Kapitel 1

Elementare Nonstandard-Analysis

Überblick: Wir beginnen mit einem historischen Abriß. Danach stellen wir das wichtigste Hilfsmittel für die Konstruktion eines hyperrellen Körpers ${}^*\mathbb{R} \supset \mathbb{R}$ mit unendlich kleinen und großen Zahlen bereit. Im dritten Abschnitt konstruieren wir einen solchen Körper ${}^*\mathbb{R}$ in voller Analogie dazu wie man $\mathbb{R}$ aus $\mathbb{Q}$ durch Cauchyfolgen konstruiert. Im vierten Abschnitt zeigen wir, wie man Funktionen und Relationen auf $\mathbb{R}$ auf den hyperreellen Körper ${}^*\mathbb{R}$ unter Erhalt der wesentlichen Eigenschaften fortsetzen kann. In den Abschnitten 5 und 6 wenden wir dann dies alles auf die Differential- und Integralrechnung an. Wir erhalten so unter anderem den Existenzsatz von Peano für die Lösung gewöhnlicher Differentialgleichungen.

1.1 Einleitung

Überblick: Wir skizzieren kurz einige wesentliche Entwicklungsschritte der Analysis einschließlich der Nonstandard-Analysis.

Die Entwicklung der Differential- und Integralrechnung begann mit G.W. Leibniz (1646-1716) (wesentliches Manuskript schon 1675) und I. Newton (1642-1727). Leibniz erweiterte geometrische Ideen von B. Pasquale, Newton entwickelte die Mechanik mit seiner Form der Analysis. Beide benutzten unendlich kleine und unendlich große Zahlen. Die damit verbundenen Schwierigkeiten waren beiden bewußt.

Hier ist ein einfaches Beispiel für solche Schwierigkeiten: *Was ist der Quotient zweier unendlich kleiner Zahlen?*
Sei x vom Betrag "unendlich klein". Dann ist auch $\sin(x)$ unendlich klein, der Quotient ergibt $\frac{\sin x}{x}$„="1. Aber auch $1 - \cos(x)$ ist dann unendlich klein wegen $\cos(x) = \sqrt{1 - \sin(x)^2}$. Aber $\frac{1 - \cos(x)}{x}$„="0. Ja, auch $\sqrt{|x|}$ ist unendlich klein, aber $\frac{\sqrt{|x|}}{|x|}$„="∞.

Diese und viele weitere Probleme waren beiden Mathematikern bekannt. Beide lösten sie auf verschiedene Weise: Newton argumentierte physikalisch, Leibniz berief sich darauf, daß man jeden seiner Beweise auch so führen könne wie Archimedes dies bei der Berechnung des Kreisinhalts z.B. getan hatte. Trotz der bleibenden Ungereimtheiten entwickelte sich die Analysis im 18. Jahrhundert zu ihrer ersten Blüte, vor allem dank Euler (1707-1783), der eine hohe Intuition beim Umgang mit unendlich kleinen Größen hatte. Er berechnete beispielsweise

$$e = \lim(1+\frac{1}{n})^n = 1 + \binom{n}{1}\cdot\frac{1}{n} + \binom{n}{2}\frac{1}{n^2} + \cdots + \binom{n}{n}\frac{1}{n^n}$$

und argumentierte nun, daß für unendlich großes n

$$\binom{n}{k}\frac{1}{n^k} = \frac{1}{k!}\,\frac{n(n-1)\cdots(n-(k+1))}{n\cdot n\cdots n} \text{ „}{=}\text{“ } \frac{1}{k!}$$

sei wegen $\frac{n}{n} = 1$ „$=$“ $\dfrac{n-1}{n}$ „$=$“ $\ldots$ „$=$“ $\dfrac{n-(k+1)}{n}$,
woraus er $e = \sum_{k=0}^{\infty}\frac{1}{k!}$ folgerte. Dies ist nur ein kleines Beispiel aus Eulers epochalem Werk.

A. Cauchy (1789-1857) gilt als der erste Mathematiker, der die unendlich kleinen und großen Zahlen aus der Analysis eliminiert und z.B. durch sein Cauchy-Kriterium der Konvergenz bzw. sein Stetigkeitskriterium ersetzt hat. Aber dies trifft nicht zu. Er hat nach wie vor auch mit solchen Zahlen operiert und argumentiert. Erst R. Dedekind (1831-1916) und vor allem K. Weierstrass (1815-1897) haben diese Zahlen durch die heute übliche $\varepsilon - n_0-$ bzw. $\varepsilon - \delta$–Argumentation ersetzt. Diese ihrerseits ist nichts anderes als eine systematisierte Anwendung der oben erwähnten Argumentationsweise von Archimedes und damit auch von Leibniz.

A. Robinson (1918 - 1974) entdeckte um 1955-1960, daß man mit Hilfe der modernen Logik unendlich kleine und große Zahlen widerspruchsfrei so einführen kann, daß damit die ganze Analysis aufgebaut werden kann. Die damit verbundenen Vorteile sind größere Anschaulichkeit, weniger komplizierte Beweise und die Möglichkeit auch andere Gebiete der Mathematik darauf aufzubauen. Insbesondere gibt die Robinsonsche Nonstandard-Analysis in der Weiterentwicklung von W.A. J. Luxemburg, C.W. Henson und P.A. Loeb einen hervorragenden Zugang zur Stochastik und Funktionsanalysis.

1.2 Ein fundamentales Hilfsmittel zur Konstruktion unendlich kleiner und großer Zahlen

Überblick: Wir werden im nächsten Hauptabschnitt die hyperreellen Zahlen ähnlich aus den reellen Zahlen konstruieren wie man diese aus den rationalen Zahlen mit Cauchyfolgen konstruiert. Dazu brauchen wir eine Funktion m, die jeder Teilmenge A der natürlichen Zahlen einen Wert $m(A) \in \{0,1\}$ zuordnet, und gewisse weitere Eigenschaften erfüllt, die im Hauptsatz dieses Abschnitts formuliert sind (Satz 1.5). Für die Konstruktion von m benötigen wir den Satz von Zorn.

1.2.1 Der Satz von Zorn

Dieser Satz leistet hervorragende Dienste auf allen Gebieten der Mathematik, in denen es um die Existenz maximaler Elemente einer teilweise geordneten Menge geht. Er ist äquivalent zum Auswahlaxiom. Wir benutzen eine besonders leicht verständliche Version. Sei $\emptyset \neq X$ eine beliebige nichtleere Menge, zum Beispiel $X = \{1,2,3\}$, $X = \mathbb{N}$ oder $X = \mathbb{R}$. Wir betrachten die Menge $\{0,1\}^X = F(X)$ aller Funktionen von X in die zweielementige Menge $\{0,1\}$. $F(X)$ ist in natürlicher Weise geordnet: $f \leq g$ gilt nach Definition genau dann, wenn $f(x) \leq g(x)$ für alle $x \in X$ ist.

Ist zum Beispiel $X = \{1,2,3\}$ so ist $f(x) = \begin{cases} 0 & x = 1,3 \\ 1 & x = 2 \end{cases} \quad \leq g(x) = \begin{cases} 0 & x = 1 \\ 1 & x = 2,3 \end{cases}$

Die Ordnung ist aber *nicht* **total**: es gibt Elemente f und h, die nicht vergleichbar sind. Wählen Sie zum Beispiel f wie oben, $h(x) = \begin{cases} 1 & x = 1,3 \\ 0 & x = 2 \end{cases}$

Sie kennen solche Phänomene natürlich von den üblichen reellen Funktionen auf einer Menge. Natürlich - und das haben Sie oben gesehen - gibt es total geordnete Teilmengen A in $F(X)$. Zum Beispiel wählen Sie oben $A = \{f, g, \mathbf{1}_X\}$ wo $\mathbf{1}_X$ die Funktion ist, die konstant gleich 1 ist.

Genauer heißt $A \subset F(X)$ **total geordnet**, wenn je zwei Elemente $f, g \in A$ **vergleichbar** sind, d.h. entweder $f \leq g$ oder $g \overset{\neq}{<} f$ gilt.

Deffnition 1.1 *Sei $G \subset F(X)$. Eine Funktion $g \in G$ heißt* **maximales Element** *von G, wenn aus $g \leq h \in G$ stets $g = h$ folgt.*

Als Beispiel stellen wir fest, daß die konstante Funktion $\mathbf{1}_X$ maximal in $F(X)$ ist. Wählen wir wieder $X = \{1,2,3\}$ und $G = \{f \in F(X) : f \neq \mathbf{1}_X\}$, so hat G *drei verschiedene* maximale Elemente, nämlich

$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ (wo $f = \begin{pmatrix} f(1) \\ f(2) \\ f(3) \end{pmatrix}$ geschrieben wurde).

Im folgenden Satz von Zorn wird nur die *Existenz* eines maximalen Elements behauptet, nicht jedoch seine Eindeutigkeit. Das kann man eben auch gar nicht immer.

Theorem 1.2 (Satz von Zorn) *Sei $G \subset F(X)$ eine Teilmenge. Für jede total geordnete Teilmenge $H \subset G$ gebe es eine* Majorante *$f_H \in G$ von H, d.h. es gebe ein $f_H \in G$ mit $f_H \geq f$ für alle $f \in G$. Dann besitzt G ein maximales Element g_G.*

Für einen Beweis dieses Satzes verweisen wir auf Bücher über Mengenlehre.

1.2.2 Anwendung auf Mengenfunktionen

Im Folgenden sei I eine unendliche Menge; am wichtigsten für den ersten Teil der Vorlesung ist $I = \mathbb{N}$. Wir setzen $X = \mathcal{P}(I) = \{A : A \subset I\}$. Statt f, g schreiben wir jetzt m bzw. m' usw.
G sei die Menge aller $m \in F(X)$ mit den drei Eigenschaften

1. $m(A) = 0$ falls A eine *endliche* Teilmenge ist.
2. $m(A \cap B) = 1$ genau dann, wenn $m(A) = 1 = m(B)$.
3. $m(I) = 1$.

G ist nicht leer. Denn setze zum Beispiel $m(A) = \begin{cases} 1 & A = I \\ 0 & \text{sonst} \end{cases}$

Aufgabe: Beweisen Sie, daß dies m zu G gehört, d.h. die Eigenschaften 1) - 3) hat.

Wir zeigen nun, daß G maximale Elemente besitzt, indem wir beweisen, daß G die Voraussetzungen des Satzes von Zorn erfüllt.

Satz 1.3 *Es gibt maximale Elemente in G.*

Beweis: Sei H eine total geordnete Teilmenge von G. Wir setzen

$$m_H(A) = \max\{m(A) : m \in H\}.$$

Offensichtlich gilt $m_H \geq m$ für alle $m \in H$. Außerdem erfüllt m_H natürlich die Eigenschaften (1) und (3), da ja jedes $m \in H$ dies tut.
Behauptung: m_H hat die Eigenschaft (2(. Damit ist dann m_H eine Majorante von H in G.
Beweis: Sei zunächst $m_H(A \cap B) = 1$. Nach Definition von m_H muß es dann ein $m \in H$ mit $m(A \cap B) = 1$ geben. Dann ist aber wegen $H \subset G$ auch $m(A) = m(B) = 1$, also erst recht $m_H(A) = m_H(B) = 1$.
Sei umgekehrt $m_H(A) = m_H(B) = 1$. Dann gibt es ein $m_1 \in H$ mit $m_1(A) = 1$ und ein $m_2 \in H$ mit $m_2(B) = 1$. Nach Voraussetzung ist H total geordnet, es gilt also $m_1 \leq m_2$ oder $m_2 \leq m_1$.

O. B. d. A. sei $m_1 \leq m_2$. Dann ist also $m_2(A) = 1 = m_2(B)$, also $m_2(A \cap B) = 1$ und damit $m_H(A \cap B) = 1$. m_H hat also tatsächlich Eigenschaft (2).
Damit hat jede total geordnete Teilmenge H eine Majorante in G Nach dem Satz von Zorn hat G maximale Elemente. □

Deffnition 1.4 *Ein maximales Element der Menge G heißt* **Ultrafunktional** *auf $\mathcal{P}(I)$.*

Die Eigenschaften eines Ultrafunktionals sind im nächsten Satz beschrieben.

Satz 1.5 *Sei m ein Ultrafunktional auf $X = \mathcal{P}(I)$. Dann gilt:*
a) Für alle $A \subset I$ (also $A \in X$) ist entweder $m(A) = 1$ oder $m(A^c) = 1 \quad (A^c = I \setminus A)$
b) $m(A \cup B) = 1 \Leftrightarrow m(A) = 1$ oder *$m(B) = 1$*
c) Sind $A_1, \cdots, A_n$ paarweise disjunkt und ist $m(A_1 \cup \cdots \cup A_n) = 1$, so gibt es genau ein $k \leq n$ mit $m(A_k) = 1, \quad m(A_j) = 0$ für $j \neq k$.

Beweis: a) Zunächst kann nicht $m(A) = m(A^c) = 1$ gelten. Denn sonst wäre nach Eigenschaft (2) auch $m(A \cap A^c) = 1$. Wegen $A \cap A^c = \emptyset$ würde $m(\emptyset) = 1$ gelten, ein Widerspruch zu Eignschaft (1). Ferner gilt $m(A) = 1$ für alle A, für die A^c endlich ist. Denn sonst setze

$$m_1(D) = \begin{cases} 1 & \exists A \exists C [A^c \,\text{ist endlich und}\, m(C) \;=\; 1 \,\text{und}\, D \supset C \cap A] \\ 0 & \text{sonst} \end{cases}$$

Dann ist $m < m_1 \in G$ im Widerspruch zur Maximalität von m.
Wir nehmen nun an, es gebe ein $A \subset I$ mit $m(A) = m(A^c) = 0$. Wir zeigen, daß dann m nicht maximal ist. Nach dem bisher Gezeigten ist sowohl A als auch A^c unendlich.
(I) Es gelte $C \cap A \neq \emptyset$ für alle C mit $m(C) = 1$.
Wir setzen dann $\overline{m}(D) = \begin{cases} 1 & : \text{Es gibt}\; C \,\text{mit}\, m(C) = 1 \,\text{und}\, D \supset C \cap A \\ 0 & : \text{sonst} \end{cases}$
Es ist nicht schwer zu sehen, daß $\overline{m} \in G$ ist. Außerdem gilt $\overline{m}(C) = m(C)$, falls $m(C) = 1$ war. Schließlich ist $\overline{m}(A) = 1$, also $\overline{m} > m$, ein Widerspruch zur Maximalität von m.
(II) Es gebe ein C_0 mit $m(C_0) = 1$ aber $C_0 \cap A = \emptyset$. Dann ist $C_0 \subset A^c$. Ist $m(C) = 1$ und $C \cap A^c = \emptyset$, so ist $C \subset A$, also $C \cap C_0 = \emptyset$ und damit $0 = m(\emptyset) = m(C \cap C_0) = 1$ nach Eigenschaft 2., ein Widerspruch. Also ist jetzt $C \cap A^c \neq \emptyset$ für alle C mit $m(C) = 1$ und man kann (I) anwenden für A^c statt A.
b) "$\Rightarrow$" (indirekt): Angenommen, $m(A) = m(B) = 0$. Nach a) ist dann $m(A^c) = m(B^c) = 1$, also mit Eigenschaft 2. $m((A \cup B)^c) = m(A^c \cap B^c) = 1$ und damit wieder nach a) $m(A \cup B) = 0$, ein Widerspruch zur Voraussetzung.
"$\Leftarrow$": ganz analog.
c) Wir beweisen dies durch Induktion: Sei $m(A_1 \cup A_2) = 1$ mit $A_1 \cap A_2 = \emptyset$. Nach b) ist mindestens eine der beiden Größen gleich 1.
Angenommen $m(A_1) = 1 = m(A_2)$. Sei $A_3 = (A_1 \cup A_2)^c$. Dann ist nach b) $m(A_2 \cup A_3) = 1$ und $A_1 = (A_2 \cup A_3)^c$, also $m(A_1) = 0$ nach a), ein Widerspruch. Der Rest ist nun klar. □

Wir liefern nun den Anschluß an die vorhandene Literatur:

Satz 1.6 *a) Sei $m \in G$ beliebig und $\mathcal{F}_m = \{A : m(A) = 1\}$. Dann gilt:*
(i) $A \cap B \in \mathcal{F} \Leftrightarrow A$ und $B \in \mathcal{F}_m$.
(ii) $\emptyset \notin \mathcal{F}_m$ und $I \in \mathcal{F}_m$.
(iii) A endlich impliziert $A \notin \mathcal{F}_m$.
b) m ist genau dann ein Ultrafunktional, wenn $\mathcal{F}_m \subset \mathcal{P}(I)$ maximal ist bezüglich (i) - (iii), d. h. wenn gilt: Ist $\mathcal{G} \subset \mathcal{P}(I)$ mit (i) - (iii) und $\mathcal{F}_m \subset \mathcal{G}$, so ist $\mathcal{F}_m = \mathcal{G}$.

Wir unterdrücken den einfachen Beweis und notieren nur: Eine Teilmenge $\mathcal{F} \subset \mathcal{P}(I)$ mit den Eigenschaften (i) und (ii) heißt **Filter**. Ein maximaler Filter mit der zusätzlichen Eigenschaft (iii) heißt **freier Ultrafflter**. Davon ist unsere Namensgebung "Ultrafunktional" für maximale Elemente aus G beeinflußt.

Tom Lindstrom ("An invitation to nonstandard analysis") benutzt wie wir ein Ultrafunktional, nennt es abe wegen der Eigenschaft c) in Satz 1.5 ein *Maß*. Ich habe davon Abstand genommen, weil viele Angst vor diesem Wort haben.

Aufgaben:

1. Sei m eine beliebige Funktion aus der Menge G, also mit den drei Eigenschaften 1. bis 3. auf S. 4 (s. a. die ersten drei in der Wiederholung). Zeigen Sie bitte: Ist $A \subset B$ so ist $m(A) \leq m(B)$.
   Tipp: Dann ist $A = A \cap B$.

2. Beweisen Sie bitte Satz 1.6.

Wiederholung:
Auf der Potenzmenge $\mathcal{P}(I)$ einer beliebigen unendlichen Indexmenge I gibt es ein Ultrafunktional, das heißt, eine Funktion m mit Werten in $\{0, 1\}$, die die folgenden Eigenschaften besitzt (die ersten drei stammen aus der Definition der Menge G, die letzten drei aus Satz 1.5):

1. *$m(A) = 0$ falls A eine* endliche *Teilmenge ist.*

2. *$m(A \cap B) = 1$ genau dann, wenn $m(A) = 1 = m(B)$.*

3. *$m(I) = 1$.*

4. *Für alle $A \subset I$ (also $A \in X$) ist* entweder $m(A) = 1$ oder $m(I \setminus A) = 1$.

5. $m(A \cup B) = 1 \Leftrightarrow m(A) = 1$ oder $m(B) = 1$.

6. *Sind $A_1, \cdots, A_n$ paarweise disjunkt und ist $m(A_1 \cup \cdots \cup A_n) = 1$, so gibt es genau ein $k \leq n$ mit $m(A_k) = 1, \quad m(A_j) = 0$ für $j \neq k$.*

Dies ist alles, was wir in Zukunft brauchen.

1.3 Ein hyperreeller Körper

1.3.1 Rückblick: Konstruktionen von $\mathbb{R}$

In Analysis I lernt man oft, wie man $\mathbb{R}$ aus $\mathbb{Q}$ erhalten kann. Es gibt Cauchyfolgen in $\mathbb{Q}$, die in $\mathbb{Q}$ nicht konvergieren, zum Beispiel $a_0 = 3$, $a_{n+1} = \frac{1}{2}(a_n + \frac{2}{a_n})$. Man sucht eine Körpererweiterung $\mathbb{R}$, in der jede Cauchyfolge konvergiert. Dazu sei $\mathcal{L}$ die Menge aller Cauchyfolgen $\xi = (x_n)_n$. Die Summe und das Produkt zweier Cauchyfolgen ist wieder eine Cauchyfolge. Ferner ist mit $\xi = (x_n)_n$ auch $|\xi| = (|x_n|)_n$ eine Cauchyfolge.
Man definiert: zwei Cauchyfolgen $\xi = (x_n)_n$ und $\eta = (y_n)_n$ heißen äquivalent, in Zeichen $\xi \sim \eta$, wenn die Differenzfolge $\xi - \eta = (x_n - y_n)_n$ eine Nullfolge ist. Das ist tatsächlich eine Äquivalenzrelation in $\mathcal{L}$, das heißt es gilt:

1. $\xi \sim \xi$
2. $\xi \sim \eta \Rightarrow \eta \sim \xi$
3. $\xi \sim \eta$ und $\eta \sim \tau \Rightarrow \xi \sim \tau$

Diese Äquivalenzrelation ist *verträglich* mit den algebraischen Operationen, das heißt:

1. $\xi \sim \xi'$ und $\eta \sim \eta' \Rightarrow \xi + \eta \sim \xi' + \eta'$
2. $\xi \sim \xi'$ und $\eta \sim \eta' \Rightarrow \xi \cdot \eta \sim \xi' \cdot \eta'$
3. $\xi \sim \xi' \Rightarrow |\xi| \sim |\xi'|$

Man bildet nun die Menge A der Äquivalenzklassen. $A = \{\dot{\xi} : \xi \in \mathcal{L}\}$. Dabei ist $\dot{\xi} = \{\eta \in \mathcal{L} : \eta \sim \xi\}$.
Wegen der Verträglichkeit mit den algebraischen Operationen können wir Addition, Multiplikation und Absolutbetrag für die Äquivalenzklassen eindeutig erklären:
$\dot{\xi} + \dot{\eta}$ ist die Äquivalenzklasse, die $\xi + \eta$ enthält, also $\dot{\xi} + \dot{\eta} = (\xi + \eta)^{\cdot}$
Analog:

$$\begin{aligned} \dot{\xi} \cdot \dot{\eta} &= (\xi \cdot \eta)^{\cdot} \\ |\dot{\xi}| &= (|\xi|)^{\cdot} \end{aligned}$$

In Analysis I zeigt man, daß A mit der Addition und Multiplikation ein Körper wird, der durch

$$\dot{\xi} \leq \dot{\eta} \Leftrightarrow |\dot{\eta} - \dot{\xi}| = \dot{\eta} - \dot{\xi}$$

total geordnet wird. Man zeigt ferner, daß in A jede Cauchyfolge konvergiert und A $\mathbb{Q}$ als dichte Teilmenge enthält. Kurz: A ist gleich $\mathbb{R}$.

1.3.2 Konstruktion eines hyperreellen Körpers

Überblick: Wir wiederholen den Begriff einer Äquivalenzrelation. Wir führen auf der Menge $\mathbb{R}^{\mathbb{N}}$ *aller* reellen Zahlenfolgen die üblichen algebraischen Operationen ein. Dann wählen wir ein Ultrafunktional m auf $\mathbb{N}$ und definieren mit dessen Hilfe eine Äquivalenzrelation, die mit den algebraischen Operationen verträglich ist. Die Menge aller Äquivalenzklassen wird damit auf natürliche Weise zu einem geordneten Körper $^*\mathbb{R}$ der "unendlich kleine" und "unendlich große" Zahlen enthält. Man kann $\mathbb{R}$ in kanonischer Weise in $^*\mathbb{R}$ einbetten, und in gewisser Weise wieder aus $^*\mathbb{R}$ zurückgewinnen.

Äquivalenzrelationen

Deffnition 1.7 *Sei $\emptyset \neq X$ eine Menge und $E \subset X \times X$ eine Teilmenge mit den Eigenschaften*

1. $\Delta = \{(x,x) : x \in X\} \subset E$.
2. *Ist $(x,y) \in E$ so auch (y,x).*
3. *Sind (x,y) und $(y,z) \in E$, so ist auch $(x,z) \in E$.*

Dann heißt E eine Äquivalenzrelation auf X.

Statt $(x,y) \in E$ schreibt man einfach $x \sim y$, oder falls man mehrere Äquivalenzrelationen betrachtet, $x \sim_E y$.

Beispiele:

1. Sei $\emptyset \neq X$ beliebig. Die übliche Gleichheit $x = y$, also $E = \Delta$ ist natürlich eine Äquivalenzrelation.
2. Sei $X = \mathbb{Z}$. $E_{12} = \{(x,y) : x - y \text{ ist durch 12 teilbar }\}$ ist eine Äquivalenzrelation. Wir benutzen sie im täglichen Leben.
3. $X = \mathbb{Z}$, $\quad E_{24} = \{(x,y) : x - y \text{ ist durch 24 teilbar }\}$ wird ebenfalls im täglichen Leben benutzt, wo?
4. $X = \mathcal{L}$ (s. den Rückblick auf $\mathbb{R}$) $\qquad E = \{(\xi,\eta) : \xi - \eta \text{ ist Nullfolge }\}$
5. Sei V ein Vektorraum über $\mathbb{R}$ und $W \subset V$ ein Untervektorraum, $E = E_W = \{(x,y) : x - y \in W\}$.

6. $X = \{f : \mathbb{R} \to \mathbb{R} | f$ ist meßbar $\}$, $E = \{(f,g) : f - g$ ist Nullfunktion $\}$ (nur für Hörer/Innen von Analysis IV)

Ist E (bzw. $\sim$) eine Äquivalenzrelation auf X, so zerfällt X in Äquivalenzklassen:
Sei $x \in X$ und $\dot{x} = \{y : (x,y) \in E\}$. Dann gilt für $x, y \in X$ entweder $\dot{x} = \dot{y}$ oder $\dot{x} \cap \dot{y} = \emptyset$.

Aufgabe: Beweisen Sie bitte diese Aussage.

Sei $\dot{x}$ eine Äquivalenzklasse. Jedes Element $y \in \dot{x}$, also jedes zu x äquivalente Element y heißt ein **Repräsentant dieser Äquivalenzklasse**. Zum Beispiel sind die Ziffern auf einer Uhr Repräsentanten verschiedener Äquivalenzklassen der Äquivalenzrelation E_{12} auf $\mathbb{Z}$ (s.o.)
Äquivalenzklassen sind neue mathematische Objekte, gebildet aus alten mit Hilfe von einer Äquivalenzrelation. Die Menge aller Äquivalenzklassen ist also an sich eine Menge von Mengen; betrachtet man jedoch jede Äquivalenzklasse als ein neues Objekt, vergißt also die Herkunft, so wird wieder alles "ganz normal".

Beispiele:

1. Führt man die reellen Zahlen wie wir in unserem Rückblick ein, so ist eine reelle Zahl eine Äquivalenzklasse von Cauchyfolgen. Wie gut, daß wir das wieder vergessen und eine reelle Zahl unmittelbar als (neues) Objekt betrachten (dürfen).

2. Schon die ganzen Zahlen werden korrekt aus den natürlichen Zahlen über eine Äquivalenzrelation eingeführt (siehe die Aufgabe unten).

3. Hat man die ganzen Zahlen, so erhält man eine beliebige rationale Zahl $\mathbb{Q}$ wieder als Äquivalenzklasse (siehe die Aufgabe unten).

Aufgaben:

1. Sei $\mathbb{N}_0 = \{0, 1, 2, \cdots\}$ die Menge der natürlichen Zahlen einschließlich der 0. Sei $X = \mathbb{N}_0 \times \mathbb{N}_0$. Auf X definieren wir $E = \{((x,y),(u,v)) : x + v = y + u\}$.
   Die Äquivalenzklasse $(5,0)^{\cdot}$ entspricht dann der natürlichen Zahl 5, die Äquivalenzklasse $(0,5)^{\cdot}$ der negativen ganzen Zahl -5. Das wird erst klar, wenn wir auf der Menge der Äquivalenzklassen eine Addition definieren. Zeigen Sie bitte:
   a) E ist eine Äquivalenzrelation.
   b) Durch $(x,y)^{\cdot} + (x',y')^{\cdot} = (x + x', y + y')^{\cdot}$ wird eindeutig eine assoziative, kommutative Addition auf der Menge $\dot{X}$ aller Äquivalenzklassen definiert. $(0,0)^{\cdot}$ ist das neutrale Element und $(y,x)^{\cdot}$ ist das Inverse zu $(x,y)^{\cdot}$
   c) $\mathbb{N}_0 \ni n \to (n,0)^{\cdot}$ ist injektiv mit $(n+m) \to (n,0)^{\cdot} + (m,0)^{\cdot}$, und jedes Element $(x,y)^{\cdot}$ ist gleich $(x,0)^{\cdot} + (0,y)^{\cdot}$. Schließen Sie hieraus, daß wir damit die ganzen Zahlen konstruiert haben.

