}{n}\} \tag{14}$$

vilken är icketom. Således kommer denna mängd att ha ett minsta element m_n. Då går det att visa att följderna $x_n = a + \frac{m_n}{n}$ och $y_n = a + \frac{m_n - 1}{n}$ är ekvivalenta Cauchyföljder (det vill säga att de konvergerar mot samma tal - och detta tal är det sökta $sup(A)$). □

4 Icke-standardutvidgning av $\mathbb{R}$

De hyperreella talen, betecknade ${}^*\mathbb{R}$, är en utvidgning av de reella talen och kan, som tidigare nämnt, konstrueras på fler än ett sätt. Här nedan visas först hur konstruktionen går till med ultraprodukter och sedan hur den kan utföras axiomatiskt.

4.1 Konstruktion av ${}^*\mathbb{R}$ med ultraprodukter

Det vanligaste sättet att konstruera de hyperreella talen är att göra det med ultraprodukter. Konstruktionen genomförs här i enlighet med hur den genomförs i [3].

Tanken är att först förstora $\mathbb{R}$ till $\mathbb{R}^{\mathbb{N}}$ så att varje tal kan skrivas som en följd

$$r = \langle r_1, r_2, ... \rangle = \langle r_n : n \in \mathbb{N} \rangle \tag{15}$$

för att sedan reducera $\mathbb{R}^{\mathbb{N}}$ genom att definiera

$$^*\mathbb{R} = \mathbb{R}^{\mathbb{N}}/\mathcal{U} \tag{16}$$

där $\mathcal{U}$ är ett ultrafilter. Det vill säga att de hyperreella talen är ringkvoten av $\mathbb{R}^{\mathbb{N}}$ och ett ultrafilter på $\mathbb{R}^{\mathbb{N}}$. Målet är att det resulterande, reducerade, talsystemet $^*\mathbb{R}$ skall vara större än $\mathbb{R}$ och inkludera oändligt små och stora tal utan att förlora de egenskaper som de reella talen har. Först följer dock en diskussion kring $\mathbb{R}^{\mathbb{N}}$.

4.1.1 Om $\mathbb{R}^{\mathbb{N}}$

En avbildning $f : \mathbb{N} \to \mathbb{R}$ utvidgar $\mathbb{R}$ till $\mathbb{R}^{\mathbb{N}}$ så att varje tal kan skrivas som en följd som i (15). Multiplikation och addition definieras som

$$r \oplus s = \langle r_n + s_n : n \in \mathbb{N}\rangle$$
$$r \otimes s = \langle r_n * s_n : n \in \mathbb{N}\rangle$$

$\mathbb{R}^{\mathbb{N}}$ är en ring men inte en kropp. En kropp kräver att alla tal förutom noll har en invers. Att så inte är fallet följer av att det existerar nolldelare, till exempel

$$\langle 0, 1, 0, 1, ...\rangle * \langle 1, 0, 1, 0, ...\rangle = \langle 0, 0, 0, 0, ...\rangle$$

Nolldelare har ingen invers för om $ab = 0$ där a är en inverterbar nolldelare och b ett tal skilt från noll så är

$$0 = a^{-1}0 = a^{-1}ab = b$$

vilket är en motsägelse.

$\mathbb{R}^{\mathbb{N}}$ är en partiellt ordnad mängd givet att

$$r \leq s \iff s_n \leq t_n \ , \ n \in \mathbb{N}$$

Ett reellt tal r representeras av en konstant följd $\langle r : n \in \mathbb{N}\rangle$. I enlighet med detta kan ringens neutrala element avseende multiplikation och addition skrivas som

$$\mathbf{1} = \langle 1, 1, 1, 1, ...\rangle$$
$$\mathbf{0} = \langle 0, 0, 0, 0, ...\rangle$$

Det är alltså de icke-konstanta följderna som skiljer $\mathbb{R}^{\mathbb{N}}$ ifrån $\mathbb{R}$. Detta medför att vi nu, efter utvidgningen, har ett för stort talsystem. Det saknas ett bra sätt att jämföra följder. Även en följd som

$$r = \langle 1, 0, 0, 0, ...\rangle$$

är skild från noll i $\mathbb{R}^{\mathbb{N}}$. Det var inte tanken utan det vore lämpligare om alla följder som skiljer sig på ett ändligt antal platser är lika. Det är problematiskt att $\mathbb{R}^{\mathbb{N}}$ är en partiellt ordnad mängd och inte en totalt ordnad mängd. Som det är nu kan vi inte avgöra huruvida

$$\langle 0, 1, 0, 1, ...\rangle \leq \langle 1, 0, 1, 0, ...\rangle \text{ eller } \langle 0, 1, 0, 1, ...\rangle \geq \langle 1, 0, 1, 0, ...\rangle \tag{17}$$

och alltså kommer ett filter som tar hänsyn till ovanstående problem att väljas.

4.1.2 Ultrafiltret

Att flera ultrafilter $\mathcal{U}$ på $\mathbb{N}$ existerar så att (16) går att genomföra ges av Zorns Lemma. Beroende på vilket ultrafilter man väljer så erhåller man olika kroppar av hyperreella tal. Det gäller att $\mathcal{U}$ är icke-principialt, det innehåller alltså inga ändliga mängder (följder). För ett ultrafilter på $\mathbb{N}$ så tillhör en delmängd eller delmängdens komplement filtret men inte både och. Fréchet-filtret (komplementet till alla ändliga mängder) tillhör alltså $\mathcal{U}$.

Med hjälp av ett sådant filter kan vi definiera en ekvivalensrelation mellan två följder. Vi säger att

$$\langle r_n \rangle \sim \langle s_n \rangle \Longleftrightarrow \{n \in \mathbb{N} : r_n = s_n\} \in \mathcal{F} \tag{18}$$

där $\mathcal{F}$ är ett sådant ultrafilter som beskrevs. Innebörden är att alla följder som skiljer sig i ett ”få” antal termer är lika, där den precisa betydelsen av ”få” är att mängden inte tillhör $\mathcal{F}$. Om $\langle r_n \rangle \sim \langle s_n \rangle$ kan man säga att $r_n = s_n$ för nästan alla n. På samma sätt är till exempel

$$\langle r_n \rangle \leq \langle s_n \rangle \Longleftrightarrow \{n \in \mathbb{N} : r_n \leq s_n\} \in \mathcal{F}$$

vilket hanterar svårigheten i (17). Om $A = \{1, 0, 1, 0, ...\}$ så gäller att antingen $A \in \mathcal{F}$ eller $A^c \in \mathcal{F}$. I det första fallet är $\langle 0, 1, 0, 1, ... \rangle \leq \langle 1, 0, 1, 0, ... \rangle$ och i det andra tvärtom. Huruvida det är A eller A^c som tillhör $\mathcal{F}$ beror på vilket ultrafilter vi valt.

4.1.3 Vad hade hänt om ultrafiltret var principialt?

Ovan så ställer vi krav på att ultrafiltret skall vara icke-principialt. Det är ett måste för om det hade funnits ett tal x så att

$$\mathcal{U} = \{A \subseteq \mathbb{N} \mid x \in A\}$$

så hade följderna

$$\langle s_n \rangle = \langle s_1, s_2, s_3, s_4, ... \rangle$$
$$\langle s_x \rangle = \langle s_x, s_x, s_x, s_x, ... \rangle$$

haft åtminstone en gemensam term, för $n = x$. För filtret gäller också att

$$B \in \mathcal{U} \text{ om } \{x\} \subseteq B \subseteq \mathbb{N}$$

Så att alla mängder innehållandes x finns i filtret. Eftersom det också finns minst en term som överensstämmer så gäller att $\langle s_n \rangle \sim \langle s_x \rangle$. Notera att konstanta följder, så som $\langle s_x \rangle$, motsvarar reella tal. Alla följder $\langle s_n \rangle$ är alltså lika med ett reellt tal när ett principialt ultrafilter används och ingen utvidgning kan åstadkommas.

4.1.4 De hyperreella talen

Det återstår nu att konstruera de hyperreella talen utifrån den tidigare givna definitionen (16) och att visa att ${}^*\mathbb{R}$ har de egenskaper som vi sökte.

Ekvivalensklassen för en följd $\langle r_n : n \in \mathbb{N} \rangle$ med avseende på (18) definieras som

$$[r] = \{s \in \mathbb{R}^{\mathbb{N}} : r \sim s\}$$

och är mängden av alla följder som är lika med r. De hyperreella talen är mängden av alla sådana ekvivalensklasser.

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

Operationerna på $^*\mathbb{R}$ definieras i likhet med hur vi har gjort tidigare, till exempel

$$[r] + [s] = [\langle r_n + s_n \rangle]$$

$$[r] < [s] \iff \{n \in \mathbb{N} : r_n < s_n\} \in \mathcal{F} \tag{19}$$

Man kan visa att $\langle ^*\mathbb{R}, +, \cdot, < \rangle$ utgör en totalt ordnad kropp och att $^*\mathbb{R}$ innehåller oändligt små och stora tal. Om vi låter $\varepsilon = \langle 0, \frac{1}{2}, \frac{1}{3}, ... \rangle$ ger (19) att $[\mathbf{0}] < [\varepsilon]$. Men för ett reellt tal $r = \langle r, r, r, ... \rangle$, oavsett hur litet, är $[\varepsilon] < [r]$. På samma sätt är $\Omega = \varepsilon^{-1} = \langle 1, 2, 3, ... \rangle$ ett oändligt stort tal då $[\Omega] > [r]$ för alla r. Eftersom varken $[\Omega]$ eller $[\varepsilon]$ motsvaras av ett reellt tal så gäller att $[\Omega], [\varepsilon] \in {}^*\mathbb{R} \setminus \mathbb{R}$ vilket visar att vi utvidgat $\mathbb{R}$.

4.2 Axiomatisk utvidgning av $\mathbb{R}$

Vi har visat hur man gör en konkret utvidgning av de reella talen till de hyperreella talen med hjälp av ultrafilter, på ett liknande sätt som man utvidgar rationella talen till reella talen med hjälp av Cauchy-följder. En konkret utvidgning är viktig eftersom det bevisar existensen av hyperreella tal, och det kan underlätta förståelsen. Man får dock inte glömma att hyperreella tal är just *tal*, och på samma sätt som man ofta inte tänker på hur reella tal är konstruerade (till exempel Cauchyföljder eller Dedekindsnitt), så går det att bortse från själva konstruktionen då talen används. Det enda viktiga att veta är vilka operationer man kan utföra på talen och vilka algebraiska lagar de uppfyller. Detta är det axiomatiska tillvägagångssättet. Satser, bevis och definitioner i detta avsnitt följer förfarandet i [1].

Målet är att utöka $\mathbb{R}$ till en ny kropp $^*\mathbb{R}$ så att $\mathbb{R} \subset {}^*\mathbb{R}$. Vi kräver även att $^*\mathbb{R}$ ska vara ordnat, dvs. elementen ska kunna jämföras (detta gäller inte för exempelvis de komplexa talen som också är en utvidgning av $\mathbb{R}$). Starkare kommer vi kräva att *nästan alla* egenskaper, funktioner och relationer över $\mathbb{R}$ kommer ha en motsvarighet i $^*\mathbb{R}$ (detta kommer definieras precist snart). Observera att relationer kan uttryckas som mängder, till exempel kan vi göra en mängd $A \subset \mathbb{R}^2$ med alla par (x, y) där $x < y$, så att $x < y \iff (x, y) \in A$. Även funktioner kan beskrivas med mängder, vi kan till exempel beskriva en funktion $f : \mathbb{R} \mapsto \mathbb{R}$ med dess graf, dvs. alla punkter (x, y) där $y = f(x)$. Eftersom både funktioner och relationer ska ha en utökad motsvarighet, kommer definitionen av icke-standardutvidgning innefatta inte bara $\mathbb{R}$, utan alla delmängder till kartesiska produkter $\mathbb{R}^n$.

Definition 11. Icke-standardutvidgning av en mängd. *Låt $\mathbb{X}$ beteckna en icke-tom mängd. Då är en icke-standardutvidgning av $\mathbb{X}$ en avbildning som avbildar varje $A \subseteq \mathbb{X}^m$, $m \geq 0$ på $^*A$ så att $^*\mathbb{X}$ är icke-tom och följande gäller för alla $m, n \geq 0$:*

1. *Utvidgningen är distributiv över mängdoperationer. Om $\mathbb{A}, \mathbb{B} \subseteq \mathbb{X}^m$, gäller*

   $^*\mathbb{A}, {}^*\mathbb{B} \subseteq ({}^*\mathbb{X})^m$
   $^*(\mathbb{A} \cap \mathbb{B}) = {}^*\mathbb{A} \cap {}^*\mathbb{B}$
   $^*(\mathbb{A} \cup \mathbb{B}) = {}^*\mathbb{A} \cup {}^*\mathbb{B}$
   $^*(\mathbb{A} \setminus \mathbb{B}) = {}^*\mathbb{A} \setminus {}^*\mathbb{B}.$

2. *Utvidgningen bevarar elementära basdiagonaler. Om $1 \leq i < j \leq m$ och*

   $$\Delta = \{(x_1, ..., x_m) \in \mathbb{X}^m \mid x_i = x_j\}$$

   så är

   $$^*\Delta = \{(x_1, ..., x_m) \in ({}^*\mathbb{X})^m \mid x_i = x_j\}.$$

3. *Utvidgningen bevarar kartesiska mängdprodukter. Om $\mathbb{A} \subseteq \mathbb{X}^m$ och $\mathbb{B} \subseteq \mathbb{X}^n$ så är $^*(\mathbb{A} \times \mathbb{B}) = {}^*\mathbb{A} \times {}^*\mathbb{B}$.*

4. *Utvidgningen bevarar projektioner där sista koordinaten utelämnats. Låt π beteckna denna projektion, dvs.*

   $$\pi\left((x_1, x_2, ..., x_n, x_{n+1})\right) = (x_1, x_2, ..., x_n).$$

   För $\mathbb{A} \subseteq \mathbb{X}^{n+1}$ gäller då

   $$^*(\pi(A)) = \pi({}^*A).$$

Dessa regler är ganska självklara för vad vi förväntar oss av ${}^*\mathbb{R}$. Första reglerna säger till exempel att Booleska operationer är bevarade, detta är ett krav om vi ska kunna föra över relationer på ett vettigt sätt.

Det är inte uppenbart i nuläget att denna definition kommer leda någonstans, hur vet vi ens att vi fått med infinitesimaler och oändligt stora tal i vår utökning? Det kommer visa sig att denna definition räcker väldigt långt, men först visar vi hur denna definition är tillräcklig (och nödvändig) för att bevisa några enkla följdsatser.

Sats 4. *${}^*\emptyset = \emptyset$.*

Bevis. Använd (1) i definition 11

$$ {}^*\emptyset = {}^*\left(\emptyset \setminus \emptyset\right) = {}^*\emptyset \setminus {}^*\emptyset = \emptyset. $$

□

Sats 5. *Om $\mathbb{A} \subseteq \mathbb{X}^m$ är icke-tom är även ${}^*\mathbb{A}$ icke-tom.*

Bevis. För enkelhets skull visar vi det för $m = 2$, beviset är analogt för större m. Vi använder projektionerna π och eftersom $\mathbb{A} \subseteq \mathbb{X}^2$ är icke-tom kan vi skriva $\mathbb{X} = \pi_2\left(\pi_3(\mathbb{X} \times \mathbb{A})\right)$ Nu använder vi (3) och (4) i definitionen, så

$$ {}^*\mathbb{X} = \pi_2\left(\pi_3({}^*\mathbb{X} \times {}^*\mathbb{A})\right). $$

${}^*\mathbb{X}$ är icke-tom per definition, så ${}^*\mathbb{A}$ måste också vara icke-tom för annars

$$ {}^*\mathbb{X} = \pi_2\left(\pi_3({}^*\mathbb{X} \times \emptyset)\right) = \pi_2\left(\pi_3(\emptyset)\right) = \emptyset \neq {}^*\mathbb{X}. $$

□

Sats 6. *Låt $\mathbb{A} \subseteq \mathbb{X}^m$. Då gäller*

$$ \mathbb{A} = \mathbb{B} \iff {}^*\mathbb{A} = {}^*\mathbb{B}. $$

Bevis. Uppenbart har vi

$$ \mathbb{A} = \mathbb{B} \Rightarrow {}^*\mathbb{A} = {}^*\mathbb{B} $$

det är andra hållet som behöver en förklaring. Om ${}^*\mathbb{A} = {}^*\mathbb{B}$ använder vi (1) i definitionen och får

$$ {}^*\left(\mathbb{A} \setminus \mathbb{B}\right) = {}^*\mathbb{A} \setminus {}^*\mathbb{B} = \emptyset. $$

Nu ger sats 5

$$ {}^*\left(\mathbb{A} \setminus \mathbb{B}\right) = \emptyset \Rightarrow \mathbb{A} \setminus \mathbb{B} = \emptyset. $$

På samma sätt implicerar ${}^*\mathbb{A} = {}^*\mathbb{B}$ att $\mathbb{B} \setminus \mathbb{A} = \emptyset$, och tillsammans ger det

$$ {}^*\mathbb{A} = {}^*\mathbb{B} \Rightarrow \mathbb{A} = \mathbb{B}. $$

□

Sats 7. *Om $x \in \mathbb{X}$ har ${}^*\{x\}$ exakt ett element.*

Bevis. Vi använder oss av basdiagonalerna

$$ \Delta = \{(u, u) \mid u \in \mathbb{X}\} $$

och har uppenbart

$$ \{x\} \times \{x\} = \{(x, x)\} \subseteq \Delta. $$

Nu använder vi (2) och (3) i definitionen och får

$$ {}^*\{x\} \times {}^*\{x\} = {}^*\left(\{x\} \times \{x\}\right) = {}^*\{(x, x)\} \subseteq {}^*\Delta = \{(u, u) \mid u \in {}^*\mathbb{X}\}. $$

Om nu ${}^*\{x\}$ innehåller mer än ett element, har vi två skilda $a, b \in {}^*\{x\}$, och då skulle ${}^*\{x\} \times {}^*\{x\}$ innehålla (a, b) vilket inte tillhör ${}^*\Delta$. Därmed har ${}^*\{x\}$ max ett element, och enligt sats 5 innehåller den åtminstone ett element. □

Med dessa satser kan vi definiera en avbildning av $\mathbb{X}$ på ${}^*\mathbb{X}$.

Definition 12. *Låt ${}^*x$ beteckna det unika elementet i ${}^*\{x\}$ för alla $x \in \mathbb{X}$.*

Enligt sats 6 får vi ${}^*x = {}^*y \iff x = y$. Det är denna avbildning som låter oss kalla $\mathbb{R}$ en delmängd av ${}^*\mathbb{R}$. Vi kallar ett element i ${}^*\mathbb{R}$ standard om det är av formen ${}^*x$ för något $x \in \mathbb{R}$.

Som snart kommer visas med hjälp av transfer, kan alla vanliga operationer och relationer för de reella talen utökas till de hyperreella talen, så att ${}^*\mathbb{R}$ blir en ordnad kropp med $\mathbb{R}$ som delmängd. Därmed kan vi definiera följande.

Definition 13. *Låt $x \in {}^*\mathbb{R}$, vi kallar*

1. *x infinitesimal om $|x| < \epsilon$ för alla positiva $\epsilon \in \mathbb{R}$,*
2. *x ändlig om $|x| < r$ för något reellt $r \in \mathbb{R}$,*
3. *x obegränsad om den inte är ändlig,*
4. *$x \simeq y$ om $x - y$ är infinitesimal.*

Observera att 0 är en infinitesimal (den enda reella), och att alla infinitesimaler är ändliga.

Sats 8. Standarddelsatsen. *Om $x \in {}^*\mathbb{R}$ är ändlig finns ett unikt $r \in \mathbb{R}$ så att $x \simeq r$, dvs. x har en unik representation $x = r + \sigma$, där σ är infinitesimal.*

Bevis. Låt $\mathbb{A}$ vara den delmängd av $\mathbb{R}$ som är mindre eller lika med x, denna mängd är begränsad uppåt eftersom x ändlig så

$$r = \sup \mathbb{A} = \sup\{a \leq x \mid a \in \mathbb{R}\}$$

existerar. Observera att trots detta kan $r > x$. Om vi låter ϵ beteckna ett positivt reellt tal, har vi att $r - \epsilon$ är ett reellt tal som därmed omöjligt kan vara den övre gränsen för $\mathbb{A}$, så det existerar $a \in \mathbb{A}$ med $r - \epsilon < a \leq x$, och därmed $-\epsilon < x - r$. Samtidigt har vi att $r + \epsilon$ är ett reellt tal större än r, så det kan omöjligt vara medlem i $\mathbb{A}$ eftersom r är den övre gränsen, och därmed $x < r + \epsilon$, eller $x - r < \epsilon$. Följaktligen

$$-\epsilon < x - r < \epsilon$$

så $\sigma = x - r$ är per definition infinitesimal. Det kvarstår att visa unikheten, anta att

$$x = r_1 + \delta_1 = r_2 + \delta_2.$$

Då är $r_1 - r_2 = \delta_2 - \delta_1$ en reell infinitesimal, och därmed lika med noll. □

Definition 14. *Det unika reella talet $r \simeq x$ för ändliga x, kallas* **skuggan** *eller* **standarddelen** *av x, och betecknas $r = st(x) = {}^o x$.*

Observera att vi ännu inte bevisat existensen av vare sig obegränsade eller infinitesimala tal i ${}^*\mathbb{R}$.

Sats 9. *Om ${}^*\mathbb{R}$ är en äkta utökning (${}^*\mathbb{R} \setminus \mathbb{R} \neq \emptyset$) innehåller ${}^*\mathbb{R}$ infinitesimaler skilda från noll och obegränsade tal.*

Bevis. Låt $x \in {}^*\mathbb{R} \setminus \mathbb{R}$. Om x ändlig, låt $\sigma = x - {}^o x$. Då är σ en infinitesimal skild från noll, och $1/\sigma$ är ett obegränsat tal. Om däremot x är obegränsat, så är $1/x$ en infinitesimal skild från noll. □

5 Transfer

Transferprincipen säger vilka påståenden om de reella talen som är överförbara till de hyperreella. För att utvidgningen av $\mathbb{R}$ ska vara användbar så måste man veta vilka egenskaper som kan överföras från $\mathbb{R}$ till $^*\mathbb{R}$ och tvärtom. Det är denna överföring som kallas transfer. Utvidgningen $^*\mathbb{R}$ uppfyller inte alla egenskaper som $\mathbb{R}$ har. Transferprincipen ger som sagt oss en metod för att urskilja de överförbara egenskaperna, men i detta kapitel dock endast de egenskaper som kan formuleras i det första ordningens logiska språk som vi begränsar oss till. Det finns generellare versioner av transfer som gäller för högre ordningens objekt vilket behövs för mer komplexa tillämpningar än de som tas upp här.

I denna framställning begränsar vi oss inte till $\mathbb{R}$, utan tittar på en mängd $\mathbb{X}$ och dess utvidgning $^*\mathbb{X}$ som uppfyller de axiom som introducerats i ett tidigare kapitel. Först måste det logiska språket som kommer användas definieras eftersom det i denna version av transfer endast är korrekt formulerade satser i detta språk som kan överföras. Därefter går vi igenom den transform som ligger till grund för transfer den så kallade *-transformen som tar oss från en formel över $\mathbb{R}$ till en *-transformerad formel över $^*\mathbb{R}$. Satser, bevis och definitioner i detta avsnitt följer [1].

5.1 Logik

Definition 15. Logiska konnektiv och kvantifikatorer.

$\neg$ *negation - "icke"*

$\wedge$ *konjunktion - "och"*

$\vee$ *disjunktion - "eller"*

$\rightarrow$ *implikation - "medför"*

$\leftrightarrow$ *ekvivalens - "om och endast om"*

$\exists$ *existenskvantifikatorn - "det finns"*

$\forall$ *allkvantifikatorn - "för alla"*

En formel i ett logiskt språk är en utsaga som kan innehålla variabler, både bundna och fria. Givet variablerna x och y, där $x \in A$ och $y \in B$, låt $\varphi(x,y)$ vara en formel, dvs. ett villkor på (x,y) som definierar en delmängd Φ av $A \times B$. Man skiljer på bundna och fria variabler, bundna variabler är av formen $\exists x \in A$ och $\forall y \in B$, de är bundna av en kvantifikator fria variabler däremot är fria för substitution. Om man tittar på följande formel $(\forall x \in A) \vee (y \leq x)$ så ser man att x är en bunden variabel och att y är en fri variabel eftersom x är inom räckvidden för en kvantifikator i detta fall $\forall$. I formeln $(x < 1) \vee (\forall x \in A) \vee (y \leq x)$ så är den första instansen av x fri eftersom den inte är i kvantifikatorns räckvidd och den andra bunden. Räckvidden för en kvantifikator är för en formel $(\forall x \in A)\varphi(x)$ eller $(\exists x \in A)\varphi(x)$ alla förekomster av x i delformeln φ med undantag av de instanser av x som i sin tur tillhör en delformel φ' av φ på formen $(\forall x \in A')\varphi'(x)$ eller $(\exists x \in A')\varphi'(x)$. Givet en formel $\varphi(x_1, \ldots, x_n)$ menar man med $x_1, \ldots, x_n$ alla fria distinkta variabler som ingår i formeln och de bundna variablerna ingår inte i dessa.

Med satser avses formler som inte har några fria variabler det vill säga endast bundna eller inga alls. En sats kan betecknas φ och är antingen sann eller falsk oberoende av några variabler till skillnad från en formel. Relationer så som $<$, $\leq, =$ kan alla uttryckas med mängder, till exempel låt C utgöra mängden av ordnade par (x,y) sådana att $x < y$ då kan man skriva $x < y$ som $(x,y) \in C$ där C utgör en delmängd av $A \times B$. En atomär formel, byggstenen för mer komplexa formler utgörs av $(x_1, \ldots, x_n) \in B$, för en mängd B där $x_1, \ldots, x_n$ är konstanter eller variabler. Atomer innehåller inga logiska konnektiv eller kvantifikatorer.

För att kunna använda sig av transferprincipen på ett korrekt sätt, behövs en mer precis definition av vad en formel är. Givet en formel med variabler som varierar över mängden $\mathbb{X}$ kallas den en formel över $\mathbb{X}$. De formler som definieras nedan är första ordningens formler vilket innebär att kvantifikatorer går över $\mathbb{X}$, ($\forall x \in \mathbb{X}$, $\exists x \in \mathbb{X}$), och inte över delmängder av $\mathbb{X}$ eller andra typer av högre ordningens objekt som är baserade på $\mathbb{X}$.

Definition 16. Formler över mängden $\mathbb{X}$. *För en formel över mängden $\mathbb{X}$ varierar variablerna i formeln över $\mathbb{X}$ och om variablerna kallas för $x_1, \ldots, x_m, y_1, \ldots, y_n$ och x så gäller följande:*

1. *För varje mängd $A \subseteq \mathbb{X}^m$, så är uttrycket $(x_1, \ldots, x_m) \in A$ en atomär formel över $\mathbb{X}$.*
2. *För varje funktion $f : A \to B$, där $A \subseteq \mathbb{X}^m$ och $B \subseteq \mathbb{X}^n$ så är $f(x_1, \ldots, x_m) = (y_1, \ldots, y_n)$ en formel över $\mathbb{X}$, vilket kan skrivas som $(x_1, \ldots, x_m, y_1, \ldots, y_n) \in \Gamma$, där Γ är grafen till f.*
3. *Om $\varphi(x_1, \ldots, x_m, y_1, \ldots, y_n)$ är en formel över $\mathbb{X}$, och om man tar $a_1, \ldots, a_n \in \mathbb{X}$ så är även $\varphi(x_1, \ldots, x_m, a_1, \ldots, a_n)$ en formel över $\mathbb{X}$.*
4. *Om φ och ψ är formler över $\mathbb{X}$ så är även $\neg\varphi, \varphi \wedge \psi, \varphi \vee \psi, \varphi \to \psi, \varphi \leftrightarrow \psi, \exists x\varphi$ och $\forall x\psi$ formler över $\mathbb{X}$.*

På så sätt kan mer komplexa formler byggas upp av enklare atomära formler. För att se kopplingen mellan formler och de axiom som gäller för icke-standardutvidgningen av en mängd så är det viktigt att inse att en formel definierar en mängd, låt x $\in A$ och $y \in B$ då definierar $\varphi(x,y)$ en delmängd av $A \times B$, bestående av par $(x,y) \in A \times B$ för vilka $\varphi(x,y)$ är sann.

Givet $\varphi(x,y)$ som definierar en mängd Φ och $\psi(x,y)$ som definierar en mängd Ψ, $x \in A$ och $y \in B$, så gäller följande:

$\neg\varphi(x,y)$ är komplementet till Φ,

$\varphi(x,y) \wedge \psi(x,y)$ motsvarar $\Phi \cap \Psi$,

$\varphi(x,y) \vee \psi(x,y)$ motsvarar $\Phi \cup \Psi$,

$\exists x\varphi(x,y)$ definierar projektionen $\pi(\Phi)$, där $\pi(x,y) = y$ är projektionen på den andra koordinaten,

$\forall y\varphi(x,y)$ definierar $\{x \in A | \{x\} \times B \subseteq \Phi\}$.

Dessa samband kommer användas i beviset utav transferprincipen då vi med hjälp av axiomen visar hur egenskaper kan överföras från $\mathbb{X}$ till $^*\mathbb{X}$ och tvärtom.

5.2 *-transform

Givet en formel över $\mathbb{X}$ så fås via *-transformen motsvarande *-transformerade formel över $^*\mathbb{X}$. Med $^*\mathbb{X}$ avses den icke-standardutvidgning av en mängd $\mathbb{X}$ som definierades i föregående kapitel.

Definition 17. *-transform av en formel över $\mathbb{X}$. *$^*\mathbb{X}$ är en icke-standardutvidgning av mängden $\mathbb{X}$, om $\varphi(x_1, \ldots, x_m, a_1, \ldots, a_n)$ är en formel över $\mathbb{X}$, där $a_1, \ldots, a_n \in \mathbb{X}$ så är dess *-transform $^*\varphi(X_1, \ldots, X_m, ^*a_1, \ldots, ^*a_n)$, där $X_1, \ldots, X_m$ är variabler med värden i $^*\mathbb{X}$. Låt φ och ψ vara formler över $\mathbb{X}$ då gäller följande:*

1. *För en atomär formel $(x_1, \ldots, x_m) \in A$, där $A \subseteq \mathbb{X}^m$ så är dess *-transform $(X_1, \ldots, X_m) \in {}^*A$.*
2. *Om $\varphi(x_1, \ldots, x_m, y_1, \ldots y_p)$ är en atomär formel och $a_1, \ldots, a_p \in \mathbb{X}$ så är dess *-transform $^*\varphi(X_1, \ldots, X_m, ^*a_1, \ldots ^*a_p)$.*
3. *För de logiska operatorerna och kvantifikatorerna som bygger upp formler gäller följande, där x är en variabel över $\mathbb{X}$:*

   $^*(\neg\varphi) = \neg^*\varphi$

$$^*(\varphi \vee \psi) = {}^*\varphi \vee {}^*\psi$$
$$^*(\varphi \wedge \psi) = {}^*\varphi \wedge {}^*\psi$$
$$^*(\exists x\varphi) = \exists X^*\varphi$$
$$^*(\forall x\varphi) = \forall X^*\varphi.$$

Därmed har vi definierat en transform som tar oss från egenskaper formulerade i formler över $\mathbb{X}$ till egenskaper formulerade i *-transformerade formler över $^*\mathbb{X}$. I nästa avsnitt visas att transferprincipen ger att en sats över $\mathbb{X}$ och dess *-transform över $^*\mathbb{X}$ är ekvivalenta, vilket gör att via *-transformen kan egenskaper i $\mathbb{X}$ överföras till motsvarande egenskaper i $^*\mathbb{X}$ och tvärtom.

5.3 Transferprincipen

Egenskaper som kan formuleras med satser definierade enligt ovan är överförbara mellan $\mathbb{R}$ och $^*\mathbb{R}$. För en sats φ över $\mathbb{X}$ så gäller det enligt transferprincipen att

$$\varphi \text{ sann i } \mathbb{X} \Leftrightarrow {}^*\varphi \text{ sann i } {}^*\mathbb{X}.$$

Sats 10. Transferprincipen. *Givet $\mathbb{X}$ en icke-tom mängd och $^*\mathbb{X}$, dess icke-standardutvidgning.*

1. *Låt $\varphi(x_1, ..., x_m)$ vara en formel över $\mathbb{X}$ och $^*\varphi(X_1, ..., X_m)$ dess *-transform. Låt $B \subseteq \mathbb{X}^m$ vara mängden definierad av $\varphi(x_1, ..., x_m)$:*

$$B = \{(x_1, ..., x_m) \in \mathbb{X}^m | \varphi(x_1, ..., x_m) \textit{ är sann i } \mathbb{X}\}. \tag{20}$$

   Därmed är $^*B$ mängden definierad av $^*\varphi(X_1, ..., X_m)$:

$$^*B = \{(X_1, ..., X_m) \in {}^*\mathbb{X}^m | {}^*\varphi(x_1, ..., x_m) \textit{ är sann i } {}^*\mathbb{X}\}$$

2. *Låt φ vara en godtycklig sats över $\mathbb{X}$ och låt $^*\varphi$ vara dess *-transform, då gäller följande*

$$\varphi \textit{ sann i } \mathbb{X} \Leftrightarrow {}^*\varphi \textit{ sann i } {}^*\mathbb{X}.$$

Beviset av transferprincipen sker via induktion över delarna som bygger upp formlerna. Först bevisas att del 1 i transfersatsen ovan medför del 2, sedan bevisas del 1 med hjälp av induktion. Nedan följer en skiss av hur induktionsbeviset ser ut.

Basfallet innebär att man bevisar att för en atomär formel φ_0 så gäller att $\varphi_0 \Leftrightarrow {}^*\varphi_0$. Antag sedan att transfer är sann för delar av formeln φ som är uppbyggda av atomer, logiska konnektiv och kvantifikatorer. Om två av dessa delar benämns för φ_k och φ_h, så visas att satsen även är sann för meningen (φ_k) (logisk konnektiv) (φ_h). Man visar också att om satsen gäller för $\varphi(x)$ så gäller den också för (logisk kvantifikator)x$\varphi(x)$. Därmed gäller transfer för hela formler uppbyggda av atomer, logiska kvantifikatorer och konnektiv. Om vi kallar en sådan formel för φ så gäller det att φ sann i $\mathbb{X} \Leftrightarrow {}^*\varphi$ sann i $^*\mathbb{X}$.

För att bevisa transferprincipen är ett par satser som följer av de axiom som presenterats i ett tidigare kapitel nödvändiga. Den första av dessa satser bevisas eftersom den är essentiell för förståelsen av varför transferprincipen gäller för basfallet i induktionsbeviset.

Sats 11. *Låt σ vara en godtycklig funktion från {1,...,m} till $\{1, ..., n\}$. Givet ett $A \in \mathbb{X}^m$ definiera*

$$B = \{(x_1, ..., x_n) \in \mathbb{X}^n \mid (x_{\sigma(1)}, ..., x_{\sigma(m)}) \in A\}.$$

Då gäller

$$^*B = \{(x_1, ..., x_n) \in ({}^*\mathbb{X})^n \mid (x_{\sigma(1)}, ..., x_{\sigma(m)}) \in {}^*A\}.$$

Bevis. För att göra notationen enklare så betraktas fallet då $m = 3$ och $n = 2$, då är $A \subseteq \mathbb{X}^3$. Definiera $B \subseteq \mathbb{X}^2$ och $C \subseteq \mathbb{X}^5$ som

$$B = \{(x_1,x_2) \in \mathbb{X}^2 \mid (x_2,x_1,x_2) \in A\}$$

och

$$C = \{(x_1,x_2,x_3,x_4,x_5) \in \mathbb{X}^5 \mid (x_3 = x_2) \wedge (x_4 = x_1) \wedge (x_5 = x_2) \wedge (x_3,x_4,x_5) \in A\}.$$

B är därmed resultatet av att projicera C på de två första koordinaterna. C är snittet mellan tre diagonala delmängder av $\mathbb{X}^5$ och mängden $\mathbb{X}^2 \times A$. Från Axiom 1-4 följer det att

$${}^*C = \{(x_1,x_2,x_3,x_4,x_5) \in {}^*\mathbb{X}^5 \mid (x_3 = x_2) \wedge (x_4 = x_1) \wedge (x_5 = x_2) \wedge (x_3,x_4,x_5) \in {}^*A\}.$$

$B = \pi(C)$ där π är projektionen på de två första koordinaterna, det följer av Axiom 4 att ${}^*B = {}^*(\pi(C)) = \pi({}^*C)$ därmed är

$${}^*B = \{(x_1,x_2) \in {}^*\mathbb{X}^2 \mid (x_2,x_1,x_2) \in {}^*A\}.$$

$\square$

Detta behövs som sagt för att visa transferprincipen för basfallet, det vill säga för atomer vilka som bekant kan skrivas på formen $(x_{\sigma(1)}, ..., x_{\sigma(m)}) \in A$ där A och σ är definierade som i sats 11.

Sats 12. Formler med konstanter. *Låt $A \in \mathbb{X}^{m+n}$ och $a = (a_1, \ldots, a_m) \in \mathbb{X}^m$. Definiera*

$$A(a) = \{(x_1, \ldots, x_n) \in \mathbb{X}^n \mid (a_1, \ldots, a_m, x_1, \ldots, x_n) \in A\}$$

och

$$({}^*A)({}^*a) = \{(x_1, \ldots, x_n) \in ({}^*\mathbb{X})^n \mid ({}^*a_1, \ldots, {}^*a_m x_1, \ldots, x_n) \in {}^*A\}.$$

Då gäller det att

$${}^*A(a) = ({}^*A)({}^*a).$$

Sats 12 behövs om en formel innehåller specifika element i $\mathbb{X}$ och inte bara variabler som varierar över $\mathbb{X}$.

Sats 13. Förstärkning av Axiom 4. *Axiom 4 gäller för alla projektioner π från m-tuplar till n-tuplar, där $n \leq m$ och π är definierad enligt följande*

$$\pi(x_1, \ldots, x_m) = (x_{i(1)}, \ldots, x_{i(n)})$$

där $1 \leq i(1) \leq \cdots \leq i(n) \leq m$. Detta innebär att givet $A \subseteq \mathbb{X}^m$ så gäller det att ${}^*(\pi(A)) = \pi({}^*A)$.

Denna förstärkning behövs om $\exists$ inte syftar på den sista variabeln i en formel $\varphi(x_1, \ldots, x_{n+1})$ eftersom det inte ger en projektion på de n första koordinaterna vilket är det fall som Axiom 4 behandlar. Vi behöver även känna till dualiteten mellan de logiska kvantifikatorerna, dvs. att $\forall$ kan skrivas med hjälp av $\exists$ och $\neg$ enligt följande

$$\forall y \varphi(y) \Leftrightarrow \neg \exists y \neg \varphi(y).$$

I beviset används även att $\rightarrow$ och $\leftrightarrow$ kan uttryckas som kombinationer av $\neg$, $\wedge$ och $\vee$. För implikation, $\rightarrow$, kan det göras enligt följande

$$\varphi(x,y) \rightarrow \psi(x,y)$$

vilket är det samma som

$$(\neg\varphi(x,y)) \vee \psi(x,y).$$

Med dessa satser och samband givna kan nu transferprincipen bevisas.

Bevis. Antag att φ är en sats över $\mathbb{X}$ och att ${}^*\varphi$ är dess *-transform. Låt A vara mängden definierad av φ enligt ekvation (20) och ${}^*A$ motsvarande för ${}^*\varphi$. Därmed är A antingen $\mathbb{X}^0$ eller $\emptyset$, beroende på om φ är sann eller ej i $\mathbb{X}$. ${}^*A$ är antingen $({}^*\mathbb{X})^0$ eller $\emptyset$, beroende på om ${}^*\varphi$ sann eller ej i ${}^*\mathbb{X}$. Axiom 1 ger att ${}^*\emptyset ={}^* (\emptyset \setminus \emptyset) ={}^* \emptyset \setminus{}^* \emptyset = \emptyset$ och Axiom 3 att ${}^*(\mathbb{X}^0) = ({}^*\mathbb{X})^0$. Detta tillsammans ger att om $A = \emptyset$ så måste också ${}^*A = \emptyset$, eller $A = \mathbb{X}^0$ och ${}^*A = ({}^*\mathbb{X})^0$, då följer att $\varphi \Leftrightarrow {}^*\varphi$. Därmed gäller det att del 1 av transfersatsen medför del 2. Då återstår det att visa del 1 av satsen.

Basfallet: Atomer, här benämnda $\varphi_0(x_1, ..., x_n)$, utgör basfallet, de kan skrivas på formen $(x_{\sigma(1)}, ..., x_{\sigma(m)}) \in A$ där $\{\sigma(1), ..., \sigma(m)\} \subseteq \{1, ..., n\}$ och $A \subseteq \mathbb{X}^m$. *-transformen av dessa formler är $(X_{\sigma(1)}, ..., X_{\sigma(m)}) \in {}^*A$. Låt $B = \{(x_1, ..., x_n) \in \mathbb{X}^n \mid (x_{\sigma(1)}, ..., x_{\sigma(m)}) \in A\}$, vilket är det samma som $B = \{(x_1, ..., x_n) \in \mathbb{X}^n \mid \varphi_0(x_1, \ldots, x_n)$ sann i $\mathbb{X}\}$. Sats 11 ger då att ${}^*B = \{(x_1, ..., x_n) \in ({}^*\mathbb{X})^n \mid (x_{\sigma(1)}, ..., x_{\sigma(m)}) \in {}^*A\}$ och del 1 av transfersatsen är därmed sann för atomära formler.

I formeln ovan kan variabler också inneha ett specifikt värde i $\mathbb{X}$ det vill säga $B(a) = \{(x_1, ..., x_n, a_1, \ldots, a_m) \in \mathbb{X}^n \mid (x_{\sigma(1)}, ..., x_{\sigma(m)}, a_1, \ldots, a_m) \in A\}$, av sats 12 följer att ${}^*(B(a)) = {}^*B({}^*a)$ och del 1 är sann i detta fall också.

Induktion: Givet två mängder C och D definierade enligt ekvation (20) av $\varphi_c(x_1, ..., x_n)$ respektive $\varphi_d(y_1, ..., y_m)$ för vilka del 1 av satsen antas vara sann. Dessa formler är uppbyggda av atomer, logiska konnektiv och kvantifikatorer. Vi vill visa att del 1 av satsen fortfarande gäller om man binder samman φ_c och φ_d med logiska konnektiv. $\varphi_c(x_1, ..., x_n) \wedge \varphi_d(y_1, ..., y_m)$ ger mängden $C \cap D$, enligt Axiom 1 så är ${}^*(C \cap D) = ({}^*C) \cap ({}^*D)$ eftersom del 1 av satsen antas gälla för C och D så gäller den därmed också för $C \cap D$. På samma sätt kan man använda Axiom 1 för de andra logiska konnektiven.

Tack vare dualiteten mellan all- och existenskvantifikatorn räcker det med att visa att del 1 av satsen är sann för existenskvantifikatorn. Givet formeln $\exists x_j \varphi(x_1, ..., x_n)$ för något $1 \leq j \leq n$, där $\varphi(x_1, ..., x_n)$ antas uppfylla del 1 och definiera mängden B enligt ekvation (20) så är $\exists x_j \varphi(x_1, ..., x_n)$ en projektion av mängden B definierad av $\varphi(x_1, ..., x_n)$ på koordinaterna $\{x_1, ..., x_n\} \setminus x_j$. Låt $\pi(B)$ beteckna denna projektion, enligt sats 13 gäller det att ${}^*\pi(B) = \pi({}^*B)$, därav är del 1 av satsen sann även för formler uppbyggda med kvantifikatorer.

$\square$

5.4 Tillämpningar av transfer

För att visa hur transferprincipen kan användas följer här ett par exempel på tillämpningar av principen först hur definitionen av kontinuerliga funktioner och likformigt kontinuerliga funktioner kan formuleras på icke-standardsidan, sedan hur riemannintegralen kan ses som en summa.

5.4.1 Kontinuerliga funktioner

Man kan med hjälp av ϵ, δ-definitionen av kontinuerliga funktioner nå en ekvivalent icke-standarddefinition via transferprincipen. I definitionen nedan har vi variabler som går över mängden $\mathbb{R}^+$ vilket är en delmängd av $\mathbb{R}$ detta är egentligen inte tillåtet i ett första ordningens logiskt språk, men detta är en förkortning för den tillåtna formuleringen $(\forall \epsilon \in \mathbb{R}) \wedge (\epsilon > 0)$.

Sats 14. Icke-standard kontinuitet. *För $f : \mathbb{R} \to \mathbb{R}$, f kontinuerlig i $c \in \mathbb{R}$, där c är en konstant så gäller det att:*

$$(\forall \epsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)(\forall x \in \mathbb{R})(|x - c| < \delta \Rightarrow |f(x) - f(c)| < \epsilon). \tag{21}$$

Vilket är ekvivalent med

$$(\forall x \in {}^*\mathbb{R})(x \simeq c \Rightarrow f(x) \simeq f(c)).$$

Bevis. Beviset för att dessa definitioner är ekvivalenta går via transferprincipen. För några konstanter $\epsilon, \delta \in \mathbb{R}^+$ så gäller via transfer av standarddefinitionen av kontinuitet, sats (21), att

$$(\forall x \in {}^*\mathbb{R})(|x - c| < \delta \Rightarrow |f(x) - f(c)| < \epsilon).$$

Om $x \simeq c$ så är $|x - c| < \delta, \forall \delta \in \mathbb{R}^+$. Så om $x \simeq c$ så gäller $|f(x) - f(c)| < \epsilon$ för godtyckligt reellt ϵ. Eftersom transfer av $(\forall \epsilon \in \mathbb{R}^+)(|f(x) - f(c)| < \epsilon)$ ger att $f(x) \simeq f(c)$, så har vi att $x \simeq c \Rightarrow f(x) \simeq f(c)$. Därmed har vi att kontinuitet medför $(x \simeq c \Rightarrow f(x) \simeq f(c)), \forall x \in {}^*\mathbb{R}$.

Utgående ifrån att $(x \simeq c \Rightarrow f(x) \simeq f(c))$, så måste man för att bevisa ekvivalens visa att definitionen av kontinuitet (21) håller. Givet $x \simeq c \Rightarrow f(x) \simeq f(c)$ så gäller det att

$$(|x - c| < \delta \Rightarrow |f(x) - f(c)| < \epsilon)$$

för något $\delta \simeq 0, \epsilon \simeq 0$. Låt ϵ vara ett positivt reellt tal och sök ett δ sådant att (21) är sann. Om δ är en infinitesimal så är $x \simeq c$ då gäller utifrån antagandet också att $f(x) \simeq f(c)$. Därmed gäller att $|f(x) - f(c)| < \epsilon$ för godtyckligt reellt ϵ. Därmed gäller det att

$$(\exists \delta \in {}^*\mathbb{R}^+)(\forall x \in {}^*\mathbb{R})(|x - c| < \delta \Rightarrow |f(x) - f(c)| < \epsilon)$$

där c, ϵ är reella. Via transfer av ovanstående så gäller det att

$$(\exists \delta \in \mathbb{R}^+)(\forall x \in \mathbb{R})(|x - c| < \delta \Rightarrow |f(x) - f(c)| < \epsilon).$$

Vilket därmed ger att $(x \simeq c \Rightarrow f(x) \simeq f(c), \forall x \in {}^*\mathbb{R})$ medför ϵ, δ-kontinuitet i c. □

Sats 15. Icke-standard likformig kontinuitet. *Om $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$, så är f likformigt kontinuerlig i A om:*

$$(\forall \epsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)(\forall x,y \in A)(|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon). \tag{22}$$

En ekvivalent icke-standarddefinition fås av

$$(\forall x,y \in {}^*A)(x \simeq y \Rightarrow f(x) \simeq f(y)). \tag{23}$$

Bevis. Beviset för att dessa definitioner är ekvivalenta är likt det för kontinuitet. Om vi väljer ϵ och δ, positiva reella konstanter, så att (22) är sann och använder transfer så fås

$$(\forall x,y \in {}^*A)(|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon).$$

Om $x \simeq y$ så gäller det att $|x - y| < \delta, \forall \delta$, eftersom δ är reellt. Därmed kan reellt ϵ väljas godtyckligt, vilket innebär, via transfer, att $f(x) \simeq f(y)$. Likformig kontinuitet medför därmed att $(\forall x,y \in {}^*A)(x \simeq y \Rightarrow f(x) \simeq f(y))$. Vi börjar nu istället med att anta att $(\forall x,y \in {}^*A)(x \simeq y \Rightarrow f(x) \simeq f(y))$ är sant, tag ϵ positivt reellt då gäller

$$(\exists \delta \in {}^*\mathbb{R}^+)(\forall x,y \in {}^*A)(|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon).$$

eftersom om $f(x) \simeq f(y)$ så är $|f(x) - f(y)| < \epsilon$ för alla reella ϵ. Transfer av ovanstående ger

$$(\exists \delta \in \mathbb{R}^+)(\forall x,y \in A)(|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon).$$

□

Likformig kontinuitet är ett starkare krav på så sätt att

$$(\forall x \in {}^*A)(x \simeq y \Rightarrow f(x) \simeq f(y)).$$

måste gälla för alla hyperreella y, dvs. för alla $y \in {}^*A$, kontra kontinuitet som endast kräver

$$(\forall x \in {}^*A)(x \simeq y \Rightarrow f(x) \simeq f(y))$$

där $y \in A$.