2. Genau so konstruiert man die Brüche.

$$X = \mathbb{Z} \times \mathbb{Z}, \qquad E = \{((x,y),(u,v)) : xv = yu, x,y,u,v \in \mathbb{Z}\}.$$

Schauen Sie in [?] nach.

Zur Vorbereitung der Konstruktion eines hyperreellen Zahlkörpers wollen wir an einem ganz einfachen Beispiel das Vorgehen aufzeigen.
Wir betrachten $\mathbb{Z}$ mit der Addition und Multiplikation und setzen $x \sim y \Leftrightarrow x - y$ ist durch 3 teilbar. Es gibt genau drei Äquivalenzklassen, nämlich
$\dot{0} = \{x : x$ ist durch 3 teilbar $\}$,
$\dot{1} = \{x : x : 3$ hat Rest $1\} = \{x : x = 3k(x) + 1$ für ein $k(x) \in \mathbb{Z}\}$,
$\dot{2} = \{x : x : 3$ hat Rest $2\} = \{x : x = 3k(x) + 2$ für ein $k(x) \in \mathbb{Z}\}$.

Für 2 beliebige Mengen $M, N \neq \emptyset$ in $\mathbb{Z}$ setzen wir

$$\begin{aligned} M + N &= \{x + y : x \in M, y \in N\} \\ M \cdot N &= \{xy : x \in M, y \in N\} \end{aligned}$$

Dann ergibt sich $\dot{0} + \dot{0} = \dot{0}$ und allgemeiner

$$\begin{aligned} \dot{x} + \dot{y} &= (x+y)^{\cdot} \\ \dot{x} \cdot \dot{y} &= (xy)^{\cdot} \end{aligned}$$

Mit diesen Verknüpfungen ist $\{\dot{0}, \dot{1}, \dot{2}\}$, also die Menge der Äquivalenzklassen ein Körper.
Aufgabe: Beweisen Sie dies bitte.

Geordnete Körper

Deffnition 1.8 *Sei $\mathcal{K}$ ein Körper, auf dem eine totale Ordnung $\leq$ mit den folgenden Verträglichkeitseigenschaften erklärt ist:*

1. $x \leq y \Rightarrow x + a \leq y + a$ *für alle* $a \in \mathcal{K}$.
2. $x \leq y$ *und* $a \geq 0 \Rightarrow ax \leq ay$.

Dann heißt $\mathcal{K}$ **geordneter Körper**. *Er heißt* **archimedisch geordnet**, *wenn zu jedem $0 < x < y$ stets eine natürliche Zahl n existiert mit*

$$\underbrace{x + x + \cdots + x}_{n \text{ mal}} = n \cdot x > y.$$

Gilt dies nicht, so heißt $\mathcal{K}$ **nichtarchimedischer Körper**.

Typische Beispiele für geordnete Körper sind $\mathbb{Q}$ und $\mathbb{R}$. Beide sind archimedisch geordnet. "Unendlich kleine´´ und "unendliche große´´ Zahlen sind in archimedisch geordneten Körpern unmöglich. Denn wie kann man sinnvoll solche Begriffe definieren? Dazu definieren wir genau wie im Reellen in einem beliebigen geordneten Körper $\mathcal{K}$ den Absolutbetrag einer Zahl durch $|x| = \begin{cases} x & 0 \leq x \\ -x & x < 0 \end{cases}$

Deffnition 1.9 *Sei $\mathcal{K}$ ein geordneter Körper mit Einselement* 1. *$x \in \mathcal{K}$ heißt* **unendlich groß**, *wenn für alle n in $\mathbb{N}$ stets*

$$\underbrace{1 + 1 + \cdots + 1}_{n \text{ mal}} = n \cdot 1 < x$$

gilt.
x heißt **(betragsmäßig) unendlich klein**, *wenn $x = 0$ oder $|1/x|$ unendlich groß ist.*

x ist also unendlich klein wenn für alle $n \in \mathbb{N}$ stets $n|x| < 1$ ist.

Aus dieser einzig möglichen Definition folgt automatisch, daß Körper mit unendlich großen Elementen nichtarchimedisch sind.

Bevor wir solch einen Körper konstruieren, wollen wir kurz festlegen, was wir unter einer Körpereinbettung verstehen wollen.

Deffnition 1.10 *a) Seien $\mathcal{K}$ und $\mathcal{L}$ zwei Körper. Eine Abbildung γ von $\mathcal{K}$ in $\mathcal{L}$ heißt eine* **Körpereinbettung**, *wenn $\gamma(x + y) = \gamma(x) + \gamma(y)$, $\gamma(xy) = \gamma(x)\gamma(y)$ und $\gamma(1) \neq 0$ gilt. $\gamma(\mathcal{K})$ ist dann ein zu $\mathcal{K}$ isomorpher Teilkörper von $\mathcal{L}$, der selbst auch Einbettung heißt.*
b) Sind zusätzlich $\mathcal{K}$ und $\mathcal{L}$ geordnet, und gilt

$$x \leq y \Rightarrow \gamma(x) \leq \gamma(y),$$

so sprechen wir von einer **geordneten Körpereinbettung**.

Aus der Algebra ist bekannt, daß γ automatisch injektiv ist, wenn es additiv und multiplikativ ist und $\gamma(1) \neq 0$ gilt.

Aufgabe: Beweisen Sie dies.
Tipp: Zeigen Sie, daß andernfalls $\gamma(1) = \gamma(x)\gamma(1/x) = 0$ wäre für ein $x \neq 0$ mit $\gamma(x) = 0$.

Schließlich wollen wir natürlich unseren wohlbekannten Körper $\mathbb{R}$ in unseren "neuen´´ Körper einbetten können. Dann kann man ihn ja mit seinem Bild identifizieren.

Deffnition 1.11 *Ein geordneter Körper heißt* **hyperreell**, *wenn er eine geordnete Einbettung von $\mathbb{R}$ als echten Teilkörper enthält.*

Die Konstruktion eines hyperreellen Körpers

Sei $X = \mathbb{R}^{\mathbb{N}}$ die Menge aller reellen Zahlenfolgen. Zwei Folgen$\xi = (x_n)_n$ und $\eta = (y_n)_n$ werden wie üblich gliedweise addiert und multipliziert.
Für die konstante Folge $\mathbf{1} = (1, 1, \cdots)$ gilt $\mathbf{1}\xi = \xi\mathbf{1} = \xi$.
Für die konstante Folge $\mathbf{0} = (0, 0, \cdots)$ gilt $\mathbf{0} + \xi = \xi + \mathbf{0} = \xi$ und $\mathbf{0}\xi = \xi\mathbf{0} = \mathbf{0}$. Die Operationen sind assoziativ und $-\xi = (-x_n)_n$ ist das additive Inverse zu ξ.
Sei nun m ein fest gewähltes Ultrafunktional auf $\mathbb{N}$. Wir definieren $(x_n)_n \sim (y_n)_n$ wenn $m(\{n : x_n = y_n\}) = 1$ gilt.
Um uns das Leben zu erleichtern, schreiben wir $[\xi = \eta]$ für $\{n : x_n = y_n\}$, wo $\xi = (x_n)_n, \eta = (y_n)_n$. Ferner sagen wir A ist eine 1–Menge, wenn $m(A) = 1$ bzw. A ist eine 0–Menge, wenn $m(A) = 0$ gilt. Damit lautet die Definition:

$$\xi \sim \eta \text{ falls } [\xi = \eta] \text{ eine 1–Menge ist.}$$

Lemma 1.12 *“$\sim$“ ist eine Äquivalenzrelation auf* $\mathbb{R}^{\mathbb{N}}$.

Aufgabe: Beweisen Sie bitte dieses Lemma.

Bemerkung: Für die, die Algebra II gehört haben: $X = \mathbb{R}^{\mathbb{N}}$ ist ein kommutativer Ring mit Einselement. $J = \{\xi : \xi \sim 0\}$ ist ein maximales Ideal, und X/J ist der gewünschte Körper, dem nur noch die Ordnung fehlt.

Sei ${}^*\mathbb{R}$ die Menge der Äquivalenzklassen.

$${}^*\mathbb{R} = \{\dot{\xi} : \xi \in \mathbb{R}^{\mathbb{N}}\}.$$

Wie im Eingangsbeispiel mit $\mathbb{Z}$ und $E = E_3$ setzen wir für nichtleere Mengen M, N von X

$$\begin{aligned} M + N &= \{\xi + \eta : \xi \in M, \eta \in N\} \\ M \cdot N &= \{\xi\eta : \xi \in M, \eta \in N\} \end{aligned}$$

Damit erhalten wir

Lemma 1.13
a) Es gilt $\dot{\xi} + \dot{\eta} = (\xi + \eta)^{\cdot}$
b) Es gilt $\dot{\xi}\dot{\eta} = (\xi\eta)^{\cdot}$

Beweis:
(I) a) Sei $\lambda \sim \xi$, $\mu \sim \eta$, wir zeigen $\lambda + \mu \sim \xi + \eta$. Damit wäre $\dot{\xi} + \dot{\eta} \subset (\xi + \eta)^{\cdot}$ gezeigt.
Die Mengen $[\lambda = \xi]$ und $\mu = \eta]$ sind 1–Mengen, also auch $[\lambda = \xi] \cap [\mu = \eta]$. Diese Menge liegt aber in $[\lambda + \mu = \xi + \eta]$, wie man sofort sieht. Also ist auch diese letztere Menge 1–Menge, was $\lambda + \mu \sim \xi + \eta$ bedeutet.
(II) Sei $\lambda \sim \xi + \eta$. Gesucht sind Folgen ξ', η' mit $\xi' \sim \xi$, $\eta' \sim \eta$ und $\lambda = \xi' + \eta'$. Nach Voraussetzung ist $[\lambda = \xi + \eta]$ eine 1–Menge. Sei $\lambda = (\ell_n)_n$, $\xi = (x_n)_n$, $\eta = (y_n)_n$. Setze $y'_n = \ell_n - x_n$. Für $n \in [\lambda = \xi + \eta]$ ist $y'_n = y_n$, also ist $\eta' = (y'_n) \sim \eta$ und $\lambda = \xi + \eta' \in \dot{\xi} + \dot{\eta}$. Damit ist $(\xi + \eta)^{\cdot} \subset \dot{\xi} + \dot{\eta}$ gezeigt, insgesamt ist a) bewiesen.
b) $\dot{\xi}\dot{\eta} \subset (\xi\eta)^{\cdot}$ wird genau so bewiesen. Sei wieder $\xi = (x_n)_n$, $\eta = (y_n)_n$ und $\lambda = (\ell_n)_n \in (\xi\eta)^{\cdot}$. Zeige $\lambda \in \dot{\xi}\dot{\lambda}\cdot$ ist $\lambda \sim \xi\eta$, also $\{n : \ell_n = x_n y_n\}$ eine 1–Menge. Für $\ell n \notin [\lambda = \xi\eta]$ ist $\ell_n = x_n y_n + a_n$. Setze hier einfach $x'_n = 1$, $y'_n = \ell_n$. Für $n \in [\lambda = \xi\eta]$ setze $x'_n = x_n, y_n = y_n$. Dann gilt $\xi' = (x'_n)_n \sim \xi$, $\eta' = (y'_n) \sim \eta$ und $\lambda = \xi'\eta' \in \dot{\xi}\dot{\eta}$. □

Wir kommen zum ersten Hauptergebnis:

Theorem 1.14 *$^*\mathbb{R}$ ist versehen mit der Addition*
$$(\dot{\xi}, \dot{\eta}) \mapsto \dot{\xi} + \dot{\eta}$$
und der Multiplikation
$$(\dot{\xi}, \dot{\eta}) \mapsto \dot{\xi}\dot{\eta}$$
ein Körper.

Beweis: a) Die Addition ist assoziativ, denn nach dem Lemma 1.13
$$\begin{aligned}(\dot{\xi} + \dot{\eta}) + \dot{\xi} &= (\xi + \eta)^{\circ} + \dot{\xi} = (\xi + \eta + \xi)^{\circ} \\ &= \dot{\xi} + (\eta + \xi)^{\cdot} = \dot{\xi} + (\dot{\eta} + \dot{\xi}).\end{aligned}$$
Genau so beweist man die Kommutativität, daß $\dot{\gamma}_0$ das Nullelement und $(-\xi)^{\cdot}$ das additive Inverse zu $\dot{\xi}$ ist.
b) Entsprechend ist die Multiplikation assoziativ, kommutativ und besitzt $\dot{\gamma}_1$ als Einselement. Auch die Distributivgesetze folgen sofort aus Lemma 1.13
c) Der einzig aufregende Punkt ist, daß zu jedem $\dot{\xi} \neq \dot{\gamma}_0$ ein multiplikatives Inverses existiert. Das geht so: $\dot{\xi} \neq \dot{\gamma}_0 \Rightarrow \xi \not\sim \gamma_0$. Das bedeutet für $\xi = (x_n)_n$, daß $\{n : x_n = 0\} = [\xi = \gamma_0]$ keine 1–Menge und daher eine 0–Menge ist nach Satz 1.5 a).
Wir setzen also $x'_n = \begin{cases} 1 & x_n = 0 \\ x_n & x_n \neq 0 \end{cases}$
Dann ist $\xi' = (x'_n) \sim \xi$, weil $[\xi' \neq \xi] = [\xi = \gamma_0]$ eine 0–Menge und damit $[\xi' = \xi]$ eine 1–Menge ist. Nun ist $x'_n \neq 0$ für alle n. Setze $y_n = \dfrac{1}{x'_n}$. Dann gilt für $\eta = (y_n)$ $\quad \eta\xi \sim \gamma_1$, also $\dot{\eta}\dot{\xi} = \dot{\gamma}_1$. □

Wir erhalten unmittelbar:

Korollar 1.15 *Für $x \in \mathbb{R}$ sei $\gamma_x = (x, x, x, \cdots)$ die konstante Folge. Die Abbildung $x \mapsto \dot{\gamma}_x$ ist eine Einbettung von $\mathbb{R}$ in $^*\mathbb{R}$.*

Die Ordnung auf ${}^*\mathbb{R}$

Satz 1.16 *Durch $\dot\xi \le \dot\eta \Leftrightarrow [\xi \le \eta]$ ist 1–Menge wird auf ${}^*\mathbb{R}$ eine totale Ordnung definiert mit den Eigenschaften:*
1) $\dot\xi \le \dot\eta \Leftrightarrow \dot 0 \le \dot\eta - \dot\xi$
2) $\dot\xi \le \dot\eta \Leftrightarrow \dot\xi + \dot\alpha \le \dot\eta + \dot\alpha$
3) $\dot\xi \le \dot\eta$ und $\dot 0 \le \dot\alpha \Rightarrow \dot\alpha\dot\xi \le \dot\alpha\dot\eta$
4) Auf $\mathbb{R} \subset {}^*\mathbb{R}$ wird die alte Ordnung induziert.

Kurz: ${}^*\mathbb{R}$ *ist ein hyperreeller Körper*

Beweis: (I) Wir müssen die Eigenschaften einer totalen Ordnung nachweisen.
(i) $\dot\xi \le \dot\xi$ ist klar.
(ii) $\dot\xi \le \dot\eta$ und $\dot\eta \le \dot\xi$ besagt, daß $[\xi \le \eta]$ und $[\eta \le \xi]$ 1–Mengen sind. Dann ist aber auch ihr Durchschnitt $[\xi \le \eta] \cap [\eta \le \xi] = [\xi = \eta]$ eine 1–Menge, also $\xi \sim \eta$ nach Definition, d.h. $\dot\xi = \dot\eta$
(iii) Sei $\dot\xi \le \dot\eta$ und $\dot\eta \le \dot\tau$. Dann sind $[\xi \le \eta]$ und $[\eta \le \tau]$ 1–Mengen, also ist auch $[\xi \le \eta] \cap [\eta \le \tau]$ eine 1–Menge. Auf ihr gilt aber $\xi \le \tau$, also ist nach Definition $\dot\xi \le \dot\tau$
(iv) Es gelte *nicht* $\dot\xi \le \dot\eta$. Dann ist $[\xi \le \eta]$ keine 1–Menge, daher 0–Menge und damit ist ihr Komplement $[\xi > \eta] = [\eta < \xi]$ eine 1–Menge nach Satz 1.5 a). Damit ist $\dot\eta \le \dot\xi$ und $\dot\eta \ne \dot\xi$, also $\dot\eta < \dot\xi$ und damit ist auch die Totalität der Ordnung bewiesen.

(II) 1) $\dot\xi \le \dot\eta \Rightarrow [\xi \le \eta]$ ist 1–Menge, aber $[\xi \le \eta] = [0 \le \eta - \xi]$, also $\dot 0 \le (\eta - \xi)^{\cdot} = \dot\eta - \dot\xi$. Genau so beweist man 2) - 4). □

Nun kommen wir zu den unendlich kleinen und großen Zahlen. Wir betonen noch einmal: $\mathbb{R}$ und damit $\mathbb{Q}, \mathbb{Z}$ und $\mathbb{N}$ sind Teilmengen von ${}^*\mathbb{R}$. In Einklang mit unserer früheren Definition 1.9 erklären wir:

Deffnition 1.17 *Im Folgenden sei $\dot\xi = (x_n)^{\cdot}_n$ usw.*
a) $\dot\xi$ heißt **hypernatürlich**, *wenn $\{n : x_n \in \mathbb{N}\}$ eine 1–Menge ist. Ist ξ konstant und hypernatürlich, also $x_n = k \in \mathbb{N}$ für alle n, so wird $\dot\xi$ wieder mit k bezeichnet und heißt* **natürliche Standard-Zahl**.
b) $\dot\xi$ heißt **hyperganz**, *wenn $\{n : x_n \in \mathbb{Z}\}$ eine 1–Menge ist. Die gewöhnlichen ganzen Zahlen (repräsentiert durch konstante Folgen) heißen* **ganze Standard-Zahl**.
c) $\dot\xi$ heißt **hyperrational**, *wenn $\{n : x_n \in \mathbb{Q}\}$ eine 1–Menge ist. Die gewöhnlichen rationalen Zahlen (repräsentiert durch konstante Folgen mit Werten in $\mathbb{Q}$) heißen* **rationale Standard-Zahl**.

Bezeichnungen:

$$\left\{\begin{array}{c} {}^*\mathbb{N} \\ {}^*\mathbb{Z} \\ {}^*\mathbb{Q} \\ {}^*\mathbb{R} \end{array}\right\} \text{ ist die Menge der } \left\{\begin{array}{r} \text{hypernatürlichen} \\ \text{hyperganzen} \\ \text{hyperrationalen} \\ \text{hyperreellen} \end{array}\right\} \text{ Zahlen.}$$

Beispiele:

1. Sei $x_n = n$, also $\xi = (1, 2, 3, \cdots)$. $\dot{\xi} \in^* \mathbb{N}$. Es ist $\dot{\xi} > k (= \dot{\gamma}_k)$ für jede natürliche Standardzahl, weil$[\xi \geq k]$ eine 1Menge ist.
2. Sei $y_n = n^2$, also $\dot{\eta} = \dot{\xi}^2$. Dannist $\dot{\eta} > \dot{\xi}$. Sei $z_n = 2^n$. Dann ist $\dot{\zeta} > \dot{\xi}^k$ für jede feste natürliche Standardzahl k. Wir schreiben $\dot{\zeta} = 2^{\dot{\xi}}$
3. $\frac{1}{\dot{\xi}} \in^* \mathbb{Q}$. Ist $r > 0$ eine beliebige reelle Zahl, so ist $\dot{0} < \frac{1}{\dot{\xi}} < \dot{\gamma}_r (= r)$. $\frac{1}{\dot{\xi}}$ ist also eine unendlich kleine rationale Zahl.
4. $-\dot{\xi}$, $-\dot{\xi}^2$, $-2^{\dot{\xi}}$ sind hyperganze Zahlen, die negativ unendlich groß sind.
5. Betrachten Sie die monoton fallende Folge $a_0 = 3$, $a_{n+1} = \frac{1}{2}(a_n + \frac{2}{a_n})$. $\dot{\alpha} = (a_n)^{\cdot}_n$ ist unendlich dicht bei $\sqrt{2}$, genauer $\dot{\alpha} - \sqrt{2}$ ist "unendlich klein".

Wie wir weiter oben (s. S. 11) gesehen haben, kann man in ${}^*\mathbb{R}$ als geordnetem Körper den Absolutbetrag definieren. Es ergibt sich unmittelbar

$$|\dot{\xi}| = (|x_n|)^{\cdot}_n.$$

Im Einklang mit Definition 1.9 erklären wir:

Deffnition 1.18 *a) Eine Zahl $\dot{\xi}$ heißt* **positiv unendlich groß**, *wenn $\dot{\xi}$ größer als jede Standardzahl ist, d.h. wenn $\dot{\xi} > r$ für jedes $r \in \mathbb{R}$ gilt.*
b) $\dot{\xi}$ heißt **negativ unendlich groß**, *wenn $-\dot{\xi}$ positiv unendlich groß ist.*
c) $\dot{\xi}$ heißt **(betragsmäßig) unendlich klein** *oder* **inffnitesimal**, *wenn für jede Standardzahl $r > 0$ stets $|\dot{\xi}| < r$ gilt.*
d) $\dot{\xi}$ und $\dot{\eta}$ heißen **inffnitesimal benachbart** *oder* **fast gleich**, *in Zeichen $\dot{\xi} \approx \dot{\eta}$, wenn $\dot{\xi} - \dot{\eta}$ infinitesimal ist.*

Aufgabe: Zeigen Sie bitte, daß $\approx$ eine Äquivalenzrelation ist.

Bezeichnungen: Mit F in $({}^*\mathbb{R})$ bezeichnen wir die Menge aller $\dot{\xi} \in {}^*\mathbb{R}$, für die es eine Standardzahl $r = r(\dot{\xi})$ mit $|\dot{\xi}| \leq r$ gibt. $\dot{\xi}$ aus F in $({}^*\mathbb{R})$ heißt **endlich**. Mit $\mu(\dot{\xi})$ bezeichnen wir die Menge aller $\dot{\eta} \approx \dot{\xi}$. $\mu(\dot{\xi})$ heißt auch **Monade** von $\dot{\xi}$.

Vor dem Hauptsatz wollen wir noch einige Rechenregeln für das Rechnen mit Ungleichungen zusammenstellen.

Satz 1.19 (Rechenregeln für Ungleichungen) *a)* $|\dot{\xi} + \dot{\eta}| \leq |\dot{\xi}| + |\dot{\eta}|$, $|\dot{\xi}\dot{\eta}| = |\dot{\xi}|\,|\dot{\eta}|$
b) Sind $\dot{\xi}, \dot{\eta} \approx 0$, so ist auch $\dot{\xi} + \dot{\eta} \approx 0$.
c) Sind $\dot{\xi}, \dot{\eta}$ endlich, so auch $\dot{\xi} + \dot{\eta}$ und $\dot{\xi} \cdot \dot{\eta}$.

d) Ist $0 \approx \dot{\xi}$ und $\dot{\eta}$ endlich, so ist $\dot{\xi}\dot{\eta} \approx 0$.
e) Ist $0 \neq \dot{\xi} \approx 0$, so ist $\frac{1}{|\dot{\xi}|}$ unendlich groß.
f) Ist $\dot{\xi}$ betragsmäßig unendlich groß, so ist $\frac{1}{\dot{\xi}} \approx 0$.

Beweis: a) gilt in jedem geordneten Körper (Beweis wie für $\mathbb{R}$).
b) Sei $r > 0$ eine beliebige Standardzahl. Dann ist $|\dot{\xi}|, |\dot{\eta}| \leq \frac{r}{2}$, also $|\dot{\xi} + \dot{\eta}| \leq |\dot{\xi}| + |\dot{\eta}| < r$. Da r beliebig war, folgt $\dot{\xi} + \dot{\eta} \approx 0$.
c) Sind $\dot{\xi}, \dot{\eta}$ endlich, so gibt es Standardzahlen r, s mit $|\dot{\xi}| \leq r$, $\dot{\eta} \leq s$, also $|\dot{\xi} + \dot{\eta}| \leq |\dot{\xi}| + |\dot{\eta}| \leq r + s \in \mathbb{R}$ und damit ist $\dot{\xi} + \dot{\eta}$ endlich. Analog: $|\dot{\xi}\dot{\eta}| = |\dot{\xi}|\,|\dot{\eta}| \leq rs \in \mathbb{R}$, also ist $\dot{\xi}\dot{\eta}$ endlich.
d) Da $\dot{\eta}$ endlich ist, gibt es eine Standardzahl r mit $|\dot{\eta}| < r$. Sei $\varepsilon > 0$ eine beliebige Standardzahl. Dann ist $|\dot{\xi}| < \frac{\varepsilon}{r+1}$, also $|\dot{\xi}\dot{\eta}| < \frac{r\varepsilon}{r+1} < \varepsilon$. Da $\varepsilon > 0$ eine beliebige Standardzahl war ist $\dot{\xi}\dot{\eta} \approx 0$.
e) Sei $0 \neq \dot{\xi}$. Sei $0 < r$ eine beliebige Standardzahl. Dann ist $|\dot{\xi}| \leq \frac{1}{r}$, also $\frac{1}{|\dot{\xi}|} > r$. Da r beliebig war, folgt die Behauptung.
f) geht analog. □

Aufgaben:

1. Sei $\dot{\xi} = (1, 2, 3, \cdots)$. Zeigen Sie bitte:
   a) $\frac{1}{\dot{\xi}} \approx 0$, b) $2^{-\dot{\xi}} \approx 0$, c) $\dot{\xi}2^{-\dot{\xi}} \approx 0$. Dabei ist $2^{-\dot{\xi}} = (2^{-1}, 2^{-2}, 2^{-3}, \cdots)^{\cdot}$

2. Sei $0 < q < 1$ eine Standardzahl und $\dot{\xi}$ wie oben. Wir definieren $\sum_{k=0}^{\dot{\xi}} q^k = (1 + q, 1 + q + q^2, 1 + q + q^2 + q^3, \cdots)^{\cdot}$.
   Zeigen Sie bitte: $\sum_{k=0}^{\dot{\xi}} q^k \approx \frac{1}{1-q}$

3. Sei $0 \neq \dot{\xi}$ unendlich klein und $\dot{\eta}$ eine endliche Zahl. Zeigen Sie bitte $\frac{(\dot{\eta} + \dot{\xi})^2 - \dot{\eta}^2}{\dot{\xi}} \approx 2\dot{\eta}$. Das bedeutet in der "Standard-Analysis": Die Funktion $f(x) = x^2$ ist auf jedem beschränkten Intervall gleichmäßig differenzierbar.

4. Verallgemeinern Sie die vorige Aufgabe auf die Funktion $f(x) = x^n$, (n eine Standard-natürliche Zahl).

Der Hauptsatz lautet

Theorem 1.20 *a) Zu jeder endlichen Zahl $\dot{\xi}$ gibt es genau eine Standardzahl s mit $\dot{\xi} \approx s$. s heißt der* **Standardteil** *$s = {}^{\circ}\dot{\xi}$ von $\dot{\xi}$.*
b) Die Abbildung $\dot{\xi} \mapsto {}^{\circ}\dot{\xi}$ ist additiv und multiplikativ von F (in ${}^{*}\mathbb{R}$) auf $\mathbb{R}$.