5.4.2 Integralen blir en summa

I icke-standardanalysen kan integraler betraktas som summor. Nedan demonstreras det som i [10].

Sats 16. *För en kontinuerlig funktion $f : \mathbb{R} \to \mathbb{R}$ ges Riemannintegralen av*

$$\int_0^1 f(x)dx = {}^o(\sum_{i=1}^{H} {}^*f(\frac{i}{H})\frac{1}{H})$$

där $H \in {}^*\mathbb{N} \setminus \mathbb{N}$ är ett oändligt stort tal.

Bevis. För en uniform partition av $[0, 1]$ i n intervall gäller att Riemannintegralen kan betecknas som gränsvärdet av den högra Riemannsumman $R_n(f)$ när $n \to \infty$.

$$\int_0^1 f(x)dx = \lim_{n\to\infty} R_n(f(x)) = \lim_{n\to\infty} \sum_{i=1}^{n} f(\frac{i}{n})\frac{1}{n}$$

så att problemet reduceras till ett bevis för att $\lim_{n\to\infty} R_n(f) = {}^oR_H(f)$.

Låt a vara det reella tal för vilket det gäller att

$$\int_0^1 f(x)dx = a$$

Eftersom $R_n(f)$ konvergerar mot a så existerar för ett positivt reellt tal ε ett naturligt tal $N = N(\varepsilon)$ så att

$$(\forall n \in \mathbb{N})(n \geq N \Rightarrow \mid R_n - a \mid < \varepsilon)$$

vilket efter transfer blir

$$(\forall n \in {}^*\mathbb{N})(n \geq N \Rightarrow \mid R_n - a \mid < \varepsilon)$$

i ${}^*\mathbb{R}$. Eftersom $H > N$ oavsett N gäller då att också $\mid R_H - a \mid < \varepsilon$ så att $\mid R_H - a \mid \simeq 0$ och därmed ${}^oR_H = a$. □

6 Definitioner av hyperreella begrepp och regler

6.1 Nya begrepp

Definition 18. Galaxer, Halos och innebördes avstånd. *Ett hyperreellt tal x ligger oändligt nära ett hyperreellt tal y om differensen $x-y$ är infinitesimal. Detta betecknas $x \simeq y$*

Två hyperrella tal x och y sägs vara på ett begränsat avstånd ifrån varandra om differensen $x - y$ är begränsad , det vill säga ligger i $\mathbb{R}$. Relationen betecknas $\sim$.

De två relationerna $\simeq$ och $\sim$ är ekvilanesrelationer på ${}^*\mathbb{R}$, (att vilkoren för detta är uppfyllda kontrolleras enkelt via räknelagarna för hyperreella tal) de delar således in ${}^*\mathbb{R}$ i ekvivalensklasserna :

$$hal(x) = \{c \in {}^*\mathbb{R} : x \simeq y\} \tag{24}$$

$$gal(x) = \{c \in {}^*\mathbb{R} : x \sim y\} \tag{25}$$

Vi säger att ett hypperreellt tal x halo utgörs av mängden av alla tal som uppfyller (24) *och dess Galax utgörs av mängden av tal som uppfyller* (25)*.*

Alltså är hyperreelt tal x halo alla tal som ligger onändligt nära x, och dess galax är de tal som ligger på ändligt avstånd från x. Alternativa namn för ett tals halo är *monad* eller *atom*.

Sats 17. *Varje begränsat hyperreellt tal x ligger oändligt nära ett, och endast ett, reellt tal.*

Bevis. Låt $A = \{r \in \mathbb{R} : r < b\}$. Eftersom b är begränsat finns det reella tal r och s s.a. $r < b < s$. A är alltså icke-tom samt uppåt begränsad (av s) i $\mathbb{R}$. Av fullständigheten för $\mathbb{R}$ följer det att A har ett supremum $c \in \mathbb{R}$.

För att visa att b $\simeq$ c , ta något positivt $\epsilon \in \mathbb{R}$. Eftersom c är supremum till A så kan vi inte ha något c + $\epsilon \in A$; varför b $\leq c + \epsilon$. Vidare ; om $b \leq c - \epsilon$ så skulle detta vara en övre begränsing för A , men enligt antagande är c minsta övre begränsning, alltså kan inte denna olikhet vara uppfylld. Sammantaget gäller att $c - \epsilon < b \leq c + \epsilon$, alltså är $|b - c| \leq \epsilon$. Eftersom detta är sant för alla positiva reella tal epsilon så är b oändligt nära c.

Det återstår bara att visa entydigheten. Om b $\simeq d \in \mathbb{R}$, så gäller (ty $b \simeq c$) $c \simeq d$, och därmed (ty reella) $c = d$. □

Definition 19. Skuggan av x. *Det reella tal vilket ett hyperreellt tal x ligger oändligt nära kalls för skuggan av x och betecknas $sh(x)$.*

6.2 Räkneregler för hyperreella tal

Räknelagarna för hyperreella tal stämmer väl överrens med vad som intuivt förväntas. För infinitesimala tal δ, ϵ , begränsade tal b,c och oändligt stora tal Ω,Λ gäller :

- **Summor**

  $\epsilon + \delta$ är infinitesimal

  $b + \epsilon$ är begränsat

  $b + c$ är begränsat (kan vara infinitesimalt)

  $\Omega + b + \epsilon$ är oändligt stort

- **Additativa Inverser**

  $-\epsilon$ är infinitesimalt

  $-b$ är begränsat

  $-\Omega$ är oändligt stort

- **Produkter**

  $\epsilon \cdot \delta$ och

  $\epsilon \cdot b$ är infinitesimala

  $b \cdot c$ är begränsat

  $\Omega \cdot b$ och $\Omega \cdot \Lambda$ är oändligt stora

- **Multiplikativa Inversioner**

  $\frac{1}{\epsilon}$, $\epsilon \neq 0$ är oändligt stort

  $\frac{1}{b}$ är begränsat

  $\frac{1}{\Omega}$ är infinitesimalt

- **Kvoter**

  $\frac{\epsilon}{b}$,$\frac{\epsilon}{\Omega}$ samt $\frac{b}{\Omega}$ är infinitesimala

  $\frac{b}{\epsilon}$,$\frac{\Omega}{b}$ samt $\frac{\Omega}{\epsilon}$ är oändligt ϵ,$b \neq 0$ stora

  $\frac{b}{c}$ $c \neq 0$ är begränsat

- **Rötter**

  $\sqrt[n]{\epsilon}$ är infinitesimalt

  $\sqrt[n]{b}$ är begränsat

  $\sqrt[n]{\Omega}$ är oändligt stort

$\epsilon, b, \Omega > 0$

- **Obestämda**

  $\frac{\epsilon}{\delta}, \frac{\Omega}{\Lambda}, \epsilon \cdot \Omega,\ \Omega + \Lambda$

7 Exempel 1: Komplex analys

Detta kapitel behandlar den första av de två fördjupningar som kommer göras, denna inom komplex analys. Speciellt en del av den komplexa analysen kommer ligga i fokus i detta avsnitt; nämligen teori berörande analytiska funktioner.

7.1 Syfte

Syftet med detta kapitel är att jämföra icke-standard och standardversionen av Picards stora sats, detta för att visa på fördelarna med icke-standardanalysen. Först så kommer satser och definitioner införas som är nödvändiga för att bevisa Picards stora sats. Det kommer att visa sig att det faktiskt inte är så mycket som behöver tilläggas utöver en grundläggande kurs i analytiska funktioner (som läsaren alltså förväntas ha genomfört) - åtminstone inte i icke-standardfallet.

Ett avgörande resultat är fundamentalt för de bevis som kommer att genomföras i denna text; att enhetsskivan, $D = \{z : |z| \leq 1\}$, avbildas på hela $\mathbb{C}$ förutom på 0 och 1, det vill säga $\mathbb{C}\backslash\{0{,}1\}$. Detta görs med en elliptisk modulärfunktion som är relativt komplicerad. Det går att bevisa Picards stora sats (bara Picards sats i fortsättningen om inget missförstånd kan uppstå) utan att använda sig av denna men av ett antal anledningar har bevis som utnyttjar denna ändå valts. Exempelvis går det att använda "Bloch-Laundau varianten" för att bevisa Picard och på så sätt går det att undvika att använda den elliptiska modulärfunktionen. Priset att betala för att slippa denna är att använda sig av Schottkys och Bloch-Laundaus satser. Dessa reducerar förvisso beviset av Picard till ett kort, elementärt bevis - men först efter drygt 20 sidors introduktion som är allt annat än trivial. Beviset på detta sätt går att läsa om i [9].

Ett ytterligare argument för valet att använda förfarandet med den elliptiska modulärfunktionen i beviset är att detta bevis mer liknar icke-standardvarianten och en mer direkt jämförelse blir på så sätt enklare och betydligt mer givande. De två bevisens tillvägagångssätt är väldigt lika varandra, men de tar sig skilda uttryck i standard respektive icke-standardfallet eftersom olika matematiska verktyg är tillgängliga.

Slutligen så vill vi poängtera att då den elliptiska modulärfunktionen förvisso är viktig för detta bevis av Picards sats så är det långt ifrån nödvändigt att helt förstå funktionen i sig - det som är viktigt att veta är egentligen bara att enhetsskivan är en övertäckande mängd till $\mathbb{C} \setminus \{0, 1\}$ (det vill säga en mängd som via en övertäckande avbildning avbildas på hela $\mathbb{C} \setminus \{0, 1\}$) och vad denna har för egenskaper. Att inte i detalj förstå hur själva funktionen fungerar och varför är ovidkommande vad gäller förståelse av beviset till Picard.

7.2 Nödvändig teori inom komplex analys

Nedan introduceras några grundläggande topologiska begrepp som kommer att användas för att bevisa delar av Picards sats. Sedan definieras övertäckningsavbildning och lyft, det visas även hur den elliptiska modulärfunktionen kan konstrueras. En avvägning har gjorts inför detta kapitel - vissa begrepp var nödvändiga för beviset i senare kapitel (såväl standardvarianten som icke-standardvarianten) men bakomliggande teori var såväl omfattande som abstrakt. Målet var att åstadkomma en så kortfattad framställning som möjligt utan att göra avkall på stringens och fullständighet. Det visade sig dock rätt snart att om ingen definition eller sats skulle utelämnas skulle det ta alldeles för mycket tid, utrymme och uppmärksamhet i anspråk - därför valdes en slags mellanväg där de viktigaste koncepten presenteras och definieras abstrakt, men med viss förklarande text efteråt som pekar på vad som är av intresse i denna text. Vissa detaljer utelämnas dock (till exempel definieras inte den fundamentala gruppfunktorn) men vi ansåg att denna framställning är mer informativ än den som många författare använder då Picards sats skall bevisas: att bara i förbigående nämna att övertäckningsmängden till $C \setminus 0{,}1$ är enhetsskivan utan bevis. Det är onekligen en delikat balansgång, men vi anser att detta kapitlet är lagom djupgående för att inte kännas som totalt handviftande men samtidigt tillräckligt överskådligt för att inte ta för mycket fokus från den huvudsakliga uppgiften.

7.2.1 Topologiska begrepp

Följande topologiska begrepp kommer från Rotman [12] om ingen annan källa nämns.

Definition 20. Kategori. *En kategori $\mathcal{K}$ består av:*

1. *En uppsättning objekt (till exempel mängder eller grupper), som vi kallar obj $\mathcal{K}$.*
2. *En uppsättning morfismer, hom(X,Y), en för varje ordnat par X,Y $\in$ obj $\mathcal{K}$. Om $f \in$ hom(X,Y) skriver vi $f : X \longrightarrow Y$.*
3. *Sammansättningar av morfismer hom$(Y,Z) \times$ hom$(X,Y) \longrightarrow$ hom(X,Z) till varje ordnad trippel $(X,Y,Z) \in$ obj $\mathcal{K}$. Om $f \in$ hom(X,Y) och $g \in$ hom(Y,Z) skrivs avbildningen av (f,g) på hom(X,Z) som $f \circ g$.*

För att vara en kategori behöver dessa objekt, morfismer och sammansättningar av morfismer uppfylla följande axiom:

1. *Låt $f \in$ hom(X,Y), $g \in$ hom(Y,Z) och $h \in$ hom(Z,W). Om $h \circ (g \circ f)$ eller $(h \circ g) \circ f$ är definierad så är båda definierade och $h \circ (g \circ f) = (h \circ g) \circ f$.*
2. *För varje $Y \in$ obj $\mathcal{K}$ så finns en identitetsfunktion $1_Y \in$ hom(Y,Y) som uppfyller $1_Y \circ g = g$ för alla $g \in$ hom(X,Y), $X \in$ obj $\mathcal{K}$ och $h \circ 1_Y = h$ för varje $h \in$ hom(Y,Z), för alla $Z \in$ obj $\mathcal{K}$.*

Det går att visa att 1_Y är unik.

Speciellt kallas en homomorfi mellan kategorier för funktor.

Definition 21. Topologisk ekvivalens - homeomorfi. *Två objekt X och Y sägs vara topologiskt ekvivalenta eller homeomorfa om $\exists f \in$ hom(X,Y) och $g \in$ hom(Y,X) sådant att $g \circ f = 1_X$ och $f \circ g = 1_Y$. f och g kallas ekvivalenser.*

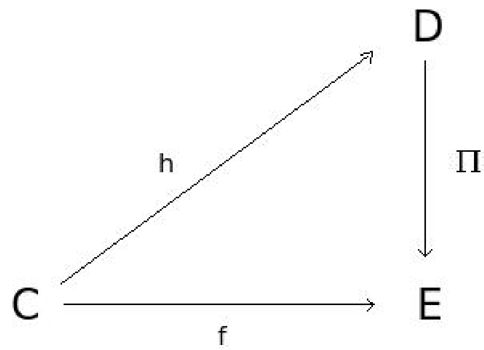
Sammantaget säger ovanstående satser och definitioner att en homeomorfi är en kontinuerlig och bijektiv avbildning som har en kontinuerlig invers. Det vill säga att två stycken objekt är homeomorfa om det finns en kontinuerlig, inverterbar avbildning som kan deformera det ena objektet till det andra. Vi behöver detta för att kunna definiera en jämnt täckande avbildning, vilken i sin tur kommer behövas för att kunna definiera övertäckningsavbildningen.

Definition 22. Jämnt övertäckt. *Låt X och $\tilde{X}$ vara topologiska rum och låt $p : \tilde{X} \to X$ vara en kontinuerlig avbildning. En öppen mängd $U \subset X$ är jämnt övertäckt av p om $p^{-1}(U)$ är en disjunkt union av öppna mängder $S_i \subset \tilde{X}$, för varje i så gäller det att $p : S_i \to U$ är en homeomorfi.*

7.2.2 Övertäckningsavbildning och lyft

Definition 23. Övertäckningsavbildning. *Om X och $\tilde{X}$ är topologiska rum, då är $\tilde{X}$ en övertäckande mängd till X om följande är uppfyllt:*

1. *$\tilde{X}$ är bågvis sammanhängande,*
2. *p: $\tilde{X} \to X$ är kontinuerlig,*
3. *varje $x \in X$ har en öppen omgivning U_x som är jämnt övertäckt av p.*



Figur 1: Givet en analytisk övertäckningsavbildning Π som avbildas på E så kan alla analytiska funktioner f som avbildas på E lyftas via en funktion h till mängden som utgör definitionsmängden till Π.

Avbildningen p kallas för en övertäckningsavbildning. Det som egentligen är intressant för beviset av Picards Sats är en specifik analytisk övertäckningsavbildning; avbildningen från den öppna enhetsskivan till det komplexa talplanet förutom punkterna (0,0) och (1,0). Vad som är av störst vikt där är att avbildningen p är kontinuerlig, och det faktum att p är en övertäckningsavbildning kommer att utnyttjas för att göra ett *lyft* - detta är en av grundpelarna för beviset av icke-standardversionen av Picards Stora Sats. Existensen av ett sådant lyft garanteras om villkoren i följande sats är uppfyllda.

Sats 18. Existens av lyft. *Låt $f : X \longrightarrow B$ vara en kontinuerlig avbildning, där X är en sammanhängande och lokalt bågvis sammanhängande mängd. $x \in X$ och $e \in E$ är punkter sådana att $f(x) = p(e)$. Givet E, en övertäckande mängd till B, och $p : E \longrightarrow B$ den tillhörande övertäckningsavbildning så kan f lyftas till en avbildning $\tilde{f} : X \longrightarrow E$ med $\tilde{f}(x) = e$ om och endast om $\pi_1(f)$ avbildar $\pi_1(X,x)$ helt in i bilden $\pi_1(p)(\pi_1(E,e))$. Där π_1 betecknar den fundamentala gruppfunktorn.*

I något lösa ordalag säger satsen att om två relativt "snälla" (kontinuerliga) avbildningar avbildar varsin mängd på precis samma mängd så är de såpass lika varandra att det borde finnas en slags "översättning" mellan de båda funktionerna och det är den man kallar lyftet av f.

Figur 7.2.2 ger en överskådlig bild av hur de olika avbildningarna och mängderna förhåller sig till varandra. I fallet med komplexanalys blir satsen mer överskådlig och vi slipper villkoret i sats 18, eftersom här har vi det ytterligare kravet att mängden X, där funktionen som ska lyftas är definierad, är enkelt sammanhängande detta innebär att $\pi_1(X,x) = 1$ och alla kontinuerliga avbildningar från X till B kan lyftas [13].

Sats 19. Lyftsatsen. *Låt $\Pi : D \longrightarrow E$ vara en analytisk övertäckningsavbildning, och C vara ett enkelt sammanhängande område i $\mathbb{C}$. Då gäller att varje analytisk funktion $f : C \longrightarrow E$ kan lyftas via h. Det vill säga $\exists g : E \longrightarrow D$ sådan att $\Pi = g \circ h$.*

7.2.3 Konstruktion av övertäckningsavbildningen

I detta avsnitt behandlas hur övertäckningsavbildningen från D till $\mathbb{C} \setminus \{0,1\}$ kan konstrueras, det är denna övertäckningsavbildning som kommer användas i följande kapitel om Picards sats. Innan själva beskrivningen av funktionens uppbyggnad kommer ett par nödvändiga resultat att presenteras. Dessa är inte särskilt komplicerade eller tunga, men för att underlätta för läsaren så har de lyfts ut ur beskrivningen av själva konstruktionen av den elliptiska modulärfunktionen. Konstruktionen av övertäckningsavbildningen utgår främst ifrån källa nummer [7].

Ett resultat som kommer att användas under konstruktionen av övertäckningsavbildningen är Schwarz Lemma. Läsaren förväntas vara bekant med detta men en formulering följer här - dock utan bevis [3].

Sats 20. Schwarz Lemma. *Låt $D = \{z : |z| < 1, z \in \mathbb{C}\}$ vara den öppna enhetsskivan. Givet en funktion f som är analytisk och uppfyller att $f(0) = 0$ samt $|f(z)| \leq 1, \forall z \in D$. Då gäller : $|f(z)| \leq |z|$ samt $|f'(0)| \leq 1, z \in D$. Vidare gäller att om $|f(z)| = |z|$ för något $z \neq 0, z \in D$ eller om $f'(0) = 1$ så är $f(z) = az$ för något $a \in D$.*

För att definiera Schwarz reflektionsprincip behövs analytisk fortsättning, precis som i fallet med Schwarz lemma så är detta något de flesta läsare antagligen redan känner till så därför följer här endast en kort genomgång av vad det är. Antag att en funktion $f(z)$ är analytisk på en öppen mängd $A \in \mathbb{C}$ och $F(z)$ är analytisk på $B, A \subset B$, B öppen med $F(z) = f(z), \forall z \in A$. I detta fall kallas F för en *analytisk fortsättning* av f.

En analytisk fortsättning är unik i bemärkelsen att om B är den sammanhängande unionen av definitionsmängderna till två funktioner f_1, f_2 samt att $A \subset B$ så gäller att $f_1(z) = f_2(z) = F(z), \forall z \in A \Longrightarrow f_1(z) = f_2(z), \forall z \in B$. Detta gäller eftersom differensen av två analytiska funktioner också är en analytisk funktion och enligt satsen om isolerade nollställen så måste $f_1 - f_2 \equiv 0$ eftersom $f_1 - f_2 = 0$ på hela A, detta givetvis för värden för vilka f_1, f_2 är definierade. Detta resultat kommer att visa sig vara avgörande för konstruktionen av övertäckningsavbildningen som i sin tur används för att bevisa Picards Stora Sats.

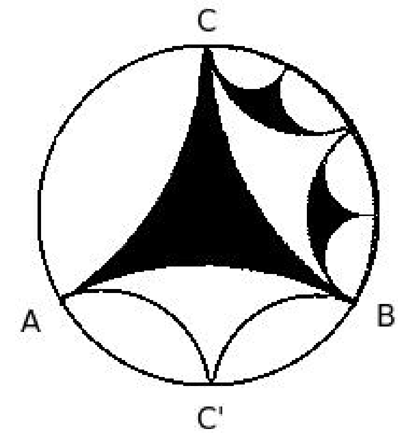
Nu är det dags att presentera nästa resultat som kommer att användas; Schwarz reflektionsprincip. Schwarz reflektionsprincip säger att det går att utvidga definitionsmängden för en funktion $f(z)$ med z i övre halvplanet (öhp) om f antar reella värden på den reella axeln. Eftersom f är definierad på det över halvplanet är alltså (den än så länge förmodade) utvidgningen till f dess komplexkonjugat. Reella tal saknar som bekant imaginärdel (eller rättare sagt den är 0) och är därmed lika med sitt konjugat. Om $\bar{z}$ betecknar komplexkonjugatet till z kommer alltså $f(\bar{z}) = \bar{f(z)}$ att överensstämma längs den reella axeln. Satsen säger följande;

Sats 21. Schwarz reflektionsprincip. *Låt f(z) vara en funktion av en komplex variabel och som är kontinuerlig på det (halv)slutna övre halvplanet och analytisk på det öppna övre halvplanet, samt att $f(z) \in \mathbb{R}, \forall z \in R$. Då är $f(\bar{z}) = \bar{f(z)}$ en analytisk fortsättning till hela $\mathbb{C}$.*

Nu är det dags att visa att det finns en funktion, en *övertäckningsavbildning*, som avbildar enhetsskivan på hela $\mathbb{C} \setminus \{0, 1\}$. Betrakta figur 7.2.3. Området Ω_1 utgörs av en hyperbolisk triangel som är inskriven i enhetsskivan (som vi i vanlig ordning betecknar D). Randen till Ω_0 (cirkelbågsegmenten som är triangelns sidor) definieras av cirklar som skär D vinkelrätt i punkterna A, B och C. Cirkelbågarna möts alltså i dessa tre punkter och har relativ vinkel 0 där. Enligt Schwarz existerar det en analytisk funktion som avbildar det övre halvplanet bijektivt på Ω_0. Funktionen är bijektiv och analytisk - därmed har funktionen en invers som vi väljer att kalla $\mu(z)$, som alltså kommer att avbilda Ω_0 bijektivt på det övre halvplanet. Speciellt så avbildar $\mu(z)$ de tre gränsbågarna till Ω_0 på den utvidgade reella tallinjen $\mathbb{R} \cup \infty$. Vidare så kan funktionen modifieras så att de tre punkterna A, B och C (där cirkelbågarna skärs) skickas till $0, 1$ och till ∞. Nästa steg är att utveckla $\mu(z)$ till hela D.

För att utveckla området Ω_0 till hela C' kommer Ω_0s sidor reflekteras i C' och detta kommer upprepas successivt så att en triangulering av C' erhålles. Låt områdena som bildas av de första reflektionerna kallas för Ω_1, Ω_2 och Ω_3. $\mu(z)$ är konstruerad så att den endast antar reella värden på randen till Ω_0 och detta erbjuder möjligheten att tillämpa Schwarz reflektionsprincip. Schwarz reflektionsprincip tillåter oss nu att låta $\mu(z)$ utvecklas analytiskt till det nya området $\bar{\Omega} = \Omega_0 \cup \Omega_1 \cup \Omega_2 \cup \Omega_3$. Funktionen $\mu(z)$ avbildar varje nytt område Ω_i bijektivt och analytiskt på det *undre* halvplanet - följaktligen avbildas $\bar{\Omega}$ analytiskt på hela $\mathbb{C}$ dock är $\mu(z)$ nu flervärd eftersom Ω_1, Ω_2 och Ω_3 vart och ett avbildas på det undre halvplanet av $\mu(z)$. Eftersom spegling är en konform avbildning så kommer vinklarna att bevaras, och cirkelbågarna som utgör gränserna för de nya områdena $\Omega_{1,2,3}$ är alltså vinkelräta mot enhetscirkeln i punkterna A, B och C. Vidare så antar $\mu(z)$ reella värden på randen till vart och ett av de nya områdena. Upprepa nu samma förfarande och reflektera $\Omega_{1,2,3}$ i enhetscirkeln, de nya områden som bildas kommer vart och ett att avbildas analytiskt på öhp av μ med reella värden på randen. Alltså kommer hela C' att fyllas så småningom; en

[3] Bevis går att läsa i exempelvis [13]



Figur 2: Bild av hur enhetsskivan fylls ut genom reflektion. Varje par av mörka och ljusa områden avbildas på $\mathbb{C} \setminus \{0{,}1\}$ av övertäckningsavbildningen.

triangulering av C' fås. μ är alltså analytiskt på hela enhetsskivan och alla begränsade värden utom $0, 1$ antas.