Beweis: a) Wir benutzen das Resultat aus Analysis I, daß jede nach oben beschränkte nicht leere Menge $A \subset \mathbb{R}$ eine (eindeutig bestimmte) obere Grenze (Suprenum $\sup(A)$) besitzt.
(I) Existenz von s: Sei $\dot{\xi}$ endlich. Dann gibt es ein $r \in \mathbb{R}$ mit $|\dot{\xi}| < r$. Die Menge $A = \{t \in \mathbb{R} : t < \dot{\xi}\}$ ist also wegen $-r \in A$ nicht leer und wegen $|\dot{\xi}| < r$ nach oben beschränkt. Sei $s = \sup(A)$.
Behauptung: $s \approx \dot{\xi}$. *Beweis:* Angenommen nicht, dann gibt es eine Standardzahl $u > 0$ mit $0 < u < |s - \dot{\xi}|$ also $s - \dot{\xi} < -u$ oder $u < s - \dot{\xi}$. Im ersten Fall ist $s + u < \dot{\xi}$, also $s < s + u \in A$ im Widerspruch zu $s = \sup(A)$, im zweiten Fall ist $\dot{\xi} < s - u$, also $s - u$ obere Schranke von A, ebenfalls im Widerspruch zu $s = \sup(A)$.
(II) Eindeutigkeit von s. Seien s und t zwei Standardzahlen mit $s \approx \dot{\xi} \approx t$. Dann ist $|s - t| \leq |s - \dot{\xi}| + |\dot{\xi} - t| \approx 0$ nach Satz 1.19. Das bedeutet aber: Für jedes Standard $n \in \mathbb{N}$ ist $|s - t| < \dfrac{1}{n}$.
Daraus folgt (in $\mathbb{R}$, nicht in $^*\mathbb{R}$) $\quad s = t$.
b) Es ist $s = {}^\circ\dot{\xi} \approx \dot{\xi}$, $t = {}^\circ\dot{\eta} \approx \dot{\eta}$, also

$$\begin{aligned} |s + t - (\dot{\xi} + \dot{\eta})| &= |(s - \dot{\xi}) + (t - \dot{\eta})| \\ &\leq |s - \dot{\xi}| + |t - \dot{\eta}| \approx 0. \end{aligned}$$

Also ist

$${}^\circ\dot{\xi} + {}^\circ\dot{\eta} = (s + t) \;=\; {}^\circ(\dot{\xi} + \dot{\eta}).$$

Analog beweist man ${}^\circ\dot{\xi}{}^\circ\dot{\eta} = {}^\circ(\dot{\xi}\dot{\eta})$. Genauer (mit den Bezeichnungen von oben)

$$\begin{aligned} |st - \dot{\xi}\dot{\eta}| &= |s(t - \dot{\eta}) + (s - \dot{\xi})\dot{\eta}| \\ &\leq |s|\,|t - \dot{\eta}| + |s - \dot{\xi}|\,|\dot{\eta}|. \end{aligned}$$

Aus dem Satz 1.19 d) folgt $st \approx \dot{\xi}\dot{\eta}$. Da st eine Standardzahl ist, folgt $st = {}^\circ(\dot{\xi}\dot{\eta})$.
Daß die Abbildung surjektiv auf ganz $\mathbb{R}$ ist, folgt aus $s = \dot{\gamma}_s$. $\square$

1.4 Funktionen und Relationen und ihre Erweiterung

Überblick: Funktionen und Relationen lassen sich als Teilmengen von $\mathbb{R}^p$ für $p = 1, 2, \cdots$ auffassen. Wir konstruieren in Analogie zu $^*\mathbb{R}$ bzw. $^*\mathbb{Q}$, usw. Erweiterungen in $^*\mathbb{R}^p$ von Mengen $D \subset \mathbb{R}^p$ und untersuchen speziell Erweiterungen von Abbildungen.

Das cartesische Produkt $^*\mathbb{R}^p$

Überblick: Wir können $^*\mathbb{R}^p$ auf 2 Weisen konstruieren, einmal in Analogie zu $^*\mathbb{R}$ als Quotientenstruktur $(\mathbb{R}^p)^{\mathbb{N}}/\sim_p$, einmals als gewöhnliches cartesisches Produkt von $^*\mathbb{R}$ mit sich. Es ergibt sich beide Male das gleiche.

Der Raum $(\mathbb{R}^p)^{\mathbb{N}}$ aller vektorwertigen Folgen $\vec{\xi} = (\vec{x}_n)_n$ (mit $\vec{x}_n = (x_{1,n}, \cdots, x_{p,n}) \in \mathbb{R}^p$) ist kanonisch isomorph zu $(\mathbb{R}^{\mathbb{N}})^p$. Der Isomorphismus ist gegeben durch $(\vec{x}_n)_n \to ((x_{1,n})_n, \quad (x_{2,n})_n, \cdots, \quad (x_{p,n})_n)$. $(\mathbb{R}^p)^{\mathbb{N}}$ ist ein Vektorraum über $\mathbb{R}$, der sogar Multiplikation mit skalaren Folgen zuläßt:

$$(\alpha_n)_n \cdot (\vec{x}_n)_n = (\alpha_n \vec{x}_n)_n.$$

Wir setzen

$$\vec{\xi} \sim_p \vec{\eta} \Leftrightarrow \{n : \vec{x}_n = \vec{y}_n\} \quad \text{ist eine 1-Menge.}$$

Es gilt

Lemma 1.21 *a) $\sim_p$ ist eine mit den Vektorraumoperationen verträgliche Äquivalenzrelation.*
b) $\vec{\xi} \sim_p \vec{\eta}$ gilt genau dann, wenn für die Koordinatenfolgen $\xi_j \sim \eta_j$ für $j = 1, \cdots, p$ gilt.
c) $\vec{\xi} \sim_p \vec{\eta}$ und $\alpha \sim \beta$ impliziert $\alpha\vec{\xi} \sim_p \beta\vec{\eta}$.

Beweis: a) Es genügt, daß $W = \{\vec{\xi} : \vec{\xi} \sim_p \vec{0}\}$ ein Untervektorraum ist, wo $\vec{0} = (\vec{0}, \vec{0}, \vec{0}, \cdots)$ ist. Aber das ist klar.
b) Es ist $\{n : \vec{x}_n = \vec{y}_n\} = \cap_{j=1}^{p} \{n : x_{j,n} = y_{j,n}\}$ Also ist die linke Seite genau dann eine 1-Menge, wenn jede der p–Mengen $[x_{j,n} = y_{j,n}]$ eine solche ist, und das ist die Behauptung
c) Es ist $[\alpha_n \vec{x}_n = \beta_n \vec{y}_n] \supset [\alpha_n = \beta_n] \cap [\vec{x}_n = \vec{y}_n]$. Nach Voraussetzung ist die rechte, also auch die linke Seite eine 1-Menge. $\square$

Satz 1.22 *Sei $W = \{\vec{\xi} : \vec{\xi} \sim_p \vec{0}\}$. Dann ist der Quotientraum $(\mathbb{R}^p)^{\mathbb{N}}/W$ ein $^*\mathbb{R}$–Vektorraum, der kanonisch isomorph zu $(^*\mathbb{R})^p$ ist.*

Der Satz besagt, daß man ${}^*\mathbb{R}^p$ auf zwei verschiedene Weisen konstruieren kann; wir machen von beiden Verfahren Gebrauch und identifizieren die beiden Räume. Offensichtlich ist $\mathbb{R}^p \subset {}^*\mathbb{R}^p$.

Beweis: Es ist $\vec{\xi}+W = \{\vec{\eta} : \vec{\eta} \sim \vec{\xi}\} =: \dot{\vec{\xi}}$. Nach dem vorangegangenen Lemma ist die Zuordnung $\dot{\vec{\xi}} \to (\dot{\xi}_1, \cdots, \dot{\xi}_p)$ wohldefiniert und offensichtlich ein Vektorraum-Isomorphismus über $\mathbb{R}$. Teil c) des Lemmas besagt, daß $(\mathbb{R}^p)^{\mathbb{N}}/W$ ein ${}^*\mathbb{R}$–Vektorraum ist und

$$\begin{aligned} \dot{\alpha}\dot{\vec{\xi}} &= (\alpha\vec{\xi})^{\cdot} = ((\alpha\xi_1)^{\cdot}, \cdots, (\alpha\xi_p)^{\cdot}) \\ &= (\dot{\alpha}\dot{\xi}_1, \cdots, \dot{\alpha}\dot{\xi}_p) \qquad \text{gilt.} \end{aligned}$$

Damit ist die Zuordnung sogar ${}^*\mathbb{R}$–linear. □

Erweiterungen (Enlargements) von Teilmengen

Überblick: In Analogie zu ${}^*\mathbb{N}, {}^*\mathbb{Z}$ bzw. ${}^*\mathbb{Q}$ als Erweiterungen von $\mathbb{N}, \mathbb{Z}$, bzw. $\mathbb{Q}$ erweitern wir jede Teilmenge D von $\mathbb{R}^p$ und untersuchen die Eigenschaften dieses Erweiterungsprozesses.

Deffnition 1.23 *Sei $D \subset \mathbb{R}^p$ beliebig.* ${}^*D = \{\dot{\vec{\xi}} : \vec{\xi} \in D^{\mathbb{N}}\}$ *heißt* **Enlargement** *oder* **Erweiterung** *von D.*

Beispiele:

1. ${}^*\mathbb{N}, {}^*\mathbb{Z}$ und ${}^*\mathbb{Q}$ sind solche Erweiterungen
2. ${}^*]a,b[= \{\dot{\xi} : a < \dot{\xi} < b\}$
3. ${}^*[a,b] = \{\dot{\xi} : a \leq \dot{\xi} \leq b\}$
4. Für $a \in \mathbb{R}$ ist $\mu(a) = \cap_{0<\varepsilon\in\mathbb{R}} {}^*]a-\varepsilon, a+\varepsilon[= \cap_{n\in\mathbb{N}} {}^*]a-\frac{1}{n}, a+\frac{1}{n}[$
5. $D = \{(x, \sin(x)) : x \in \mathbb{R}\}$. ${}^*D$ ist wieder Graph einer Funktion, der Fortsetzung des Sinus, s.u.

Bevor wir die Eigenschaften von Erweiterungen diskutieren, erinnern wir an die Verknüpfung zweistelliger Relationen:
Seien M, N, P Mengen, $D_1 \subset M \times N, D_2 \subset N \times P$. Dann ist $D_2 \circ D_1 = \{(x,z) \in M \times P : \exists y \in N[(x,y) \in D_1$ und $(y,z) \in D_2]$ die Verknüpfung von D_1 mit D_2. Für Graphen von Funktionen erhalten wir:
Sind $D_1 = \{(x, f_1(x)) : x \in A \subset M\}$, $D_2 = \{(y, f_2(y)) : y \in B \supset f_1(A)\}$, so ist $D_2 \circ D_1 = \{(x, f_2 \circ f_1(x)) : x \in A\}$.

Ist $M = N = \mathbb{R}$ und $D = \{(x,y) : x \le y\}$, so ist nach dem Transitivitätspostulat $D \circ D \subset D$.
Für eine beliebige Relation $D \subset M \times N$ ist $D^{-1} = \{(y,x) \in N \times M : (x,y) \in D\}$. Angewendet auf $D = \{(x, f(x)) : x \in A \subset M\}$ ergibt D^{-1} die Umkehrfunktion, falls f injektiv ist. Ist D die obige Ordnungsrelation und $\Delta = \{(x,x) : x \in \mathbb{R}\}$, so besagt das Antisymmetriepostulat gerade $D \cap D^{-1} \subset \Delta$ und die Totalität der Ordnung gerade $D \cup D^{-1} = \mathbb{R}^2$.

Wir erhalten den folgenden grundlegenden Satz

Satz 1.24 *Für die Erweiterung von Mengen gilt:*
a) $D \subset {}^*D$ *und* $D = {}^*D \Leftrightarrow D$ *ist endlich.*
b) ${}^*(D \cap E) = {}^*D \cap {}^*E$.
c) ${}^*(\mathbb{R}^p \setminus D) = {}^*\mathbb{R}^p \setminus {}^*D$.
d) ${}^*(D \cup E) = {}^*D \cup {}^*E$.
e) $D \subset E \Rightarrow {}^*D \subset {}^*E$.
f) ${}^*(D \times E) = {}^*D \times {}^*E$.
g) ${}^*(E \circ D) = {}^*E \circ {}^*D$.
h) ${}^*(D^{-1}) = ({}^*D)^{-1}$.
i) Ist $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$ *eine Funktion mit Graphen* G, *so ist* ${}^*G$ *Graph einer Funktion* ${}^*f : {}^*D \to {}^*\mathbb{R}^q$ *und es gilt* ${}^*f({}^*D) = {}^*(f(D))$.
k) Sei $1 \le i < j \le p$ *und* $\Delta = \{\vec{x} = (x_1, \dots, x_p) \in \mathbb{R}^p : x_i = x_j\}$. *Dann ist* ${}^*\Delta = \{\vec{\xi} \in {}^*\mathbb{R}^p : \xi_i = \xi_j\}$.

Beweis: *Vorbemerkung:* $\dot{\xi} \in {}^*D \Leftrightarrow \{n : x_n \in D\}$ ist 1-Menge.
Beweis: $\dot{\xi} \in {}^*D \Rightarrow \exists \eta \sim \xi$ und $\eta \in D^{\mathbb{N}} \Rightarrow \quad \{n : x_n \in D\} \supset [x_n = y_n]$. Letztere ist 1-Menge, also folgt die Behauptung.
Umgekehrt: $\{n : x_n \in D\}$ sei 1-Menge. Sei $d = x_{n_0} \in D$ und $x'_n = \begin{cases} x_n & : x_n \in D \\ d & : \text{sonst} \end{cases} \Rightarrow \xi' \sim \xi$ und $\xi' \in D^{\mathbb{N}} \Rightarrow \dot{\xi} = \dot{\xi}' \in {}^*D$.
a) $x \in D \Rightarrow \gamma_x \in D^{\mathbb{N}} \Rightarrow \quad x(= \dot{\gamma}_x) \in {}^*D$.
Sei $D = \{d_1, \cdots, d_q\}$ und $\dot{\xi} \in {}^*D$. Sei o. B. d. A. (s. Vorbem.) $\xi \in D^{\mathbb{N}}$ und $A_j = \{n : x_n = d_j\}$. Dann ist $A_j \cap A_k = \phi$ für $j \ne k$ und $\cup_1^q A_j = \mathbb{N}$, also existiert genau ein j mit $m(A_j) = 1, \quad m(A_k) = 0$ für $k \ne j$. Also ist $\dot{\xi} = d_j$.
Sei nun D unendlich. Dann gibt es eine Folge $(x_n) \subset D$ mit $x_m \ne x_n$ für $m \ne n \Rightarrow (x_n)^{\cdot}_n \in {}^*D$, aber $(x_n)^{\cdot} \ne y$ für alle $y \in D$, also $D \ne {}^*D$.
b) $\dot{\xi} \in {}^*(D \cap E) \Rightarrow [x_n \in D] \cap [x_n \in E] = [x_n \in D \cap E]$ und letztere ist 1-Menge, also ist $\dot{\xi} \in {}^*D \cap {}^*E$.
Ist umgekehrt $\dot{\xi} \in {}^*D \cap {}^*E$, so ist $[x_n \in D] \cap [x_n \in E] = [x_n \in D \cap E]$ 1-Menge, also $\dot{\xi} \in {}^*(D \cap E)$
c) $\dot{\vec{\xi}} \in {}^*(\mathbb{R}^p \setminus D) \Leftrightarrow \{n : \vec{x}_n \in \mathbb{R}^p \setminus D\}$ ist 1–Menge $\Leftrightarrow \{n : \vec{x}_n \in D\}$ ist 0–Menge $\Leftrightarrow \dot{\vec{\xi}} \notin {}^*(D) \Leftrightarrow \dot{\vec{\xi}} \in {}^*\mathbb{R}^p \setminus {}^*D$.

d) sei $D \subset \mathbb{R}^p$, $E \subset \mathbb{R}^p$. Dann ist $D \cup E = \mathbb{R}^p \setminus ((\mathbb{R}^p \setminus D) \cap (\mathbb{R}^p \setminus E))$ und die Behauptung folgt leicht aus b) und c).
e): klar.
f) Sei $\dot{\xi} \in {}^*(D \times E)$, dann ist $\{n : x_n \in D \times E\}$ 1-Menge. Es gilt $x_n \in D \times E \Leftrightarrow x_n = (y_n, z_n)$ mit $y_n \in D, z_n \in E$. Wähle (y_0, z_0) in $D \times E$ beliebig.
Setze $y'_n = \begin{cases} y_n & : (y_n, z_n) \in D \\ y_0 & : \text{sonst} \end{cases}$ $z'_n = \begin{cases} z_n & : (y_n, z_n) \in D \\ z_0 & : \text{sonst} \end{cases}$ Dann ist für $x'_n = (y'_n, z'_n)$ $[x'_n \in D \times E]$ eine 1-Menge und $\dot{\xi} = \dot{\xi}' = (\dot{\eta}, \dot{\varphi}) \in {}^*D \times {}^*E$.
Sei umgekehrt $\dot{\xi} \in {}^*D \times {}^*E$. Dann ist $\dot{\xi} = (\dot{\eta}, \dot{\tau})$ mit $\dot{\eta} \in {}^*D, \dot{\tau} \in {}^*E$, also ist $[(y_n, t_n) \in D \times E] \supset [y_n \in D] \cap [t_n \in E]$ 1-Menge also $\dot{\xi} \in {}^*(D \times E)$.
g) Sei $\dot{\tau} \in {}^*(E \circ D)$ und oBdA $\tau \in (E \circ D)^{\mathbb{N}}$. Nach Definition ist $\tau = ((x_n, z_n))_n$ wobei zu jedem $n \in \mathbb{N}$ ein y_n existiert mit $(x_n, y_n) \in D$ und $(y_n, z_n) \in E$. Also gilt: $(\dot{\xi}, \dot{\eta}) \in {}^*D$ und $(\dot{\eta}, \dot{\varphi}) \in {}^*E$ und damit $\dot{\tau} \in {}^*E \circ {}^*D$. Sei umgekehrt $\dot{\tau} \in {}^*E \circ {}^*D$. $\dot{\tau} = (\dot{\xi}, \dot{\varphi})$. Dann gibt es $\dot{\eta}$ mit $(\dot{\xi}, \dot{\eta}) \in {}^*D$ und $(\dot{\eta}, \dot{\varphi}) \in {}^*E$, also ist $[(x_n, y_n) \in D] \cap [(y_n, z_n) \in E] \subset [(x_n, z_n) \in E \circ D]$ eine 1-Menge. Damit ist $\dot{\tau} \in {}^*(E \circ D)$.
h) ist klar.
i) Zunächst ist ${}^*G \subset {}^*\mathbb{R}^p \times {}^*\mathbb{R}^q$.
Behauptung: ${}^*G$ ist Graph einer Funktion. *Beweis:* Sind $(\dot{\xi}, \dot{\eta})$ und $(\dot{\xi}, \dot{\tau})$ aus ${}^*G$, so ist $\{n : (x_n, y_n) \in G\} \cap \{n : (x_n, t_n) \in G\} \subset A := \{n : y_n = t_n\}$, weil G Graph von der Funktion f ist. Also ist A nach Konstruktion von ${}^*G$ eine 1-Menge und damit $\dot{\eta} = \dot{\tau}$, ${}^*G$ also tatsächlich Graph einer Funktion.
Sei nun für $\dot{\xi} \in {}^*D$ einfach ${}^*f(\dot{\xi}) := (f(x_n)_n)^\bullet$, wo $(x_n)_n \in D$ o. B. d. A. angenommen werden kann. Wegen $(x_n, f(x_n))_n \in G^{\mathbb{N}}$ ist $(\dot{\xi}, {}^*f(\dot{\xi})) \in {}^*G$. Ist umgekehrt $(\dot{\xi}, \dot{\eta}) \in {}^*G$, so ist $\{n : (x_n, y_n) \in G\}$ eine 1-Menge und damit sind $[x_n \in D]$ und $[y_n = f(x_n)]$ 1-Mengen, also $\dot{\eta} = {}^*f(\dot{\xi})$.
k) ist wiederum offensichtlich. □

Wir können nun wichtige Beispiele bringen. Vorher treffen wir einige erleichternde Bezeichnungsänderungen. Statt $\dot{\xi}$ schreiben wir in Zukunft $\xi (= (x_n)_n^\bullet)$. Bekannte Funktionen wie $f(x) = x^k$, $f(x) = \exp(x)$ schreiben wir ohne den Stern davor. Aus dem Zusammenhang geht hervor, was wir meinen.

Beispiele:

1. Polynome: $f(x) = \sum_0^n a_k x^k$, $f(\xi) = \sum_0^n a_k \xi^k$.
2. $\exp(\xi) = (\exp(x_n))_n^\bullet$.
3. Es ist $\sin(\xi + 2\pi) = \sin(\xi)$ wegen $\sin(x_n + 2\pi))_n = (\sin(x_n))_n$.
4. Die Addition auf $\mathbb{R}^2$ hat ihre eindeutige Fortsetzung auf ${}^*\mathbb{R}^2$. Die komplexe Multiplikation lautet ja - in Ausdrücken von $\mathbb{R}^2$ geschrieben - $\mathbb{R}^2 \times \mathbb{R}^2 \ni ((u, v), (w, x)) \longrightarrow (uw - vx, ux + vw)$. Die Fortsetzung gemäß Satz 1.24 hat offenbar dieselben Eigenschaften. Wir erhalten somit ${}^*\mathbb{C}$, wo man genau so rechnet wie in $\mathbb{C}$.

5. In ${}^*\mathbb{C}$ gilt die Eulersche Formel
   $\exp(i\tau) = \cos(\tau) + i\sin(\tau)$
   und $\exp(i(\tau+\sigma)) = \exp(i\tau)\exp(i\sigma)$
   wie man sofort über die repräsentierenden Folgen ausrechnet. Damit gelten alle trigonometrischen Formeln auch2 in ${}^*\mathbb{R}$ bzw. ${}^*\mathbb{C}$.

6. Wir haben auf $\mathbb{R}^p$ die verschiedensten Normen. Besonders wichtig sind $\|x\|_2 = (\sum_1^p |x_j|^2)^{1/2}$
   $\|x\|_1 = \sum_1^p |x_j|$ und $\|x\|_\infty = \max_{1\le j\le p}(|x_j|)$.
   Für sie gelten die Abschätzungen

   $$\|x\|_\infty \le \|x\|_1 \le \sqrt{p}\|x\|_2 \le p\sqrt{p}\|x_1\| \le p^2\sqrt{p}\|x\|_\infty,$$

   eine Beziehung, die sich sofort auf ${}^*\mathbb{R}^p$ überträgt.

7. $\Delta = \{(x,x) : x \in \mathbb{R}^p\}$. Es ist ${}^*\Delta = \{(\xi,\xi) : \xi \in {}^*\mathbb{R}^p\}$.

8. Sei $a \in \mathbb{R}$ fest gewählt und $B(a) = \{(x_1,\ldots,x_p) \in \mathbb{R}^p : x_j = a\}$. Dann ist ${}^*B(a) = \{(\xi_1,\ldots,\xi_p) \in {}^*\mathbb{R}^p : \xi_j = a\}$.

9. Sei $\sigma : \{1,\ldots,q\} \to \{1,\ldots,p\}$. Sei $A \subset \mathbb{R}^p$ und

   $$B = \{(x_1,\ldots,x_q) \in \mathbb{R}^q : (x_{\sigma(1)},\ldots,x_{\sigma(q)}) \in A\}.$$

   Dann ist

   $${}^*B = \{(\xi_1,\ldots,\xi_q) \in {}^*\mathbb{R}^q : (\xi_{\sigma(1)},\ldots,\xi_{\sigma(q)}) \in {}^*A\}.$$

1.5 Mathematische Sätze und das elementare Transferprinzip

Überblick: Wir entwickeln eine Formelsprache, in der nur Elemente und Mengen vorkommen. Viele Sachverhalte der Analysis lassen sich damit ausdrücken. Wir zeigen, daß die Sätze, die man damit formulieren kann, in $\mathbb{R}$ genau dann gelten, wenn sie in ${}^*\mathbb{R}$ gelten. Das ist das Transferprinzip.

Wir haben uns in Satz 1.24 überzeugt, daß eine Reihe elementarer Beziehungen zwischen Mengen genau dann in $\mathbb{R}$ (bzw. $\mathbb{R}^p$) gelten, wenn die entsprechenden "gesternten" Beziehungen in ${}^*\mathbb{R}$ (bzw. ${}^*\mathbb{R}^p$) gelten.
Um uns nicht immer mit Einzelfällen herumzuschlagen, entwickeln wir ein allgemeines Schema zur Formulierung von Sätzen der Analysis und beweisen dann, daß ein solcher Satz in $\cup_{p=1}^\infty \mathbb{R}^p$ genau dann gilt, wenn seine "gesternte" Version in $\cup_{p=1}^\infty {}^*\mathbb{R}^p$ gilt.
Dazu setzen wir $\cup_{p=1}^\infty \mathbb{R}^p = \mathcal{S}$ und $\cup_{p=1}^\infty {}^*\mathbb{R}^p = {}^*\mathcal{S}$.

Wir wählen ein möglichst einfaches Formelschema, von dem wir ganz schnell beweisen können, was wir oben gesagt haben. Wir zeigen dann, daß wir alle uns zunächst interessierenden Sätze der Analysis in dieser Formelsprache ausdrücken können - wenn auch oft nur sehr kompliziert.

1.5.1 Entwicklung der Formeln

I) **Zeichensatz**
Wir benutzen die Zeichen $\neg$ (nicht), $=$ (gleich) $\wedge$ (und), ferner $\exists$ (es existiert), sowie Klammern (,), [,], , , $\in$, ferner Variable $(x, y, z, u, v, w, x_1, x_2, \cdots)$ und schließlich die reellen Zahlen selbst, sowie Elemente in $\mathbb{R}^p$ (für jedes p) und Teilmengen von $\mathbb{R}^p$ (für jedes p).
Konkrete reelle Zahlen und Elemente aus $\mathbb{R}^p$ bezeichnen wir mit kleinen lateinischen Buchstaben vom Anfang des Alphabetes, (es sei denn, sie haben schon eine Bezeichnung wie π oder 5). Konkrete Teilmengen aus $\mathbb{R}$ bzw. $\mathbb{R}^p$ bezeichnen wir mit großen lateinischen Buchstaben des Alphabetes, (es sei denn, sie haben schon eine Bezeichnung wie z.B. $[0,1]$ (Einheitsintervall). Dann wählen wir gelegentlich auch diese.

II) **Atomare Formeln und Sätze**
Es sind dies

1. $s = t$, wo s, t Konstante oder Variable sein können;

2. $(s_1, \ldots s_p) = (t_1, \ldots, t_p)$, wo die s_j, t_k Konstanten oder Variablen sein können;

3. $s \in A$, wo s eine Konstante oder eine Variable sein kann und $A \subset \mathbb{R}^p$ eine Menge ist;

4. $(s_1, \cdots, s_p) \in B$, wo s_j wieder für eine Konstante oder Variable stehen kann und B wieder eine Teilmenge von $\mathbb{R}^p$ ist.

Kommen in einer atomaren Formel nur Konstanten und keine Variablen vor, so nennen wir dies einen **atomaren Satz.**

III) **Zusammengesetzte Formeln und Sätze**
(i) Ist Φ eine Formel, so ist $\neg\Phi$ wieder eine Formel, ihre **Verneinung (Negation)**
Zum Beispiel ist $\neg(\pi \in [0,2])$ einfach dasselbe wie $\pi \notin [0,2]$.

(ii) Sind Φ und Ψ Formeln, so ist $\Phi \wedge \Psi$ eine Formel, ihre Konjuktion (Φ "und" Ψ).