Det kommer att finnas ett z i vart och ett av de mörka områdena som uppfyller $\mu(z) = \zeta \in$ öhp . Två på varandra efterföljande reflektioner kommer att avbilda $z \to \bar{z}$ med $\zeta = \mu(z) = \mu(\bar{f}z)$. Alltså är $\mu(z)$ invariant vid jämnt antal reflektioner av argumentet. Den till μ inversa funktionen $v(w) = \mu^{-1}(w)$ är flervärd men till varje kurva som undviker $\{0, 1, \infty\}$ så finns det en funktion (som alltså är ett element i μ) som kan utvidgas analytiskt längs kurvan. Speciellt är $v(w)$s prinicpalvärden de som antas i tetragonen $\Omega_0 \cup \Omega_1$ inklusive randen på den övre halvan (det vill säga den slutna tetragonen ADBC minus kurvan ADB). Varje par av skuggade och ljusa områden utgör ett fundamentalt område (det vill säga ett område som avbildas på hela $\mathbb{C} \setminus \{0{,}1\}$.

Avslutningsvis vill vi upprepa att detaljerna i detta kapitel inte är av enorm vikt när det kommer till att förstå Picard. Icke desto mindre är det vissa saker som är viktiga att minnas och det är just att *det finns* en avbildning från enhetsskivan, D till $\mathbb{C} \setminus \{0{,}1\}$ denna är en *analytisk övertäckande avbildning* vilket innebär att enhetsskivan alltså är en *övertäckande mängd*. Detta gör att de villkor som ställs på den ena funktionen då ett lyft genomförs är uppfyllda - och det är detta som kommer att utnyttjas i beviset av Picard.

7.3 Icke-standardanalys och Picards stora sats

Innan vi kan introducera icke-standardformuleringen av Picards stora sats och Robinson-Callots sats så måste S-kontinuitet definieras. Vi tittar även på hur S-kontinuitet hänger samman med väsentliga singulariteter vilket kommer visa sig viktigt för att förstå icke-standardformuleringen av satsen. Satser, bevis och definitioner i detta avsnitt följer i stora drag tillvägagångssättet i Dieners bok [2].

7.3.1 S-kontinuitet

Det är en skillnad mellan vanlig kontinuitet och S-kontinuitet vilken försökes belysas i detta avsnitt. Här presenteras först de två olika begreppens definitioner, där vi ser att S-kontinuitet innebär att en funktion är kontinuerlig i icke-standardpunkter.

Definition 24. S-kontinuerlig. *Låt E,$F \subseteq {}^*\mathbb{C}$ och $f : E \to F$. f kallas S-kontinuerlig i $x_0 \in E$ omm för alla $x \simeq x_0$ så gäller $f(x) \simeq f(x_0)$.*

Definitionen för vanlig kontinuitet lyder enligt följande:

Definition 25. Kontinuerlig. *Låt E,$F \subseteq \mathbb{C}$ och $f : E \to F$. f kallas kontinuerlig i $x_0 \in E$ omm för alla $x \simeq x_0$ så gäller $f(x) \simeq f(x_0)$.*

En standardfunktion f är kontinuerlig i standardpunkten x_0 omm $f(x) \simeq f(x_0)$ för alla $x \simeq x_0$, där x är icke-standardpunkter runt x_0. Funktionen är då kontinuerlig i ett område A till x_0 om den i detta område är kontinuerlig i varje standardpunkt.

En standardfunktion f är likformigt kontinuerlig i A omm $f(x) \simeq f(y)$ för alla $x \simeq y$ där $x,y \in {}^*A$, vilket är det samma som att säga omm f är S-kontinuerlig i varje punkt i ${}^*A$.

Ta till exempel funktionen $f(x) = \frac{1}{x}$ denna funktion är kontinuerlig i alla punkter i (0,1) men den är inte S-kontinuerlig i detta intervall. Tag en infinitesimal punkt ϵ, detta ger enligt S-kontinuitet att $f(\epsilon) \simeq f(\epsilon_1)$ om $\epsilon \simeq \epsilon_1$, men om $\epsilon_1 = 2\epsilon \simeq \epsilon$ tas så får vi att $\frac{1}{\epsilon} - \frac{1}{2\epsilon} = \frac{1}{2\epsilon} \not\simeq 0$ vilket innebär att funktionen inte är S-kontinuerlig i punkter infinitesimalt nära 0. Funktionen växer obegränsat fort infinitesimalt nära 0, den är fortfarande kontinuerlig där men inte S-kontinuerlig.

Om f,x är standard så är f S-kontinuerlig i x omm f är kontinuerlig i x. Så med standardvärda variabler och standardfunktionsvärden så är S-kontinuitet och kontinuitet samma sak. Om vi har en standard funktion f på en mängd ${}^*A$ om f istället är S-kontinuerlig i varje punkt i ${}^*A$ både standard och icke-standard så är funktionen likformigt kontinuerlig där.

7.3.2 S-kontinuitet och väsentliga singulariteter

I icke-standardversionen av Picards sats blir kopplingen mellan väsentliga singulariteter och S-kontinuitet viktig för att se hur denna version av satsen hänger ihop med standardformuleringen. Om f har en väsentlig singularitet i z_0 så är funktionen inte S-kontinuerlig där, detta följer av Casorati-Weierstrass sats som säger att i varje område till en väsentlig singularitet så kommer funktionen f godtyckligt nära alla komplexa tal. Detta kan formuleras som, där $U \subset \mathbb{C}$ innehåller z_0,

$$(\forall \epsilon \in \mathbb{R})(\forall \delta \in \mathbb{R})(\forall w \in \mathbb{C})(\exists z \in U)(|z - z_0| < \delta \wedge |f(z) - w| < \epsilon)$$

vilket ger att $z \simeq z_0 \not\Rightarrow f(z) \simeq f(z_0)$. Detta visar att om en funktion har en väsentlig singularitet i en punkt så medför detta att funktionen där inte är S-kontinuerlig där.

Det motsatta är inte nödvändigtvis sant, om man tittar på exemplet från tidigare så är $\frac{1}{z}$ analytisk i $\mathbb{C} \setminus 0$ men funktionen är som sagt inte S-kontinuerlig infinitesimalt nära origo, trots detta har funktionen ingen väsentlig singularitet i $\mathbb{C}$ utan endast en pol i origo.

7.3.3 Robinson-Callot och Picards stora sats

Följande satser innehåller en beskrivning av analytiska funktioner med ett icke-standard synsätt. Robinson-Callot beskriver analytiska funktioner i områden där de är begränsade och Picard beskriver hur en analytisk funktion beter sig kring en väsentlig singularitet. Funktionerna beter sig ungefär som man kan vänta sig: så länge de är analytiska (och därmed kontinuerliga) är de rätt lugna och beter sig som förväntat, i fallet med en väsentlig singularitet å andra sidan beter de sig väldigt extremt - de kan i en oändligt liten omgivning till singulariteten anta vilket komplext tal som helst som värde utom möjligtvis ett.

Sats 22. Robinson-Callot. *Låt f vara en analytisk funktion på en mängd som innehåller x_0 och $hal(x_0)$, där x_0 och $f(x_0)$ är begränsade. Om f bara antar begränsade värden i halon av x_0 så existerar ett standard område V av x_0 på vilket:*

1. *f är S-kontinuerlig*
2. *sh(f) är analytisk*
3. *$(sh(f))^{(n)} = sh(f^{(n)})$ där $n \in \mathbb{N}$.*

Bevis. Om f är begränsad och analytisk i $hal(x_0)$ så är enligt permanensprincipen, f analytisk och begränsad i ett standardområde V som innehåller $hal(x_0)$. Det räcker att visa satsen för punkter som ligger i det S-inre av V. Där man med S-inre av V menar alla punkter x sådana att om man tar en cirkelskiva med mittpunkt i x och en skattningsbar radie så ligger hela skivan i V.

1. Tag två punkter $z_1 \simeq z_2$ i det S-inre av V, och en sluten kurva $\gamma_\rho(t) = z_1 + \rho e^{it}$ med $t \in [0, 2\pi]$ och ρ skattbar sådan att kurvan ligger i V. Cauchys formel ger då

$$f(z_1) - f(z_2) = \int_{\gamma_\rho} \frac{f(z)}{z - z_1} dz - \int_{\gamma_\rho} \frac{f(z)}{z - z_2} dz = (z_1 - z_2) \int_{\gamma_\rho} \frac{f(z)}{(z - z_1)(z - z_2)} dz.$$

Eftersom f endast antar begränsade värden och $z \not\simeq z_1$,z_2 så är termen $\int_{\gamma_\rho} \frac{f(z)}{(z-z_1)(z-z_2)} dz$ begränsad. Eftersom $z_1 \simeq z_2$ så är högerledet $\simeq 0$.

2. Ta en sluten standardskiva $D \subseteq V$. $sh(f)$ är kontinuerlig eftersom f är S-kontinuerlig. Då har vi att $f(z) \simeq f(sh(z)) \simeq sh(f(sh(z))) = sh(f(z))$. Eftersom f är analytisk har vi $\int_{\partial D} sh(f(z)) = \int_{\partial D} f(z) = 0$, och därmed är även $sh(f)$ analytisk i ett område som innehåller $hal(x_0)$

3. Med γ_ρ och x_1 definierade som ovan, så ger Cauchys formel

$$sh(f)^{(n)}(z) = \frac{1}{2\pi i} \int_{\gamma_\rho} \frac{sh(f(z))}{(z - z_1)^{n+1}} dz \simeq \frac{1}{2\pi i} \int_{\gamma_\rho} \frac{f(z)}{(z - z_1)^{n+1}} dz = f^{(n)}(z)$$

Därmed är $sh(f)^{(n)}(z) = sh(f^{(n)}(z))$. □

Innan Picards sats bevisas tas först Casorati-Weierstrass sats upp som är en svagare variant av icke-standard Picard, här har man inte att f antar alla värden utom som mest en del av halon till en punkt utan här kommer funktionen endast godtyckligt nära varje värde i $\mathbb{C}$. Denna sats är mycket enklare att bevisa, det som i huvudsak krävs är Robinson-Callots sats, till skillnad från Picards stora sats där en hel del mer teori krävs.

Sats 23. Icke-standard Casorati Weierstrass. *Antag att f inte är S-kontinuerlig i x_0. Då gäller det att $f(hal(x_0)$ möter halon av alla begränsade punkter.*

Bevis. Antag det motsatta att det existerar en punkt w_0 vars halo inte innehåller något värde av $f(hal(x_0))$. Då är $g(x) = \frac{1}{f(x) - w_0}$ analytisk och begränsad i $hal(x_0)$. Så enligt Robinson-Callots sats gäller att g är S-kontinuerlig $x_1 \simeq x_0 \Rightarrow g(x_1) \simeq g(x_0)$.

Därmed är $g(x)$ antingen skattbar eller infinitesimal $\forall x \in hal(x_0)$. Eftersom $f(x) = w_0 + 1/g(x)$, så har vi att om $g(x)$ är skattbart så är $f(x)$ begränsad i $hal(x_0)$ och då är enligt Robinson-Callots sats f S-kontinuerlig i $hal(x_0)$ och därmed kontinuerlig i $f(x_0)$ vilket är en motsägelse. Om $g(x)$ är infinitesimal i $hal(x_0)$ så är istället $f(x_0)$ obegränsad vilket också är en motsägelse. □

Nedan följer en sats som redan vid första anblick är lik Picards stora sats på standardsidan, men vi pratar inte här om att funktioner har en väsentlig singularitet i en punkt z_0 utan om huruvida de är S-kontinuerlig där. Till skillnad från föregående sats så krävs för att visa att $f(hal(x_0))$ faktiskt innehåller alla begränsad punkter förutom möjligen en del av halon till en punkt att man känner till att den analytiska övertäckningsmängden av $\mathbb{C} \setminus \{0,1\}$ är enhetsskivan. Det är via övertäckningsavbildningen som tillhör den nämnda övertäckningsmängden som vi visar att f är S-kontinuerlig i x_0 om den undviker två värden i $\mathbb{C}$. Vi kommer senare att förklara varför satsen inte bara ser ut som Picard på standardsidan utan faktiskt är en helt analog sats i icke-standardutförande.

Sats 24. Icke-standard Picard. *Låt f vara en analytisk funktion på en mängd som innehåller $hal(x_0)$, där x_0 och $f(x_0)$ är begränsade. Om f inte är S-kontinuerlig i x_0, så gäller att $f(hal(x_0))$ innehåller alla begränsade komplexa tal förutom möjligen en del av halon till en punkt.*

Bevis. Antag att $a \not\simeq b$ och att $a,b \notin f(hal(x_0))$. Permanensprincipen ger att det finns ett område V till x_0 som innehåller $hal(x_0)$ och där f inte är antar a eller b. Låt

$$T : \mathbb{C} \to \mathbb{C}, z \to \frac{z - a}{b - a},$$

och

$$g = T \circ f : V \to \mathbb{C} \setminus \{0,1\}.$$

Om $g(hal(x_0))$ helt ligger i $hal(0)$ eller $hal(1)$ så är g S-kontinuerlig i x_0, och då är också $f(x_0)$ det. Eftersom f är kontinuerlig och $f(x_0)$ är begränsad finns det $x_1 \simeq x_0$ sådant att $g(x_1)$ är begränsat och $g(x_1) \not\simeq 0, 1$.

Nu använder vi resultatet om holomorfa övertäckande mängder, låt D beteckna den öppna enhetsskivan dvs. $D = \{\xi \in \mathbb{C} : |\xi| < 1\}$ vilket är en övertäckande mängd av $\mathbb{C} \setminus \{0,1\}$ som enligt definitionen har en tillhörande analytisk övertäckningsavbildning (standardavbildning).

$$\Pi : D \to \mathbb{C} \setminus \{0,1\},$$

som är surjektiv och lokalt injektiv.

Π och g har samma värdemängd därmed existerar ett standard $\xi_0 \in D : \Pi(\xi_0) = sh(g(x_1)))$. Eftersom Π är kontinuerlig och $g(x_1) \not\simeq 0, 1$ så gäller det att $\Pi(\xi_1) = g(x_1)$, för något $\xi_0 \simeq \xi_1 \in D$. Benämn lyftet av g för $G = \Pi^{-1} \circ g$ och låt detta vara lyftet definierat av $G(x_1) = \xi_1$. G är analytisk i ett område som innehåller $hal(x_1)$, begränsad eftersom $|\xi| < 1$ och därmed S-kontinuerlig i x_1. ξ_1 är en icke-standard punkt och $G(x_1) = \Pi(\xi_1)$, därmed är Π S-kontinuerlig i ξ_1.

Eftersom $\Pi \circ G = g$ och $g = T \circ f$ så har vi att $f = T^{-1} \circ \Pi \circ G$, detta ger att f är S-kontinuerlig i x_1 och därmed även i x_0, vilket är en motsägelse.

Antagandet att f inte är S-kontinuerlig i x_0 ger enligt ovanstående att det inte finns två godtyckliga begränsade värden, a och b som uppfyller $a \not\simeq b$, som f inte antar, så antingen antar f alla begränsade värden eller alla förutom ett begränsat värde. □

De två tidigare nämnda satserna beskriver analytiska funktioners beteenden kring en punkt x_0 i fallen då funktionen är begränsad i en omgivning kring denna, respektive när den inte är S-kontinuerlig men antar åtminstone ett begränsat värde i halon till x_0.

Ett fall som inte tas upp här (eftersom det antagits att funktionen avbildar x_0 på ett begränsat värde) är att funktionen är S-kontinuerlig och inte antar något begränsat värde i halon till x_0. Detta innebär naturligtvis att funktionen har en pol i x_0 och därmed går mot oändligheten i halon. Trots att en funktion far iväg mot oändligheten nära en pol så är den trots allt mycket mer lik en analytisk funktion än en funktion som har en väsentlig singularitet. En funktion med en pol av ordning n_0 är helt enkelt en analytisk funktion delat med $(z - z_0)^{n_0}$ där z_0 är polen. Polen kan kancelleras genom att multiplicera med $(z - z_0)^{n_0}$ och på så sätt erhålls en analytisk funktion.

Det sista fallet är en icke-standardmotsvarighet till satsen om öppna avbildningar och beskriver egenskaperna för en analytisk funktion som är S-kontinuerlig i x_0.

Sats 25. Icke-standardvarianten av satsen om öppna avbildningar. *Om f är S-kontinuerlig i x_0 och om $sh(f)$ inte är konstant så gäller att $f(hal(x_0)) = hal(f(x_0))$.*

Som avslutning till detta avsnitt presenteras här ytterligare ett exempel på hur elegant icke-standardanalysen kan tillämpas för att bevisa kända klassiska resultat. Vi har valt ett kort bevis av Liouvilles sats.

Sats 26. Liouvilles sats. *En begränsad hel funktion är konstant.*

Bevis. Låt f vara en hel begränsad (standard) funktion. Ta sedan ett tal $M \in {}^*\mathbb{R}_\infty$, och ansätt funktionen $g(z) = f(Mz)$, då är g begränsad och därmed är även dess derivata begränsad i $\mathbb{C}$. Dess derivata är $g'(z) = Mf'(Mz)$, eftersom g' är begränsad så antar f' endast infinitesimala värden. Men eftersom $f' = sh({}^*f')$, dvs antar standardvärden, så måste $f' = 0$ i $\mathbb{C}$ och därmed är f konstant. □

7.4 Standardanalys och Picards stora sats

Beviset för icke-standard Picard har en motsvarighet på standardsidan där man istället för Robinsons sats går via Montels sats, och teori rörande normalfamiljer. Det är detta bevis och kringliggande teori som kommer behandlas i detta avsnitt. Satser, bevis och definitioner i detta kapitel kommer om inte annat nämns från [7].

Först har vi Riemanns sats om hävbar singularitet som används i beviset. Antagandet som vi utgår ifrån är att funktionen inte antar två skilda värden och utifrån det visas att funktionen måste ha en hävbar singularitet, via denna sats så räcker det med att visa att f är begränsad i en omgivning kring singulariteten.

Sats 27. Riemann hävbar singularitet. *Låt $U \subset \mathbb{C}$, $z_0 \in U$, och låt $f : U \setminus \{z_0\} \to \mathbb{C}$ vara en analytisk funktion. Då är följande påståenden ekvivalenta:*

1. *f kan utvidgas till en analytisk funktion $f : U \to \mathbb{C}$.*
2. *f kan utvidgas till en kontinuerlig funktion $f : U \to \mathbb{C}$.*
3. *f är begränsad i en omgivning till z_0.*
4. *$\lim_{z\to a}(z - z_0)f(z) = 0$.*

Bevis. (1) ⇒ (2) ⇒ (3) ⇒ (4) följer av definitionerna av analytiska och kontinuerliga funktioner. Vi visar sedan att (4) ⇒ (1), och då är ekvivalensrelationen bevisad.

Utgående från (4) så har vi att för alla $n \in \mathbb{N}^+$

$$\mathrm{Res}((z - z_0)^n f(z), z_0) = \frac{1}{2\pi i} \lim_{\delta\to 0^+} \int_{|z-z_0|=\delta} (z - z_0)^n f(z)dz,$$

där denna integral kan uppskattas enligt följande

$$|\frac{1}{2\pi i} \int_{|z-z_0|=\delta} (z - z_0)^n f(z)| \leq \frac{1}{2\pi} 2\pi\delta \max_{|z-z_0|=\delta} |(z - z_0)^n f(z)|$$

$$\delta \max_{|z-z_0|=\delta} |(z - z_0)^n f(z)| = \delta^n \max_{|z-z_0|=\delta} |(z - z_0)f(z)|$$

där $\lim_{z\to a}(z - z_0)f(z) = 0$. Eftersom residyn av $(z - z_0)^n f(z)$ motsvaras av de negativa koefficienterna i Laurent serien, så är dessa noll och f har en hävbar singularitet i z_0. □

Nedan följer en sats som är en svagare variant av Picards stora sats där funktionen kommer godtyckligt nära vilket komplext tal som helst i varje omgivning till en väsentlig singularitet, beviset för denna sats är mycket lik det för icke-standardformuleringen av satsen.

Sats 28. Casorati-Weierstrass. *Givet en analytisk funktion f med en väsentlig singularitet i z_0 så finns det för alla $\epsilon > 0$, $\delta > 0$ och ett komplext tal w ett tal z sådant att $|z - z_0| < \delta$ och $|f(z) - w| < \epsilon$. Vilket med andra ord säger att f kommer godtyckligt nära vilket komplext tal som helst i varje område till z_0.*

Bevis. Antag att f är analytisk i $V \setminus \{z_0\}$ och att z_0 är en väsentlig singularitet. Antag att det finns ett tal b som f inte kommer godtyckligt nära. $|f(z) - b| \geq \epsilon$ för alla $z \in V$.

Då är funktionen

$$g(z) = \frac{1}{f(z) - b}$$

analytisk i $V \setminus \{z_0\}$ begränsad av $1/\epsilon$. Det existerar därmed en analytisk fortsättning av g till hela V via Riemanns sats om hävbar singularitet.

$$f(z) = \frac{1}{g(z)} + b$$

för alla $V \setminus \{z_0\}$. Då finns två möjligheter antingen $\lim_{z\to z_0} g(z) = 0$ då har f en pol i z_0, eller så är gränsvärdet något annat tal och då är f begränsad i z_0. Både möjligheter ger att f inte har en väsentlig singularitet i z_0. □

I beviset av Picards sats visas att f är begränsad på randen till en omgivning kring z_0 via följande sats vet man då också att funktionen är begränsad i hela omgivningen. Denna sats antar vi läsaren är bekant med sedan innan och bevisas inte här.

Sats 29. Maximumprincipen. *Låt $U \subseteq \mathbb{C}$ vara en begränsad mängd, och låt f vara en kontinuerlig funktion på $\bar{U}$ som är analytisk på U. Det största värdet av $|f|$ på $\bar{U}$ finns på $\partial\bar{U}$.*

Ovanstående satser kommer användas för att visa att f måste ha en hävbar singularitet under antagandet att det finns två skilda tal i $\mathbb{C}$ som f inte antar. För att nå fram till den punkt då dessa satser kan tillämpas så går vi här via normalfamiljer. I beviset visas att f tillhör en normalfamilj under det nämnda antagandet och sedan använder man de egenskaper som en normalfamilj av funktioner har för att visa att den är begränsad. Vi börjar här med att definiera en normalfamilj, för att sedan närmare utforska dess egenskaper och villkor på familjen $\mathcal{F}$ som medför att den är normal.

Definition 26. Normalfamilj. *En familj $\mathcal{F}$ av analytiska funktioner som är definierade på ett område $U \subseteq \mathbb{C}$, sådana att $f : U \to Y$, där $Y \subseteq \mathbb{C}$, där varje sekvens av funktioner i $\mathcal{F}$ innehåller en delsekvens som konvergerar likformigt till en analytisk funktion på kompakta delmängder av X, (denna funktion får även vara $f \equiv \infty$).*

Ett exempel på en normalfamilj av funktioner fås om vi tar en funktion f som är analytisk och S-kontinuerlig på U, tag sedan en funktionsserie av denna funktion definierad som $\{f_n(z) = f(z + z/n)\}$ på $(z + z/n) \in U$ för $n = 1, 2, 3, \ldots$. Det existerar då delsekvenser av denna serie som konvergerar likformigt mot f och vi har en normalfamilj av funktioner $\{f_n(z + z/n)\}$.

Det följande nya begreppet ekvikontinuitet och satsen av Arzelà-Ascoli används för att bevisa Montels sats som vi behöver i beviset av Picards sats.

Definition 27. Ekvikontinuitet. *En familj $\mathcal{F}$ sägs vara ekvikontinuerlig i en punkt z om för varje $\epsilon > 0$ så existerar ett δ (oberoende av n,z och z') sådant att $|f_n(z) - f_n(z')| < \epsilon$ då $|z - z'|\delta$ för varje $f_n \in \mathcal{F}$.*

Ekvikontinuiteten används nu tillsammans med villkoret att funktionerna är punktvis begränsade som ett tillräckligt krav för normalitet för en familj funktioner.

Sats 30. Arzelà-Ascoli. *Givet kompakta mängder X,$T \subseteq \mathbb{C}$ och $\mathcal{F}$ en familj av kontinuerliga funktioner från X till T antag att $\mathcal{F}$ är punktvis begränsad och lokalt ekvikontinuerlig då är $\mathcal{F}$ en normalfamilj på X.*

Bevis. Givet en serie funktioner $\{f_n\} \subseteq \mathcal{F}$ så måste här visas att det existerar en delserie som konvergerar likformigt på varje kompakt delmängd av X. Definiera en tät delmängd till X som $\{x_k\}, k = 1, 2, 3, \ldots$. Tag en serie $\{f_n(x_0)\}_{n\in\mathbb{N}}$, denna serie är begränsad enligt antagandena i satsen, därmed finns det enligt Bolzano-Weierstrass sats en konvergent delserie där index varierar över $N_0 \subset \mathbb{N}$. Denna delserie är begränsad i x_1 och därmed finns det ytterligare en konvergent delserie definierad av $N_1 \subseteq N_0$, fortsätter man på samma sätt för alla k så fås en sekvens av oändliga mängder $N_0 \supset N_1 \supset N_2 \ldots$. Det finns därmed en konvergent delserie för varje x i $\{x_k\}_{k\in\mathbb{N}}$. Betrakta nu serien av strikt växande naturliga tal $\{n_l\}_{l\in\mathbb{N}}$, där n_l är element l i mängden N_l. Den diagonala serien $\{f_{n_l}(x_k)\}_{l\in\mathbb{N}}$ konvergerar därmed för varje $k \in \mathbb{N}$.

Nu återstår att visa likformig konvergens av denna serie på kompakta delmängder av X. Här använder vi oss av ekvikontinuiteten, om vi tar $K \subseteq X$ en kompakt delmängd och $\epsilon > 0$ så har vi via ekvikontinuitet att det för alla $a \in K$ och alla $n \in \mathbb{N}$ finns $\delta > 0$ sådant att

$$|f_n(a) - f_n(b)| < \epsilon$$

då $b \in X$ och $|a - b| < \delta$, låt D_a beteckna den enhetsskiva med mittpunkt i a och med tillhörande radie δ. Eftersom K är kompakt så finns det ett begränsat antal punkter $a_1, \ldots, a_m \in K$ sådana att $D_{a_1}, \ldots, D_{a_m}$ täcker K. Eftersom $\{x_k\}_{k\in\mathbb{N}}$ är en tät delmängd av X så finns det till varje $j \in 1, \ldots, m$ ett tal $x_{k_j} \in \{x_k\}_{k\in\mathbb{N}}$ sådant att $x_{k_j} \in D_{a_j}$. $\{f_{n_l}(x_k)\}_{l\in\mathbb{N}}$ konvergerar och är därmed en Cauchy följd för alla $j \in \{1, \ldots, m\}$, det finns därmed ett $l_0 \in \mathbb{N}$ sådant att $|f_{n_i}(x_{k_j}) - f_{n_l}(x_{k_j})| < \epsilon$ för alla i,$l \geq l_0$ och för alla $j \in \{1, \ldots, m\}$.

Om vi nu tar ett tal p, sådant att $p \in K$ så existerar det ett $j \in \{1, \ldots, m\}$ sådant att $p \in U_{a_j}$. För i,$l > l_0$ så har vi då att

$$\begin{aligned} |f_{n_i}(p) - f_{n_l}(p)| &\leq |f_{n_i}(p) - f_{n_i}(x_{k_j})| + |f_{n_i}(x_{k_j}) - f_{n_l}(x_{k_j})| + |f_{n_l}(x_{k_j}) - f_{n_l}(p)| \\ &< \epsilon/3, \end{aligned}$$

vilket innebär att $\{f_{n_l}\}_{l\in\mathbb{N}}$ konvergerar likformigt på K. Därmed existerar delserier av $\mathcal{F}$ som konvergerar likformigt på kompakta delmängder av X.

$\square$

Montels sats ger ett tillräckligt villkor på en familj av analytiska funktioner för att denna ska vara normal. Detta kan jämföras med Robinsons sats som säger att om f är analytisk och begränsad i ett område så är f S-kontinuerlig där. Denna sats används alltså för att klassificera funktionsfamiljer som normala.

Sats 31. Montel. *Om $\mathcal{F}$ är en lokalt begränsad familj av analytiska funktioner på en mängd Ω så är $\mathcal{F}$ en normalfamilj i Ω.*

Bevis. Givet förutsättningarna i satsen så visas att $\mathcal{F}$ måste vara ekvikontinuerlig och därmed via Arzelà-Ascolis sats en normalfamilj.

Tag $z_0 \in \Omega$ och välj ett $\epsilon > 0$. Eftersom $\mathcal{F}$ är lokalt begränsad så existerar en konstant $M > 0$ och $r > 0$ där $D(\bar{z_0},r) \subset \Omega$ sådana att $|f(z)| < M$ för alla $z \in D(\bar{z_0},2r)$ och $f \in F$. Cauchys formel ger för alla $f \in F$ och $z,w \in D(z_0,2r)$

$$\begin{aligned}(f(z)-f(w))2\pi i &= \int_{\partial D(z_0,2r)} \frac{f(\zeta)}{\zeta - z} d\zeta - \int_{\partial D(z_0,2r)} \frac{f(\zeta)}{\zeta - w} d\zeta \\ &= (z-w)\int_{\partial D(z_0,2r)} \frac{f(\zeta)}{(\zeta - z)(\zeta - w)} d\zeta.\end{aligned}$$

Om vi sedan begränsar $z,w \in D(z_0,r)$ så gäller det att $|(\zeta - z)(\zeta - w)| > r^2$.

$$|f(z)-f(w)| \leq |z-w| \frac{2r\|f\|_{\partial D(z_0,2r)}}{r^2} < |z-w|\frac{2M}{r}$$

Vi söker nu ett δ så att denna olikhet gäller för alla ϵ i ett område kring z_0. Om vi nu väljer $\delta = \min\{r, \frac{r\epsilon}{4M}\}$

$$|f(z)-f(w)| < \epsilon \ \forall z,w \in D(z_0,\delta)$$

$\square$

Innan vi kan bevisa Fundamentala normalitets testet så behöver vi en till sats som nog inte alla läsare är bekanta med, denna sats bevisas dock inte här [4].

Sats 32. Hurwitz. *Låt $\{f_n\}$ vara en serie analytiska funktioner på en sammanhängande öppen mängd Ω som konvergerar likformigt på kompakta delmängder av Ω till en icke-konstant analytisk funktion $f(z)$. Om $f(z_0) = 0$ för något $z_0 \in \Omega$ så gäller för alla $r > 0$ tillräckligt små att det finns ett N som beror på r sådant att för alla $n > N$ så har $f_n(z)$ samma antal nollställen i $D(z_0, r)$ som $f(z)$.*

Nu kommer vi till en sats som tar oss ett stort steg närmare resultatet i Picards stora sats, denna sats karaktäriserar analytiska funktionsfamiljer för vilka det existerar två olika värden i $\mathbb{C}$ som funktionerna inte antar på ett område $\Omega \subseteq \mathbb{C}$.

Sats 33. Fundamentala normalitets testet. *Låt $\mathcal{F}$ vara en familj av analytiska funktioner definierad på en mängd $\Omega \subseteq \mathbb{C}$. Om det finns två komplexa värden a och b som varje $f \in \mathcal{F}$ inte antar så är $\mathcal{F}$ en normalfamilj på Ω.*

Bevis. Det räcker att anta att $a = 0$ och $b = 1$ ty detta kan uppnås med en linjär avbildning. Vi kan anta att Ω är enhetsskivan dvs att $|z| < 1$ eftersom normalitet är en lokal egenskap.Vi antar alltså att

$$f : D \to \mathbb{C} \setminus \{0,1\}.$$

Elliptiska modulärfunktionen är en övertäckningsavbildning av $\mathbb{C} \setminus \{0,1\}$ från D därmed kan vi definiera en begränsad funktion $\tilde{f}$ som beror på $f \in \mathcal{F}$, det faktum att denna funktion är begränsad används sedan i kombination med Montels sats för att visa att serier av $\tilde{f}$

[4] Se [7] för bevis.

konvergerar mot en begränsad analytisk funktion och utifrån detta visas det att även f måste vara begränsad om den undviker två värden a och b.

Låt $\Pi(\zeta) = w$ vara elliptiska modulärfunktionen och ta ett $z_0 \in \Omega$. För ett godtyckligt $f \in \mathcal{F}$ så väljes den enkelvärda analytiska grenen av $\nu(w) = \Pi^{-1}(w)$ för vilken $\nu(f(z_0))$ är principalvärdet, då kan $\nu(f(z))$ definieras entydigt i ett område kring z_0. Därmed kan $\nu(f(z))$ ges en analytisk fortsättning i hela Ω enligt sats (19), kan en enkelvärd analytisk funktion definieras som $\tilde{f}(z) = \nu(f(z))$, denna funktion uppfyller $|\tilde{f}(z)| < 1$, för $z \in \Omega$ och $f \in \mathcal{F}$. Låt $\{f_n\} \subseteq F$ vara en godtycklig serie i $\mathcal{F}$ och låt $\alpha \in \hat{\mathbb{C}}$ vara en hopningspunkt till serien $\{f_n(z_0)\}$. Beroende på värdet av detta α så betraktar vi fyra olika fall.

1. $\alpha \neq 0{,}1{,}\infty$

   Låt $\{f_{n'}\} \subseteq \{f_n\}$ sådan att $\{f_{n'}(z_0)\} \to \alpha$ då $n' \to \infty$. Eftersom $\tilde{f}$ är begränsad så ger Montels sats att $\{\tilde{f}_{n'}\}$ har en delserie $\{\tilde{f}_{n_k}\}$ som konvergerar likformigt på kompakta delmängder av Ω mot en analytisk funktion $F(z)$ och $|F(z)| \leq 1$ enligt ovan. Om likhet gäller dvs. $|F(z)| = 1$ så innebär det att $|\tilde{f}_{n_k}(z)| \to 1$ då $k \to \infty$ men detta sker endast om $f_{n_k} \to 0{,}1$ och ∞, så vi har alltså en strikt olikhet $|F(z)| < 1$ i Ω. Givet en godtycklig kompakt mängd $K \subseteq \Omega$ och en begränsad konstant m så har vi $|F(z)| \leq m < 1$ på K. Likformig konvergens ger att det finns m' sådant att $|\tilde{f}_{n_k}(z)| \leq m' < 1, z \in K$ för alla k tillräckligt stora. $\mu(\zeta)$ är analytisk i D och därmed likformigt begränsad i $|\zeta| \leq m' < 1$, ta M sådant att $|\mu(\zeta)| \leq M$. Då gäller för tillräckligt stora k och $z \in K$ att $|f_{n_k}(z)| = |\mu(\nu(f_{n_k}(z)))| \leq M$, så $\{f_{n_k}\}$ är begränsad på K dvs. kompakta delmängder av Ω. Därmed har $\{f_{n_k}\}$ en delsekvens som konvergerar likformigt på kompakta delmängder till en analytisk funktion.

2. $\alpha = 1$

   Låt $\{f_{n'}\} \subseteq \{f_n\}$ sådan att $f_{n'}(z_0) \to 1$ då $n' \to \infty$ och definiera $g_{n'}(z) = \sqrt{f_{n'}(z)}$ sådan att $\lim_{n' \to \infty} g_{n'}(z_0) = -1$. Då är $g_{n'}$ där $n' = 1, 2, 3, ...$ analytiska i Ω och antar inte värdena 0 och 1. Enligt 1 så finns det $\{g_{n_k}\} \subseteq \{g_{n'}\}$ som konvergerar likformigt på kompakta delmängder av Ω till en analytisk funktion och då gör även $\{g_{n_k}^2\}$ det.

3. $\alpha = 0$

   Välj en delserie $\{f_{n'}\} \subseteq \{f_n\}$ sådan att $f_{n'}(z_0) \to 0$ då $n' \to \infty$ och definiera $g_{n'}(z) = 1 - f_{n'}(z)$, $n' = 1{,}2{,}3{,}...$ då är funktionerna $g_{n'}(z)$ analytiska och $\lim_{n' \to \infty} g_{n'} = 1$ och därmed hamnar detta fall under 2. Via del 2 så konvergerar delserier av $\{g_{n'}(z)\}$ likformigt på kompakta delmängder av Ω och då gäller det även för den ursprungliga funktionsserien $\{f_{n'}\}$.

4. $\alpha = \infty$

   Välj en delserie $\{f_{n'}\} \subseteq \{f_n\}$ sådan att $f_{n'}(z_0) \to \infty$ då $n' \to \infty$. Ansätt funktionerna $g_{n'}(z) = 1/f_{n'}(z)$, $n' = 1{,}2{,}3, \ldots$, dessa är analytiska i Ω och undivker 0 och 1, med $\lim_{n' \to \infty} g_{n'} = 0$ och faller in under fall (3). Därmed konvergerar en delserie till en analytisk funktion g som har värdet $g(z_0) = 0$, men eftersom funktionerna $g_{n'}$ inte har några nollställen i Ω så ger Hurwitz sats att g inte heller har det eller så är g konstant, så $g \equiv 0$ i Ω och därmed $f_{n_k} \to \infty$ likformigt(oberoende av z) på kompakta delmängder av Ω.

□

Följande bevis av Picards sats går via normalfamiljer och följer samma väg som om man går via Robinsons sats på icke-standardsidan. Beviset bygger i stort på fundamentala normalitets testet, som i sin tur bygger på Montels sats. Den väg vi gått i detta kapitel har låtit oss nå fram till fundamentala normalitets testet vilken nu kan ta oss igenom beviset av Picards stora sats.

Sats 34. Picards stora sats. *$f(z)$ är analytisk i ett område $K_\rho = \{z : 0 < |z - z_0| < \rho\}$ och z_0 är en isolerad väsentlig singularitet till $f(z)$. Då gäller det att $f(z)$ antar alla värden förutom som mest ett i K_ρ.*

Bevis. Vi antar att f är analytisk i ett område $D = \{z \in \mathbb{C} : |z| < R\}$ och att $z_0 = 0$, detta antagande ger inga inskränkningar utan kan åstadkommas via en affin avbildning.

Vi antar att $f : D \to \mathbb{C} \setminus \{a{,}b\}$ dvs. att det existerar två skilda värden $\{a{,}b\}$ som f inte antar i D, vi vill då visa att detta antagande leder till att f har en hävbar singularitet eller en pol i z_0.

Familjen av funktioner $\mathcal{F}$ definierade som

$$f_n = f(\frac{z}{2^n}),\ n = 1, 2, 3, ...$$

är analytisk i annulusen $A = \{\frac{R}{2} < |z| < R\}$ och funktionerna antar inte a eller b där. Därmed är via fundamentala normalitets testet (sats (33)) $\mathcal{F}$ en normalfamilj på A. Normaliteten innebär att det på kompakta delmängder existerar en delserie som konvergerar likformigt. Mängden $\{|z| = r : \frac{R}{2} < r < R\}$ är kompakt och på denna existerar en likformigt konvergent serie $\{f_{n_k}\}$, låt $F(z)$ antingen vara den analytiska funktion eller $F(z) \equiv \infty$ som serien konvergerar mot i A.

Om $F(z)$ är en analytisk funktion på A så innebär det att den är begränsad på den kompakta mängden $\{|z| = r : \frac{R}{2} < r < R\}$ och därmed är $\{f_{n_k}\}$ likformigt begränsad där. Därmed har vi att

$$|f_{n_k}(z)| \le M,\ |z| = r,\ n = 1, 2, 3, ...$$

och då följer det att

$$|f(z)| \le M,\ |z| = \frac{r}{2^{n_k}},\ n = 1, 2, 3, ...$$

vilket innebär att f är begränsad på en serie cirklar centrerade runt origo vars radie går mot 0. Maximumprincipen ger att det största värdet antas på randen till ett område, vilket innebär att f är begränsad i annulusarna som två följande cirklarna definierar. Detta medför att vi har följande resultat

$$|f(z)| \le M,\ 0 < |z| < \frac{r}{2^{n_1}},\ n = 1, 2, 3, ...$$

vilket via Riemanns sats om hävbar singularitet (sats (27)) ger att f har en hävbar singularitet i 0, vilket är en motsägelse.

Om vi istället antar att $F(z) \equiv \infty$, så kommer $\frac{1}{f_{n_k} - a}$ konvergera likformigt mot 0 på kompakta delmängder av A. Som ovan drar vi då slutsatsen att funktionen $\frac{1}{f(z) - a}$ är begränsad i ett område till origo och därav följer det att $f(z)$ har en hävbar singularitet där, även i detta fall så fås en motsägelse. □

7.5 Jämförelse mellan icke-standard och standard Picards stora sats

Detta avsnitt berör likheter såväl som betydande skillnader mellan Picards stora sats och tillhörande bevis av satsen på icke-standard och standardsidan.

Vi har alltså två satser med likartade resultat som antar att en analytisk funktion har en väsentlig singularitet i en punkt respektive att funktionen är begränsad i en punkt och inte S-kontinuerlig där. Dessa båda antaganden visar sig vara ekvivalenta. Att en funktion inte är S-kontinuerlig i en punkt måste i sig inte innebära att den har en väsentlig singularitet där. Är däremot funktionen dessutom begränsad i en punkt i halon så kan funktionen inte ha en pol där, ty i en infinitesimal omgivning till en pol kommer funktionen att anta oändligt stora värden.

Om en funktion har en väsentlig singularitet så är den ej S-kontinuerlig där och det finns begränsad värden i halon till den väsentliga singulariteten detta innebär att om villkoren i standard Picards sats är uppfyllda så är de även det i icke-standardversionen.

I beviset av Picards stora sats så ansätter man på standardsidan serier av funktioner medan man med icke-standardmetodik kan titta direkt på funktionen och dess direkta egenskaper. För att närmare förstå hur funktionsserierna är relaterade till icke-standardsidan så tittar vi här närmare på sambandet mellan likformig konvergens av en funktionsserie och S-kontinuitet av den funktion som man kan bygga en sådan serie utav.

Sats 35. Likformig konvergens av en funktionsserie. $\{f_n : n \in \mathbb{N}\}$ *konvergerar likformigt till funktionen* $f : A \to \mathbb{C}$, $A \subseteq \mathbb{C}$ *omm för varje* $z \in {}^*A$ *och varje* $n \in {}^*\mathbb{N} \setminus \mathbb{N}$ *så är* $f_n(z) \simeq f(z)$. [3]

Här följer ett enkelt exempel för att belysa sambandet mellan S-kontinuitet och ovanstående sats (35). Givet f analytisk och begränsad i V, ansätt funktionen $f_n(z) = f(z + \frac{z}{n})$, $\{f_n | n \in \mathbb{N}\}$ utgör då en familj av funktioner. För $n \in {}^*\mathbb{N}^+ \setminus \mathbb{N}$ så är $\frac{z}{n} \simeq 0$ givet att denna familj är normal så är $f(z + \frac{z}{n}) \simeq f(z)$ detta ger att f är S-kontinuerlig. Om vi istället ansätter samma funktionsserie som ovan och antar att f är S-kontinuerlig så fås att familjen som serien utgör är normal.

I standardframställningen så används begreppet ekvikontinuitet som ett verktyg för att bevisa Montels sats. Detta begrepp och Arzelà-Ascolis sats gör att det räcker att visa att funktionsfamiljen är ekvikontinuerlig för att nå slutsatsen att den måste vara en normalfamilj. Det finns ett tydligt samband mellan S-kontinuitet och ekvikontinuitet som berörts tidigare men här formulerar vi detta formellt i en sats enligt [8].

Sats 36. S-kontinuitet och ekvikontinuitet. *Familjen* $\mathcal{F} = \{f_n\}, n = 1{,}2{,}3\ldots$ *är ekvikontinuerlig på* Ω *omm* f_n *är S-kontinuerlig på* ${}^*\Omega$ *för varje* $n \in {}^*\mathbb{N}^+$.

Bevis. Antag att $\{f_n\}$ är ekvikontinuerlig på Ω och tag ett godtyckligt $\epsilon > 0$, då existerar $\delta > 0$ sådant att

$$(\forall n \in \mathbb{N}^+)(\forall p{,}q \in \Omega)(|p - q| < \delta \Rightarrow |f_n(p) - f_n(q)| < \epsilon).$$

Används transfer på denna formel så fås att $(|p - q| < \delta \Rightarrow |f_n(p) - f_n(q)| < \epsilon)$ för $p{,}q \in {}^*\Omega$ och alla $n \in {}^*\mathbb{N}^+$. Eftersom ϵ är godtyckligt litet så gäller det att $p \simeq q \Rightarrow f_n(p) \simeq f_n(q)$. f_n är därmed S-kontinuerlig om $\{f_n\}$ är ekvikontinuerlig. Antages istället att alla funktioner i familjen $\{f_n\}$ är S-kontinuerliga i ${}^*\Omega$ så har vi via transfer att familjen är ekvikontinuerlig. □

Sats 36 och följande sats om kontinuerlig skugga kan tillsammans användas för att bevisa Arzelà-Ascoli, dessa kan ses som en icke-standardmotsvarighet till denna sats. Värt att notera är att dessa satser inte behövdes för beviset på icke-standardsidan.

Sats 37. Kontinuerlig skugga. *Låt* $X \subseteq \mathbb{C}$ *vara en kompakt mängd. Låt* f *vara en S-kontinuerlig funktion på* ${}^*X$ *och låt* $f(p)$ *vara begränsad för alla* $p \in X$. *Låt* $F(p) = sh(f(p))$ *för alla* $p \in X$. *Då är* F *kontinuerlig på* X *och* $F(p) \simeq f(p)$ *för alla* $p \in {}^*X$. [8]

Sambandet mellan sats 36, 37 och Arzelà-Ascoli kan tydligt ses om man ser på följande icke-standardbevis av Arzelà-Ascoli som helt bygger på dessa två satser.

Icke-standardbeviset av Arzelà-Ascoli [8]: Låt $\{f_n\}_{n\in\mathbb{N}}$ vara en ekvikontinuerlig familj, ta $m \in {}^*\mathbb{N} \setminus \mathbb{N}$. Enligt ekvikontinuitets satsen (sats 36) så har vi att f_m är S-kontinuerlig, $F(p) = sh(f_m(p))$ för alla $p \in X$. Vi har därmed för $m{,}k \in \mathbb{N}^+$ att

$$(\exists q \in {}^*\mathbb{N})(q > m)(\forall x \in {}^*X(|F(x) - f_q(x)| < 1/k)).$$

Transfer av detta ger sedan att om vi låter $q : N^2 \to N$, $q(m{,}k) > m$ och sätter $n_0 = 1$, $n_{i+1} = q(n_i, i+1)$. Så får vi en strikt växande serie tal $\{n_i\}_{i\in\mathbb{N}}$ och $|F(x) - f_{n_i}(x)| < 1/i$ för $i \geq 1$. Därmed konvergerar denna delserie likformigt mot F på X och Arzelà-Ascolis sats är därmed bevisad med icke-standardmetodik.

Istället för att titta på serier av funktioner så tittar man på icke-standardsidan på en unik funktion och dess kontinuerliga skugga. Den egenskap som i sammanhanget blir viktig på icke-standardsidan för att beskriva funktioner är S-kontinuiteten.

Arzelà-Ascoli ger att en ekvikontinuerlig sekvens av funktioner är en normalfamilj. Vi ser på så sätt likheten mellan S-kontinuitet och normalfamiljer. Denna koppling gör att vi ser att Montel och Robinsons sats är varandras motsvarighet på standard och icke-standardsidan. Om en analytisk funktion f är begränsad på ett område så är den S-kontinuerlig där, och om en analytisk funktionsserie är begränsad så utgör den en normalfamilj. Båda satserna utnyttjar det faktum att f är begränsad och analytisk för att säga något om dess normalitet. Som vi såg ovan om vi ansätter en funktionsserie av $f_n(z) = f(z + z/n)$ så får vi att om f är begränsad och analytisk så är den S-kontinuerlig och dess serie $\{f_n\}$ utgör en normalfamilj.

I beviset av Picards sats så visar man att om en analytisk funktion f inte antar två värden a,b så är f S-kontinuerlig detta är analogt med satsen om fundamentala normalitets testet, men beviset och resultatet i den satsen ingår alltså i icke-standardversion av beviset av Picards sats. I båda fallen använder vi oss av ett lyft under antagandet att funktionen f undviker två värden, detta lyft gör att vi lyckas visa att f måste vara begränsad med skillnaden att vi i standardfallet ansätter en funktionsserie utgående från f. Resultatet att f är begränsad ger sedan den motsägelse som gör att Picards sats är bevisad.

Med hjälp av icke-standardanalys har vi fått ett nytt sätt att se på analytiska funktioner, detta kan sammanfattas i en sammansättning av tidigare nämnda satser nämligen Robinson-Callots sats [2].