(iii) Sei Φ eine Formel. In Φ komme x als Variable vor, aber nirgends in der Form $\exists(x \in D)$ wo D eine beliebige Menge aus $\mathbb{R}^p$ ist (p beliebig). Man bezeichnet dies mit $\Phi(x)$ und sagt, x ist eine **freie Variable** in Φ. Sei B eine Teilmenge von $\mathbb{R}^p$. Dann ist

$$\Psi \equiv \exists(x \in B)\Phi(x)$$

wieder eine Formel (Lies: Es gibt $x \in B\Phi(x)$). Man sagt, daß jetzt x in Ψ als **gebundene Variable** vorkommt.

Unsere zulässige Formelmenge $\mathcal{L}$ erhält man aus den atomaren Formeln durch (beliebig oft) wiederholte Anwendungen der Regeln aus (III).

Ein **mathematischer Satz** ist eine Formel, die nur gebundene Variable enthält.

Wir haben absichtlich solch einfaches Rezept zum Aufbau der Formeln gewählt, um das Transfer-Prinzip beweisen zu können. Dafür müssen wir uns beim Nachweis, daß ein mathematischer Satz gleichwertig mit solchen Formeln ausgedrückt werden kann, etwas mehr anstrengen. Dabei benutzen wir zu allererst einmal die folgenden logischen Äquivalenzen, die wir oft in der Mathematik zum Beispiel bei indirekten Beweisen benutzen.

(i) $\Phi \vee \Psi$ ist äquivalent zu (i') $\neg(\neg\Phi \wedge \neg\Psi)$,
(ii) $\forall(x \in B)\Phi(x)$ ist äquivalent zu (ii') $\neg(\exists(x \in B)(\neg\Phi(x))$,
(iii) $\Phi \Rightarrow \Psi$ ist äquivalent zu (iii') $\neg(\Phi \wedge \neg\Psi)$.
(iv) $\forall(x \in B)\Phi(x)$ ist äquivalent zu (iv') $\neg(\exists(x \in B)(\neg\Phi(x)))$.

Das bedeutet nichts anderes als daß ein Satz der Form (i) bis (iv) genau dann gilt, wenn sein äquivalenter Satz (i'), bzw. (ii') bzw. (iv') gilt.

Wann ist ein mathematischer Satz wahr oder anders ausgedrückt, wann gilt er?
An sich ist da klar. Wir können es aber auch anhand seines Aufbaus induktiv definieren:

(i) Ein atomarer Satz von der Form $a = b$ gilt genau dann, wenn a gleich b ist. Analog ist $(a_1, \ldots, a_p) = (b_1, \ldots, b_p)$wahr, wenn $a_j = b_j$ für $j = 1 \ldots p$ ist. $a \in A$ gilt, wenn a tatsächlich ein Element von A ist. Analoges gilt für $(a_1, \ldots, a_p) \in B$.
Damit ist die Gültigkeit atomarer Sätze festgelegt.

(ii) $\neg\Phi$ gilt genau dann, wenn Φ nicht gilt.
$\Phi \wedge \Psi$ gilt genau dann, wenn sowohl Φ als auch Ψ gilt.
$\exists(x \in B)\Phi(x)$ gilt genau dann, wenn es ein Element $b \in B$ gibt, so daß der Satz $\Phi(b)$ den man erhält, wenn man an allen Stellen, an denen x vorkommt, dies durch das Element b ersetzt, gilt.

Damit haben wir induktiv nach dem Aufbau eines Satzes seine Gültigkeit erklärt. Ich betone noch einmal, daß dies genau unserer üblichen Festlegung entspricht.

1.5.2 Die *–Transformierte von Formeln

Sei Φ eine Formel. Dann ist $^*\Phi$ die Formel, die man aus Φ erhält, indem man jede Menge B, die (als Konstante) in Φ vorkommt, durch $^*B$ in $^*\mathcal{S}$ ersetzt. Damit erhält man eine Formel $^*\Phi$ in $^*\mathcal{L}$, die *–Transformierte von Φ. Ist Φ ein mathematischer Satz (in $\mathcal{L}$), so ist $^*\Phi$ ein solcher in $^*\mathcal{L}$ (und umgekehrt).

Beispiele: Wir benutzen zur Erleichterung die oben angegebenen äquivalenten Formeln.

1. $\Pi \in [0,2] \equiv \Phi \cdot \qquad {}^*\Phi \equiv \Pi \in {}^*[0,2]$.

2. Sei $P = \{(x,y) \in \mathbb{R}^2 : x \leq y\}$. Dann ist

   $$\Phi \equiv \forall(x \in \mathbb{R})\forall(y \in \mathbb{R})[(x,y) \in P \vee (y,x) \in P]$$

   ein mathematischer Satz, der nichts anderes aussagt, als daß die Ordnung in $\mathbb{R}$ total ist. $^*\Phi \equiv \forall(x \in {}^*\mathbb{R})\forall(y \in {}^*\mathbb{R})[(x,y) \in {}^*P \vee (y,x) \in {}^*P]$ besagt dasselbe in $^*\mathbb{R}$.

3. Wie beschreibt man den Sachverhalt, daß zu jedem $x \geq 0$ ein $n \in \mathbb{N}$ mit $n \geq x$ existiert?
   Der Satz lautet

   $$\forall(x \in \mathbb{R})[(0,x) \in P \Rightarrow \exists(n \in \mathbb{N})[(x,n) \in P]].$$

   Der *–transformierte Satz lautet unter der Berücksichtigung der oben formulierten äquivalenten Formeln

   $$\forall(x \in {}^*\mathbb{R})[(0,x) \in {}^*P \Rightarrow \exists(n \in {}^*\mathbb{N})((x,n) \in {}^*P)]$$

   Übersetzen wir ihn in die Alltagssprache, lautet dieser Satz in $^*\mathbb{R}$:
   Für jedes $\xi \in {}^*\mathbb{R}$ mit $\xi \geq 0$ gibt es ein $n \in {}^*\mathbb{N}$ (**nicht** $\in \mathbb{N}$) mit $\xi \leq n$.

1.5.3 Das Transferprinzip

Wir benötigen zunächst den folgenden Satz, eine eingeschränkte Version eines Theorems von Lós.

Satz 1.25 *Sei $\Phi = \Phi(x_1, \cdots, x_\tau)$ eine Formel mit $x_1, \cdots, x_\tau$ als einzigen freien Variablen. Seien $\alpha_j \in {}^*\mathbb{R}, \alpha_j = (a_{j,n})_n^\bullet$.*
Gilt $^*\Phi(\alpha_1, \cdots, \alpha_r)$, so ist $\{n : \Phi(a_{1,n}, \cdots, a_{r,n})$ gilt in $\mathcal{S}\}$ eine 1-Menge.

Beweis: (durch Induktion über den Aufbau der Formel)
(I) Aufgrund von Beispiel 8 und 9 im Anschluß an Satz 1.24 genügt es, die Aussage für die atomaren Formeln vom Typ

$$x_1 = b, \quad x_1 = x_2,$$

$$(x_1, \cdots, x_p, b_1, \cdots, b_q) = (x_{p+1}, \cdots, x_{2p}, b_1, \cdots, b_q) \quad (2p = r)$$

und

$$(x_1, \cdots, x_r, b_1, \cdots, b_q) \in B$$

zu beweisen.
Nach Konstruktion von ${}^*\mathbb{R}$ gilt $\alpha_1 = b$ genau dann wenn $[a_{1,n} = b]$ eine 1-Menge ist. Ebenso gilt $\alpha_1 = \alpha_2$ genau dann wenn $[a_{1,n} = a_{2,n}]$ eine 1-Menge ist. Auch die Aussage für die nächste Formel ist evident. Schließlich ist für $\Phi \equiv (x_1, \cdots, x_r, b_1, \cdots, b_q) \in B$ die gesternte Formel ${}^*\Phi \equiv (x_1, \cdots, x_r, b_1, \cdots, b_q) \in {}^*B$. Nach Definition gilt ${}^*\Phi(\alpha_1, \cdots, \alpha_r, b_1, \cdots, b_q)$ genau dann wenn $(\alpha_1, \cdots, \alpha_r, b_1, \cdots, b_q) \in {}^*B$ und dies ist genau dann der Fall (s.a. das oben zitierte Beispiel 8) wenn

$$\{n : (a_{1,n}, \cdots, a_{r,n}, b_1, \cdots, b_q) \in B\}$$

eine 1-Menge ist. Dies ist aber äquivalent zu der Aussage

$$\{n : \Phi(a_{1,n}, \cdots, a_{r,n}) \text{ gilt in } \mathcal{S}\}$$

ist eine 1-Menge. Damit ist der Satz für alle atomaren Formeln bewiesen.

(II) Sei der Satz schon für $\Phi(x_1, \cdots, x_r)$ bewiesen. Es ist ${}^*(\neg\Phi) = \neg{}^*\Phi$. Also gilt ${}^*(\neg\Phi(\alpha_1, \cdots, \alpha_r))$ genau dann, wenn $\neg{}^*\Phi(\alpha_1, \cdots, \alpha_r)$ gilt, das heißt, wenn ${}^*\Phi(\alpha_1, \cdots, \alpha_r)$ falsch ist. Da m ein Ultrafunktional ist, ist dies genau dann der Fall, wenn $\{n : \Phi(a_{1,n}, \cdots, a_{r,n})$ gilt $\}$ eine 0-Menge, also $\{n : \neg\Phi(a_{1,n}, \cdots, a_{r,n})$ gilt$\}$ eine 1-Menge ist. Das beweist die Aussage für $\neg\Phi$.

(III) Sei die Aussage für Φ und Ψ bewiesen. Die Variablen $(x_1, \cdots, x_r)$ mögen sich aufteilen in $(y_1, \cdots, y_p)$, die in Φ und $(z_1, \cdots, z_q)$, die in Ψ frei vorkommen.
Es ist ${}^*(\Phi(y_1, \cdots, y_p) \wedge \Psi(z_1, \cdots, z_q)) = {}^*\Phi(y_1, \cdots, y_p) \wedge {}^*\Psi(z_1, \cdots, z_q)$.
${}^*\Phi(\alpha_1, \cdots, \alpha_p) \wedge {}^*\Phi(\beta_1, \cdots, \beta_q)$ gilt nun genau dann, wenn sowohl ${}^*\Phi(\alpha_1, \cdots, \alpha_p)$ als auch ${}^*\Phi(\beta_1, \cdots, \beta_q)$ gilt. Nach Voraussetzung ist die genau dann der Fall, wenn sowohl
$\{n : \Phi(a_{1,n}, \cdots, a_{p,n})$ gilt$\}$ als auch $\{n : \Psi(b_{1,n}, \cdots, b_{q,n})$ gilt$\}$ 1-Mengen sind, also genau dann, wenn

$$\{n : \Phi(a_{1,n}, \cdots, a_{p,n}) \wedge \Psi(b_{1,n}, \cdots, b_{q,n}) \text{ gilt}\} = \{n : \Phi(a_{1,n}, \cdots, a_{p,n}) \text{ gilt}\} \cap \{n : \Psi(b_{1,n}, \cdots, b_{q,n}) \text{ gilt}\}$$

eine 1-Menge ist. Also ist der Satz auch für $\Phi \wedge \Psi$ gezeigt (falls er für Φ und Ψ gilt).

(IV) Sei $\Phi \equiv (\exists y \in B)\Psi(x_1, \cdots, x_r, y)$, wo der Satz für Ψ bereits bewiesen ist.
Es ist ${}^*\Phi \equiv (\exists y \in {}^*B){}^*\Psi(x_1, \cdots, x_r, y)$.
Es ist nun

$$\exists(y \in {}^*B){}^*\Psi(\alpha_1, \cdots, \alpha_r, y)$$

genau dann wahr in ${}^*\mathcal{L}$, wenn es ein $\beta = (b_n)_n^\bullet \in {}^*B$ gibt, so daß

$${}^*\Psi(\alpha_1, \cdots, \alpha_r, \beta)$$

gilt. Dies ist nach Induktionsvoraussetzung genau dann wahr, wenn

$$\{n : b_n \in B \wedge \Psi(a_{1,n}, \cdots, a_{r,n}, b_n) \,\text{gilt}\}$$

eine 1-Menge ist. Daraus folgt aber, daß

$$\{n : (\exists y \in B)\Psi(a_{1,n}, \cdots, a_{r,n}, y) \,\text{gilt}\}$$

erst recht eine 1-Menge ist; damit ist der Satz für Φ bewiesen. □

Aus diesem Satz folgt sofort das gewünschte elementare Transferprinzip:

Theorem 1.26 (Elementares Transferprinzip) *Sei Φ ein mathematischer Satz in $\mathcal{L}$. Φ gilt genau dann in $\mathcal{S}$, wenn $^*\Phi$ in $^*\mathcal{S}$ gilt.*

Beweis: Sei $^*\Phi$ wahr in $^*\mathcal{S}$. Nach Satz 1.25 ist dann $\{n : \Phi \,\text{gilt}\}$ eine 1-Menge. Aber da Φ keine Variablen hat, ist $\{n : \Phi \,\text{gilt}\}$ entweder $\mathbb{N}$ oder leer; in unserem Fall muß die Menge gleich $\mathbb{N}$ sein, also gilt Φ in $\mathcal{S}$. Sei $^*\Phi$ falsch in $^*\mathcal{S}$. Dann ist $\neg{}^*\Phi = {}^*(\neg\Phi)$ wahr in $^*\mathcal{S}$, also nach dem bisher Bewiesenen $\neg\Phi$ wahr in $\mathcal{S}$, also Φ falsch in $\mathcal{S}$. □

Korollar 1.27 *Sei Ψ ein Satz, der aus den atomaren Formeln und zusätzlich aus "$\vee$", "$\Rightarrow$", "$\forall$" (neben "$\wedge$", "$\neg$", "$\exists$") aufgebaut ist. Ψ gilt genau dann in $\mathcal{S}$, wenn $^*\Psi$ in $^*\mathcal{S}$ gilt.*

Beweis: Man kann den Satz aufgrund der Äquivalenz auf S. 24 in einen Satz umwandeln, der nur die Zeichen "$\wedge$", "$\neg$", und "$\exists$" enthält. Die *–Transformation respektiert diese Umformulierung, weil der * nur bei Mengen angewandt wird. Daraus folgt die Behauptung. □

Wir liefern einige Beispiele, um zu demonstrieren, daß unser Transfer-Prinzip liefert, was wir möchten. Dabei beweisen wir zunächst Dinge, die wir früher direkt gezeigt haben.

Beispiele:

1. Es gilt $\sin(\xi + 2\pi) = \sin(\xi)$ für alle $\xi \in {}^*\mathbb{R}$. Dabei schreiben wir sin statt $^*\sin$.
   Beweis: Sei $b = 2\pi$ (das ist eine Konstante in $\mathbb{R}$). Dann gilt in $\mathbb{R}$

   $$\forall(x \in \mathbb{R})(\sin(x + b) = \sin(x)).$$

   Sei G der Graph des Sinus und S der Graph der Addition. (Wir müssen das so umständlich machen, weil wir nur "$\in$" und "$=$" in unserer Sprache haben). Dann lautet die Formel

   $$\forall(x \in \mathbb{R})\forall(y \in \mathbb{R})\forall(z \in \mathbb{R})[(x, b, z) \in S \wedge (x, y) \in G \Rightarrow (z, y) \in G].$$

   Dieser Satz gilt in $\mathbb{R}$, also gilt der gesternte in $^*\mathbb{R}$, aber das war gerade, was wir beweisen wollten.

2. Genauso beweist man für komplexe Zahlen ξ, $\eta \in {}^*\mathbb{C}$ die Formel

$$\exp(\xi + \eta) \;=\; \exp(\xi)\exp(\eta).$$

Den Graphen des Produkts in $\mathbb{C}$ bezeichnen wir mit $P_{\mathbb{C}}$, den Graphen der Addition mit $S_{\mathbb{C}}$. Dann erhalten wir den in $\mathbb{C}$ gültigen Satz
$\forall x, y, z, u, v, w \in \mathbb{C}$

$$[\underbrace{(x,y,z) \in S}_{z=x+y} \wedge \underbrace{(x,v) \in \exp}_{v=\exp(x)} \wedge \underbrace{(y,w) \in \exp}_{w=\exp(y)} \wedge \underbrace{(z,u) \in \exp}_{u=\exp(z)} \Rightarrow \underbrace{(v,w,u) \in P_{\mathbb{C}}}_{u=vw}].$$

Der gesternte Satz liefert die Behauptung.

3. Wir wissen: ist G_f Graph einer Funktion f, so ist ${}^*G_f$ Graph einer eindeutig bestimmten Fortsetzung. Statt $(x, y) \in G_f$ schreiben wir in Zukunft $y = f(x)$ usw. Damit erhalten wir viel einfacher:

$$\cos(\xi)^2 \,+\, \sin(\xi)^2 \;=\; 1.$$

Denn mit der erleichterten Schreibweise haben wir in $\mathbb{R}$

$$\forall(x \in \mathbb{R})(\cos(x)^2 \,+\, \sin(x)^2 \;=\; 1).$$

4. Sei $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$ und M eine feste Konstante. Es gelte $\|f(x)\| \leq M$ für alle $x \in D$, d.h.: $\forall(x \in D)(\|f(x)\| \leq M)$. Der *–transformierte Satz heißt dann

$$\forall(x \in {}^*D)(\|{}^*f(x)\| \leq M)$$

und besagt, daß die fortgesetzte Funktion beschränkt mit der gleichen Schranke ist.

5. Sei wieder $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$ und ${}^*f$ die fortgesetzte Funktion. Es gebe eine (unendlich große) Zahl ξ_0 mit $\|{}^*f(\xi)\| \leq \xi_0$ für alle $\xi \in {}^*D$. Dann gibt es eine Standardzahl M mit $\|{}^*f(\xi)\| \leq M$ für alle $\xi \in {}^*D$. Denn unsere Voraussetzung sagt:

$$\exists(x \in {}^*\mathbb{R})\forall(y \in {}^*D)(\|f(y)\| \leq x)$$

gilt in ${}^*\mathcal{S}$ (dort ist das existierende x gerade ξ_0). Nach dem Transferprinzip muß der Satz auch in $\mathcal{S}$ gelten. Also gibt es $M \in \mathbb{R}$ mit $\|f(x)\| \leq M$ für alle $x \in D$, d.h. es gilt

$$\forall(x \in D(\|f(x)\| \leq M).$$

Erneute Anwendung des Transferprinzips liefert die Behauptung.

Als eine weitere Anwendung des Transferprinzips notieren wir das folgende kleine, aber hilfreiche Spillover-Prinzip:

Lemma 1.28 (Spillover-Prinzip) *Sei A eine Teilmenge von $\mathbb{N}$ und ${}^*A$ enthalte ${}^*\mathbb{N}\backslash\mathbb{N}$, d.h. alle unendlich großen natürlichen Zahlen. Dann gibt es ein $n_0 \in \mathbb{N}$, so daß A alle $n \geq n_0$ enthält.*

Beweis: Sei $N_0 \approx \infty$. Dann gilt

$$\forall(N \in {}^*\mathbb{N})\,(N \geq N_0 \Rightarrow N \in {}^*A).$$

Erst recht gilt also

$$\exists(n \in {}^*\mathbb{N})\,(\forall n \in {}^*\mathbb{N})\,(n \geq n_0 \Rightarrow n \in {}^*A).$$

Nach dem Transferprinzip gilt der ungesternte Satz in $\mathcal{S}$, und das war die Behauptung. $\square$

Korollar 1.29 *Sei $A \subset \mathbb{N}$ und $\mathbb{N} \subset {}^*\mathbb{N} \setminus {}^*A$. Dann ist A endlich.*

Ebenso wie das Lemma 1.28 kann man zeigen:

Lemma 1.30 (2-dimensionales Spillover-Prinzip) *Sei $B \subset \mathbb{N} \times \mathbb{N}$. Enthält ${}^*B$ die Menge $({}^*\mathbb{N} \setminus \mathbb{N})^2$, so gibt es ein n_0, so daß $(n,m) \in B$ für alle (n,m) mit $n \geq n_0$ und $m \geq n_0$.*

1.6 Folgen und Reihen

Wir fassen Folgen in $\mathbb{R}$ als Funktionen von $\mathbb{N}$ in $\mathbb{R}$ auf, also $a = (a_n)_n$ ist nichts anderes als $n \mapsto a_n$. Die auf $^*\mathbb{N}$ fortgesetzte Funktion wird wieder als Folge bezeichnet. Der Ausdruck a_N für $n \in {}^*\mathbb{N} \setminus \mathbb{N}$ ist also eindeutig erklärt.

Deffnition 1.31 *Sei $a = (a_n)_n$ eine Folge. $b \in \mathbb{R}$ heißt Häufungspunkt von a, wenn es ein $N \approx \infty$ (d.h. $N \in {}^*\mathbb{N} \setminus \mathbb{N}$) gibt mit $b = {}^\circ a_N$. a heißt konvergent gegen b, wenn $a_N \approx b$ für alle $N \approx \infty$ gilt.*

Von den vielen Anwendungen dieser simplen Definition geben wir die folgende:

Satz 1.32 (Bolzano-Weierstraß) *Jede beschränkte Folge $(a_n)_n$ hat einen Häufungspunkt.*

Beweis: Sei $|a_n|_n \leq M$ (in $\mathbb{R}$) für alle $n \in \mathbb{N}$. Nach dem Transferprinzip ist dann $|a_N| \leq M$ für alle $N \in {}^*\mathbb{N}$. Nach Theorem 1.20 hat a_N dann einen Standardteil $b = {}^\circ a_N$. $\square$

Der größte Häufungspunkt $\limsup a_n$ ist nichts anderes als

$$\limsup a_n = \begin{cases} \infty & \exists (N \approx \infty) a_N \approx \infty \\ -\infty & \forall (N \approx \infty) a_N \approx -\infty \\ \sup & \{{}^\circ a_N : N \approx \infty\} \text{ sonst} \end{cases}$$

Ist dieser Wert endlich, so ist er tatsächlich ein Häufungspunkt der Folge, wie der folgende Satz sagt.

Satz 1.33 *Ist $-\infty < b := \limsup a_n < \infty$, so ist b ein Häufungspunkt der Folge (a_n).*

Beweis: Ist $N \approx \infty$, so ist nach Definition ${}^\circ a_N \leq b$, bzw. a_N negativ unendlich groß. Daraus folgt für alle $N \approx \infty$ $a_N \approx b$ oder $a_N < b$. Wäre b kein Häufungspunkt, so gäbe es ein infinitesimales $\delta > 0$ mit $a_N < \delta$ für alle $N \approx \infty$ insbesondere für alle $N \geq N_0 \approx \infty$, wo N_0 fest gewählt ist. Damit gilt

$$\exists (d > 0) \exists (n_0 \in {}^*\mathbb{N}) \forall n \geq n_0 [a_n < b - d]$$

in $^*\mathbb{R}$, also nach dem Transferprinzip auch in $\mathbb{R}$. Seien $d > 0$ aus $\mathbb{R}$ und $n_0 \in \mathbb{N}$ solche Größen. Dann ist also $a_n < b - d$ für alle $n \geq n_0$. Erneute Anwendung des Transferprinzips liefert $a_N \leq b - d$ für alle $N \approx \infty$, also ${}^\circ a_N \leq b - d$ für alle $N \approx \infty$, also $b \leq b - d$, ein Widerspruch. $\square$

Deffnition 1.34 *Eine Folge $a = (a_n)_n$ heißt Cauchyfolge, wenn für alle $N, N' \approx \infty$ stets $a_N \approx a_{N'}$ ist.*

Ist (a_n) konvergent, also $s \approx a_N$ für alle $N \approx \infty$, so ist $a_N \approx a_{N'}$, für alle $N, N' \approx \infty$.

Satz 1.35 *Jede Cauchyfolge konvergiert.*

Beweis: (I) Sei (a_n) eine Cauchyfolge. Dann sind alle a_N $(N \approx \infty)$ endlich. Denn sei $B = \{(m,n) : |a_m - a_n| < 1\}$. ${}^*B$ enthält nach Voraussetzung ${}^*\mathbb{N}_\infty \times {}^*\mathbb{N}_\infty$ mit ${}^*\mathbb{N}_\infty = {}^*\mathbb{N} \setminus \mathbb{N})$. Nach dem Spillover-Prinzip 1.30 folgt, daß $B \supset A_{n_0} \times A_{n_0}$, wo $A_{n_0} = \{n \in \mathbb{N} : n \geq n_0\}$ ist. Dann ist aber für $n \geq n_0$ $|a_n| \leq |a_n - a_{n_0}| \leq 1 + |a_{n_0}|$, also für beliebige n

$$|a_n| \leq 1 + \max\{|a_k| : k \leq n_0\}.$$

Aufgrund von Theorem 1.20 existiert $s = {}^\circ a_N$ für ein festes $N \approx \infty$. Aber dann gilt für alle $M \approx \infty$ stets $s \approx a_N \approx a_M$, also nach unserer Definition $s = \lim_{n\to\infty} a_n$. □

Als nächstes steuern wir die Césaro-Limitierung an.
Für eine beliebige Folge $a = (a_n)_n$ sei $c(a) = (c_n(a))$ mit $c_n = \dfrac{1}{n} \sum_{k=1}^{n} a_k$. Mit a ist auch $c(a)$ beschränkt. Sei $\ell^\infty(\mathbb{N}) = \{(a_n)_n : \sup_n(|a_n|) < \infty\}$. Zur Abkürzung schreiben wir $\|a\| := \sup\{|a_n| : n \in \mathbb{N}\}$.

Satz 1.36 *Sei $N \approx \infty$ fest gewählt und $\gamma : \ell^\infty(\mathbb{N}) \to \mathbb{R}$ gegeben durch $\gamma(a) = {}^\circ c_N(a)$. Dann gilt:*
(i) γ ist linear.
(ii) $|\gamma(a)| \leq \gamma(|a|)) \leq \|a\|$.
(iii) Ist a konvergent, so ist $\gamma(a) = \lim a_n$.
(iv) Sei (a'_n) die um k Stellen verschobene Folge also $a'_n = a_{n+k}$. Dann ist $\gamma(a')) = \gamma(a)$.

Eine Abbildung $\gamma : \ell^\infty(\mathbb{N}) \to \mathbb{R}$ mit den Eigenschaften (i) - (iv) heißt Banach-Limes. Der klassische Existenzbeweis ist erheblich aufwendiger.

Beweis: (i) Es ist $c(\lambda a + \mu b) = \lambda c(a) + \mu c(b)$, daher nach dem Transferprinzip $c_N(\lambda a + \mu b) = \lambda c_N(a) + \mu c_N(b)$, also $\gamma(\lambda a + \mu b) = {}^\circ(\lambda c_N(a) + \mu c_N(b)) = \lambda\gamma(a) + \mu\gamma(b)$, weil "°" linear ist.
(ii) Es ist $|c_n(a)| \leq c_n(|a|) \leq \sup_k |a_k|$ für alle n. Nach dem Transferprinzip folgt

$$|c_N(a)| \leq c_N(|a|) \leq \|a\|.$$

Aus den Eigenschaften von "°" folgt die Behauptung.
(iii) Sei $h(n) = \max\{k \in \mathbb{N} : k \leq \sqrt{n}\}$. Dann gilt $h(n) \leq \sqrt{n} < h(n) + 1$. Transfer liefert ${}^*h(N) \leq \sqrt{N} \leq h(N) + 1$. Wir schreiben wie üblich $h(n) = [\sqrt{n}]$.
Sei $s_n(a) = \frac{1}{n} \sum_{k=1}^{[\sqrt{n}]} a_k$, $t_n(a) = \frac{1}{n} \sum_{k=[\sqrt{n}]+1}^{n} a_k$.
Dann gilt $c_n(a) = s_n(a) + t_n(a)$, ferner $|s_n(a)| \leq \dfrac{1}{\sqrt{n}-1}\|a\|$, und $|t_n(a)| \leq \max\{|a_k| : [\sqrt{n}] \leq k \leq n\}$. Wir beachten, daß das Maximum angenommen wird, was sich durch die folgende Formel ausdrückt:

$$\forall(n \in \mathbb{N})\exists(k \in \mathbb{N})\left(([\sqrt{n}] \leq k \leq n) \wedge \forall(l \in \mathbb{N})([\sqrt{n}] \leq l \leq n \Rightarrow |a_l| \leq |a_k|)\right),$$

eine Formel deren gesternte Variante dann auch in $^*\mathbb{R}$ gilt. Sei nun a eine Nullfolge. Dann ist für unser festes $N \approx \infty$

$$\begin{array}{rcl} |c_N(a)| & \leq & |s_N(a)| + |t_N(a)| \\ & \leq & \dfrac{\|a\|}{\sqrt{N}-1} + \max\{|a_k| : [\sqrt{N}] \leq k \leq N\} \\ & \approx & 0 + 0, \end{array}$$

weil $a_M \approx 0$ für alle $M \approx \infty$; also ist $\gamma(a) = 0 = \lim a_n$. Ist b konstant, $b = (u, u, u, \cdots)$, so ist $c_n(b) = u$. Ist also $\lim a_n = u$, so ist $\lim(b_n - a_n) = 0$ und damit $u - \gamma(a) = \gamma(b - a) = 0$, also $\gamma(a) = \lim a_n$.
(iv) Wir zeigen die Formel für $k = 1$. Für beliebiges k folgt sie durch Induktion.
Es ist $c_n(a) - c_n(a') = \frac{a_1}{n} - \frac{a_{n+1}}{n}$, also $|c_N(a) - c_N(a')| \leq \dfrac{2\|a\|}{N} \approx 0$ und damit $\gamma(a) = \gamma(a')$.
$\square$

Brücke zur Standard-Mathematik

Wir müssen zeigen, daß unsere Definitionen des Häufungspunktes, der Konvergenz und der Cauchyfolge mit dem aus Analysis I übereinstimmen. Dies geschieht mit dem Transferprinzip und dem Spill-Over-Prinzip 1.28 bzw. 1.30.