Sats 38. Robinson-Callot. *Låt f vara en analytisk funktion på en mängd som innehåller x_0 och $hal(x_0)$, där x_0 och $f(x_0)$ är begränsade.*

1. *Om f bara antar begränsade värden i halon av x_0 så existerar ett standardomgivning V av x_0 på vilken:*
   (a) *f är S-kontinuerlig*
   (b) *$sh(f)$ är analytisk*
   (c) *$(sh(f))^{(n)} = sh(f^{(n)})$ där $n \in \mathbb{N}$.*
2. *Om f är S-kontinuerlig i x_0 och om $sh(f)$ inte är konstant så gäller att $f(hal(x_0)) = hal(f(x_0))$.*
3. *Om f inte är S-kontinuerlig i x_0, så gäller att $f(hal(x_0))$ innehåller alla begränsade komplexa tal förutom möjligen en del av halon till en punkt.*

Denna sats ger en beskrivning av hur en analytisk funktion beter sig i en given punkt och dess halo, den är antingen helt oregelbunden eller helt regelbunden. Satsen sammanfattar en del av teorin kring analytiska funktioner och beskriver hur dessa beter sig på ett sätt som det inte finns någon direkt motsvarighet till på standardsidan.

7.6 Slutsats

Robinsons sats visade sig motsvara teorin om normala familjer. Att bevisa Picards sats blir lättare med hjälp av icke-standardanalys, tack vare att man kan se på funktionen direkt och att man inte behöver ansätta funktionsserier för att kunna säga något om hur funktionen beter sig. Sammanfattningsvis kan vi alltså konstatera att Picards sats på standard- och icke-standardsidan är ekvivalenta. Icke-standardförfarandet tycks lämpa sig bättre för att bevisa Picards sats och detta motiveras främst av att det är enklare att förstå beviset, beviset blir kortare och det är enklare att se samband mellan andra delar av teorin för analytiska funktioner.

Nu är det inte så att det alltid lönar sig att tillämpa ett icke-standardförfarande, men i vissa fall har det visat sig mycket kraftfullt (av bland annat de anledningar som nämnts här). Författarna vill då poängtera att naturligtvis handlar inte icke-standardanalys om att välja sida, eller att helt överge de välkända metoderna - utan icke-standardanalysen bör snarare ses som ett verktyg, ett verktyg som kan vara mycket användbart om det används på rätt sätt.

8 Exempel 2: Finansiell matematik

Icke-standardanalys har använts framgångsrikt inom matematisk statistik och i den andra fördjupningen studerar vi grundläggande tillämpningar inom finansiell matematik.

Finansiell matematik handlar till stor del om att modellera värdet på finansiella tillgångar samt prissätta derivat på dessa. Mycket av teorin kretsar kring hur man finner en teoretisk kurs för optioner på aktier, vilket är ett fundamentalt fall.

Black och Scholes (1973) [19] var först med att publicera en formel för värdering av optioner. Deras arbete ansågs vara banbrytande och ett stort genombrott inom fältet. Cox, Ross och Rubinstein (1979) [18] använde ett annat tillvägagångssätt och presenterade den så kallade binomialmodellen. Skillnaden mellan Black-Scholes modell och binomialmodellen består främst i att den förstnämnda är kontinuerlig medan den sistnämnda är diskret. Binomialmodellen belyser tydligare de ekonomiska principerna och är enklare att ta till sig. För att kunna användas måste dock antalet tidsperioder i den diskreta prisformeln gå mot oändligheten varpå den sammanfaller med Black-Scholes formel.

Vi vill med icke-standardanalys visa hur lika och tätt sammankopplade modellerna egentligen är. Inom ramen för den här fördjupningen kommer vi att introducera binomialmodellen för att sedan skapa den hyperfinita binomialmodellen. I samband med det kommer nya begrepp inom icke-standardanalysen såsom interna mängder och Loeb-rum att presenteras. I slutändan vill vi visa att den hyperfinita binomialmodellen rymmer prisformeln för både binomialmodellen och Black-Scholes modell.

8.1 Introduktion till finans

Först och främst skall vi nu introducera Cox, Ross och Rubinsteins binomialmodell [18] och tillhörande ekonomiska begrepp. Notera att vi inledningsvis använder begränsat med matematik och istället tydliggör principerna för att senare definiera begreppen på ett mer precist sätt.

Binomialmodellen är som nämnt en diskret modell för prissättning av optioner. Den bygger på ett antal antaganden om en finansiell marknad där tre typer av instrument existerar. Dessa är obligationer, aktier och optioner.

8.1.1 Instrument och antaganden

Obligationer ställs ut av bland annat stater och företag som ett medel för att låna pengar. Köparen av en obligation får rätten att vid ett senare tillfälle T lösa in obligationen i utbyte mot pengarna han lånat ut plus en ränta. Obligationens värde B som en funktion av tiden kan skrivas som

$$B(t+1) = rB(t),\ t = 0, 1, \ldots, T-1 \tag{26}$$

där r är räntan per tidsenhet och antas vara konstant. $B(0)$ är följaktligen den summa som obligationen ställs ut för.

Aktier hanteras i modellen som stokastiska processer och värdet vid tiden t betecknas i modellen $S(t)$. En akties värde kan skrivas som

$$S(t+1) = X_{t+1}S(t),\ t = 0, 1, \ldots, T-1 \tag{27}$$

där $X_1, X_2, \ldots, X_T$ är oberoende och likafördelade slumpvariabler som följer en bernoullifördelning där utfallen är de reella talen u och d där $u > d$. Vi har

$$q = P[X_t = u]$$

Förändringen i aktiens värde mellan $S(t)$ och $S(t+1)$ kan grafiskt representeras som

$$\text{S(t)} \quad \begin{cases} uS(t) \text{ med sannolikhet } q \\ dS(t) \text{ med sannolikhet } 1-q \end{cases}$$

En option är ett finansiellt derivat. Det finns olika typer av optioner och vi intresserar oss här för en europeisk köpoption med en aktie som underliggande tillgång. Ägaren av en sådan option har rätten men inte skyldigheten att vid en framtida tidpunkt ("lösendagen") köpa aktien till ett förutbestämt pris ("lösenpriset") oberoende av vad aktiekursen då står i. Det finns också säljoptioner, som ger rätten att sälja aktien. Optioner kan vara amerikanska vilka skiljer sig från europeiska genom att ägaren har rätt att lösa in optionen när som helst fram till och med den förutbestämda tidpunkten istället för enbart vid tidpunkten. Hädanefter så antas optionen alltid vara europeisk.

En köpoption betecknas c efter engelskans ”call” och har fyra parametrar $c(t, T, S(t), K)$ där t är nuvarande tid, T lösendagen, $S(t)$ aktiens pris vid t och K lösenpriset. Ofta anges bara en parameter, tiden. En köpoptions utbetalningsfunktion (värdet på slutdagen) är

$$c(T) = max(0, S(T) - K) = (S(T) - K)^+$$

Om aktiens pris på slutdagen är högre än lösenpriset, så kommer ägaren av optionen att lösa in sin option mot rätten att köpa en aktie för lösenpriset och han tjänar då mellanskillnaden mot vad köpet annars hade kostat. Är priset däremot lägre, så väljer han att inte lösa in sin option. En options värde på lösendagen är alltid känt, och förändringen i värde mellan $T-1$ och T kan illustreras som

$$\begin{array}{lll} & & (uS(T-1) - K)^+ \text{ med sannolikhet } q \\ & \nearrow & \\ \text{c(T-1)} & & \\ & \searrow & \\ & & (dS(T-1) - K)^+ \text{ med sannolikhet } 1 - q \end{array}$$

Frågan som binomialmodellen så väl som Black-Scholes modell försöker besvara är vad värdet vid andra tidpunkter än T är. Intuitivt kan man tycka att en sådan prissättning borde vara beroende av en mängd parametrar såsom till exempel aktiernas sannolikhetsfördelning och investerarnas riskvillighet. Den möjligtvis viktigaste insikten i finansiell matematik är att det inte krävs någon sådan information, givet några troliga antaganden om marknaden.

Definition 28. Antaganden om marknaden.

I: Aktiepriset $S(t)$ är en binomialprocess.

II: Räntan är konstant. Man kan låna och låna ut obegränsade summor på denna räntenivå.

III: Marknaden är avgiftsfri. Inga skatter, courtage eller likvärdigt existerar.

IV: Det är tillåtet att äga bråkdelar av aktier och det är även tillåtet att låna ut aktier (äga ett negativt antal aktier).

V: Marknaden är arbitragefri, det går inte att göra riskfria vinster.

Vad det innebär att marknaden är arbitragefri kommer senare att definieras precist, men innebörden är att inga strategier på marknaden existerar så att riskfria vinster kan erhållas. En följd av detta antagande är till exempel att $u > r > d$.

Sats 39. *I en arbitragefri marknad så gäller att $u > r > d$.*

Bevis. Antag att $r < d$. Om man då ställer ut obligationer (lånar pengar) och köper aktier för de pengar det inbringar så har man en portfölj bestående av en skuld och aktier där aktiernas värde motsvarar skulden.

$$S(0) - B(0) = 0$$

Skulden skall betalas med räntan r efter en tidsperiod och aktiens värde kan då antingen ha utvecklats till $uS(0)$ eller $dS(0)$. Men om $r < d$ betyder det att portföljens värde efter en tidsperiod då aktien gått dåligt är

$$dS(0) - rB(0) > 0$$

Alltså finns möjlighet till en riskfri vinst och påståendet är bevisat för den ena olikheten. På samma sätt visar man att $u > r$ genom att skapa en portfölj där man går kort i aktier (”blankar”) och köper obligationer. □

Innan modellens lösning på problemet presenteras så följer här ett exempel på varför prissättning av optioner är ett i praktiken relevant problem.

Exempel 3. *Ett skeppsvarv i Göteborg har fått i uppdrag att bygga en stor och dyr båt åt en kund i Amerika. Efter att beställningen bekräftats och priset bestämts så kommer bygget att pågå i ett år men betalningen sker vid leverans och i dollar. Eftersom båten byggs i Sverige så är alla utgifter i kronor. Det här gör varvets finanschef orolig, för han tror att dollarn kommer att sjunka i värde gentemot kronan under året. Om priset när det sattes innebar en vinstmarginal på 30% efter att betalningen växlats till kronor så skulle en nedgång i dollarkursen med lika mycket innebära att hela vinsten raderades ut. Finanschefen går därför till banken och frågar om de kan ställa ut säljoptioner på dollarkursen, en så kallad valutaförsäkring. Genom säljoptionerna kan skeppsvarvet få möjligheten att växla sin betalning i dollar till kronor om ett år till samma växelkurs som när kontraktet slöts. På så sätt kan man utan risk bibehålla vinstmarginalen på minst 30% minus kostnaden för optionerna. Hur mycket skall banken ta betalt för att ställa ut optionerna åt skeppsvarvet?*

Ovanstående är en situation där optioner kommer till användning men om inte företaget vill ta risken, varför skulle banken göra det? Banken måste på något sätt eliminera risken. Innan Black och Scholes presenterade sin modell hade man ingen kunskap om vilket pris på optionerna som krävdes för att risken skulle kunna elimineras. Sedan 1973 har banken kunnat använda sig av deras formel för att bestämma vilket pris optionerna skall säljas för när de ställs ut. Det är inte någon större skillnad mellan att värdera optioner på aktier och valutor men låt oss för framställningens skull återgå till optioner på aktier, där blir priset vid tidpunkten för utställandet enligt Black-Scholes formel

$$\begin{gathered}
c(0, S(0), K, T) = \\
S(0)\Phi(d_1) - Ke^{-rT}\Phi(d_2) \\
d_1 = \frac{\ln(\frac{S(0)}{K}) + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \\
d_2 = \frac{\ln(\frac{S(0)}{K}) + (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}
\end{gathered} \tag{28}$$

där σ är aktiens volatilitet och kan uppskattas från historisk data.

Binomialmodellen beskriver precis som Black-Scholes modell hur banken kan eliminera sin risk och vilket pris på optionerna ett sådant förfarande leder till.

8.1.2 Binomialmodellen

Idén med binomialmodellen är att den som ställer ut optionerna, till exempel en bank, samtidigt köper en portfölj med aktier och obligationer som motsvarar värdet på optionen oavsett utfall. Vi börjar med att visa hur resultatet av en sådan strategi blir för en tidsperiod, $T = 1$, för att sedan utvidga den till fler perioder. Man vill skapa en portfölj som vid $t = 0$ är lika mycket värd som optionen

$$c(0) = \theta_0 S(0) + \theta_1 B(0)$$

för några θ_0 och θ_1 som är antalet aktier respektive obligationer. Optionens utbetalning vid T är känd enligt tidigare. Portföljens värde efter en period är på samma sätt

$$\begin{tikzcd}[column sep=large]
 & \theta_0 S(0)u + \theta_1 B(0)r \text{ med sannolikhet } q \\
\theta_0 S(0) + \theta_1 B(0) \arrow[ur, dash] \arrow[dr, dash] & \\
 & \theta_0 S(0)d + \theta_1 B(0)r \text{ med sannolikhet } 1 - q
\end{tikzcd}$$

Om optionens värde och portföljens värde skall överensstämma måste alltså följande ekvationssystem lösas

$$\begin{cases} \theta_0 S(0)u + \theta_1 B(0)r = c_u \\ \theta_0 S(0)d + \theta_1 B(0)r = c_d \end{cases}$$

Där c_u och c_d är optionens utbetalning vid uppgång respektive nedgång. Detta ger

$$\theta_0 = \frac{c_u - c_d}{S(0)(u-d)}$$
$$\theta_1 = \frac{uc_d - dc_u}{rB(0)(u-d)}$$

Om optionen inte prissätts efter en portfölj konstruerad på detta sätt uppstår arbitragemöjligheter. Om optionen skulle ha ett lägre värde än portföljen så skulle en handlare kunna sälja portföljen och köpa optionen. Värdet på en korrekt prissatt option över en period är därför

$$c(0, S(0), K, 1) = \theta_0 S(0) + \theta_1 B(0) = \frac{c_u - c_d}{u-d} + \frac{uc_d - dc_u}{r(u-d)}$$

vilket också kan skrivas

$$c(0, S(0), K, 1) = \frac{pc_u + (1-p)c_d}{r} \tag{29}$$

där

$$p = \frac{r-d}{u-d}$$

I binomialmodellen med en period kan en options värde alltså beräknas enligt (29) som är en funktion av c_u, c_d, r, u och d. Nästa steg är att använda den här metoden på ett sådant sätt att den fungerar för godtyckligt många tidsperioder. Vi har visat hur man kan värdera optionen en period före slutdag tack vare kunskap om värdet på slutdagen (c_u och c_d). På precis samma sätt som värdet på optionen vid $T-1$ kan avgöras när värdet på T är känt så kan värdet på $T-2$ avgöras när värdet på $T-1$ är känt. Genom att upprepa ovanstående förfarande kan man genom rekursion värdera en option vid valfri tidpunkt. Man finner att lösningen för n perioder blir

$$c(0, S(0), K, n) = \frac{1}{r^n} \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} (u^j d^{n-j} S(0) - K)^+ \tag{30}$$

vilket löser problemet om hur en option bör värderas i en diskret modell.

Med (30) kan problemet i exempel 3 hanteras. Låt oss tänka på den underliggande tillgången som en aktie och inte en valutakurs. Man kan anta att skeppsvarvet inte kommer att handla med optionerna vid något annat tillfälle än slutdagen, vilket gör att man kan se på situationen som att optionerna ställs ut enligt en binomialmodell med $T = 1$. Banken använder (30) för att räkna ut hur mycket en option bör kosta (eller i vårt enkla fall, (29)) och säljer den till skeppsvarvet för detta pris plus en avgift. Med de pengar försäljningen inbringar köper banken sedan en portfölj med aktier och obligationer i de kvantiteter som föreslås av modellen. I slutet av året när skeppsvarvet har rätt att lösa in sina optioner kan två saker hända. Om dollarkursen har sjunkit så löser skeppsvarvet in sina optioner och får kompensation för det minskade värdet på betalningen av skeppet. Om dollarkursen ökat är optionerna värdelösa och förfaller. Oavsett så säljer banken sin portfölj som är värd lika mycket som de utställda optionerna oberoende av utfall. I de fall där skeppsvarvet förväntar sig en utbetalning bekostas den alltså inte av bankens eget kapitel utan av portföljens värdeökning. På detta sätt kan banken utan risk erbjuda sig att ställa ut optioner.

8.1.3 Vad är u och d?

Som nämnt beror värdet av en option på c_u, c_d, r, u och d. Medan de två första är kända och r ges av marknaden så vet vi inte vad u och d är. Som tur är går det att uppskatta.

Efter n perioder varav j resulterade i en uppåtrörelse u för aktien är dess värde

$$S(n) = u^j d^{n-j} S(0).$$

Genom att ta väntevärdet och variansen på $\ln(\frac{S(n)}{S(0})$ när $n \to \infty$ kan man beräkna vilka värden u och d måste ha för att vår modell skall överensstämma med en aktie i verkligheten. Man får då att

$$u = e^{\sigma\sqrt{\frac{t}{n}}} \tag{31}$$

$$d = e^{-\sigma\sqrt{\frac{t}{n}}} \tag{32}$$

där σ är den empiriskt uppmätta volatiliteten (standardavvikelsen) hos aktien.

8.1.4 En diskret och en kontinuerlig formel

För att ytterligare kunna jämföra den diskreta formel som följer av binomialmodellen med Black-Scholes formel så kan man fördelaktigt skriva om (30).

Notera att minsta antalet utfall j som gör att $(u^j d^{n-j} S(0) - K)^+$ är skilt från noll ges av olikheten $u^j d^{n-j} S(0) > K$. Vilket är samma som att

$$j > \frac{\log \frac{K}{S(0)d^n}}{\log \frac{u}{d}} \tag{33}$$

Om man låter a vara det minsta j som gör att (33) uppfylls så kan optionens värde uttryckas som

$$c(0, S(0), K, n) =$$
$$S(0) \sum_{j=a}^{n} \binom{n}{j} p^j (1-p)^{n-j} (\frac{u^j d^{n-j}}{r^n}) - K r^{-n} \sum_{j=a}^{n} \binom{n}{j} p^j (1-p)^{n-j}$$

genom att (30) bryts upp i två termer och $(u^j d^{n-j} S(0) - K)^+$ ersätts med $(u^j d^{n-j} S(0) - K)$ då summationsgränsen ändras.

Om binomialprocessen $\Phi(a, n, p)$ definieras som

$$\Phi(a, n, p) = \sum_{j=a}^{n} \binom{n}{j} p^j (1-p)^{n-j}$$

når vi den slutgiltiga formuleringen

$$c(0, S(0), K, n) = S(0)\Phi(a, n, \frac{u}{r}p) - K r^{-n} \Phi(a, n, p) \tag{34}$$

som kan jämföras med Black-Scholes formel, upprepad nedan, där Φ istället betecknar normalfördelning.

$$c(0, S(0), K, T) =$$
$$S(0)\Phi(d_1) - K e^{-rT} \Phi(d_2)$$
$$d_1 = \frac{\ln(\frac{S(0)}{K}) + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$
$$d_2 = \frac{\ln(\frac{S(0)}{K}) + (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

Att u och d i det kontinuerliga alternativet är ersatt av σ är inte förvånande med tanke på (31) och (32). Motsvarigheten till (26) och (27) i Black-Scholes modell, det vill säga hur man modellerar aktier och obligationer, är

$$S(t) = S(0)e^{\alpha t + \sigma W(t)}$$
$$B(t) = B(0)e^{rt}$$

där $W(t)$ är en Wienerprocess, σ standardavvikelsen och α förhåller sig till aktiens väntevärde μ som $\alpha = \mu - \frac{\sigma^2}{2}$.

När vi så småningom går över till den hyperfinita binomialmodellen så kommer sambandet mellan formlerna att framstå klarare.

8.2 Grundläggande stokastiska processer

Innan vi går vidare behöver vi definiera några grundläggande stokastiska processer, vilka eventuellt redan är bekanta för läsaren. Dessa processer är centrala inom sannolikhetsläran för både standard och icke-standard.

8.2.1 Martingal

En martingal är en stokastisk process där framtida väntevärden är nuvarande värdet. Det är ett väldigt vanligt begrepp inom spelteori och används ofta som modell för ett rättvist spel. För att ge en korrekt definition av martingaler behöver vi först förklara begreppet filtration. Anledningen är att framtida värden av processen tillåts bero på den tillgängliga historiska informationen, och filtrationen är ett sätt att representera denna information. Formellt är filtrationen en funktion $\mathcal{P} : t \mapsto \mathcal{P}_t$, där $\mathcal{P}_s \subseteq \mathcal{P}_t$ för alla $s \leq t$, och en stokastisk process ξ anpassad till $\mathcal{P}$ uppfyller $\xi(t) \in \mathcal{P}_t$ för alla tidpunkter t. Det betingade väntevärdet $\mathrm{E}_{\mathcal{P}_t}$ tar hänsyn till all historisk information tillgänglig fram till tidpunkt t. Vi är nu redo att definiera martingal

Definition 29. Martingal. *En process ξ är martingal om följande gäller för alla tidpunkter $s \leq t$*

1. $\mathrm{E}_{\mathcal{P}_t}\left[\,|\xi(t)|\,\right] < \infty$,
2. $\mathrm{E}_{\mathcal{P}_s}[\xi(t)] = \xi(s)$.

Det är (2) som är den centrala egenskapen. I vårt fall kommer processerna att ha oberoende förändringar, så att den enda historiska information de använder är nuvarande värdet

$$\mathrm{E}_{\mathcal{P}_s}[\xi(t)] = \mathrm{E}\left[\xi(t) \mid \xi(s) = \xi_s\right] = \xi_s$$

där ξ_s alltså är känt vid tidpunkt s. På grund av att vårt fall är ett enkelt specialfall kommer vi inte nämna filtrationer igen.

8.2.2 Wienerprocess och geometrisk Brownsk rörelse

Den stokastiska process som ligger till grund för modelleringen av aktier i det kontinuerliga fallet är Wienerprocessen. Man kan tänka på Wienerprocessen som en kontinuerlig slumpvandring, eller brus. Ibland kallas den även för Brownsk rörelse.

Definition 30. Wienerprocess. *$W(t)$ är en Wienerprocess för $t \geq 0$ om*

5. $W(0) = 0$.
   $E[W(t)] = 0$.

5.5. *Om $t > s$ så är $X(t) - X(s)$ normalfördelad.*

5. *$W(t)$ har stationära förändringar, det vill säga $X(t) - X(s)$ har samma sannolikhetsfördelning som $X(t+h) - X(s+h)$ oberoende av t, s och h.*

5. *$W(t)$ har oberoende förändringar. $W(0), W(t_1) - W(0), W(t_2) - W(t_1), \ldots$ är oberoende av varandra.*

5. *$W(t)$ är kontinuerlig nästan säkert (sannolikhet 1).*

För en standard Wienerprocess W(t) är variansen

$$\mathrm{Var}\,[W(t)] = 1.$$

I Black-Scholes-modellen modelleras en aktie S som en **geometrisk Brownsk rörelse**

$$S(t) = S(0)e^{\mu t + \sigma W(t)}$$

där W är den drivande Wienerprocessen, eller motsvarande

$$\sigma W(t) = \ln \frac{S(t)}{S(0)} - \mu t.$$

Parametern μ är en exponentiell trend i priset och motsvarar ett företags organiska tillväxt.

Det är viktigt att poängtera att parametrarna σ och μ inte går att beräkna matematiskt, utan måste uppskattas. Parametern σ går att uppskatta med stor noggrannhet givet historiken för S. Däremot kan μ för en aktie inte mätas och därför inte heller användas vid till exempel värdering av optioner.

Anledningen till att det blir så här är för att σ kan uppskattas godtyckligt bra genom att mäta processen tillräckligt ofta, till skillnad från uppskattningen av μ som inte beror på antalet mätningar överhuvudtaget. Man kan visa att den uppskattning $\hat{\mu}$ av μ som ger minst varians ges av

$$\hat{\mu} = \frac{\ln S(t) - \ln S(s)}{t - s}, \text{ där } S(s) \text{ och } S(t) \text{ är första och sista kända värde.}$$

Observera att

$$\hat{\mu} - \mu \sim \mathcal{N}\left(0, \frac{\sigma^2}{t - s}\right)$$

så

$$P\left[|\hat{\mu} - \mu| \leq \frac{\sigma x}{\sqrt{t - s}}\right] = 2\Phi(x) - 1.$$

Ett vanligt värde på den årliga volatiliteten är 30% ($\sigma = 0.3$). Anta att vi vill ha en noggrannhet på $|\hat{\mu} - \mu| < 0{,}02$ med 95% sannolikhet (x = 1,96). Det är en väldigt grov uppskattning eftersom storleksordningen av μ är ungefär 0,02 (2% årlig utveckling). Ändå innebär det att vi behöver observera processen i minst $t - s = (\frac{\sigma x}{0{,}02})^2 \approx 864$ år! Därför är det en stor framgång inom finansmatematik att μ försvinner ur beräkningarna av optionspriser.

8.3 Matematiska definitioner av de finansiella begreppen

När grunderna för den finansiella modell vilken vi arbetar med nu har presenterats är vi redo för att på ett mer exakt sätt definiera några av de begrepp vi använt och därmed ge en precis matematisk innebörd.

8.3.1 Portföljer och strategier

En portfölj är en kombination av innehav i olika aktier S_i och ränteobligationen B. Vi har valt att begränsa oss till endast en aktie S, det förenklar notationen och allt vi skriver går att utöka analogt till en marknad med många aktier. Så i det här sammanhanget består en portfölj θ av θ_0 andelar i aktien S och θ_1 andelar i ränteobligationen B. Vi betecknar denna portfölj med

$$\theta = (\theta_0, \theta_1)$$

med värdet

$$V_\theta = \theta_0 S + \theta_1 B.$$

En strategi består av en portfölj där andelarna i innehaven tillåts förändras i varje tidssteg, med tillgång till endast historisk information, dvs. inga orakel eller insideraffärer tillåts.

En strategi kallas *självfinansierande* om man aldrig tar ut eller sätter in pengar i portföljen, så att värdeförändringen av portföljen är ett resultat enbart av värdeförändringen av de underliggande instrumenten. En strategi θ är alltså självfinansierande om

$$V_\theta(t+1) - V_\theta(t) = \theta_0(t)\left(S(t+1) - S(t)\right) + \theta_1(t)\left(B(t+1) - B(t)\right).$$

8.3.2 Räntejusterade värden

För att kunna jämföra värden över tiden måste vi räntejustera dem. Tanken är att en krona idag ska vara lika mycket värd som en räntejusterad krona i framtiden. Om $y(t)$ är ett värde vid tidpunkt t, låter vi $\bar{y}(t)$ beteckna dess räntejusterade värde

$$\bar{y}(t) = y(t)\frac{B(0)}{B(t)}.$$

Vi får då speciellt att

$$\bar{y}(0) = y(0)$$

och

$$\bar{B}(t) = B(0).$$

8.3.3 Arbitrage

En arbitragestrategi är en självfinansierande strategi som initialt är gratis, men som efter en tid har en möjlighet att anta ett positivt värde, men aldrig ett negativt värde. En sådan strategi skulle naturligtvis vara populär inom finansvärlden, men man kan argumentera för att en effektiv marknad med tiden skulle gå mot att eliminera sådana möjligheter. Ett vanligt antagande är därför att marknaden är arbitragefri.

Definition 31. Arbitrage.
En möjlighet till arbitrage är en självfinansierande strategi θ där t är nuvarande tiden, T är ett framtida slutdatum $(T > t)$, och

1. $V_\theta(0) = 0$*, den är initialt gratis*
2. $V_\theta(T) \geq 0$*, vi blir aldrig skyldiga pengar vid slutdatumet (riskfritt)*
3. $\mathrm{E}\left[V_\theta(T)\right] > 0$*, vi har en chans att erhålla vinst vid slutdatumet.*

8.3.4 Hedging

Hedging är ett generellt begrepp för en investeringsstrategi som minskar eller eliminerar risk, lite som en försäkring. Vi kommer använda hedging i en mer specifik bemärkelse här, nämligen en strategi som eliminerar *all* risk förknippat med att äga något riskfyllt derivat c. Tanken är att om vi finner en sådan strategi ger det oss möjlighet att ge en entydig och objektiv prissättning av derivatet.

Tekniskt går det till så här: anta att vi finner en självfinansierande strategi θ som vid lösendatumet T har värdet $V_\theta(T) = c$ för alla möjliga utfall. Ett derivat som har en sådan strategi kallas *åtkomlig*, och strategin som uppfyller detta sägs *imitera* c. En marknad där alla derivat är åtkomliga kallas komplett.

En första insikt är att i en arbitragefri marknad är värdeprocessen entydig för strategier som imiterar c.

Bevis. Låt θ och ψ vara två självfinansierande strategier som imiterar c i en arbitragefri marknad. Anta att de *inte* har samma värdeprocess, dvs. det existerar en tidpunkt t där $V_\theta(t) \neq V_\psi(t)$. Vi kan utan inskränkning anta att θ är den dyrare av de två, dvs. $V_\theta(t) > V_\psi(t)$. Vid denna tidpunkt, köp den billiga portföljen, sälj den dyra (blanka), och köp ränteobligationer för överflödet $x = V_\theta(t) - V_\psi(t) > 0$.

Kalla denna strategi för η. Vi har därmed $V_\eta(t) = 0$, men vid lösendatumet T vet vi att att $V_\phi(T) = V_\psi(T) = c$ och därmed $V_\eta(T) = x > 0$. Därmed är η en arbitragestrategi vilket motsäger antagandet.

□

Låt oss kalla den entydiga värdeprocessen V_c. Entydigheten antyder att detta måste vara det objektiva priset för c. Anta att någon aktör på marknaden erbjuder att köpa/sälja c för ett pris som är högre respektive lägre än V_c. Detta är en arbitragemöjlighet på precis samma sätt som i beviset ovan. Köp helt enkelt den billigare av de två, sälj den dyra, invänta lösendatumet där det går att kvitta dessa mot varandra, och erhåll överflödet som vinst. I en arbitragefri marknad måste alltså $V_c(t)$ vara det entydigt korrekta värdet av c vid tidpunkt t. Vi kommer fokusera på dagsvärdet ($t = 0$) av c, vilket vi betecknar $\pi(c)$ där alltså $\pi(c) = V_c(0)$. Med hjälp av hedgingtekniken har vi nu ett sätt att prissätta derivat i arbitragefria marknader: hitta en självfinansierande strategi θ som imiterar c, och beräkna $\pi(c) = V_\theta(0)$. Det var på detta sätt vi beräknade binomialpriset i föregående avsnitt.

8.3.5 Riskneutrala måttet

Det är viktigt att poängtera att sannolikheten q för uppgång respektive nedgång i binomialmodellellen inte dök upp i optionspriset (30), eftersom vi i varje utfall hedgat oss, så att sannolikheten inte spelade någon roll. Detta leder till idén att vi kan *välja* en sannolikhetsfördelning som på något sätt är praktisk. Vi behöver alltså inte använda den *sanna* sannolikhetsfördelningen, vilket är tur eftersom den inte går att mäta. Den sannolikhetsfördelning Q som visar sig vara praktisk är den fördelning som gör $\bar{S}$ till en martingal. Detta innebär att vid tidpunkt $t = 0$ gäller $\mathrm{E}_Q\left[\bar{S}(T)\right] = \bar{S}(0)$. $\bar{B}$ är alltid martingal eftersom $\bar{B}(T) = B(0)$ så $\mathrm{E}_Q\left[\bar{B}(T)\right] = B(0) = \bar{B}(0)$ trivialt. Eftersom portföljerna består av enbart dessa två basinstrument blir deras värden V_θ summan av två martingaler och därmed själva martingaler. Vi får alltså även $\mathrm{E}_Q\left[\bar{V}_\theta(T)\right] = \bar{V}_\theta(0)$ för alla självfinansierande strategier θ. Detta är praktiskt eftersom vi vet värdet av $\bar{V}_\theta$ vid tidpunkt T då $\bar{V}_\theta(T) = \bar{c}(T)$. Vi kan nu beräkna $\pi(c)$ med

$$\pi(c) = V_\theta(0) = \bar{V}_\theta(0) = \mathrm{E}_Q\left[\bar{V}_\theta(T)\right] = \mathrm{E}_Q\left[\bar{c}(T)\right].$$

Vi har nu ett enkelt sätt att prissätta derivat c. Vi beräknar helt enkelt det räntejusterade väntevärdet av c under ett martingalmått Q, som brukar kallas det *riskneutrala måttet*. Beroendet av en explicit hedgingkonstruktion θ försvinner, vi behöver bara existensen av en sådan.

Värdet $\pi(c)$ blir alltså *inte*, som man skulle kunna tro, det *sanna* väntevärdet $\mathrm{E}_P\left[\bar{c}\right]$, där P är den sanna sannolikhetsfördelningen.

(Att värdeprocesserna är martingaler är ett tillräckligt villkor för att marknaden ska vara arbitragefri, eftersom första villkoret för arbitrage är $V_\theta(0) = 0$ vilket för martingaler direkt ger $\mathrm{E}\left[\bar{V}_\theta(T)\right] = \bar{V}_\theta(0) = 0$ vilket bryter mot tredje villkoret för arbitrage. För diskreta modeller visar det sig även vara ett nödvändigt villkor.)

8.4 Interna och externa mängder

Vi kommer att behöva bekanta oss med ytterligare begrepp inom icke-standardanalysen för att ha de verktyg till hands som vi behöver. Därför presenteras här den mest grundläggande teorin för interna mängder med extra fokus på ω_1-mättnad.

Vid konstruktion av ${}^*\mathbb{R}$ skapas med hjälp av ultrafilter hyperreella motsvarigheter till de reella talen. På samma sätt kan man skapa motsvarigheter till mängder. Redogörelsen nedanför följer [3].

8.4.1 Interna mängder

En intern mängd är en utökad mängd i samma bemärkelse som $[r] \in {}^*\mathbb{R}$ utökar $r \in \mathbb{R}$. Man bildar interna mängder med samma metod som de hyperreella talen. Istället för följder av tal så använder man följder av mängder.

Definition 32. *En följd av delmängder $A_n \subseteq \mathbb{R}$ definierar en mängd i ${}^*\mathbb{R}$ genom*

$$[r_n] \in [A_n] \Longleftrightarrow \{n \in \mathbb{N} : r_n \in A_n\} \in \mathcal{F}$$

där $\mathcal{F}$ är det använda ultrafiltret. En mängd i $^*\mathbb{R}$ kallas **intern** *om den har en sådan representation, annars* **extern**.

Ett hyperreellt tal $[r_n]$ tillhör alltså mängden $[A_n]$ om r_n tillhör $\mathcal{F}$-nästan alla A_n. Låt oss ta ett exempel på en intern mängd.

Exempel 4. *Utgå från ett tal* $[\,\langle T_n : n \in \mathbb{N}\rangle\,] \in {}^*\mathbb{N}$ *och en följd av mängder* A_n *som ges av*

$$A_n = \left\{ \frac{k}{T_n} : k \in \mathbb{N} \text{ och } k \leq T_n \right\} = \left\{ \frac{1}{T_n}, \frac{2}{T_n}, \cdots, \frac{T_n}{T_n} \right\}$$

Om då $T = [T_n]$ så innebär det att (32) uppfylls av mängden

$$[A_n] = \left\{ \frac{1}{T}, \frac{2}{T}, \cdots, \frac{T}{T} \right\}$$

eftersom

$$\frac{[r_n]}{T} \in [A_n] \Longleftrightarrow \{n \in \mathbb{N} : \frac{r_n}{T_n} \in A_n\} = \{n \in \mathbb{N} : r_n \in \mathbb{N} \text{ och } r_n \leq T_n\} \in \mathcal{F} \Longleftrightarrow$$
$$\Longleftrightarrow [r_n] \leq T \text{ och } [r_n] \in {}^*\mathbb{N}$$

Om vi väljer $T \in {}^*\mathbb{N} \setminus \mathbb{N}$ får vi en uppdelning av intervallet (0,1] *där det går att urskilja punkter som inte är möjliga att urskilja på en reell tallinje. Det kan man använda sig av för att skapa en så kallad hyperfinit tidslinje.*

Snittet, unionen och mängddifferensen av två interna mängder kommer uppfylla

$$[A_n] \cap [B_n] = [A_n \cap B_n], \tag{35}$$
$$[A_n] \cup [B_n] = [A_n \cup B_n], \tag{36}$$
$$[A_n] \setminus [B_n] = [A_n \setminus B_n], \tag{37}$$

och är därmed också interna mängder. Man kan också visa att

$$[A_n] \subseteq [B_n] \Longleftrightarrow \{n \in \mathbb{N} : A_n \subseteq B_n\} \in \mathcal{F} \quad , \tag{38}$$
$$[A_n] \neq [B_n] \Longleftrightarrow \{n \in \mathbb{N} : A_n \neq B_n\} \in \mathcal{F} \quad . \tag{39}$$

8.4.2 Exempel på externa mängder

En intressant och viktig egenskap hos interna mängder ges av följande sats vars bevis återfinns i [3] men utelämnas här.

Sats 40. *För alla interna mängder $M \subseteq {}^*\mathbb{N}$ där $M \neq \emptyset$ gäller att M har ett minimum, det vill säga ett minsta element.*

Som följd av detta kan man direkt dra slutsatsen att mängden $^*\mathbb{N} \setminus \mathbb{N}$ är extern eftersom den inte har ett minsta element.

I andra fall kan man använda sig av (35)-(37). $^*\mathbb{N}$ är ett trivialt exempel på en intern mängd vilket ges av att A_n kan vara $\mathbb{N}$ för alla n. Med hjälp av (37) ser vi då att $\mathbb{N}$ måste vara extern eftersom $^*\mathbb{N} \setminus \mathbb{N}$ annars skulle varit intern, vilket enligt tidigare inte stämmer. Vidare så är $\mathbb{R} \cap {}^*\mathbb{N} = \mathbb{N}$ varför det följer av (35) att $\mathbb{R}$ är extern, eftersom $^*\mathbb{N}$ inte är det.

8.4.3 Mättnad hos interna mängder

Den för vårt arbete mest relevanta egenskapen hos interna mängder är ω_1-mättnad, som är viktig för många av de mer avancerade konstruktionerna i icke-standardmatematik. Den kommer visa sig vara kritisk i skapandet av Loeb-rummet som vi kommer se senare. I enlighet med hur det utförs i [3] bevisas här tre satser som är av betydelse för vår förmåga att hantera senare svårigheter.

Sats 41. *För en avtagande följd bestående av icke-tomma interna mängder*

$$X^1 \supseteq X^2 \supseteq \cdots \supseteq X^k \supseteq \cdots \tag{40}$$

gäller att snittet är icke-tomt

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

Bevis. Låt $[A_n^k]$ vara följden som definierar X^k. Beviset är ganska trickartat och tekniskt. Vi kommer bevisa satsen genom att explicit konstruera ett hyperreellt tal $[s_n]$ som tillhör varje X^k. För detta ändamål behöver vi först skapa tre nya mängdföljder J^k, G^k och H^n, vars syften inte är uppenbara till en början.

Enligt (38) respektive (39) vet vi att

$$X^k \supseteq X^{k+1} \iff \{n \in \mathbb{N} : A_n^k \supseteq A_n^{k+1}\} \in \mathcal{F}, \tag{41}$$

$$X^k \neq \emptyset \iff \{n \in \mathbb{N} : A_n^k \neq \emptyset\} \in \mathcal{F}. \tag{42}$$

Definiera sedan J^k som motsvarigheten till de första k termerna i (40) tillsammans med villkoret att A_n^k inte är tomt

$$J^k = \{n \in \mathbb{N} : A_n^1 \supseteq A_n^2 \supseteq \cdots \supseteq A_n^k \text{ och } A_n^k \neq \emptyset\}. \tag{43}$$

Vi kan även skriva om J^k som

$$J^k = \{n \in \mathbb{N} : A_n^k \neq \emptyset\} \cap \{n \in \mathbb{N} : A_n^1 \supseteq A_n^2\} \cap \cdots \cap \{n \in \mathbb{N} : A_n^{k-1} \supseteq A_n^k\}$$

och då blir det uppenbart att J^k är ett snitt av ändligt många mängder ur $\mathcal{F}$ (enligt (41), (42)), och tillhör därmed själv $\mathcal{F}$

$$J^k \in \mathcal{F}.$$

Observera även att vi har $J^1 \supseteq J^2 \supseteq \cdots$. Vi skapar två snarlika mängder

$$G^k = \left\{n \in \mathbb{N} : k \leq n \text{ och } n \in J^k\right\},$$

$$H^n = \left\{k \in \mathbb{N} : k \leq n \text{ och } n \in J^k\right\}.$$

Observera specifikt att om $n \in G^k$ så hör $k \in H^n$. Vi kan även skriva G^k som

$$G^k = \{n \in \mathbb{N} : k \leq n\} \cap J^k$$

och då blir det uppenbart att

$$G^k \in \mathcal{F}$$

eftersom $\{n \in \mathbb{N} : k \leq n\}$ är komplementet till den ändliga mängden $\{n \in \mathbb{N} : 1 \leq n \leq k\}$ och därmed hör till $\mathcal{F}$.

För alla $n \in J^1$ har vi att $1 \in H^n \implies H^n \neq \emptyset$, så för dessa n kan vi definiera $k_n = \max(H^n)$. Eftersom $k_n \in H^n$ följer att $n \in J^{k_n}$, och enligt konstruktionen av J^{k_n} följer att $A_n^{k_n} \neq \emptyset$. Vi är nu äntligen redo att konstruera $[s_n]$. För $n \notin J^1$, låt s_n vara ett godtyckligt tal, och för $n \in J^1$ låt s_n vara något av talen i $A_n^{k_n}$ (som vi vet inte är tom), så $s_n \in A_n^{k_n}$. Eftersom $n \in J^{k_n}$ har vi $s_n \in A_n^{k_n} \subseteq \cdots \subseteq A_n^1$.

För att visa att $[s_n]$ tillhör det oändliga snittet har vi per definition

$$[s_n] \in \bigcap_{k\in\mathbb{N}} X^k \iff \left\{n \in \mathbb{N} : s_n \in A_n^k\right\} \in \mathcal{F} \text{ för alla } k.$$

Om vi kan bevisa

$$G^k \subseteq \left\{n \in \mathbb{N} : s_n \in A_n^k\right\}$$

är vi klara eftersom det skulle innebära att $\left\{n \in \mathbb{N} : s_n \in A_n^k\right\}$ är supermängden till en mängd i $\mathcal{F}$. G^k är en delmängd eftersom

$$n \in G^k \implies n \in J^k \subseteq J^1 \text{ och } k \in H^n \implies k \leq \max(H^n) = k_n \implies$$
$$\implies s_n \in A_n^{k_n} \subseteq \cdots \subseteq A_n^k \subseteq \cdots \subseteq A_n^1 \implies s_n \in A_n^k.$$

Beviset är därmed klart. □

Sats 41 leder till flera följdsatser varav två är särskilt viktiga för oss och kommer att användas senare.

Sats 42. *Låt $A = \{X_n : n \in \mathbb{N}\}$ vara en samling interna mängder för vilka det gäller att snittet av varje ändlig delmängd J uppfyller*

$$\bigcap_{i \in J} X_i \neq \emptyset$$

det vill säga att snittet av alla element i en ändlig delmängd till A alltid är icke-tomt. Vi säger att A har fip [5]*. Då gäller att*

$$\bigcap_{n \in \mathbb{N}} X_n \neq \emptyset$$

Bevis. Definiera snittet av de k första mängderna i A med

$$Y^k = X_1 \cap X_2 \cap \cdots\cdots X_k$$

Enligt (44) är $Y^k \neq \emptyset$ och som följd av (35) är Y^k en intern mängd. Observera att

$$Y_1 \supseteq Y_2 \supseteq \cdots \supseteq Y_k \supseteq \cdots$$

varför sats 41 är tillämpbar och medför att det finns ett element som tillhör varje Y^k för $k \in \mathbb{N}$. En snabb blick på (44) försäkrar oss att det även måste finnas i X_k. □

Sats 43. *Låt X vara en intern mängd och som innan $\{X_n : n \in \mathbb{N}\}$ en samling interna mängder. Om $X \subseteq \bigcup_{n \in \mathbb{N}} X_n$ så är $X \subseteq \bigcup_{n \le k} X_n$ för något $k \in \mathbb{N}$.*

Bevis. Motsägelsebevis, antag att $X \nsubseteq \bigcup_{n \le k} X_n$ för alla $k \in \mathbb{N}$.

$$X \nsubseteq \bigcup_{n \le k} X_n \Longrightarrow \bigcap_{n \le k} (X \setminus X_n) = X \setminus \bigcup_{n \le k} X_n \neq \emptyset$$

Mängden $B_n = \{X \setminus X_n : n \in \mathbb{N}\}$ är intern enligt (37) och den har *fip* eftersom det finns ett element i X som inte finns i X_n för något $n \in \mathbb{N}$. Enligt sats 42 finns det då ett element $x \in \cap_{n \in \mathbb{N}} B_n$ vilket innebär att

$$x \in \{X \setminus X_n : n \in \mathbb{N}\} = X \setminus \bigcup_{n \in \mathbb{N}} X_n \neq \emptyset$$

men eftersom $X \subseteq \bigcup_{n \in \mathbb{N}} X_n$ så är det en motsägelse. □

8.5 Mått och sannolikhetsrum

Detta avsnitt ger en kort introduktion till mått och sannolikhetsrum.

8.5.1 Hyperfinita sannolikhetsrum

Ett hyperfinit sannolikhetsrum är en trippel $(\Omega, \mathcal{A}, Q)$ där

1. Ω är utfallsrummet,
2. $\mathcal{A}$ är en algebra av de interna delmängderna av Ω,
3. Q är ett (ändligt additivt) sannolikhetsmått som mappar mätbara mängder av utfall till deras (hyperreella) sannolikhet, $Q : \mathcal{A} \mapsto {}^*[0, 1]$.

[5] Finite Intersection Property

Utfallsrummet Ω är helt enkelt mängden av alla möjliga utfall. Det är ofta ointressant att mäta sannolikheten för enskilda utfall $\omega \in \Omega$, så istället mäter sannolikhetsmåttet Q en mängd av utfall, men bara mängder som tillhör $\mathcal{A}$. Dessa mängder kallas mätbara. Läsare nya till måtteori undrar kanske varför inte alla mängder är mätbara. Det är för att man kan skapa (märkliga) mängder med hjälp av till exempel urvalsaxiomet vars struktur är så komplicerad att man inte kan tilldela ett meningsfullt mått till dem. Genom att definiera Q endast på mätbara mängder kommer man undan problemet. Att Q är ett ändligt additivt sannolikhetsmått innebär att det uppfyller

1. $Q(\emptyset) = 0$,
2. $Q(\Omega) = 1$,
3. $Q(A \cup B) = Q(A) + Q(B) - Q(A \cap B)$.

Att $\mathcal{A}$ är en algebra (av mängder) innebär att den innehåller Ω och är sluten under parvisa mängdoperationer, dvs. om $A, B \in \mathcal{A}$ gäller

1. $\Omega \in \mathcal{A}$,
2. $A \cup B \in \mathcal{A}$,
3. $A \setminus B \in \mathcal{A}$,
4. $A \cap B \in \mathcal{A}$, följer av 3 eftersom $A \cap B = A \setminus (A \setminus B)$,
5. $A \triangle B \in \mathcal{A}$, följer av 2 och 3 eftersom $A \triangle B = (A \setminus B) \cup (B \setminus A)$,
6. $A^C \in \mathcal{A}$, följer av 1 och 3 och eftersom $A^C = \Omega \setminus A$,
7. $\emptyset \in \mathcal{A}$, följer av 1 och 3 eftersom $\emptyset = \Omega \setminus \Omega$.

När vi ändå pratar om algebror kan vi passa på att definiera σ-algebra vilket vi behöver senare. Det är en starkare variant av algebra, så en σ-algebra är en algebra, med det starkare villkoret att även *uppräkneliga* unioner är slutna, dvs.

$$\cup_{j \in \mathbb{N}} A_j \in \mathcal{A}, \text{ om alla } A_j \in \mathcal{A}.$$

Ur detta följer att även uppräkneliga snitt är sluta eftersom

$$\cap_{j \in \mathbb{N}} A_j = A_1 \setminus \cup_{j \in \mathbb{N}} (A_1 \setminus A_j).$$

8.5.2 Loeb-rummet

Med hjälp av ett hyperfinit sannolikhetsrum $(\Omega, \mathcal{A}, Q)$ som beskrivet ovan kan man skapa ett standard (σ-additivt) sannolikhetsrum $(\Omega, \mathcal{A}_L, Q_L)$, där

1. $\mathcal{A}_L$ är en σ-algebra med $\mathcal{A} \subseteq \mathcal{A}_L$,
2. $Q_L : \mathcal{A}_L \mapsto [0, 1]$ är σ-additiv, och på $\mathcal{A}$ är $Q_L = {}^oQ$.