Satz 1.37 *Sei $a = (a_n)_n$ eine Folge reeller Zahlen.*
a) Genau dann gilt $s = {}^\circ a_N$ für ein $N \approx \infty$, wenn s ein Häufungspunkt von a im Sinne der Standardmathematik ist.
b) Genau dann ist a eine Cauchyfolge in unserem Sinn, wenn für alle $\varepsilon > 0$ ein $n(\varepsilon)$ mit $|a_m - a_n| < \varepsilon$ für alle $m, n \geq n(\varepsilon)$ gilt.
c) Genau dann ist $s = \lim a_n$, wenn zu jedem $\varepsilon > 0$ ein $n(\varepsilon)$ mit $|a_n - s| < \varepsilon$ für alle $n \geq n(\varepsilon)$ gilt.

Beweis: a) (i) Sei $s = {}^\circ a_N$, $k \in \mathbb{N}$ und $\varepsilon > 0$ aus $\mathbb{R}$, also eine Standardzahl. $N \approx \infty$ erfüllt $N > k$ und $|a_N - s| < \varepsilon$. Also gilt $\exists(n \in {}^*\mathbb{N})((N > k) \wedge (|a_n - s| < \varepsilon))$ in $^*\mathbb{R}$ und nach dem TFP (Transferprinzip) in $\mathbb{R}$. Das bedeutet aber: zu $\varepsilon > 0$ und $k \in \mathbb{N}$ beliebig existiert ein $n(\varepsilon, k) > k$ mit $|a_{n(\varepsilon,k)} - s| < \varepsilon$. s ist Häufungspunkt im Standardsinn.
(ii) Sei s Häufungspunkt im Standardsinn. Dann gilt

$$\forall(\varepsilon \in \mathbb{R})\forall(n \in \mathbb{N})\exists(k \in \mathbb{N})(\varepsilon > 0 \Rightarrow (k \geq n) \wedge |s - a_k| < \varepsilon).$$

Nach dem TFP gilt dieser Satz in $^*\mathbb{R}$. Wählt man also $\varepsilon \approx 0$ und $N \approx \infty$, so erhält man ein $k \geq N$ mit $|s - a_k| < \varepsilon$, also $s = {}^\circ a_k$.
b) (i) Sei a eine Cauchyfolge, d.h. $a_M \approx a_N$ für $M, N \approx \infty$. Sei $\varepsilon > 0$ aus $\mathbb{R}$ und $A(\varepsilon) = \{(m, n) : |a_m - a_n| < \varepsilon\}$ $^*A(\varepsilon)$ enthält $(^*\mathbb{N} \setminus \mathbb{N})^2$. Nach dem Spilloverprinzip Lemma 1.30

enthält $A(\varepsilon)$ eine Menge der Form $A^2_{n(\varepsilon)}$ mit $A_{n(\varepsilon)} = \{k : k \geq n(\varepsilon)\}$, also folgt die Behauptung.
(iii) Es gelte umgekehrt die angegebene Bedingung. Seien $M, N \approx \infty$ fest gewählt und sei $\varepsilon > 0$ aus $\mathbb{R}$ beliebig, ferner $n(\varepsilon) \in \mathbb{N}$ wie angegeben. Unsere Voraussetzung gilt nach dem TFP auch in ${}^*\mathbb{R}$, d.h. $\forall(m \in {}^*\mathbb{N})\forall(n \in {}^*\mathbb{N})[m, n \geq n(\varepsilon) \Rightarrow |a_m - a_n| < \varepsilon]$. M, N sind größer als die Standard-Zahl $n(\varepsilon)$ also gilt $|a_M - a_N| < \varepsilon$. Da $\varepsilon > 0$ aus $\mathbb{R}$ beliebig war, folgt $|a_M - a_N| \approx 0$, also $a_M \approx a_N$.
c) Übungsaufgabe. □

1.7 Funktionen, Grenzwerte und Stetigkeit

Überblick: Wir werden Grenzwert und Stetigkeit zunächst intuitiv, das mit Hilfe der Nonstandard-Analysis definieren und damit in einfacher Weise die üblichen Sätze der analysis beweisen. Dann schlagen wir wieder die Brücke zur Standard-Mathematik. Im letzten Teil behandeln wir die gleichmäßgige Stetigkeit und die gleichmäßige Konvergenz von Folgen von Funktionen.

Häufungspunkte und Grenzwerte

Sei $\emptyset \neq D \subset \mathbb{R}$.

Deffnition 1.38 *$c \in \mathbb{R}$ heißt* **Häufungspunkt** *von D, wenn es ein $x \in {}^*D$ mit $c \neq x \approx c$ gibt.*

c ist also genau dann Häufungspunkt von D wenn die punktierte Monade $\mu(c) \setminus \{c\}$ mit ${}^*D$ einen nichtleeren Durchschnitt hat, in Formeln, wenn $\mu(c) \setminus \{c\} \cap {}^*D \neq \emptyset$ gilt.

Im folgenden sei $D \subset \mathbb{R}$ und c ein Häufungspunkt von D. Sei $f : D \to \mathbb{R}$.

Deffnition 1.39 *Wir sagen, $f(x)$ konvergiert gegen d für $x \to c$ auf D wenn für alle $x \approx c$, $x \neq c$ stets ${}^*f(x) \approx d$ gilt. Wir schreiben $d = \lim_{c \neq x \to c} f(x)$.*
Wir können dies auch so ausdrücken:

$$d = \lim_{x \to c\, x \neq c} f(x) \Leftrightarrow {}^*f(\mu(c) \setminus \{c\}) \subset \mu(d).$$

Aus den Rechenregeln mit Ungleichungen in ${}^*\mathbb{R}$ (Satz 1.19) erhält man sämtliche bekannten Rechenregeln für die Grenzwerte.

Beispiel: Sei $d_j = \lim_{c \neq x \to c} f_j(x)$. Dann ist $d_1 d_2 = \lim_{c \neq x \to c} f_1(x) f_2(x)$ wegen ${}^*f_j(x) \approx d_j \Rightarrow {}^*f_1(x) {}^*f_2(x) \approx d_1 d_2$ (siehe Satz 1.19).

Stetigkeit

Deffnition 1.40 *Sei $f : D \subset \mathbb{R} \to \mathbb{R}$ eine Funktion. f heißt in $x_0 \in D$ stetig, wenn für alle $x \in {}^*D$ mit $x \approx x_0$ stets ${}^*f(x) \approx {}^*f(x_0)$ gilt; anders ausgedrückt, wenn ${}^*f(\mu(x_0) \cap {}^*D) \subset \mu(f(x_0))$ gilt.*

Satz 1.41 *Sei $f : D \subset \mathbb{R} \to \mathbb{R}$ stetig in x_0, ferner $f(D) \subset D'$ und $g : D' \to \mathbb{R}$ stetig in $y_0 = f(x_0)$. Dann ist $g \circ f$ stetig in x_0.*

Beweis: Sei $x \approx x_0$. Dann ist ${}^*f(x) \approx f(x_0) = y_0$, also ${}^*g({}^*f(x)) \approx g(y_0) = g(f(y_0))$. Wegen ${}^*(g \circ f) = {}^*g \circ {}^*f$ (s. Satz 1.24) folgt die Behauptung. $\square$

Für das folgende benötigen wir das Lemma:

Lemma 1.42 *Sei $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$. Dann gilt ${}^*(f(D)) = {}^*f({}^*D)$ und ${}^*(f^{-1}(B)) = ({}^*f)^{-1}({}^*B)$ für $B \subset \mathbb{R}^q$.*

Beweis: Der erste Teil folgt aus dem TFP, angewandt auf

$$\forall(y \in \mathbb{R}^q)(y \in f(D) \Leftrightarrow \exists(x \in D)(f(x) = y)),$$

der zweite wird analog bewiesen. $\square$

Es gibt drei wichtige Sätze für stetige Funktionen: der Satz über die Existenz von Maxima und Minima auf kompakten Mengen, der Zwischenwertsatz und der Umkehrsatz.
Wir definieren zunächst:

Deffnition 1.43 *$K \subset \mathbb{R}$ heißt* **kompakt**, *wenn ${}^*K \subset$ Fin $({}^*\mathbb{R})$ und für $x \in {}^*K$ stets der Standardteil ${}^\circ x \in K$ ist.*

Das folgende Lemma ist nützlich und einfach zu beweisen:

Lemma 1.44 *Sei K kompakt. Dann ist K beschränkt und die obere Grenze* $\sup(K) = s$ *liegt in K*

Beweis: Wegen ${}^*K \subset Fin({}^*\mathbb{R})$ ist K beschränkt. Aus der Definition der oberen Grenze folgt:

$$\forall(\varepsilon \in \mathbb{R})\,(\varepsilon > 0 \Rightarrow \exists(x \in K)(s - \varepsilon < x \leq s))\,.$$

Nach dem Transferprinzip findet man also zu $0 \approx \varepsilon > 0$ ein $x \in {}^*K$ mit $s - \varepsilon < x \leq s$. Aus Satz 1.19 folgt $s = {}^\circ(s - \varepsilon) \leq {}^\circ x \leq s$ und ${}^\circ x \in K$ gilt nach Voraussetzung. $\square$

Wir zeigen:

Satz 1.45 *Sei $K \subset \mathbb{R}$ kompakt und $f : K \subset \mathbb{R}$ stetig. Dann ist $f(K)$ kompakt.*

Beweis: Mit Lemma 1.42 folgt ${}^*(f(K)) = {}^*f({}^*K)$. Sei $y \in {}^*(f(K))$. Dann gibt es ein $x \in {}^*K$ mit ${}^*f(x) = y$. Nach Voraussetzung existiert ${}^\circ x$ und liegt in K. Da f stetig ist, ist wegen $x \approx {}^\circ x$ $y = {}^*f(x) \approx f({}^\circ x)$, also ${}^\circ y = f({}^\circ x) \in f(K)$. □

Korollar 1.46 *Sei K kompakt und $f : K \to \mathbb{R}$ stetig. Dann besitzt f ein Minimum und ein Maximum.*

Beweis: Da $f(K)$ kompakt ist, liegt $s := \sup(f(K))$ nach Lemma 1.44 in $f(K)$. Es gibt also ein $x_0 \in K$ mit $f(x_0) = s$. Die Existenz des Minmums folgt, wenn man $-f$ statt f betrachtet. □

Der Zwischenwertsatz lautet:

Satz 1.47 *Sei J ein Intervall und $f : J \to \mathbb{R}$ stetig. Seien $a, b \in J$ mit $f(a)f(b) < 0$. Dann gibt es eine Nullstelle von f in $]a, b[$.*

Beweis: Sei 0.B.d.A. $f(a) < 0 < f(b)$. Die Menge $K = \{x \in [a, b] : f(x) \leq 0\}$ ist kompakt (benutze das Transferprinzip). Es ist nach Lemma 1.44 $s = \sup(K) \in K$, also $f(s) \leq 0$. Sei $0 < \varepsilon \approx 0$. Dann ist nach Definition des Supremums und wegen des Transferprinzips $s + \varepsilon \notin {}^*K$, also

$$0 < {}^*f(s + \varepsilon) \approx f(s) \leq 0$$

und damit ist $f(s) = 0$. □

1.7.1 Brücke zur Standardanalysis

Satz 1.48 (Grenzwerte) *Sei $D = J \setminus \{c\}$ wo J ein offenes Intervall ist. Sei $f : D \to \mathbb{R}$. Die folgenden Aussagen sind äquivalent.*
a) $d = \lim_{x \to c} f(x)$, d.h. $x \approx c$ impliziert ${}^*f(x) \approx d$.
b) Für alle $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß für alle $x \in D$ aus $|x - c| < \delta$ stets $|f(x) - d| < \varepsilon$ folgt.

Beweis: a) $\Rightarrow$ b): Sei $\mathbb{R} \ni \varepsilon > 0$ und $\delta \in {}^*\mathbb{R}$, $\delta > 0$, $\delta \approx 0$. Nach a) ist dann $|{}^*f(x) - d| < \varepsilon$ für $|x - c| < \delta$. Daher ist der Satz

$$\exists(\delta \in {}^*\mathbb{R})(\delta > 0 \wedge \forall(x \in {}^*D)(|x - c| < \delta \Rightarrow |{}^*f(x) - d| < \varepsilon))$$

richtig in ${}^*\mathbb{R}$. Nach dem TFP gilt der ungesternte Satz in $\mathbb{R}$. Da $\varepsilon > 0$ beliebig war folgt b)
b) $\Rightarrow$ a) Sei ${}^*D \ni x \approx c$ und $\mathbb{R} \ni \varepsilon > 0$. Nach Voraussetzung gibt es $\delta > 0$ aus $\mathbb{R}$ mit $|f(y) - d| < \varepsilon$ für alle $|y - c| < \delta$, also $\forall(y \in D)(|y - c| < \delta \Rightarrow |f(y) - d| < \varepsilon)$. Nach dem TFP gilt der Satz in ${}^*\mathbb{R}$, also insbesondere für unser x wegen $|x - c| \approx 0$, also $|x - c| < \delta$. Damit ist $|f(x) - d| < \varepsilon$. Da $\mathbb{R} \ni \varepsilon > 0$ beliebig war folgt $f(x) \approx d$. □

Korollar 1.49 *Sei $f : D \to \mathbb{R}$ und $x_0 \in D$. Dann sind die folgenden Aussagen äquivalent.*
a) f ist in x_0 stetig, d.h. $x \approx x_0$ impliziert stets ${}^*f(x) \approx f(x_0)$.
b) $\forall(\varepsilon \in \mathbb{R})\exists(\delta \in \mathbb{R})\forall(x \in D)(|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon)$.

Satz 1.50 *Sei $K \subset \mathbb{R}$. Dann sind die folgenden Aussagen äquivalent:*
a) K ist kompakt.
b) Jede Folge $a = (a_n)_n \subset K$ hat einen Häufungspunkt in K.
c) K ist abgeschlossen und beschränkt.

Beweis: a)$\Rightarrow$ c): Wegen ${}^*K \subset Fin({}^*\mathbb{R})$ ist K beschränkt. Sei x_0 aus dem Komplement $\mathbb{R} \setminus K$ von K. Wäre x_0 ein Häufungspunkt von K, so gäbe es $x \in {}^*K$ mit $x \approx x_0$, also $K \ni {}^\circ x = x_0$, ein Widerspruch. Damit ist das Komplement offen und K abgeschlossen.
c) $\Leftrightarrow$ b): Das ist offensichtlich (vgl. die Vorl. Analysis I).
c) $\Rightarrow$ a): Da K beschränkt ist, liegt ${}^*K$ in $Fin({}^*\mathbb{R})$, jedes $x \in {}^*K$ hat also einen Standardteil ${}^\circ x =: y$. Liegt dieser nicht in K, so gibt es aufgrund der Abgeschlossenheit von K ein $\varepsilon > 0$ mit $]y - \varepsilon, y + \varepsilon[\cap K = \emptyset$. Wegen $x \approx y$ ist

$$x \in {}^*]y - \varepsilon, y + \varepsilon[\cap {}^*K \neq \emptyset,$$

obwohl diese Menge nach dem Transferprinzip leer sein müßte, ein Widerspruch. $\square$

Verallgemeinerung auf den $\mathbb{R}^p$

In ${}^*\mathbb{R}^p$ definieren wir eine Äquivalenzrelation "fast gleich" (infinitesimal benachbart) durch $\vec{\xi} \approx \vec{\eta} \Leftrightarrow \|\vec{\xi} - \vec{\eta}\| \approx 0$ (in ${}^*\mathbb{R}$). Ferner setzen wir $\mathrm{Fin}({}^*\mathbb{R}^p) = \{\vec{\xi} : \|\vec{\xi}\| \in \mathrm{Fin}({}^*\mathbb{R})\}$ und $\mu(\vec{\xi}) = \{\vec{\eta} : \vec{\xi} \approx \vec{\eta}\}$.

Wir erhalten den folgenden Satz

Satz 1.51 *a) "$\approx$" ist eine Äquivalenzrelation auf ${}^*\mathbb{R}^p$.*
b) $\vec{\xi} \approx \vec{\eta} \Leftrightarrow \xi_j \approx \eta_j$ in ${}^*\mathbb{R}$ für $j = 1, \cdots, p$, wo $\vec{\xi} = (\xi_1, \cdots, \xi_p)$ usw. ist.
c) $Fin({}^*\mathbb{R}^p)$ und $\mu(\vec{0})$ sind Vektorräume über $\mathbb{R}$.
d) Ist $\vec{\xi} \in Fin({}^*\mathbb{R}^p)$, so ist ${}^\circ\xi = \vec{\xi}\,{}^\circ\xi_1, \cdots, {}^\circ\xi_p)$ das eindeutig bestimmte Element aus $\mathbb{R}^p$ mit $\vec{\xi} \approx {}^{\circ *}\xi$.
Die Abbildung ${}^\circ : Fin({}^*\mathbb{R}^p) \to \mathbb{R}^p$, $\vec{\xi} \mapsto {}^\circ\vec{\xi}$ ist $\mathbb{R}$–linear und hat $\mu(\vec{0})$ als Kern. Sie heißt Standardteil-Abbildung.
e) $\vec{\xi} \approx \vec{0}$ und $\lambda \in Fin({}^*\mathbb{R}) \Rightarrow \lambda\vec{\xi} \approx \vec{0}$.

Aufgabe: Beweisen Sie bitte diesen Satz. Benutzen Sie dabei die Eigenschaften einer Norm, die wegen des TFP auch die fortgesetzte Norm, jetzt aber mit Werten in ${}^*\mathbb{R}_+$, hat.

Grenzwert, Stetigkeit und Kompaktheit werden analog wie in $\mathbb{R}$ definiert. Satz 1.45 und Korollar 1.46 gelten mutatis mutandis. Auch die Beweise ändern sich nicht. Wir behandeln das Problem der gleichmäßigen Stetigkeit und der gleichmäßigen Konvergenz.

Deffnition 1.52 *Sei $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$. f heißt* **gleichmäßig** *stetig, wenn aus $\vec{\xi}, \vec{\eta} \in {}^*D$ mit $\vec{\xi} \approx \vec{\eta}$ stets ${}^*f(\vec{\xi}) \approx {}^*f(\vec{\eta})$ folgt.*

Beispiele:

1. Sei $p = 1$ und $f(x) = x^2$. f ist *nicht* gleichmäßig stetig auf $\mathbb{R}$. Denn für $x \approx \infty$, $y = x + \frac{1}{x} \approx x$ gilt $y^2 = (x + \frac{1}{x})^2 = x^2 + 2 + \frac{1}{x^2} \approx x^2 + 2 \not\approx x^2$.

2. $f(x) = \sqrt{|x|}$ ist gleichmäßig stetig.
   Denn sei $x \approx y$.
   (i) ${}^\circ x = 0$. Dann gilt $|x| \approx |y| \approx 0$ und damit $f(x) \approx 0 = f(0) \approx f(y)$.
   (ii) $x \not\approx 0$. Dann ist $y \not\approx 0$ und damit $\sqrt{|x|} + \sqrt{|y|} \not\approx 0$, also $\dfrac{1}{\sqrt{|x|} + \sqrt{|y|}} \not\approx \infty$.
   Wegen $\mid \sqrt{|x|} - \sqrt{|y|} \mid = \dfrac{\mid |x| - |y| \mid}{\sqrt{|x|} + \sqrt{|y|}}$ folgt $\mid \sqrt{|x|} - \sqrt{|y|} \mid \approx 0$.

Sie kennen den folgenden Satz, aber vergleichen Sie diesen Beweis mit den klassischen!

Satz 1.53 *Sei $K \subset \mathbb{R}^p$ kompakt und $f : K \to \mathbb{R}^q$ sei stetig. Dann ist f gleichmäßig stetig.*

Beweis: Seien $\vec{\xi}, \vec{\eta} \in {}^*K$ mit $\vec{\xi} \approx \vec{\eta}$. Da K kompakt ist, existiert der Standardteil ${}^\circ\vec{\xi}$ und ist aus K (!). Dann ist $\vec{\xi} \approx {}^\circ\vec{\xi} \approx \vec{\eta}$ und da f stetig in ${}^\circ\vec{\xi}$ ist, ist ${}^*f(\vec{\xi}) \approx f({}^\circ\vec{\xi}) \approx {}^*f(\vec{\eta})$. □

Brücke zur Standard-Mathematik

Satz 1.54 *Sei $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$. Die folgenden Aussagen sind äquivalent:*
a) f ist gleichmäßig stetig gemäß Definition 1.52.
b) Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ mit $\|f(x) - f(y)\| < \varepsilon$ für alle $x, y \in D$ mit $\|x - y\| < \delta$.

Beweis: a) $\Rightarrow$ b) Sei $0 < \varepsilon \in \mathbb{R}$. Der folgende Satz gilt in ${}^*\mathcal{S}$ für ein beliebiges $\delta > 0$, $\delta \approx 0$:

$$\exists(\delta \in {}^*\mathbb{R})(\delta > 0 \wedge \forall(x \in {}^*D)\forall(y \in {}^*D)[\|x - y\| < \delta \Rightarrow \|{}^*f(x) - {}^*f(y)\| < \varepsilon].$$

Nach dem TFP gilt seine ungesternte Version in $\mathcal{S}$ und liefert uns ein gewünschtes δ zu unserem ε.
b) $\Rightarrow$ a) Für $\varepsilon > 0$ aus $\mathbb{R}$ sei

$$B(\varepsilon) = \{(x, y) \in D \times D : \|f(x) - f(y)\| < \varepsilon\}.$$

Nach b) liegt $C(\varepsilon) = \{(x, y) \in D \times D : \|x - y\| < \delta(\varepsilon)\}$ in $B(\varepsilon)$. Wegen des TFP gilt

$$U := \{(x, y) \in {}^*D \times {}^*D : x \approx y\} \subset {}^*C(\varepsilon) \subset {}^*B(\varepsilon).$$

Also ist $U \subset \bigcap_{\mathbb{R}\ni\varepsilon>0} {}^*B(\varepsilon)$, d.h. für $x \approx y$ ist $\|{}^*f(x) - {}^*f(y)\| < \varepsilon$ für jedes Standard $\varepsilon > 0$, also ist $\|{}^*f(x) - {}^*f(y)\| \approx 0$. $\square$

Folgen von Funktionen

Illustrativ für die Nonstandard-Analysis ist auch der Begriff der gleichmäßigen Konvergenz:

Deffnition 1.55 *Sei $(f_n)_n$ eine Folge von Funktionen $f_n : D \subset \mathbb{R}^p \to \mathbb{R}^q$, geschrieben als* **eine** *Funktion $f : \mathbb{N} \times D \to \mathbb{R}^q$, $(n, x) \mapsto f_n(x)$.*
$(f_n)_n$ konvergiert gleichmäßig gegen eine Funktion $g : D \to \mathbb{R}^q$, wenn für alle $N \approx \infty$ und $x \in {}^*D$ stets ${}^*f_N(x) \approx {}^*g(x)$ gilt.

Ist $x \in D$ und $(f_n) \to g$ gleichmäßig, so ist für $N \approx \infty$ ${}^*f_N(x) \approx g(x)$, das heißt die Folge $(f_n(x))$ konvergiert gegen $g(x)$. Damit impliziert die gleichmäßige Konvergenz die Konvergenz punktweise. Die Umkehrung gilt nicht, wie die Funktionenfolge $f_n(x) = x^n$, $g(x) = 0$ auf $[0, 1[$ zeigt. Denn für $0 < \eta \approx 0$ ist $\xi = (1 - \eta)^{1/N} < 1$, aber $f_N(\xi) \approx 1 \not\approx 0$. Es gilt aber natürlich das Theorem von Weierstraß:

Theorem 1.56 *Sei $D \subset \mathbb{R}^p$ beliebig und $(f_n)_n$ sei eine Folge stetiger Funktionen $f_n : D \to \mathbb{R}^q$, die gleichmäßig gegen die Funktion g konvergiert. Dann ist g stetig.*

Beweis: Ist g in x_0 nicht stetig, so gibt es ein $\xi \in {}^*D$, $\xi \approx x_0$ und ein $\delta > 0$ aus $\mathbb{R}$ mit $\|{}^*g(\xi) - g(x_0)\| > \delta$. Sei $A = \{n \in \mathbb{N} : \forall(x \in D)\|f_n(x) - g(x)\| < \frac{\delta}{3}\}$. Nach der Definition der gleichmäßigen Stetigkeit und dem TFP enthält die Menge ${}^*A$ ganz ${}^*\mathbb{N} \setminus \mathbb{N}$, also ist A nicht leer. Sei $m \in A$. Dann ist

$$\delta < \|g^*(\xi) - g(x_0)\| \leq \|{}^*g(\xi) - {}^*f_m(\xi)\| + \|{}^*f_m(\xi) - f(x_0)\| + \|f_m(x_0) - g(x_0)\|.$$

Nach dem TFP sind der erste und letzte Term der rechten Seite jeweils $< \frac{\delta}{3}$. Wegen der Stetigkeit von f_m ist der mittlere Term ≈ 0. Die rechte Seite ist somit $< \delta$, ein Widerspruch. □

Genau so einfach erhält man den Satz von Dini.

Theorem 1.57 *Sei $K \subset \mathbb{R}^p$ kompakt und $(f_n)_n$ eine monotone Folge stetiger Funktionen $f_n : K \to \mathbb{R}$, die punktweise gegen die stetige Funktion $g : K \to \mathbb{R}$ konvergiert. Dann konvergiert die Folge gleichmäßig gegen g.*

Beweis: Die Aussage ist äquivalent dazu, daß die monotone Folge stetiger Funktionen (h_n) mit $h_n = f_n - g$ gleichmäßig gegen 0 konvergiert.
OBdA sei (h_n) monoton fallend (sonst betrachte $-h_n$).
Behauptung: ${}^*h_n(\xi) \approx 0$ für alle $N \approx \infty$ und $\xi \in {}^*K$.
Beweis: Angenommen das gilt nicht. Dann gibt es ein $N \approx \infty$, ein $\xi \in {}^*K$ und ein $\delta > 0$ aus $\mathbb{R}$ (!) mit ${}^*h_N(\xi) > \delta$. Da K kompakt ist, existiert $x_0 = {}^\circ\xi$ und liegt in K. Sei $A = \{n \in \mathbb{N} : h_n(x_0) < \delta/2\}$. Nach Voraussetzung enthält ${}^*A$ die Menge ${}^*\mathbb{N} \setminus \mathbb{N}$, also ist A nicht leer. Sei $m \in A$. Aus der Monotonie der Folge und der Stetigkeit von h_m folgt

$$\delta < {}^*h_N(\xi) \leq {}^*h_m(\xi) \approx h_m(x_0) < \delta/2,$$

ein Widerspruch. □

1.8 Differentialrechnung

Überblick: Wir geben auch hier Definitionen mit Hilfe der Nonstandard-Analysis und beweisen nur einige wenige klassische Sätze.

1.8.1 Differentialrechnung in $\mathbb{R}$

Sei im folgenden J stets ein Intervall, dessen Endpunkte dazu gehören können aber nicht müssen.

Deffnition 1.58 *a)* $f : J \to \mathbb{R}$ *heißt in* $x_0 \in J$ *differenzierbar, wenn es eine reelle Zahl* $d =: f'(x_0)$ *gibt mit* $\dfrac{f(x) - f(x_0)}{x - x_0} \approx d$ *für* $0 \neq x - x_0 \approx 0$. $f'(x_0)$ *heißt Ableitung von* f *in* x_0.
b) f *heißt gleichmäßig differenzierbar, wenn* f *in jedem Punkt differenzierbar ist und für* $\xi \approx \eta$ *und* $0 < |\sigma| \approx$ *stets*

$$\frac{{}^*f(\xi + \sigma) - {}^*f(\xi)}{\sigma} \approx \frac{{}^*f(\eta + \sigma) - {}^*f(\eta)}{\sigma}$$

gilt.

Die einfachen Eigenschaften der Ableitung lassen sich sofort ablesen

Aufgaben:

1. Ist f in x_0 differenzierbar, so ist f dort stetig.

2. Seien f, g in x_0 differenzierbar. Dann sind auch $\alpha f + \beta g$ und fg in x_0 differenzierbar. Die Ableitung ist linear und es gilt die Leibnizsche Produktregel
$$(fg)'(x_0)g(x_0) + f(x_0)g'(x_0).$$
Tipp zur Produktregel:
$f(x)g(x) - f(x_0)g(x_0) = f(x)(g(x) - g(x_0)) + (f(x) - f(x_0))g(x_0)$ und Aufg. 1.

3. Sei f in x_0 differenzierbar. Ist $f(x) = f(x_0)$ für ein $x \approx x_0$, $x \neq x_0$, so ist $f'(x_0) = 0$

4. Sei $f : J \to J'$ in x_0 und $g : J' \to \mathbb{R}$ in $y_0 = f(x_0)$ differenzierbar. Dann ist $g \circ f$ in x_0 differenzierbar mit $(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$.
Tipp: Ist $f(x) = f(x_0)$ für ein $x \approx x_0$, $x \neq x_0$, so verwende die vorige Aufgabe. Sonst ist
$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \cdot \frac{f(x) - f(x_0)}{x - x_0}$$
Verwende nun, daß f stetig in x_0 ist.

5. Beweisen Sie den Satz von Rolle (und damit den Mittelwertsatz) mit Hilfe der Nonstandard-Analysis.
   Tipp: Sei $a < x_0 < b$, $f : [a,b] \to \mathbb{R}$ sei differenzierbar und habe in x_0 ein Maximum. Betrachte $\dfrac{{}^*f(x_0+\xi) - f(x_0)}{\xi}$ für $0 \approx \xi \neq 0$ und unterscheide die Fälle $\xi < 0$ und $\xi > 0$. Benutze außerdem für den Beweis des Satzes von Rolle Korollar 1.46.

Bevor wir die gleichmäßige Differenzierbarkeit erklären, schlagen wir die Brücke zur klassischen Analysis.

Satz 1.59 *$f : J \to \mathbb{R}$ ist genau dann differenzierbar mit Ableitung $f'(x_0)$, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit $\left| \dfrac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| < \varepsilon$ für alle $x \neq x_0$ mit $\|x - x_0\| < \delta$.*

Der Beweis folgt direkt aus Satz 1.48.

Als Konsequenz aus dem TFP haben wir die "klassische"Differenzierbarkeit auch in Nichtstandard-Punkten:

Korollar 1.60 *Sei $f : J \to \mathbb{R}$ in jedem Punkt differenzierbar und $f' : J \to \mathbb{R}$ die Ableitung. Dann gilt für jedes $\xi \in {}^*J$. Zu jedem $\varepsilon > 0$ aus ${}^*\mathbb{R}$ gibt es ein $\delta > 0$ aus ${}^*\mathbb{R}$ so daß $\left| \dfrac{{}^*f(\eta) - {}^*f(\xi)}{\eta - \xi} - {}^*f'(\xi) \right| < \varepsilon$ für alle $\eta \neq \xi$ mit $|\eta - \xi| < \delta$.*

Via TFP gilt auch der Mittelwertsatz in ${}^*\mathcal{S}$. Denn er lautet:

Für alle $x, y \in J$ mit $x < y$ gibt es ein z mit $x < z < y$, so daß

$$\frac{f(y) - f(x)}{y - x} = f'(z)$$

gilt. Die gesternte Aussage gilt dann in ${}^*\mathcal{S}$.

Die gleichmäßige Differenzierbarkeit wird an folgenden Beispielen klar:

Beispiele:

1. $f(x) = x^2$ ist nicht gleichmäßig stetig (s.o.) aber gleichmäßig differenzierbar, denn zunächst ist f wegen $\dfrac{x^2 - y_0^2}{x - y_0} = x + y_0 \approx 2y_0$ (falls $x \approx y_0$) in jedem Punkt $y_0 \in \mathbb{R}$ differenzierbar. Es ist für $\xi \approx \eta$ (auch unendlich groß!)

$$\frac{(\xi+\sigma)^2 - \xi^2}{\sigma} = 2\xi + \sigma^2 \approx 2\eta + \sigma^2 = \frac{(\eta+\sigma)^2 - \eta^2}{\sigma},$$

   also ist f gleichmäßig differenzierbar

2. Sei $f(x) = \begin{cases} x^2 \sin\frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$

   Für $x_0 \neq 0$ ist nach den bekannten Regeln $f'(x_0) = 2x_0 \sin(\frac{1}{x_0}) - \cos(\frac{1}{x_0})$. Für $x_0 = 0$, und $\xi \approx 0$ ist $\frac{{}^*f(\xi) - f(0)}{\xi} = \xi \sin(\frac{1}{\xi}) \approx 0$ weil der Sinus beschränkt ist, also $f'(0) = 0$. Wähle nun $\xi = \frac{1}{2\pi N}$ $(N \in {}^*\mathbb{N} \setminus \mathbb{N}$ fest) und $\sigma = \xi^3$. Dann ist

$$\frac{1}{\xi + \vartheta\sigma} = \frac{1}{\xi} - \frac{\vartheta\xi}{1 + \vartheta\xi^2} \approx \frac{1}{\xi} = 2\pi N$$

   für $0 \leq \vartheta \leq 1$.
   Mit dem Transferprinzip, angewendet auf den Mittelwertsatz, erhalten wir

$$\frac{{}^*f(\xi + \sigma) - {}^*f(\xi)}{\sigma} = {}^*f'(\xi + \vartheta\sigma) \approx 1 \neq 0,$$

   weil der cosinus gleichmäßig stetig und $2(\xi + \vartheta\sigma)\sin(\frac{1}{\xi + \vartheta\xi}) \approx 0$ gilt. In der Monade $\mu(0)$ kommen also Steigungen vor, die $\not\approx f'(0)$ sind.

3. Sei das Intervall J kompakt. In diesem Fall ist f genau dann gleichmäßig differenzierbar, wenn f differenzierbar ist und für alle $0 < \sigma \approx 0$ und alle $\xi \in {}^*J$ stets

$$\frac{{}^*f(\xi + \sigma) - {}^*f(\xi)}{\sigma} \approx f'({}^\circ\xi)$$

   ist. Dies ergibt sich aus Satz 1.61 unten, weil eine stetige Funktion auf einer kompakten Menge gleichmäßig stetig ist.

Die Anwendung des Mittelwertsatzes auf ${}^*f$ liefert:

Satz 1.61 *Sei $f : J \to \mathbb{R}$ differenzierbar. Dann sind die folgenden Aussagen äquivalent:*
a) f ist gleichmäßig differenzierbar.
b) f' ist gleichmäßig stetig.

Beweis: a) $\Rightarrow$ b) Sei $\xi, \eta \in {}^*J$, $\xi \approx \eta$. Nach Korollar 1.60 gibt es zu $0 < \varepsilon \approx 0$ ein $\delta(\xi) > 0$ mit

$$\left| {}^*f'(\xi) - \frac{{}^*f(\xi + \sigma) - {}^*f(\xi)}{\sigma} \right| < \varepsilon$$

für alle $0 < |\sigma| < \delta(\xi)$ und ein $\delta(\eta) > 0$ mit

$$\left| {}^*f'(\eta) - \frac{{}^*f(\eta + \tau) - {}^*f(\eta)}{\tau} \right| < \varepsilon$$

für alle $0 < \tau < \delta(\eta)$.
Für $0 < |\sigma| < \min(\delta(\xi), \delta(\eta))$ ergibt sich

$$\begin{aligned}
| {}^*f'(\xi) - {}^*f'(\eta) | &\leq | {}^*f'(\xi) - \frac{{}^*f(\xi+\sigma) - {}^*f(\xi)}{\sigma} | + | \frac{{}^*f(\xi+\sigma) - {}^*f(\xi)}{\sigma} - \frac{{}^*f(\eta+\sigma) - {}^*f(\eta)}{\sigma} | \\
&\quad + | \frac{{}^*f(\eta+\sigma) - {}^*f(\eta)}{\sigma} - {}^*f'(\eta) | \\
&\leq 2\varepsilon + | \frac{{}^*f(\xi+\sigma) - {}^*f(\xi)}{\sigma} - \frac{{}^*f(\eta+\sigma) - {}^*f(\eta)}{\sigma} | \\
&\approx 0
\end{aligned}$$

wegen der gleichmäßigen Differenzierbarkeit. Also ist f' gleichmäßig stetig.
b) $\Rightarrow$ a) Es ist für $0 < |\sigma| \approx 0$

$$\frac{{}^*f(\xi+\sigma) - {}^*f(\xi)}{\sigma} = {}^*f'(\xi + \vartheta_1\sigma) \approx {}^*f'(\xi)$$

und

$$\frac{{}^*f(\eta+\sigma) - {}^*f(\eta)}{\sigma} = {}^*f'(\eta + \vartheta_2\sigma) \approx {}^*f'(\eta),$$

also folgt die Behauptung wegen ${}^*f'(\xi) \approx {}^*f'(\eta)$. □

Bemerkung: Obwohl der Satz eine überzeugende Begründung der (gleichmäßigen) Stetigkeit der Ableitung enthält, ist er in der Standard-Analysis (nahezu) unbekannt. Das liegt daran, daß der Begriff der gleichmäßigen Differenzierbarkeit vom Standpunkt der Nonstandard-Analysis fast trivial ist, seine (äquivalente) Standard-Formulierung jedoch sehr mühevoll. Sie lautet für kompakte Intervalle J (ohne Beweis):

Sei $x \in J$ beliebig. Dann existiert für alle $\varepsilon > 0$ ein $\delta > 0$ so daß für alle ξ mit $|\xi - x| < \delta$ und für alle h mit $0 < |h| < \delta$ stets

$$|\frac{f(\xi+h) - f(\xi)}{h} - f'(x)| < \varepsilon$$

ist.

1.9 Mehrdimensionale Differentialrechnung

Überblick: Wir greifen nur ganz wenige Definitionen und Sätze auf, um die Leistungsfähigkeit der Nonstandard-Analysis zu demonstrieren.

Im folgenden sei U stets eine nicht leere offene Menge des $\mathbb{R}^p$.

Aufgabe: Sei $D \subset \mathbb{R}^p$. Zeigen Sie bitte: D ist genau dann offen, wenn mit $x \in D$ die ganze Monade $\mu(x) = \{\xi \in {}^*\mathbb{R}^p : \xi \approx x\}$ in ${}^*D$ enthalten ist.

Deffnition 1.62 *Sei $f : U \to \mathbb{R}^q$ eine Funktion*
a) f heißt in $x \in U$ **partiell nach x_j differenzierbar***, wenn es ein $a \in \mathbb{R}^q$ gibt, sodaß für alle $0 < t \approx 0$ aus $\mathbb{R}$ $\dfrac{{}^*f(x + te_j) - f(x)}{t} \approx a \in \mathbb{R}^q$ ist. Hier ist $e_j = (\delta_{jk})$ der $j \cdot te$ kanonische Basisvektor. a heißt partielle Ableitung $\dfrac{\partial f}{\partial x_j}(x)$ an der Stelle x.*
b) f heißt in x **total differenzierbar***, wenn es eine lineare Abbildung A von $\mathbb{R}^p$ in $\mathbb{R}^q$ gibt, mit $\dfrac{{}^*f(x + h) - f(x)}{\|h\|} \approx {}^*A \dfrac{h}{\|h\|}$ für alle $h \in \mathbb{R}^p$ mit $0 \neq h \approx 0$. A heißt dann* **totale Ableitung** *$f'(x)$ an der Stelle x.*

Sie können sofort die folgenden Aufgaben lösen:

Aufgaben:

1. Sei $f(x, y) = \begin{cases} \dfrac{2xy}{x^2 + y^2} & (x, y) \neq 0 \\ 0 & (x, y) = 0 \end{cases}$ Dann ist f in 0 nach x und y partiell differenzierbar; aber f ist nicht total differenzierbar.
   Tipp: Es ist $\dfrac{\partial f}{\partial x}(0) = 0 = \dfrac{\partial f}{\partial y}$. Wähle $h = te_1$ mit $0 \neq t \approx 0$

2. Sei f in x total differenzierbar. Dann ist f nach x_j partiell differenzierbar für jedes j und die totale Ableitung $f'(x)$ hat bezüglich der kanonischen Basis die Matrix $f'(x) = (\dfrac{\partial f(x)}{\partial x_1}, \cdots, \dfrac{\partial f(x)}{\partial x_p})$.
   Tipp: Wähle $h = t \cdot e_j$.

3. Zeigen Sie bitte: Ist f in x total differenzierbar, so ist f dort stetig. Reicht die partielle Differenzierbarkeit nach allen x_j hierfür aus?

Brücke zur Standard-Analysis

Wir zeigen, daß der Begriff der totalen Ableitung mit dem klassischen übereinstimmt. Genau dasselbe gilt natürlich für die partiellen Ableitungen.

Satz 1.63 *Sei $f : U \to \mathbb{R}^q$ eine Funktion und $x \in U$. Dann sind die folgenden Aussagen äquivalent:*
a) f ist in x total differenzierbar.
b) Es gibt eine lineare Abbildung A und eine Funktion $R : U \to \mathbb{R}^q$, die in x stetig ist und $R(x) = 0$ erfüllt, so daß für alle $y \in U$ gilt

$$f(y) = f(x) + A(y - x) + \|y - x\| R(x).$$

Beweis: a) $\Rightarrow$ b) Setze $R(y) = \begin{cases} 0 & y = x \\ \dfrac{f(y) - f(x)}{\|y - x\|} - f'(x)\dfrac{(y - x)}{\|y - x\|} & y \neq x \end{cases}$ Dann gilt die Gleichung unter b). Ist $y \approx x$, so ist $^*R(y) \approx 0$ nach Voraussetzung a), also ist R in x stetig.
b) $\Rightarrow$ a) Ist $0 \neq h \approx 0$, so ist $^*R(x + h) \approx 0$ nach Voraussetzung, also ist $\dfrac{^*f(x + h) - f(x)}{\|h\|} \approx {}^*A\dfrac{h}{\|h\|}$. □

Deffnition 1.64 *$f : U \to \mathbb{R}^q$ heißt gleichmäßig differenzierbar, wenn f überall differenzierbar ist und für alle $x, y \in {}^*U$ mit $0 \neq x - y \approx 0$ und alle $0 \neq h \approx 0$ gilt*

$$\frac{^*f(y + h) - {}^*f(y)}{\|h\|} \approx \frac{^*f(x + h) - {}^*f(x)}{\|h\|}$$

Bevor wir einen zu Satz 1.61 analogen Satz beweisen, brauchen wir eine Charakterisierung der Operatornorm :

Lemma 1.65 *Sei B eine lineare Abbildung von $\mathbb{R}^p$ in $\mathbb{R}^q$. Dann gibt es ein $h \in \mathbb{R}^p$ der Norm 1 mit*

$$\|B\|_o p := \sup\{\|B(x)\| : \|x\| = 1\} = B(h).$$

Beweis: Die Abbildung $x \to B(x) \to \|B(x)\|$ ist stetig und $S := \{x : \|x\| = 1\}$ ist kompakt. □

Ganz ähnlich wie Satz 1.61 erhält man den folgenden Satz:

Satz 1.66 *Sei $U \subset \mathbb{R}^p$ offen und konvex und $f : U \to \mathbb{R}^q$ eine differenzierbare Abbildung. Dann sind die folgenden Aussagen äquivalent:*
a) f ist gleichmäßig differenzierbar.
b) f' ist gleichmäßig stetig.

Beweis: a) $\Rightarrow$ b) Sei $x, y \in {}^*U$, $0 \neq x - y \approx 0$. Nach dem Transferprinzip ist ${}^*f$ in x und y differenzierbar. Das bedeutet: zu $\varepsilon > 0$, $\varepsilon \approx 0$ gibt es $\delta(x) > 0$, $\delta(y) > 0$ mit $\|{}^*f'(x)\,\frac{h}{\|h\|} - \frac{{}^*f(x+h) - {}^*f(x)}{\|h\|}\| < \varepsilon$ für alle $0 \neq \|h\| < \delta(x)$ und entsprechend für y mit $\delta(y)$.
Für $\delta = \min(\delta(x), \delta(y), \varepsilon)$ und $0 < \|h\| < \delta$ ist dann

$$
\begin{aligned}
\|({}^*f'(x) - {}^*f'(y))\frac{h}{\|h\|}\| &\leq \|{}^*f'(x)\frac{h}{\|h\|} - \frac{{}^*f(x+h) - {}^*f(x)}{\|h\|}\| \\
&+ \|\frac{{}^*f(x+h) - {}^*f(x)}{\|h\|} - \frac{{}^*f(y+h) - {}^*f(y)}{\|h\|}\| \\
&+ \|\frac{{}^*f(y+h) - {}^*f(y)}{\|h\|} - {}^*f'(y)\frac{h}{\|h\|}\| \\
&\leq 2\varepsilon + \|\frac{{}^*f(x+h) - {}^*f(x)}{\|h\|} - \frac{{}^*f(y+h) - {}^*f(y)}{\|h\|}\| \\
&\approx 0
\end{aligned}
$$

nach Voraussetzung. Angenommen die Abbildung $U \ni x \mapsto f'(x) \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q)$ sei nicht gleichmäßig stetig. Dann gibt es nach Satz 1.54 ein $\varepsilon > 0$ aus $\mathbb{R}$ (!) und zu jedem $1/n$ wegen des obigen Lemmas ein Tripel (x_n, y_n, h_n) mit $\|x_n - y_n\| < 1/n$, $\|h_n\| = 1$ und $\|(f'(x_n) - f'(y_n))(h_n)\| > \varepsilon$, ein Widerspruch zur obigen Abschätzung (1.1).
b $\Rightarrow$ a) Sei $0 \neq x - y \approx 0$ und $0 \neq h \approx 0$. Wir nehmen zunächst $q_j = 1$, also $f : U \to \mathbb{R}$ an und setzen $u_j = \sum_{k=1}^{j} h_k e_k$, wo $h = (h_1, \cdots, h_p)$ und $e_k = (\delta_{jk})$ der $k \cdot te$ kanonische Einheitsvektor ist. Dann ist nach dem Transferprinzip

$$
\begin{aligned}
{}^*f(y+h) - {}^*f(y) &= \sum_{j=0}^{p-1}({}^*f(y+u_{j+1}) - {}^*f(y+u_j)) \\
&= \sum_{j=0} {}^*(\frac{\partial f}{\partial x_{j+1}})(y + u_j + \vartheta_{j+1}h_{j+1}) \cdot h_{j+1}, \text{ wo } \quad 0 \leq \vartheta_{j+1} < 1.
\end{aligned}
$$

Also erhalten wir

$$\frac{{}^*f(y+h) - {}^*f(y)}{\|h\|} = \sum_{j=1}^{p} {}^*(\frac{\partial f}{\partial x_j})(y + u_{j-1} + \vartheta_j h_j)\frac{h_j}{\|h\|}.$$

Analog ergibt sich

$$\frac{{}^*f(x+h) - {}^*f(x)}{\|h\|} = \sum_{j=1}^{p} {}^*(\frac{\partial f}{\partial x_j})(x + u_{j-1} + \vartheta'_j h_j)\frac{h_j}{\|h\|}.$$

Wegen der gleichmäßigen Stetigkeit von f' sind die beiden rechten Seiten infinitesimal benachbart, also auch die linken.
Den allgemeinen Fall erhält man, indem man den eben behandelten Fall auf die einzelnen Koordinatenfunktionen anwendet ($f = \begin{pmatrix} f_1 \\ \vdots \\ f_q \end{pmatrix}$). $\square$

Wir wenden uns nun noch den 2. Ableitungen zu.

Satz 1.67 (H.A. Schwarz) *Sei $f : U \to \mathbb{R}^q$ zweimal stetig partiell differenzierbar. Dann gilt*

$$\frac{\partial^2 f}{\partial x_i \partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_i}.$$

Beweis: Wie üblich reduziert man den Beweis auf eine Funktion $f : U \subset \mathbb{R}^2 \to \mathbb{R}$, und den Spezialfall $x = 0$, man zeigt also $\frac{\partial^2 f}{\partial y \partial x}(0) = \frac{\partial^2 f}{\partial x \partial y}(0)$.
Wir schreiben $\frac{\partial f}{\partial x} = f_x$, $\frac{\partial f}{\partial y} = f_y$, $\frac{\partial^2 f}{\partial y \partial x} = f_{xy}$, $\frac{\partial^2 f}{\partial x \partial y} = f_{yx}$. Sei $0 \neq v \in \mathbb{R}$.
Dann ist für $0 \neq \tau \approx 0$

$$\frac{f_x(0,v) - f_x(0,0)}{v} \approx \frac{1}{v}\left(\frac{{}^*f(\tau,v) - f(0,v)}{\tau} - \frac{{}^*f(\tau,0) - f(0,0)}{\tau}\right).$$

Sei $g(y) = f(\tau, y) - f(0, y)$. Dann ist also

$$\begin{aligned}
\frac{f_x(0,v) - f_x(0,0)}{v} &\approx \frac{1}{\tau}\left(\frac{g(v) - g(0)}{v}\right) \\
&=_{TFP} \frac{1}{\tau} g'(\vartheta v) \\
&= \frac{1}{\tau}(f_y(\tau, \vartheta v) - f_y(0, \vartheta v)) \\
&\approx f_{yx}(0, \vartheta v),
\end{aligned}$$

wobei die letzte Beziehung aus Satz 1.61 folgt, da f_{yx} stetig, also f_y gleichmäßig differenzierbar nach x ist. Es ist $\vartheta \in {}^*[0,1]$ also ist wegen der Stetigkeit von f_{yx} einfach $f_{yx}(0, \vartheta v) \approx f_{yx}(0, {}^\circ\vartheta v)$ und das ist eine Standardzahl. Damit ergibt sich zunächst

$$\frac{f_x(0,v) - f_x(0,0)}{v} = f_{yx}(0, {}^\circ\vartheta v),$$

also

$$\forall (v \in \mathbb{R})(v \neq 0 \land (0,v) \in U \Rightarrow \exists \sigma \in [0,1](\frac{f_x(0,v) - f_x(0,0)}{v} = f_{yx}(0, \sigma v))).$$

Nach dem Transferprinzip folgt für $0 \neq v \approx 0$

$$f_{xy}(0,0) \approx \frac{f_x(0,v) - f_x(0,0)}{v} = f_{yx}(0, \sigma v) \approx f_{yx}(0,0).$$

□

Der Umkehrsatz in $\mathbb{R}^p$

Wir liefern einen Nonstandard-Analysis-Beweis des folgenden Theorems:

Theorem 1.68 (Satz über die Umkehrfunktion) *Sei U offen in $\mathbb{R}^p$ und $f : U \to \mathbb{R}^p$ stetig differenzierbar. Sei $x_0 \in U$ mit $\det f'(x_0) \neq 0$. Dann gibt es eine offene Menge $U_0 \subset U$ mit $x_0 \in U_0$, so daß $f(U_0) = V$ offen ist, $f\,U_0$ bijektiv auf V abbildet, die Umkehrfunktion $g : V \to U_0$ stetig differenzierbar ist und $g'(f(x)) = f'^{-1}(x)$ gilt.*

Wie üblich führt man den allgemeinen Fall durch die Transformation $g(x) = f'^{-1}(x_0)(f(x_0 + x) - f(x_0))$ auf den Fall $x_0 = 0$, $g(0) = 0$, $g'(0) = I$ zurück. Wir schreiben wieder f statt g.
Wir benötigen zunächst das folgende Lemma mit Standard-Beweis (Banachscher Fixpunktsatz) und erinnern dazu an den Schrankensatz:

Satz 1.69 (Schrankensatz) *Sei $f : U \to \mathbb{R}^q$ stetig differenzierbar und $K \subset U$ eine kompakte konvexe Menge. Sei $L = \sup\{\|f'(x)\|_{op} : x \in K\}$, wo $\|\cdot\|_{op}$ die Operatornorm in $L(\mathbb{R}^p, \mathbb{R}^q)$ bezeichnet. Dann gilt für $x, y \in K$ stets $\|f(x) - f(y)\| \leq L\|x - y\|$.*

Lemma 1.70 *Sei $0 :\in U$ und $f : U \to \mathbb{R}^p$ stetig differenzierbar mit $f(0) = 0$, $f'(0) = I$ und für ein $r > 0$ sei $\overline{B(0,r)} = \{x : \|x\| \leq r\} \subset U$ und $\sup\{\|I - f'(x)\|_{0p} : x \in \overline{B(0,r)}\} < q < 1$. Dann ist $B(0, (1-q)r) \subset f(B(0,r))$.*

Beweis: **(Standard(** Sei $y \in B(0,(1-q)r)$, also $\|y\| < (1-q)r$. Sei $\Phi : \overline{B(0,r)} \to \mathbb{R}^p$ definiert durch $\Phi(x) = y + x - f(x)$. Dann gilt $\Phi(0) = 0$, $\Phi'(x) = I - f(x)$, also nach dem Schrankensatz $\|\Phi(x)\| \leq \|y\| + \|\Phi(x) - \Phi(0)\| < (1-q)r + qr$, d.h. Φ bildet $\overline{B(0,r)}$ in $B(0,r)$ ab. Ebenfalls nach dem Schrankensatz folgt $\|\Phi(u) - \Phi(v)\| \leq q\|u - v\|$. Also hat Φ einen Fixpunkt x_0 und dieser liegt in $B(0,r)$. Für ihn gilt $f(x_0) = y$. □

Lemma 1.71 *Sei $0 \in U$ und $f : U \to \mathbb{R}^p$ sei stetig differenzierbar mit $f(0) = 0, f'(0) = I$. Dann gilt: Für $0 < q < 1$ aus $\mathbb{R}$ und $x, y \in \mu(0)$ ist $\|{}^*f(x) - {}^*f(y)\| \geq q\|x - y\|$.*

Beweis: Da $0 \in U$ offen ist, gibt es ein $\delta > 0$ mit $\overline{B(0,\delta)} \subset U$. $\overline{B(0,\delta)}$ ist kompakt, also ist f' dort gleichmäßig stetig; damit ist f auf $B(0,\delta)$ gleichmäßig differenzierbar nach Satz 1.66. Seien $0 \neq x - y \approx 0$ und $x, y \in \mu(0)$. Sei $h = y - x$. Dann ist also wegen $f'(0) = I$

$$\frac{f(y) - f(x)}{\|h\|} = \frac{f(x+h) - f(x)}{\|h\|} \approx \frac{f(h) - f(0)}{\|h\|} \approx \frac{h}{\|h\|} \approx \frac{y - x}{\|y - x\|}.$$

Daraus folgt $\dfrac{\|f(y) - f(x)\|}{\|y - x\|} \approx 1 > q$, also die Behauptung. □

Beweis des Umkehrsatzes:
Fassen wir beide vorausgegangenen Lemmata zusammen, so erhalten wir für ein festes $r \approx 0$ und $q = 1/2$
(1) $\forall(x \in {}^*\mathbb{R}^p)(\|x\| \leq r \Rightarrow x \in {}^*U)$
(2) $\forall(y \in {}^*\mathbb{R}^p)(\|y\| < r/2 \Rightarrow \exists(x \in {}^*U)(\|x\| < r \,\text{und}\, f(x) = y)$
(3) $\forall(x, y \in {}^*\mathbb{R}^p)(\|x\|\,,\ \|y\| \leq r \Rightarrow \|{}^*f(x) - {}^*f(y)\| \geq \|x - y\|/2)$,
wobei (2) aus Lemma 1.70 nach dem Transfer-Prinzip folgt. Damit gilt

$$\exists(r \in {}^*\mathbb{R})(r > 0 \wedge (1) \wedge (2) \wedge (3))$$

in ${}^*\mathcal{S}$, also gilt der ungesternte Satz in $\mathbb{R}^p$. Wir erhalten: Es gibt ein $r > 0$, so daß (1) - (3) in $\mathbb{R}^p$ gilt. Das bedeutet *im Einzelnen:*
(1') $\overline{B(0, r)} \subset U$
(2') $B(0, r/2) \subset f(B(0, r))$
(3') $\|f(x) - f(y)\| \geq \|x - y\|/2$ für $x, y \in \overline{B(0, r)}$.
Aus (3') ergibt sich, daß $f\,|_{B(0,r)}$ injektiv ist. Sei $U_0 = (f\,|_{B(0,r)})^{-1}(B(0, r/2))$. Dann ist $0 \in U_0 \subset \overline{B(0, r)}$ und $f : U_0 \to B(0, r/2)$ ist bijektiv mit wegen (3') stetiger Inverser g (setze $x = g(u)$, $y = g(v)$, dann ergibt sich $\|g(u) - g(v)\| \leq 2\|u - v\|$).