Att Q_L är σ-additiv innebär att den är additiv även för *uppräkneliga* mängder, så om alla $A_j \in \mathcal{A}_L$ är disjunkta gäller

$$Q_L \left(\cup_{j \in \mathbb{N}} A_j\right) = \sum_{j \in \mathbb{N}} Q_L(A_j).$$

Rummet $(\Omega, \mathcal{A}_L, Q_L)$ kallas *Loeb-rummet*, och att detta alltid är möjligt att skapa är ett högst icke-trivialt resultat som är centralt inom icke-standardsannolikhetsläran. Det är enkelt att visa att Q_L är ändligt additivt på $\mathcal{A}$ genom att helt enkelt definiera $Q_L = {}^oQ$ där, så

$$\begin{aligned} Q_L(A \cup B) &= {}^oQ(A \cup B) = {}^o\left(Q(A) + Q(B) - Q(A \cap B)\right) = \\ &= {}^oQ(A) + {}^oQ(B) + {}^oQ(A \cap B) = \\ &= Q_L(A) + Q_L(B) + Q_L(A \cap B), \text{ där } A, B \in \mathcal{A}. \end{aligned}$$

Det som inte är uppenbart är att den även är σ-additiv på $\mathcal{A}$, men detta går att visa enkelt med hjälp av ω_1-mättnad. Vi ska alltså visa

$$Q_L\left(\cup_{j\in\mathbb{N}}A_j\right) = \sum_{j\in\mathbb{N}} Q_L(A_j), \text{ där alla } A_j \in \mathcal{A} \text{ är disjunkta och } \cup_{j\in\mathbb{N}} A_j \in \mathcal{A}.$$

Då är alltså $\cup_{j\in\mathbb{N}}A_j$ intern och ω_1-mättnad ger enligt sats 43 att $\cup_{j\in\mathbb{N}}A_j = \cup_{j\leq k}A_j$ för något *ändligt* k och resultatet följer direkt!

Nu kan vi tillämpa Carathéodorys utökningssats eftersom vi visat att Q_L är σ-additiv, vilket ger oss att Q_L kan utökas till en σ-additiv funktion på σ-algebran $\mathcal{A}_L \supseteq \mathcal{A}$ och vi är klara.

8.5.3 $\mathcal{S}$-integrerbar

En styrka med Loeb-rum är att man relativt enkelt kan relatera hyperfinita integraler över interna funktioner $F : \Omega \mapsto {}^*\mathbb{R}$ till deras motsvarande integral i Loeb-rummet.

I ett hyperfinit sannolikhetsrum kan man alltid skapa sannolikhetsmåttet Q från en sannolikhetsmassfunktion för varje enskilt utfall $a : \Omega \mapsto {}^*[0,1]$, som om utfallsrummet vore ett ändligt antal diskreta utfall i standardmotsvarigheten, genom hyperfinit summation ${}^*\sum$

$$\tilde{Q}(A) = {}^*\sum_{w\in A} a_w, \text{ där } A \in \mathcal{A} \text{ och } a_w = Q(\ \{w\}\ )$$

där a_w alltså är den hyperfinita sannolikheten för det enskilda utfallet w. Vi är specifikt intresserade av väntevärden i det här projektet, och det hyperfinita väntevärdet för en intern funktion $F : \Omega \mapsto {}^*\mathbb{R}$ är helt enkelt

$$\mathrm{E}_Q[F] = {}^*\sum_{w\in\Omega} F(w)a_w$$

och en hyperfinit integral över ett internt A är

$$\int_A F\,\mathrm{d}Q = {}^*\sum_{w\in A} F(w)a_w.$$

Till skillnad från standardväntevärden existerar alltid hyperfinita väntevärden (men de behöver inte vara ändliga). Vi är intresserade av skuggan av det hyperfinita väntevärdet och hur det förhåller sig till motsvarande standardväntevärde i Loeb-rummet. Egenskapen hos F som vi är intresserade av kallas $\mathcal{S}$-integrerbar. Det finns flera olika ekvivalenta definitioner, och en av dem är precis den egenskap vi är ute efter, så vi väljer den som definition.

Definition 33. $\mathcal{S}$-integrerbar
En intern funktion $F : \Omega \mapsto {}^*\mathbb{R}$ kallas $\mathcal{S}$-integrerbar relativt Q om

1. ${}^oF$ *existerar nästan säkert (sannolikhet 1),*
2. $\mathrm{E}_{Q_L}[{}^oF]$ *existerar,*
3. $\mathrm{E}_{Q_L}[{}^oF] = {}^o\mathrm{E}_Q[F]$.

En annan definition är även på sin plats.

Definition 34. *Vi kallar en funktion $F : \Omega \mapsto {}^*\mathbb{R}$ en* **lifting** *av den reellvärda funktionen $f : \Omega \mapsto \mathbb{R}$ om F är intern och ${}^oF(w) = f(w)$ nästan säkert (Q_L-sannolikhet 1).*

Vi formulerar även en användbar sats i sammanhanget, men ger inget bevis, utan hänvisar istället till [1] Theorem 6.4 under "Loeb measure and probability".

Sats 44. *Låt $(\Omega, \mathcal{A}_L, Q_L)$ vara ett Loeb-rum, och $f : \Omega \mapsto \mathbb{R}$ en mätbar funktion. Då existerar $\mathrm{E}_{Q_L}[f]$ om och endast om det existerar en $\mathcal{S}$-integrerbar lifting F av f, och därmed*

$${}^o\mathrm{E}_Q[F] = \mathrm{E}_{Q_L}[{}^oF] = \mathrm{E}_{Q_L}[f].$$

8.6 Hyperfinita binomialmodellen

Vi skapar nu den hyperfinita binomialmodellen. Välj en sluttid $T \in \mathbb{R}$ och ett hyperfinit heltal $n \in {}^*\mathbb{N}$ så att tidssteget blir $h = \Delta t = T/n$ och skapa den hyperfinita tidslinjen

$$\mathbb{T} = \{0, h, 2h, ..., T\}.$$

Låt slumpvariablerna $X_0, X_h, ..., X_T$ vara oberoende och likafördelade med utfallen

$$P(X = 1) = P(X = -1) = \frac{1}{2}.$$

Vårt hyperfinita sannolikhetsrum $(\Omega, \mathcal{A}, Q)$ blir följande

1. Ω är utfallsrummet för $(X_0, X_h, ..., X_T)$, så $\Omega = \{-1, 1\}^n$,
2. $\mathcal{A}$ är algebran av de interna delmängderna i Ω,
3. Sannolikhetsmåttet $Q : \mathcal{A} \mapsto {}^*\mathbb{R}$ blir helt enkelt kardinalitetmåttet eftersom alla utfall är lika sannolika.

Vi definierar nu den hyperfinita aktieprocessen S, och den hyperfinita ränteprocessen B i detta rum. Låt σ och μ vara två reella tal och definiera aktieprocessen S i tidpunkt t som

$$S(t) = S(0) \prod_{s<t} \left(1 + \mu h + \sigma \sqrt{h} X_s\right).$$

Låt räntan r vara ett positivt reellt tal och definiera ränteobligationsprocessen B i tidpunkt t som

$$B(t) = B(0) \prod_{s<t} (1 + rh) = B(0) \left(1 + rh\right)^{t/h}.$$

I denna modell är ett räntejusterat värde

$$\bar{Y}(t) = Y(t) \frac{B(0)}{B(t)}.$$

Vi får $\bar{S}$ till en martingal genom ett lämpligt val av μ. Eftersom X är oberoende och likafördelade räcker det med att titta på ett steg. Vi väljer μ så att

$$\mathrm{E}_Q \left[\bar{S}(h)\right] = S(0).$$

Notera att $\mathrm{E}_Q[X] = 0$ så

$$\mathrm{E}_Q \left[\bar{S}(h)\right] = \frac{S(0)}{B(h)} \mathrm{E}_Q \left[1 + \mu h + \sigma \sqrt{h} X_h\right] = \frac{S(0)}{B(h)} \left(1 + \mu h\right).$$

För att uppfylla martingalvillkoret behöver alltså

$$\frac{S(0)}{B(h)} \left(1 + \mu h\right) = S(0)$$

eller

$$1 + \mu h = B(h) = 1 + rh$$

vilket ger

$$\mu = r.$$

Enligt finansdelen i föregående avsnitt kan nu ett hyperreellt derivat C prissättas med $\Pi(C)$ där

$$\Pi(C) = \mathrm{E}_Q \left[\bar{C}\left(S(T)\right)\right].$$

8.6.1 Relation till standardbinomialmodellen

För att relatera till standardbinomialmodellen som presenterades förut har vi här

$$\begin{aligned} u &= 1 + \alpha h + \sigma\sqrt{h}, \\ d &= 1 + \alpha h - \sigma\sqrt{h}, \\ p &= \frac{1}{2}, \end{aligned}$$

den enda skillnaden är att h, u, d nu generellt är infinitesimaler och n obegränsat. För att beräkna slutpriset $S(T)$ räcker det med att räkna hur många gånger priset gick upp, $J = |\{X_t : X_t = 1\}|$, eftersom resultatet är oberoende av ordningen av $\{X_t\}$. Vi kan därför skriva $S(T)$ som

$$S(T) = S(0)u^J d^{n-J}.$$

Eftersom J är binomialfördelad $J \sim \text{Binomial}(n, p)$ får vi

$$\Pi(C) = \mathrm{E}_Q\left[\bar{C}\left(S(T)\right)\right] = \frac{B(0)}{B(t)} \sum_{j=0}^{n} C\left(S(0)u^j d^{n-j}\right) B_{n,p}(j)$$

där $B_{n,p}$ är binomialkoefficienterna

$$B_{n,p}(j) = \binom{n}{j} p^j \left(1-p\right)^{n-j}.$$

Väljer vi naturligt $n \in \mathbb{N}$ får vi alltså tillbaka standardbinomialpriset. Observera att vi inte behövde skapa en explicit hedgingkonstruktion.

8.6.2 Relation till Black-Scholes-modellen

Mer intressant blir det om vi väljer $n \in {}^*\mathbb{N} \setminus \mathbb{N}$. Då blir skuggan av aktiepriset en geometrisk Brownsk rörelse som i Black-Scholes-modellen (med sannolikhet 1)

$$S(t) \simeq S(0) \exp\left((\alpha - \frac{\sigma^2}{2})t + \sigma W(t) \right)$$

där $W(t)$ är en Wienerprocess.

Bevis. Vi kommer använda uppskattningar av $\ln(1+x)$ där x är litet (infinitesimalt till och med). Uppskattningen vi använder är

$$\ln(1+x) = x - \frac{x^2}{2} + \epsilon$$

där $|\epsilon| \le |x|^3$, vilket gäller för alla $|x| < \frac{2}{3}$. Vi har därför

$$\begin{aligned} S(t) &= S(0) \prod_{s<t} \left(1 + \alpha h + \sigma\sqrt{h} X_s\right) = \\ &= S(0) \exp\left(\sum_{s<t} \ln\left(1 + \alpha h + \sigma\sqrt{h} X_s\right) \right) = \\ &= S(0) \exp\left(\sum_{s<t} \alpha h + \sigma\sqrt{h} X_s - \frac{1}{2}\left(\alpha h + \sigma\sqrt{h} X_s\right)^2 + \epsilon_s \right) = \\ &= S(0) \exp\left(\sum_{s<t} (\alpha - \frac{\sigma^2}{2} X_s^2) h + \sigma\sqrt{h} X_s - \frac{1}{2}\alpha^2 h^2 - \alpha\sigma h\sqrt{h} X_s + \epsilon_s \right) = \\ &= \left[\text{observera att } X^2 = 1, \sum_{s<t} h = t, \text{ och låt } Y_t = \sum_{s<t} X_s \right] = \\ &= S(0) \exp\left((\alpha - \frac{\sigma^2}{2})t + \sigma\sqrt{h} Y_t - \frac{1}{2}\alpha^2 h t - \alpha\sigma h^{3/2} Y_t + \sum_{s<t} \epsilon_s \right) \end{aligned}$$

där

$$|\epsilon_s| \leq \left|\alpha h + \sigma\sqrt{h}X_s\right|^3 \leq h^{3/2}\left(|\alpha| + |\sigma|\right)^3$$

så

$$\left|\sum_{s<t}\epsilon_s\right| \leq \sqrt{h}t\left(|\alpha| + |\sigma|\right)^3 \simeq 0$$

och

$$|Y_t| \leq \frac{t}{h} \Rightarrow \alpha\sigma h^{3/2}|Y_t| \leq \alpha\sigma t\sqrt{h} \simeq 0$$

och

$$\frac{1}{2}\alpha^2 ht \simeq 0.$$

Detta tillsammans med att exp är S-kontinuerlig för begränsade värden ger

$$\begin{aligned} S(t) &= S(0)\exp\left((\alpha - \frac{\sigma^2}{2})t + \sigma\sqrt{h}Y_t - \frac{1}{2}\alpha^2 ht - \alpha\sigma h^{3/2}Y_t + \sum_{s<t}\epsilon_s\right) \\ &\simeq S(0)\exp\left((\alpha - \frac{\sigma^2}{2})t + \sigma\sqrt{h}Y_t\right) \text{ med sannolikhet 1} \end{aligned}$$

eftersom Y_t är begränsad med sannolikhet 1. Det kvarstår att visa att skuggan av den hyperfinita slumpvandringen $\sqrt{h}Y_t$ är en Wienerprocess

$$^o\left(\sqrt{h}Y_t\right) = W(t).$$

Detta följer av icke-standardmotsvarigheten av centrala gränsvärdessatsen. □

Vi känner att vi fuskar lite när vi hänvisar till centrala gränsvärdessatsen, men vårt främsta intresse är Black-Scholes-prissättningen, och för detta syfte räcker det med ett svagare villkor på $\sqrt{h}Y_T$, nämligen att skuggan är normalfördelad, vilket vi ger ett bevis för.

Bevis. Observera att vi kan skriva Y_T med hjälp av binomialfördelningen J

$$Y_T = \sum_{t\in\mathbb{T}} X_t = J - (n - J) = 2J - n.$$

Eftersom vi förväntar oss att detta blir en normalfördelning tittar vi på den karakteristiska funktionen vilket brukar vara praktiskt för normalfördelningar

$$\varphi(\xi) = \mathrm{E}\left[e^{i\xi\sqrt{h}Y_T}\right] = \sum_{j=0}^{n} e^{i\xi\sqrt{h}(2j-n)}\binom{n}{j}\frac{1}{2^n}.$$

Nu använder vi binomialsatsen baklänges genom att observera att

$$\exp\left(i\xi\sqrt{h}\right)^j \exp\left(-i\xi\sqrt{h}\right)^{n-j} = e^{i\xi\sqrt{h}(2j-n)}$$

och därmed

$$\begin{aligned} \varphi(\xi) &= \frac{1}{2^n}\sum_{j=0}^{n} e^{i\xi\sqrt{h}(2j-n)}\binom{n}{j} = \\ &= \frac{1}{2^n}\left(e^{i\xi\sqrt{h}} + e^{-i\xi\sqrt{h}}\right)^n = \left(\frac{e^{i\xi\sqrt{h}} + e^{-i\xi\sqrt{h}}}{2}\right)^n = \\ &= \cos^n\left(\xi\sqrt{h}\right) = \exp\left[n\ln\left(\cos\left(\xi\sqrt{h}\right)\right)\right]. \end{aligned}$$

Vi Taylorutvecklar $\ln\left(\cos\left(\xi\sqrt{h}\right)\right)$ kring 0 och får

$$\ln\left(\cos\left(\xi\sqrt{h}\right)\right) = -\xi^2 h/2 + \mathcal{O}\left(h^2\right)$$

kom ihåg att $h = T/n$, så

$$n \ln\left(\cos\left(\xi\sqrt{h}\right)\right) = -\xi^2 T/2 + \mathcal{O}\left(\frac{1}{n}\right)$$

och slutligen

$$\varphi(\xi) \simeq \exp\left(-\xi^2 T/2\right)$$

igen eftersom exp är S-kontinuerlig, och detta är ju den karakteristiska funktionen för en normalfördelning med varians T. Eftersom en slumpvariabel bestäms helt av sin karakteristiska funktion har vi det önskade resultatet

$${}^o\left(\sqrt{h}Y_T\right) \sim \mathcal{N}(0, T).$$

□

Vi har enkelt att skuggan av B är motsvarande ränteprocess i Black-Scholes-modellen genom

$$B(t) = B(0)\left(1 + rh\right)^{t/h} = B(0)\exp\left(\frac{t}{h}\ln(1 + rh)\right) \simeq B(0)e^{\frac{t}{h}rh} = B(0)e^{rt}.$$

För att sammanfatta resultatet, vi har definierat de hyperfinita binomialmodellprocesserna S och B på det hyperfinita sannolikhetsrummet $(\Omega, \mathcal{A}, Q)$, och för motsvarande Loeb-rum $(\Omega, \mathcal{A}_L, Q_L)$ blir dessa processer samma som i Black-Scholes-modellen.

Hyperfinita derivat C kan prissättas direkt med $\Pi(C) = \mathrm{E}_Q\left[\bar{C}\left(S(T)\right)\right]$.

För reella derivat c använder vi följande icke-standardmaskineri

1. Hitta en $\mathcal{S}$-integrerbar lifting C av c,
2. Beräkna priset med $\pi(c) = {}^o\Pi(C) = \mathrm{E}_Q\left[\bar{C}\right] = \mathrm{E}_{Q_L}\left[\bar{c}\right] = e^{-rT}\mathrm{E}_{Q_L}\left[c\right]$.

Enligt sats 44 finns en $\mathcal{S}$-integrerbar lifting C av c precis då $\pi(c) = e^{-rT}\mathrm{E}_{Q_L}\left[c\right]$ existerar, så hyperfinita binomialmodellen innehåller alltså även Black-Scholes-modellen.

8.6.3 Priset för en europeisk köpoption

Vi avslutar med att visa att den hyperfinita binomialmodellen ger Black-Scholes-priset för en europeisk köpoption som vid slutdatumet ger $c(w) = (s(w, T) - K)^+$, där $s : \Omega \times \mathbb{T} \mapsto \mathbb{R}$ definieras med $s(w, {}^o t) = {}^o\left(S(w, t)\right)$ för $t \in \mathbb{T}$. Vi har tagit med beroendet på utfallet w för att understryka att det är en slumpvariabel och inte en deterministisk funktion. Kom ihåg att Black-Scholes-priset är

$$S(0)\Phi(d_1) - Ke^{-rT}\Phi(d_2)$$
$$d_1 = \frac{\ln(\frac{S(0)}{K}) + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$
$$d_2 = \frac{\ln(\frac{S(0)}{K}) + (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

där Φ är den kumulativa normalfördelningsfunktionen.

Bevis. Vi följer de tidigare givna instruktionerna för att beräkna priset. Vi har att $C(w) =$

$(S(w,T)-K)^+$ är en $\mathcal{S}$-integrerbar lifting av c. Priset blir

$$
\begin{aligned}
\pi(c) &= e^{-rT}\mathrm{E}_{Q_L}[c] = \\
&= e^{-Tr}\ \mathrm{E}\left[\left(S(0)\exp\left((r-\frac{\sigma^2}{2})T+\sigma\sqrt{T}G\right)-K\right)^+\right],\ G\sim\mathcal{N}(0,1) \\
&= e^{-Tr}\int_{\mathbb{R}}\left(S(0)\exp\left((r-\frac{\sigma^2}{2})T-\sigma\sqrt{T}x\right)-K\right)^+ e^{-\frac{x^2}{2}}\frac{\mathrm{d}x}{\sqrt{2\pi}} \\
&= e^{-Tr}\int_{x\le d_2}\left(S(0)\exp\left((r-\frac{\sigma^2}{2})T-\sigma\sqrt{T}x\right)-K\right)e^{-\frac{x^2}{2}}\frac{\mathrm{d}x}{\sqrt{2\pi}} \\
&= -e^{-Tr}K\Phi(d_2)+e^{-Tr}\int_{x\le d_2}S(0)\exp\left((r-\frac{\sigma^2}{2})T-\sigma\sqrt{T}x-\frac{x^2}{2}\right)\frac{\mathrm{d}x}{\sqrt{2\pi}} \\
&= -e^{-Tr}K\Phi(d_2)+S(0)\int_{x\le d_2}\exp\left(-\frac{1}{2}\left(x+\sigma\sqrt{T}\right)^2\right)\frac{\mathrm{d}x}{\sqrt{2\pi}} \\
&= -e^{-Tr}K\Phi(d_2)+S(0)\Phi(d_1).
\end{aligned}
$$

□

8.7 Jämförelse med standardmetoder

Det visade sig ganska svårt att göra en rättvis jämförelse mellan hyperfinita binomialmodellen och standardmetoderna. I Black och Scholes originalartikel [19] utgår de från kontinuerliga processer och använder sig av avancerade resultat inom stokastiska processer och partiella differentialekvationer för att nå sin prissättning. Deras metod anses svår, och Cox, Ross, Rubinstein presenterade sin binomialmodell [18] som ett enklare alternativ för att nå samma resultat. Deras binomialmodell som presenterades i början av projektet har självklart uppenbara likheter med den hyperfinita binomialmodellen, men de ger inget pris för generella derivat. Istället gör de en gränsövergång från binomialpriset, specifik för den europeiska köpoptionen, för att nå Black-Scholes-priset. Förmodligen går det att göra en gränsövergång för generella derivat, men hur enkel den är vet vi inte. Dessutom förlitar de sig på viktiga resultat bevisade på annat håll, något som vi också gjort på några ställen, vilket också försvårar en jämförelse.

Vi konstaterar däremot att den hyperfinita binomialmodellen är en elegant variant av standardbinomialmodellen, där man kan utnyttja välkänd diskret kombinatorik i en kontinuerlig miljö. Genombrottet som gör det möjligt är Loeb-rummet som tillåter en att bekymmerslöst lyfta ner en hyperfinit process på en kontinuerlig standardmotsvarighet. Därför blir det enkelt att konstruera och bevisa existensen av till exempel Wienerprocessen, något som tar många sidor i anspråk för standardmotsvarigheten.

Det är också tacksamt att använda den hyperfinita binomialmodellen ur en förståelseinriktad synvinkel. Trots att matematiken som används av Black och Scholes är avancerad så bygger deras modell på väldigt enkla ekonomiska antaganden. Man kommer långt, som vi har sett, med resonemang om det diskreta fallet. Därför används binomialmodellen ofta som en inkörsport till finansiell matematik, trots att man egentligen vill nå fram till Black och Scholes prisformel. Då kan det vara fördelaktigt att studera den hyperfinita binomialmodellen, som förenar enkelhet med en intuitiv koppling och tillgång till båda prisformler. Det är ett smidigare sätt än att använda sig av binomialmodellens gränsvärde, för om inte annat så belyses sambandet mellan modellerna på ett bättre sätt av det hyperfinita alternativet.

Referenser

[1] Arkeryd L, Cutland N, Henson CW, editors. Nonstandard analysis: theory and applications. Kluwer Academic Publishers; 1997.

[2] Diener F, Diener M. Nonstandard Analysis in Practice. Verlag Berlin Heidelberg; 1995.

[3] Goldblatt R. Lectures on the hyperreals - An introduction to nonstandard analysis. Springer-Verlag New York Inc; 1998.

[4] Robinson A. Non-standard analysis. Holland; 1966.

[5] Lutz R, Goze M. Nonstandard analysis - A Practical Guide with Applications. Springer-Verlag Berlin Heidelberg; 1980.

[6] Durbin JR. Modern Algebra - An Introduction. 6th ed. John Wiley and Sons Inc.; 2009.

[7] Schiff JL. Normal Families. 6th ed. John Wiley and Sons Inc.; 2009.

[8] Davis M. Applied Nonstandard Analysis. John Wiley and Sons, Inc.; 1977.

[9] Segal SL. Nine Introductions in Complex Analysis. Revised ed. Elsevier B.V.; 2008.

[10] Cutland N, Kopp E, Willinger W. A Nonstandard Approach to Option Pricing. Mathematical Finance. 1991;1(4): 1-38.

[11] Hoskins RF. Standard and nonstandard analysis - fundamental theory, techniques and applications.

[12] Rotman JJ. An Introduction to Algebraic Topology. Springer-Verlag New York Inc; 1988.

[13] Taylor JL. Complex variables. American Mathematical Society; 2011.

[14] Berkeley G. The Analyst: a Discourse addressed to an Infidel Mathematician. London edition; 1734.

[15] Berkeley G. Philosophical Commentaries. Notebooks; 1707-1708.

[16] Cantor G. Mitteilungen zur Lehre vom Transfiniten. Pfeffer; 1887.

[17] Yanovskaya S. The Mathematical Manuscripts of Karl Marx. New Park Publications; 1983.

[18] Cox J, Ross S, Rubinstein M. Option pricing: A simplified approach. Journal of Financial Economics. 1979;7(3): 229-263.

[19] Black F, Scholes M. The Pricing of Options and Corporate Liabilities. Journal of Political Economy. 1973;81(3): 637-654.

[20] Borell C. Introduction to the Black-Scholes theory. Chalmers Tekniska Högskola 412 96 Göteborg; 2011.

[21] Kopp E. Hyperfinite Mathematical Finance. L.O. Arkeryd et al. (eds.), Nonstandard Analysis: Theory and Applications: 279-307.

[22] Van Den Berg I.P., Koudjeti F. From binomial expectations to the Black-Scholes formula: the main ideas. Annales mathématiques Blaise Pascal. 1997;4(1): 93-101.