Wir müssen nur noch die Differenzierbarkeit von g und die Formel für g' zeigen:

Es ist für $v = f(x) \in B(0, r/2)$ und $0 \neq h \approx 0$ das Element $v + h$ aus ${}^*B(0, r/2)$, also gibt es $y \in {}^*U_0$ mit ${}^*f(y) = v + h$. Wegen $U_0 \subset \overline{B(0, r)}$ existiert ${}^\circ y \in \overline{B(0, r)}$. Wegen der Stetigkeit von f ist $f({}^\circ y) \approx {}^*f(y) = v + h \approx v = f(x)$, also folgt wegen (3') ${}^\circ y = x$, $y \approx x \neq y$.
Damit erhält man

$$\frac{{}^*g(v + h) - g(v)}{\|h\|} = \frac{y - x}{\|{}^*f(y) - f(x)\|}.$$

Nun ist

$$\frac{{}^*f(y) - f(x)}{\|y - x\|} \approx f'(x)\frac{y - x}{\|y - x\|},$$

also

$$\frac{y - x}{\|y - x\|} \approx f'(x)^{-1}\frac{{}^*f(y) - f(x)}{\|y - x\|}$$

und damit

$$\begin{aligned}
\frac{{}^*g(v + h) - g(v)}{\|h\|} &= \frac{y - x}{\|{}^*f(y) - f(x)\|} \\
&= \frac{y - x}{\|y - x\|} \cdot \frac{\|y - x\|}{\|{}^*f(y) - f(x)\|} \\
&\approx f'(x)^{-1}\frac{{}^*f(y) - f(x)}{\|y - x\|} \cdot \frac{\|y - x\|}{\|{}^*f(y) - f(x)\|},
\end{aligned}$$

wobei " $\approx$" gilt, weil $\dfrac{\|y-x\|}{\|^*f(y)-f(x)\|} \leq 2$ nach (3') (und TFP) ist.
Insgesamt ist $\dfrac{^*g(v+h)-g(v)}{\|h\|} \approx f'(x)^{-1}\dfrac{h}{\|h\|}$,
woraus $g'(v) = g'(f(x) = f'(x)^{-1}$ folgt.

1.10 Integralrechnung und gewöhnliche Differentialgleichungen

Überblick: Wir geben eine einfache NSA-Definition der Integrierbarkeit und beweisen damit den Existenzsatz von Peano für gewöhnliche Differentialgleichungen.

Im Folgenden sei $[a, b]$ ein kompaktes Intervall. Für $n \in \mathbb{N}$ sei

$$a_{k,n} = a + \frac{(b-a)}{2^n} \cdot k \qquad (k = 0, \cdots, 2^n) \tag{1.1}$$

die übliche Einteilung.

Deffnition 1.72 *Sei* $f : [a, b] \to \mathbb{R}^p$ *und* $J(n, f) = \sum_{k=0}^{2^n-1} f(a_{k,n}) \frac{(b-a)}{2^n}$. f *heißt* **integrierbar über** $[a, b]$, *wenn für alle* $N, N' \approx \infty$ *stets* $J(N, f) \approx J(N', f)$ *gilt, d.h. wenn* $(J(n, f))$ *eine Cauchyfolge ist. Der Grenzwert* $^\circ J(N, f)$ *heißt* **Integral** $\int_a^b f(t)dt$.

Die üblichen Rechenregeln für Integrale ergeben sich sofort aus dem Transferprinzip und den Eigenschaften von "°".

Beispiele:

1. Sei $a \leq u < v \leq b$ und $f = 1_{[u,v]}$. Dann ist $\int_a^b f(t)dt = v - u$, auch nach unserer Definition. Denn ist k die eindeutig bestimmte natürliche Zahl mit $\frac{k(b-a)}{2^n} \leq u < \frac{(k+1)(b-a)}{2^n}$ und analog ℓ mit $\frac{\ell(b-a)}{2^n} \leq v < \frac{(\ell+1)(b-a)}{2^n}$, so ist

$$\frac{(\ell-(k+1)(b-a))}{2^n} < J(n, f) < \frac{(\ell+1-k)}{2^n}(b-a)$$

und wegen $\frac{(b-a)}{2^N} \approx 0$ folgt leicht die Behauptung. Damit stimmt für Treppenfunktionen das Integral mit dem üblichen überein.

2. Sei (f_n) eine Folge von integrierbaren Funktionen, die gleichmäßig gegen die Funktion $g : [a, b] \to \mathbb{R}^p$ konvergiert. Dann ist für $N \approx \infty$
$\|J(N, g) - J(N, f_n)\| \leq \|f_n - g\|_\infty (b-a)$, also

$$\begin{aligned} \|J(N, g) - J(N', g)\| &\leq \|J(N, g) - J(N, f_n)\| + \|J(N, f_n) - J(N', f_n)\| \\ &+ \|J(N', f_n) - J(N', g)\| \\ &\lesssim_\approx 2\|f_n - g\|_\infty (b-a) \end{aligned}$$

Wählt man zu $\varepsilon > 0$ aus $\mathbb{R}$ $\quad n$ so daß $\|f_n - g\|_\infty < \varepsilon/3$ ist, so sieht man $\|J(N,g) - J(N',g)\| < \varepsilon$. Da $\varepsilon > 0$ aus $\mathbb{R}$ beliebig war, folgt die Integrierbarkeit. Außerdem erhält man

$$\int_a^b \|g(t) - f(t)_n \mid dt \leq \|g - f_n\|_\infty (b-a), \text{ also } \int_a^b g dt = \lim_n \int_a^b f_n dt.$$

3. Jede stetige Funktion ist integrierbar, da sie auf $[a,b]$ gleichmäßig stetig und damit gleichmäßiger Grenzwert einer Folge von Treppenfunktionen ist.

4. Sei f integrierbar über $[a,b]$ und $a < c < b$. Dann ist f integrierbar über $[a,c]$ und

$$\int_a^c f(x)dx = \int_a^b 1_{[0,c]}(x) f(x) dx.$$

Denn dies gilt zunächst für $t = \dfrac{k(b-a)}{2^n}$ $(n \in \mathbb{N})$, weil solch ein t Einteilungspunkt der Zerlegung für N ist, wie leicht zu sehen ist. Ein Transferargument liefert in diesem Fall die Behauptung. Sei t jetzt beliebig. Es gibt nach dem TFP genau ein k mit $a_{k,N} \leq t < a_{k+1,N}$. Damit ist der Unterschied der einzelnen Nonstandard-Summen infinitesimal und die Behauptung folgt.
Bemerkung: Mit der vollen Nonstandard-Analysis (s. Kap. 2) ist die Definition des Riemannintegrals noch einfacher und die obige Intervall-Additivität leichter nachzuweisen.

Mit diesem Integralbegriff, der ja der übliche ist (die Brücke zur Standard-Analysis ist nicht schwer), können wir schon den Satz von Peano beweisen.

Theorem 1.73 *Sei $f : [a,b] \times \mathbb{R}^p \to \mathbb{R}^p$ eine stetige und beschränkte Funktion. Dann hat das Anfangswertproblem*

$$y'(t) \;=\; f(t, y(t))\,,\; y(a) \;=\; y_0$$

mindestens eine Lösung.

Beweis: (I) Nach dem Hauptsatz der Differentialgleichung können wir statt des Anfangswertproblems die Integralgleichung

$$y(t) \;=\; y_0 + \int_a^t f(s, y(s)) ds$$

betrachten. Sei $M \geq \|f(t,x)\|$ für alle x
(II)Wir benutzen Eulers Idee der Lösung der Differentialgleichung (Eulersches Lösungsverfahren) und benutzen dazu wieder die Einteilung (1.1) für $n \in \mathbb{N}$ und setzen

$$S_n(a_{0,n}) = y_0,$$

$$S_n(a_{k+1,n}) = S_n(a_{k,n}) + (a_{k+1,n} - a_{k,n})\, f(a_{k,n}, S_n(a_{k,n}))$$

und schließlich für $a_{k,n} \leq t < a_{k+1,n}$

$$S_n(t) = S_n(a_{k,n}) + (t - a_{k,n})\, f(a_{k,n}, S_n(a_{k,n})).$$

Dann ist für $a_{k,n} \leq t < a_{k+1,n}$

$$S_n(t) = y_0 + \sum_{l=0}^{k-1}(a_{l+1,n} - a_{l,n})\, f(a_{l,n}, S_n(a_{l,n})) + (t - a_{k,n}) f(a_{k,n}, S_n(a_{k,n})). \tag{1.2}$$

Damit ist $S_n : [a,b] \to \mathbb{R}^p$ stetig und stückweise affin.
(III) Sei $a \leq s < t \leq b$. Es sei $a_{k,n} \leq s < a_{k+1,n}$ und $a_{\ell,n} \leq t < a_{(\ell+1),n}$.
Dann erhält man aus Gleichung (1.2)

$$\begin{aligned} \|S_n(t) - S_n(s)\| &\leq \sum_{r=k}^{\ell+1} \|S_n(a_{\ell,n})\| \leq M \cdot (a_{\ell+1,n} - a_{k,n}) \\ &\leq M(t-s) + \frac{2(b-a)}{2^n} \end{aligned}$$

Außerdem ist $\|S_n(t)\| \leq \sum_{r=0}^{\ell+1} \|S_n(a_{\ell,n})\| \leq M(b-a)$.
Bisher haben wir nur die klassischen Abschätzungen für das Euler-Verfahren vorgenommen.
(IV) Sei $N \approx \infty$ und $t \approx s$. Nach dem Transferprinzip ist $\|S_N(t) - S_N(s)\| \leq M(t-s) + \frac{2(b-a)}{2^N} \approx 0$. außerdem ist $\|S_N(t)\|$ endlich. Daher wird durch

$$[a,b] \ni t \to g(t) = {}^\circ S_N(t)$$

eindeutig eine Funktion g von $[a,b]$ in $\mathbb{R}^p$ erklärt.
Behauptung: g ist stetig.
Beweis: es ist

$$\|g(t) - g(s)\| \approx \|S_N(t) - S_N(s)\| \leq M \mid t - s \mid + \frac{(b-a)}{2^N}$$

Da links eine Standardzahl steht, folgt durch die Anwendung von "°" auf diese Ungleichung sofort

$$\|g(t) - g(s)\| \leq M \mid t - s \mid .$$

Ferner gilt für alle $t \in {}^*[a,b]$ stets ${}^*g(t) \approx S_N(t)$. Denn

$$\|{}^*g(t) - S_N(t)\| \leq \|{}^*g(t) - g({}^\circ t)\| + \|g({}^\circ t) - S_N({}^\circ t)\| + \|S_N({}^\circ t) - S_N(t)\| \approx 0.$$

Außerdem liegen die Werte von g nach (III) in $\overline{B(0, M(b-a))}$ und f ist auf der kompakten Menge $[a,b] \times \overline{B(0, M(b-a))}$ gleichmäßig stetig. Daraus folgt

$${}^*f(u, {}^*g(u)) \approx {}^*f(u, S_N(u))$$

für alle $u \in {}^*[a,b]$. Schließlich gilt der folgende Satz ganz offensichtlich in $\mathbb{R}^p$:

$$\forall (\varepsilon > 0)$$

$$
\begin{aligned}
&[[\forall (u \in [a,b]) \qquad \|f(u,g(u)) - f(u,S_n(u))\| < \varepsilon] \\
&\qquad\qquad \Rightarrow \qquad \|\sum_{0}^{2^n-1} \frac{b-a}{2^n}(f(a_{k,n},g(a_{k,n})) - f(a_{k,n},S_n(a_{k,n})))\| \leq \varepsilon(b-a)]. \qquad (1.3)
\end{aligned}
$$

(V) *Behauptung:* g ist Lösung der Integralgleichung.
Beweis: (i) $g(a) = y_0$ ist klar.
(ii) Wir beachten zunächst für $t \in [a,b]$ (reell): es gibt nach dem TFP genau ein k mit $a_{k,N} \leq t < a_{k+1,N}$. Damit gilt

$$
\begin{aligned}
g(t) \quad &\approx \quad S_N(t) = y_0 + \sum_{l=0}^{k} \frac{b-a}{2^N} {}^* f(a_{l,N}, S_N(a_{l,N})) + (t - a_{k,N})^* f(a_{k,N}, S_N(a_{k,N})) \\
&\approx \quad y_0 + \sum_{l=0}^{2^N-1} {}^* 1_{[0,t]}(a_{l,N})^* f(a_{l,N}, S_N(a_{l,N})) \frac{b-a}{2^N}, \qquad (1.4)
\end{aligned}
$$

und damit

$$
\begin{aligned}
\int_a^t f(s,g(s))ds \quad &= \quad \int_a^b 1_{[0,t]}(s) f(s,g(s))ds \\
&\approx \quad \sum_{k=0}^{2^N-1} 1_{[0,t]}(a_{k,N})\, {}^* f(a_{k,N}, g(a_{k,N})) \frac{(b-a)}{2^N} \\
&\underbrace{\approx}_{Formel\,(1.3)} \quad \sum_{k=0}^{2^N-1} 1_{[0,t]}(a_{k,N})\, {}^* f(a_{k,n}, S_N(a_{k,N})) \frac{(b-a)}{2^N} \\
&\underbrace{\approx}_{Formel\,(1.4)} \quad g(t) - y_0
\end{aligned}
$$

Da $\int_a^t f(s,g(s))ds$ und $g(t) - y_0$ Standardzahlen sind, folgt die Behauptung. $\square$

Kapitel 2

Allgemeine Nonstandardanalysis

Überblick: Wir führen die volle Superstruktur $\mathcal{V}(X)$ über einer Menge X ein und definieren einen NS–Monomorphismus von $\mathcal{V}(\mathbb{R})$ in $\mathcal{V}({}^*\mathbb{R})$ ein. Seine Existenz beweisen wir nicht. Statt dessen geben wir interessante Anwendungen.

2.1 Grundlagen der allgemeinen Nonstandardanalysis

Überblick: Wir motivieren erst den Bedarf an einer allgemeinen Nonstandardanalysis, führen dann die volle Superstruktur $\mathcal{V}(X)$ über einer unendlichen Menge X ein und definieren anschließend die Sprache, die sie beschreibt. Schließlich führen wir den Begriff des NS–Monomorphismus ein und geben erste Anwendungen.

2.1.1 Motivation

Bisher konnten wir das Transferprinzip auf Sätze über $\mathbb{R}$ anwenden, bei denen die Quantoren nur *Elemente* in ihrem Bereich hatten. Sätze der Form: "für alle endlichen Mengen ..." oder "es existiert ein Intervall" etc. waren noch nicht behandelt worden. Dabei sind solche Sätze sehr wichtig. Wir haben zum Beispiel beim Integral de facto über die Menge $\{1, \cdots, N\}$ $(N \approx \infty)$ summiert; um das zu tun, mußten wir aber explizit eine Funktion $J(n, f)$ einführen.
Dabei könnte man sich auch vorstellen, daß man die Funktion Σ, die jeder endlichen Menge A aus $\mathbb{R}^p$ die Summe $\Sigma(A) := \sum_{a \in A} a \in \mathbb{R}$ zuordnet, ebenfalls fortsetzen kann auf "hyperendliche" Mengen wie $\{1, \cdots, N\}$.

Eine mögliche Konstruktion hyperendlicher Mengen

Sei $N \approx \infty$ aus ${}^*\mathbb{N}$. Es gibt also eine unbeschränkte Folge $(k_n)_n$ aus $\mathbb{N}$, so daß $N = (k_n)_n^\bullet$ die Äquivalenzklasse ist, in der die Folge liegt. Für jedes n sei $A_n \subset \mathbb{R}$ eine Menge mit k_n Elementen. Wir setzen $\mathcal{A} = (A_n)_n^\bullet = \{(x_n)_n^\bullet : \{n : x_n \in A_n\}$ ist 1-Menge$\}$ $\mathcal{A}$ ist "hyperendlich". Es gibt eine bijektive Abbildung φ von $\{1, \cdots, N\}$ auf $\mathcal{A}$. Sei nämlich φ_n eine bijektive Abbildung von $\{1, \cdots, k_n\}$ auf A_n. Für $L \leq N$, also $L = (\ell_n)^\bullet$ und $\{n : \ell_n \leq k_n\}$ ist 1-Menge setze $\varphi(L) = (\varphi_n(\ell_n))_n^\bullet$. φ ist nicht die Fortsetzung einer Funktion $f : D \subset \mathbb{R} \to \mathbb{R}$ auf ${}^*D$. Trotzdem suggeriert die Konstruktion, daß $\mathcal{A}$ genau N Elemente hat.
Sei $\mathcal{P}_e(\mathbb{R})$ die Menge aller endlichen Teilmengen von $\mathbb{R}$. Dann ist $\Sigma : \mathcal{P}_e(\mathbb{R}) \to \mathbb{R}$, $A \mapsto \Sigma_A = \Sigma_{a \in A} a$ eine Abbildung. Wie sollten wir sie auf unser $\mathcal{A}$ fortsetzen? Wir setzen $\Sigma_{\mathcal{A}} = (\Sigma_{b \in A_n} b)_n^\bullet$. Man kann zeigen, daß diese Zuordnung alle Eigenschaften des Summenzeichens hat.

Beispiele:

1. Sei $N = (n)_n^\bullet$ und $A_n = \{2^{-k} : 0 \leq k \leq n-1\}$. Dann ist $\mathcal{A} = \{2^{-L} : 0 \leq L \leq N-1\}$. Es ergibt sich
$$\sum(\mathcal{A}) = (\sum_{0}^{n-1} 2^{-k})_n^\bullet = (\frac{1-2^{-n}}{1/2})_n^\bullet = \frac{1-2^{-N}}{1/2}.$$

2. Sei jetzt wieder $\mathcal{A}$ eine allgemeine hyperendliche Menge. Die Menge $\mathcal{L} = (B_n)^\bullet$ liegt in $\mathcal{A}$, wenn $\{n : B_n \subset A_n\}$ eine Einsmenge ist. Tatsächlich gilt dann
$$\forall \xi \in {}^*\mathbb{R} (\xi \in \mathcal{L} \Rightarrow \xi \in \mathcal{A}).$$
Wir erhalten im Spezialfall aus Beispiel 1 die Gleichung
$$\sum(\mathcal{A}) = \sum(\mathcal{L}) + \sum(\mathcal{A} \setminus L),$$
denn offensichtlich ist $\mathcal{A} \setminus \mathcal{L} = (A_n \setminus L_n)_n^\bullet$.

Wenn wir so weiter machen, müssen wir uns - genau so wie in Abschnitt 1.2 - alles, was wir haben möchten, in jedem Einzelfall erst einmal beweisen. Statt dessen ist es bequemer, den folgenden Weg zu gehen.

2.1.2 Superstrukturen

Deffnition 2.1 *Sei X eine unendliche Menge. Wir setzen $V_0(X) = X$, und induktiv*
$$V_{n+1}(X) = \mathcal{P}(V_n(X)) \cup V_n(X),$$

wo $\mathcal{P}(Y)$ die Potenzmenge der Menge Y, also die Menge aller Teilmengen von Y ist.

$$V(X) = \cup_{n=0}^{\infty} V_n(X)$$

heißt **volle Superstruktur** *von X.*

Die volle Superstruktur enthält alle Objekte, die uns interessieren als *Elemente*. Hier sind einige Beispiele für $X = \mathbb{R}$.

Beispiele:

1. Alle reellen Zahlen sind Elemente von $\mathbb{R}$, also wegen $\mathbb{R} \subset V_1(\mathbb{R}) \subset V(\mathbb{R})$ gilt $x \in V(\mathbb{R})$ für jede reelle Zahl.

2. Jede Teilmenge, erst recht jede endliche Teilmenge von $\mathbb{R}$ ist Element von $V(\mathbb{R})$.

3. Seien $u, v \in V_n(\mathbb{R})$ für ein n. Man definiert bekanntlich das geordnete Paar (u, v) durch $(u, v) = \{\{u\}, \{u, v\}\}$. Wegen $u, v \in V_n(\mathbb{R})$ ist $\{u\}, \{u, v\} \subset V_n(\mathbb{R})$, also $\{u\}, \{u, v\} \in V_{n+1}(\mathbb{R}) = P(V_n(\mathbb{R})) \cup V_n(\mathbb{R})$. Damit ist $(u, v) = \{\{u\}\{u, v\}\} \subset V_{n+1}(\mathbb{R})$, also $(u, v) \in V_{n+2}(\mathbb{R})$. Damit ist $\mathbb{R}^2 \subset V_2(\mathbb{R})$, also $\mathbb{R}^2 \in V_3(\mathbb{R})$.

4. Sei $\mathcal{P}_e(\mathbb{R}) = \{A \subset \mathbb{R} : A \,\text{ist endlich}\}$. Dann ist $\mathcal{P}_e(\mathbb{R}) \subset P(\mathbb{R}) \subset V_1(\mathbb{R})$, also $\mathcal{P}_e(\mathbb{R}) \in V_2(\mathbb{R})$.

5. Sei $A \subset V_n(\mathbb{R})$ und $F : A \to V_m(\mathbb{R})$ eine Abbildung. $F = \{(x, F(x)) : x \in A\}$, d.h. wir identifizieren F mit seinem Graphen G_F. Jedes $(x, F(x))$ ist Element von $V_p(X)$, wo $p = \max(m, n) + 2$ nach Beispiel 3. F ist also eine Teilmenge von $V_p(X)$, also $F \in V_{p+1}(X)$.

6. Einen Vektor $\vec{x} = (x_1, \cdots, x_p)$ kann man auffassen als Funktion von $\{1, \cdots, p\}$ in $\mathbb{R}$. Also ist $\vec{x} \in V_2(\mathbb{R})$ und damit $\mathbb{R}^p \in V_3(\mathbb{R})$ (in Übereinstimmung mit $p = 2$).

7. Sei $\mathcal{P}_e(\mathbb{R}^p) = \{A \subset \mathbb{R}^p : A \,\text{ist endlich}\}$. Ist $A \in \mathcal{P}_e(\mathbb{R}^p)$, so ist $A \subset \mathbb{R}^p$, also $A \in \mathcal{P}(\mathbb{R}^p)$ und damit $A \in V_4(\mathbb{R})$, also schließlich $\mathcal{P}_e(\mathbb{R}^p) \in V_5(\mathbb{R})$.
   $\sum : \mathcal{P}_e(\mathbb{R}^p) \to \mathbb{R}^p, A \mapsto \sum(A) := \sum_{x \in A} x$ ist eine Abbildung von $\mathcal{P}_e(\mathbb{R}^p) \to \mathbb{R}^p \subset V_2(\mathbb{R})$. $\sum$ ist damit eine Teilmenge von $V_7(\mathbb{R})$, also ein Element von $V_8(\mathbb{R})$.

8. Wo liegt $C[0, 1])$, die Menge der stetigen reellen Funktionen auf $[0, 1]$? Jedes $f \in C([0, 1])$ ist nach obigem Beispiel ein Element von $V_3(\mathbb{R})$, also ist $C[0, 1]) \in V_4(\mathbb{R})$.

Deffnition 2.2 *Eine* **Unterstruktur** *der vollen Superstruktur $V(X)$ ist eine Folge $W(X) = (W_n(X))$ mit den Eigenschaften:*
(1) $W_n(X) \subset V_n(X)$ für alle n.

(2) $W_{n+1}(X) \setminus W_n(X) \subset V_{n+1}(X) \setminus V_n(X)$.
(3) Ist $a \in W_{n+1}(X)$ *und* $b \in a$, *so ist* $b \in W_n(X)$.
(4) $a, b \in W_{n+1}(X) \setminus W_n(X)$ *impliziert* $a \cap b \in W_{n+1}(X)$ *und* $a \setminus b \in W_{n+1}(X)$.
(5) $W_n(X) \neq \emptyset$.

Wir werden im Folgenden die volle Superstruktur $V(\mathbb{R})$ erweitern zu einer echten Unterstruktur von $V({}^*\mathbb{R})$. Diese Erweiterung hat "alle Eigenschaften" von $V(\mathbb{R})$. Darüberhinaus enthält sie "auf allen Ebenen", unendlich kleine und unendlich große Elemente.

2.1.3 Die Sprache einer Superstruktur

Um den letzten Abschnitt zu präzisieren, müssen wir ein Transferprinzip formulieren können. Damit wir so ein Prinzip beweisen können, führen wir eine Sprache für eingegebene Unterstruktur $W(X)$ ein, die einerseits so "arm" wie möglich ist (s. Abschnitt 1.5), andererseits aber alle uns interessierenden mathematischen Sätze zumindest in äquivalenter Form enthält.

Aufbau der Sprache

Sei $W(X)$ eine Unterstruktur der vollen Superstruktur $V(X)$ über der unendlichen Menge X.
Der Zeichensatz der Sprache $\mathcal{L}_W$ umfaßt:

1. *Konstanten*: alle Objekte (Elemente) aus $W(X)$
2. *Variable*: $u, v, w, x, y, z,\ x_1, x_2, \cdots,$
3. *Klammern*: [,], (,), {, },
4. *Logische Symbole*: $\wedge$ (und), $\neg$ (nicht), $\exists$ (existiert),
5. *Relationssymbole*: $=$ (Gleichheit), $\in$ (Element).

Die Formeln und Sätze der Sprache werden nach folgenden Regeln induktiv aufgebaut:

(a) Sind r, s Variable oder Konstanten, so sind $r = s$ und $r \in s$ Formeln, die **atomaren Formeln**. Sind dabei r und s Konstanten, so heißen diese Formeln **atomare Sätze**.

(b) Sind Φ und Ψ Formeln, so sind $\Phi \wedge \Psi$ sowie $\neg\Phi$ Formeln.

(c) Sei x eine Variable, a eine Konstante und Φ eine Formel, die die Zeichenfolge $\exists(x \in a)$ noch nicht enthält. Dann ist $\exists(x \in a)(\Phi)$ eine Formel.

Eine **Formel in $\mathcal{L}_W$** ist eine nach den Regeln a) bis c) (unter Zuhilfenahme von weiteren Klammern, um den Aufbau zu verdeutlichen) aufgebaute Zeichenreihe.
Kommt eine Variable x in der Formel Φ nur in der Form $\exists(x \in a)\Psi(x)$ vor, wobei $\Psi(x)$ besagt, daß in dieser Formel x auftaucht, so heißt x gebunden, andernfalls frei.
Ein **mathematischer Satz** ist eine Formel, die nur gebundene Variablen enthält.

Erweiterung der Sprache

Wir möchten natürlich auch die anderen logischen Zeichen und den Allquantor benutzen. Dies geschieht in derselben Weise wie im ersten Kapitel (vergleiche Seite 24)

1. $\Phi \vee \Psi$ ist äquivalent zu (1') $\neg(\neg\Phi \wedge \neg\Psi)$,

2. $\forall(x \in a)\Phi(x)$ ist äquivalent zu (2') $\neg(\exists(x \in a)(\neg\Phi(x))$,

3. $\Phi \Rightarrow \Psi$ ist äquivalent zu (3') $\neg(\Phi \wedge \neg\Psi)$.

Das bedeutet nichts anderes als daß ein Satz der Form (1) bis (3) genau dann gilt, wenn sein äquivalenter Satz (1'), bzw. (2') bzw. (3') gilt.

Wir benutzen für die "logische" Seite der Nonstandardanalysis wie zum Beispiel im folgenden Abschnitt unsere besonders reduzierte Sprache, weil das so bequem ist. In den Anwendungen benutzen wir dann natürlich die erweiterte Sprache.

Die Interpretation der Sprache $\mathcal{L}_W$

1. Ein atomarer Satz $a = b$ bzw. $a \in b$ ist *wahr*, wenn in $W(X)$ $a = b$ bzw. $a \in b$ gilt. Andernfalls ist der Satz *falsch*.

2. Sei Φ ein mathematischer Satz. Der Satz $\neg\Phi$ ist genau dann wahr, wenn Φ falsch ist.

3. Seien Φ und Ψ mathematische Sätze. $\Phi \wedge \Psi$ ist genau dann wahr, wenn sowohl Φ als auch Ψ wahr sind.

4. Sei Φ eine Formel, in der x als einzige Variable vorkommt.
$\exists(x \in a)\Phi$ ist genau dann wahr, wenn es ein $b \in a$ gibt, so daß der mathematische Satz $\Phi(b)$, den man erhält, indem man in Φ überall x durch b ersetzt, wahr ist.

Damit kann man nach dem Aufbau des mathematischen Satzes in endlichen vielen Schritten feststellen, ob er wahr ist oder nicht.
Wir hatten schon oben erörtert, daß Formeln der Bauart $\Phi \Rightarrow \Psi, \quad \forall(x \in a)\Phi(x)$ usw. logisch äquivalent sind zu Formeln unserer "armen" mathematischen Sprache. Deshalb bringen wir jetzt Beispiele von Sätzen, die über die elementare Analysis aus Kapitel 1 hinausgehen, bereits in der Form mit $\forall(x \in a)$, $\Rightarrow$ und V.

Der Raum $C([0,1]$ aller stetigen Funktionen auf dem Einheitsintervall ist aus $V_4C\mathbb{R})$.
Der folgende Satz drückt gleichmäßige Stetigkeit aus:

$$\forall(f \in c([0,1])\,(\forall\varepsilon \in \mathbb{R}((\varepsilon > 0) \Rightarrow \exists(\delta \in \mathbb{R})(\delta > \wedge)$$

$$\forall(x \in [0,1])\forall(y \in [0,1])$$

$$|x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon)$$

Um diesen Satz in der Sprache mit $\in$ und $=$ auszudrücken müssen wir statt $\varepsilon > 0$ schreiben $\varepsilon \in \mathbb{R}_+ \wedge \neg(0 = \varepsilon)$. Dasselbe gilt für $\delta > 0$. Statt $|x-y| < \delta$ müssen wir schreiben

$$(x,y,z) \in D \wedge (z,z_1) \in Abs \wedge (\delta, z_1, u) \in D \Rightarrow (u \in \mathbb{R}_+ \wedge \neg(0 = u))$$

Dabei ist D der Graph der Funktion $z = x - y$ und Abs der Graph von $y = |x|$.
Noch komplizierter wird $|f(x) - f(y)| < \varepsilon$.

$$(x,x_1) \in f \wedge (y,y_1) \in f \Rightarrow |x_1 - y_1| < \varepsilon.$$

Die rechte Seite wird ähnlich wie $|x-y| < \delta$ umgeformt.
Das zweite Beispiel gibt die Additivität der Summation über endliche Mengen wieder.
Sei $\sum : P_e(\mathbb{R}) \to \mathbb{R}\,,\ \sum(A) = \sum_{a\in A} a$

$$\forall(x \in P_e(\mathbb{R}))\forall(y \in P_e(\mathbb{R})(x \cap y = \emptyset \Rightarrow \sum(x \cup y) = \sum(x) + \sum(y))$$

Hier muß man $x \cap y = \emptyset$ weiter umformulieren: Wir betrachten hierzu die Funktionen
$\mathcal{D} : P_1(\mathbb{R}) \times P_1(\mathbb{R}) \to P_1(\mathbb{R}) \quad \mathcal{D}(x,y) = x \cap y$
und $\mathcal{V} : P_1(\mathbb{R}) \times P_1(\mathbb{R}) \to P_1(\mathbb{R}) \quad \mathcal{V}(x,y) = x \cup y$
Für $x, y \in P_1(\mathbb{R}) \subset V_1(\mathbb{R})$ ist $(x,y,z) \in V_3(\mathbb{R})$
also $\mathcal{D}, \mathcal{V} \in V_5(\mathbb{R})$.
Statt $x \cap y = \emptyset$ schreiben wir $(x,y,\emptyset) \in \mathcal{D}$ und statt $\sum(x \cup y) = \sum(x) + \sum(y)$ schreiben wir $(S : \mathbb{R} \times \mathbb{R} \to \mathbb{R}\,,\ (r,s) \to r + s)$:

$$\forall(u \in \mathbb{R})\forall(v \in \mathbb{R})\forall(w \in \mathbb{R})\forall(z \in P_e(\mathbb{R}))$$

$$(\underbrace{(x,y,z) \in \mathcal{U}}_{z=x\cup y} \forall \underbrace{(x,u) \in \sum}_{u=\sum(x)} \wedge \underbrace{(y,v) \in \sum}_{v=\sum(y)} \wedge \underbrace{(u,v,w) \in S}_{w=u+v} \Rightarrow \underbrace{(z,w) \in \sum}_{w=\sum(z)})$$

Die beiden Beispiele zeigen, daß man alle mathematischen Sachverhalte - $\mathbb{R}$ betreffend - in unserer verarmten Sprache ausdrücken kann.

2.1.4 Monomorphismen zwischen Superstrukturen

Im Folgenden seien $V(X) \cap Y$ und $V(Y)$ die vollen Superstrukturen über X bzw. Y, wo $\mathbb{N} \subset X$ gilt. Wir betrachten eine injektive Abbildung von $V(X)$ **in** $V(Y)$. Das Bild von $a \in V(X)$ sei ${}^*a$. Eine solche induziert eine Abbildung der Formeln der Sprache $\mathcal{L}_{V(X)}$ in die Menge der Formeln in $\mathcal{L}_{V(Y)}$, indem man alle Konstanten $a, b, c, \cdots$ in der Formel Φ durch ihre Bilder ${}^*a, {}^*b, {}^*c \cdots$ ersetzt. So erhält man die *–transformierte Formel ${}^*\Phi$. Zum Beispiel geht $x \in P_e(X)$ über in $x \in {}^*P_2(X)$. Für ein Element $a \in V_{n+1}(X) \setminus V_0(X)$, also für eine Menge, sei $P_e(a)$ die Menge der endlichen Teilmengen von a. Es ist $P_e(a) \in V_{n+1}(X)$.

Zur Würdigung unserer Axiomatik definieren wir:

Deffnition 2.3 *Sei $A \subset V_{n+1}(X)$ eine Menge. A hat die endliche Durchschnittseigenschaft, wenn für jede endliche Teilmenge $B \subset A$ stets $\bigcap_{C \in B} C \neq \emptyset$ gilt.*

Damit formulieren wir:

Deffnition 2.4 *Sei $\mathbb{N} \subset X \cap Y$. Die injektive Abbildung ${}^* : V(x) \to V(Y)$ heißt Monomorphismus, wenn sie die folgenden Eigenschaften hat:*

1. ${}^*(\emptyset) = \emptyset, \quad {}^*(X) = Y, \quad {}^*n = n$ *für alle* $n \in \mathbb{N} \subset X$.
2. $a \in X$ *impliziert* ${}^*a \in Y$.
3. $a \in V_{n+1}(X) \setminus V_n(X)$ *impliziert* ${}^*a \in V_{n+1}(Y) \setminus V_n(Y)$.
4. $a \in V_{n+1}(X)$ *und* $b \in a$ *impliziert* $b \in {}^*V_n(X)$.
5. *Es gilt das Transferprinzip: Ein mathematischer Satz Φ der Sprache $\mathcal{L}_{V(X)}$ ist genau dann in $V(X)$ wahr, wenn ${}^*\Phi$ in $V(Y)$ wahr ist.*
6. *(Erweiterungseigenschaft) Für jede Menge $a \subset V_{n+1}(X)$ mit der endlichen Durchschnittseigenschaft gilt $\emptyset \neq \cap_{b \in a} {}^*b$.*
7. *(Saturiertheit) Für jede abzählbare Menge $B \subset {}^*V_{n+1}(X) \setminus {}^*V_0(X)$ mit der endlichen Durchschnittseigenschaft gilt $\emptyset \neq \cap_{b \in B} b$.*

Ob es einen solchen Monomorphismus gibt, beantwortet der folgende Satz:

Theorem 2.5 *(Zakon-Robinson) Zu jeder unendlichen Menge X gibt es eine unendliche Menge Y und einen Monomorphismus * von $V(X)$ in $V(Y)$.*

Was ist der Unterschied zum Kapitel 1? Dort hatten wir den "Monomorphismus" nur auf der Ebene $V_1(X)$ konstruiert und das Transferprinzip auf unsere elementaren Sätze eingeschränkt, die nur Quantoren über Elemente aus $V_0(\mathbb{R}) = \mathbb{R}$ zuließen. Hier haben wir das Transferprinzip in voller Allgemeinheit.
Der Beweis des Satzes läuft im Prinzip wie unser Beweis in Abschnitt 1.6, nur braucht man ein Ultrafunktional auf einer viel größeren Menge als $\mathbb{N}$. Siehe A. Hurd, P.A. Loeb: An introduction to nonstandard real analysis, Kapitel II, II.4. oder P.A. Loeb, M. Wolff, Nonstandard Analysis for the working mathematican, Kluwer 2000, S. 57 ff.

Wir stellen zunächst die einfachsten Eigenschaften eines Monomorphismus zusammen:

Satz 2.6 (Elementare Eigenschaften eines Monomorphismus)

a) Seien $a, b, a_1, \cdots a_n$ fest gewählte Elemente in $V(X)$. Dann gilt

1. $^*\{a_1, \cdots, a_n\} = \{^*a_1, \cdots, ^*a_n\}$.
2. $^*(a_1, \cdots, a_n) = (^*a_1, \cdots, ^*a_n)$.
3. $a \in b$ *gilt genau dann, wenn* $^*a \in {}^*b$ *gilt.*
4. $a = b$ *gilt genau dann, wenn* $^*a = {}^*b$ *gilt.*
5. $a \subset b$ *gilt genau dann, wenn* $^*a = {}^*b$ *gilt.*
6. $^*(\cup_{k=1}^n a_k) = \cup_{k=1}^n {}^*a_k$, $^*(\cap_{k=1}^n a_k) = \cap_{k=1}^n {}^*a_k$.

b) Ist f eine Abbildung von a nach b, so ist $^*f$ eine Abbildung von $^*a$ nach $^*b$ und für alle x aus a gilt $^*(f(x)) = {}^*f(^*x)$.

c) $f : a \to b$ *ist genau dann* $\left\{ \begin{array}{c} \textit{injektiv} \\ \textit{surjektiv} \\ \textit{bijektiv} \end{array} \right\}$ *wenn* $^*f$ $\left\{ \begin{array}{c} \textit{injektiv} \\ \textit{surjektiv} \\ \textit{bijektiv} \end{array} \right\}$ *ist.*

d) $^*(a_1 \times a_2 \times \cdots \times a_n) = {}^*a_1 \times {}^*a_2 \times \cdots \times {}^*a_n$.

Beweis: a) 1) Sei $b = \{a_1, \cdots, a_n\} \in V_{n+1}(X)$. Dann gilt

$$\forall (x \in V_n(X))(x \in b \Leftrightarrow ((x = a_1)v(x = a_2)v \cdots v(x = a_n))).$$

Der gesternte Satz sagt dann $x \in {}^*b \Leftrightarrow x = {}^*a_j$ für ein $j \leq n$.
2) Wegen $(a, b) = \{\{a\}, \{a, b\}\}$ folgt mit 2) 2-maliger Anwendung von 1):

$$^*(a, b) = \{^*\{a\}, ^*\{a, b\}\} = \{\{^*a\}, \{^*a, ^*b\}\} = (^*a, ^*b).$$

$(a_1, a_2, \cdots, a_n) = \{(1, a_1), \cdots, (n, a_n)\}$ (das ist der Graph der Abbildung $a : \{1, \cdots, n\} \ni \mapsto a_j$). Also folgt mit 1)

$$\begin{aligned} {}^*(a_1, \cdots, a_n) &= \{{}^*(1, a_1), \cdots, {}^*(n, a_n)\} \\ &= \{({}^*1, {}^*a_1), \cdots, ({}^*n, {}^*a_n)\} = ({}^*a_1, \cdots, {}^*a_n). \end{aligned}$$

3) und 4) $a \in b$ und $a = b$ sind atomare Sätze. Das TFP liefert die Behauptung.
5) $a \subset b \in V_n(X)$. Das bedeutet: $\forall(x \in V_n(X))(x \in a \Rightarrow x \in b)$. Das TFP liefert die Behauptung.
6) Beweis durch Induktion. Für $n = 2$ formuliere man einen entsprechenden Satz und wende das TFP an.

b) Sei f eine Abbildung von a nach b, identifiziert mit ihrem Graphen $\subset V_n(X) \quad (n \geq 1)$. Dann gilt

$$\begin{aligned} \forall(x \in V_n(X))(x \in f \Leftrightarrow [\exists(y \in a)\exists(z \in b)((y, z) \in f \wedge (y, z) = x)) \\ \wedge\forall(u \in b)((y, u) \in f \Leftrightarrow u = z))] \end{aligned}$$

Das TFP liefert die Behauptung, daß ${}^*f$ eine Abbildung ist. Der Rest ist klar.

c) Verwenden Sie das TFP.

d) Sei $b = \cup_{j=1}^{n} a_j \subset V_q(X)$ und $c = \{f \in V_p(X) : f(\{1, \cdots, n\}) \subset b\}$ wo $p = q + 2$ ist.

$$\forall(x \in V_p(X))(x \in a_1 \times \cdots \times a_n \Leftrightarrow x \in C \wedge (1, x(1) \in a_1 \wedge \cdots \wedge (n, x(n)) \in a_n)$$

Das Transferprinzip liefert die Behauptung. □

Sei $X = \mathbb{R}$. Nach dem vorstehenden Satz ist ${}^*(\mathbb{R}^p) = ({}^*\mathbb{R})^p$. Wir hatten schon in Kapitel 1 gesehen: Sind f, g zwei Funktionen von $D \subset \mathbb{R}^p \to \mathbb{R}^q$, so ist ${}^*(f + g) = {}^*f + {}^*g$. Wir beweisen dies um für beliebige Definitionsbereiche $d \in V_n(\mathbb{R})$. Wir wissen schon, daß die Menge F aller Abbildungen von d nach $\mathbb{R}^p$ ein Element aus einem geeigneten $V_p(\mathbb{R})$ ist. Wir setzen $S_F = \{(f, g, h) : f, g \in F \text{ und } h = f + g\}$. Dies ist also der Graph der Summenabbildung $+_F$. Sei S_p die Addition auf $\mathbb{R}^p$. Dann gilt

$$\forall(f \in F)\forall(g \in F)\forall(h \in F)((f, g, h) \in S_F \Rightarrow \forall x \in d(f(x), g(x), h(x) \in S_p) \tag{2.1}$$

und daraus folgt:

Satz 2.7 *Sind $f, g \in {}^*F$, so ist*
$(f +_F g)(x) = f(x) +_p g(x)$ für alle $x \in {}^*d$.

Beweis: Wende das TFP auf den mathematischen Satz (2.2) an. □

Auf ähnliche Weise erhalten wir nicht nur, daß ${}^*W$ ein ${}^*\mathbb{R}$–Vektorraum ist, falls W ein $\mathbb{R}$–Vektorraum ist, sondern auch: Ist U eine Menge von $\mathbb{R}$–Vektorräumen, $U \subset V_n(X)$ für ein n, so besteht ${}^*U$ nur aus ${}^*\mathbb{R}$–Vektorräumen, also auch die eventuell hinzukommenden Elemente sind ${}^*\mathbb{R}$–Vektorräume. Aber wann kommen Elemente hinzu?

Satz 2.8 *$d \in V_n(X)$ ist genau dann endlich, wenn $^*d = \{^*u : u \in d\}$ ist.*

Anders ausgedrückt: ist d unendlich, so ist $^*d \supset_{\neq} \{^*u : u \in d\}$.

Beweis: Ist $d = \{u_1, \cdots, u_n\}$, so ist $^*d = \{^*u_1, \cdots, ^*u_n\}$ nach Satz? Sei d unendlich und $c = \{a \subset d : d \setminus a$ ist endlich $\}$. c hat die endliche Durchschnittseigenschaft, also ist $\emptyset \neq \cap_{a \in c} {}^*a =: A$. Wegen $a \subset d$ ist $^*a \subset {}^*d$, also $A \subset {}^*d$. Ist $x \in d$, so ist $x \notin d \setminus \{x\} \in c$, also $^*x \notin {}^*(d \setminus \{x\}) \supset A$. Damit ist $\{^*x : x \in d\} \cap A = \emptyset$. Also kommen neue Elemente hinzu. □

2.1.5 Interne und externe Mengen

Sei $^*V(X) \to V(Y)$ ein Monomorphismus. Dann ist $\{^*V_n(X) : n \in \mathbb{N}\}$ eine Unterstruktur der vollen Superstruktur $V(Y)$. Für $n \geq 1$ heißen die Elemente aus $^*V_n(X) \setminus {}^*V_0(X)$ **interne Mengen**, diejenigen aus $V_n(Y) \setminus (^*V_n(X) \cup V_0(Y)$ heißen **externe Mengen**.

Beispiel: $^*\mathbb{N} \setminus \mathbb{N} = A$ ist extern. Denn $A \neq \emptyset$ (s.o). Nun gilt der folgende Satz in $V(X)$:

$$\forall (a \in V_1(X) \setminus V_0(X))[(a \neq \emptyset \wedge \forall (y \in V_0(X))(y \in a \Rightarrow y \in \mathbb{N}))$$
$$\Rightarrow \exists (x \in \mathbb{N})(x \in a \wedge \forall (y \in \mathbb{N})(y \in a \Rightarrow y \geq x))]$$

der besagt, daß jede nicht leere Teilmenge von $\mathbb{N}$ ein kleinstes Element hat. Wäre A intern, also $A \in {}^*V_1(X) \setminus {}^*V_0(X)$ so würde dies nach dem TFP auch für A gelten. Angenommen x sei ein kleinstes Element. Dann wäre $x-1 \notin A$, also $x-1 \in \mathbb{N}$ und damit $x = x-1+1 \in \mathbb{N}$, also $x \notin A$, ein Widerspruch.

Hier sind einige Beispiele für $X = \mathbb{R}$.

Beispiele:

1. Alle reellen Zahlen sind Elemente von $\mathbb{R}$, also wegen $\mathbb{R} \subset V_1(\mathbb{R}) \subset V(\mathbb{R})$ gilt $x \in V(\mathbb{R})$ für jede reelle Zahl.

2. Jede Teilmenge, erst recht jede endliche Teilmenge von $\mathbb{R}$ ist Element von $V(\mathbb{R})$.

3. Seien $u, v \in V_n(\mathbb{R})$ für ein n. Man definiert bekanntlich das geordnete Paar (u, v) durch $(u, v) = \{\{u\}, \{u, v\}\}$. Wegen $u, v \in V_n(\mathbb{R})$ ist $\{u\}, \{u, v\} \subset V_n(\mathbb{R})$, also $\{u\}, \{u, v\} \in V_{n+1}(\mathbb{R}) = P(V_n(\mathbb{R})) \cup V_n(\mathbb{R})$. Damit ist $(u, v) = \{\{u\}\{u, v\}\} \subset V_{n+1}(\mathbb{R})$, also $(u, v) \in V_{n+2}(\mathbb{R})$. Damit ist $\mathbb{R}^2 \subset V_2(\mathbb{R})$, also $\mathbb{R}^2 \in V_3(\mathbb{R})$.

4. Sei $\mathcal{P}_e(\mathbb{R}) = \{A \subset \mathbb{R} : A \,\text{ist endlich}\}$. Dann ist $\mathcal{P}_e(\mathbb{R}) \subset P(\mathbb{R}) \subset V_1(\mathbb{R})$, also $\mathcal{P}_e(\mathbb{R}) \in V_2(\mathbb{R})$.

5. Sei $A \subset V_n(\mathbb{R})$ und $F : A \to V_m(\mathbb{R})$ eine Abbildung. $F = \{(x, F(x)) : x \in A\}$, d.h. wir identifizieren F mit seinem Graphen G_F. Jedes $(x, F(x))$ ist Element von $V_p(X)$, wo $p = \max(m, n) + 2$ nach Beispiel 3. F ist also eine Teilmenge von $V_p(X)$, also $F \in V_{p+1}(X)$.

6. Einen Vektor $\vec{x} = (x_1, \cdots, x_p)$ kann man auffassen als Funktion von $\{1, \cdots, p\}$ in $\mathbb{R}$. Also ist $\vec{x} \in V_2(\mathbb{R})$ und damit $\mathbb{R}^p \in V_3(\mathbb{R})$ (in Übereinstimmung mit $p = 2$).

7. Sei $\mathcal{P}_e(\mathbb{R}^p) = \{A \subset \mathbb{R}^p : A \,\text{ist endlich}\}$. Ist $A \in \mathcal{P}_e(\mathbb{R}^p)$, so ist $A \subset \mathbb{R}^p$, also $A \in \mathcal{P}(\mathbb{R}^p)$ und damit $A \in V_4(\mathbb{R})$, also schließlich $\mathcal{P}_e(\mathbb{R}^p) \in V_5(\mathbb{R})$.
   $\sum : \mathcal{P}_e(\mathbb{R}^p) \to \mathbb{R}^p, A \mapsto \sum(A) := \sum_{x \in A} x$ ist eine Abbildung von $\mathcal{P}_e(\mathbb{R}^p) \to \mathbb{R}^p \subset V_2(\mathbb{R})$. $\sum$ ist damit eine Teilmenge von $V_7(\mathbb{R})$, also ein Element von $V_8(\mathbb{R})$.

8. Wo liegt $C[0, 1])$, die Menge der stetigen reellen Funktionen auf $[0, 1]$? Jedes $f \in C([0, 1])$ ist nach obigem Beispiel ein Element von $V_3(\mathbb{R})$, also ist $C[0, 1]) \in V_4(\mathbb{R})$.

Deffnition 2.9 *Eine* **Unterstruktur** *der vollen Superstruktur* $V(X)$ *ist eine Folge* $W(X) = (W_n(X))$ *mit den Eigenschaften:*
(1) $W_n(X) \subset V_n(X)$ *für alle* n.
(2) $W_{n+1}(X) \setminus W_n(X) \subset V_{n+1}(X) \setminus V_n(X)$.
(3) Ist $a \in W_{n+1}(X)$ *und* $b \in a$, *so ist* $b \in W_n(X)$.
(4) $a, b \in W_{n+1}(X) \setminus W_n(X)$ *impliziert* $a \cap b \in W_{n+1}(X)$ *und* $a \setminus b \in W_{n+1}(X)$.
(5) $W_n(X) \neq \emptyset$.

Wir werden im Folgenden die volle Superstruktur $V(\mathbb{R})$ erweitern zu einer echten Unterstruktur von $V({}^*\mathbb{R})$. Diese Erweiterung hat "alle Eigenschaften" von $V(\mathbb{R})$. Darüberhinaus enthält sie "auf allen Ebenen", unendlich kleine und unendlich große Elemente.

2.1.6 Die Sprache einer Superstruktur

Um den letzten Abschnitt zu präzisieren, müssen wir ein Transferprinzip formulieren können. Damit wir so ein Prinzip beweisen können, führen wir eine Sprache für eingegebene Unterstruktur $W(X)$ ein, die einerseits so "arm" wie möglich ist (s. Abschnitt 1.5), andererseits aber alle uns interessierenden mathematischen Sätze zumindest in äquivalenter Form enthält.

Aufbau der Sprache

Sei $W(X)$ eine Unterstruktur der vollen Superstruktur $V(X)$ über der unendlichen Menge X.
Der Zeichensatz der Sprache $\mathcal{L}_W$ umfaßt:

1. *Konstanten*: alle Objekte (Elemente) aus $W(X)$
2. *Variable*: $u, v, w, x, y, z,\ x_1, x_2, \cdots,$
3. *Klammern*: [,], (,), {, },
4. *Logische Symbole*: $\wedge$ (und), $\neg$ (nicht), $\exists$ (existiert),
5. *Relationssymbole*: $=$ (Gleichheit), $\in$ (Element).

Die Formeln und Sätze der Sprache werden nach folgenden Regeln induktiv aufgebaut:

(a) Sind r, s Variable oder Konstanten, so sind $r = s$ und $r \in s$ Formeln, die **atomaren Formeln**. Sind dabei r und s Konstanten, so heißen diese Formeln **atomare Sätze**.

(b) Sind Φ und Ψ Formeln, so sind $\Phi \wedge \Psi$ sowie $\neg\Phi$ Formeln.

(c) Sei x eine Variable, a eine Konstante und Φ eine Formel, die die Zeichenfolge $\exists(x \in a)$ noch nicht enthält. Dann ist $\exists(x \in a)(\Phi)$ eine Formel.

Eine **Formel in $\mathcal{L}_W$** ist eine nach den Regeln a) bis c) (unter Zuhilfenahme von weiteren Klammern, um den Aufbau zu verdeutlichen) aufgebaute Zeichenreihe.
Kommt eine Variable x in der Formel Φ nur in der Form $\exists(x \in a)\Psi(x)$ vor, wobei $\Psi(x)$ besagt, daß in dieser Formel x auftaucht, so heißt x gebunden, andernfalls frei.
Ein **mathematischer Satz** ist eine Formel, die nur gebundene Variablen enthält.

Erweiterung der Sprache

Wir möchten natürlich auch die anderen logischen Zeichen und den Allquantor benutzen. Dies geschieht in derselben Weise wie im ersten Kapitel (vergleiche Seite 24)

1. $\Phi \vee \Psi$ ist äquivalent zu (1') $\neg(\neg\Phi \wedge \neg\Psi)$,
2. $\forall(x \in a)\Phi(x)$ ist äquivalent zu (2') $\neg(\exists(x \in a)(\neg\Phi(x))$,
3. $\Phi \Rightarrow \Psi$ ist äquivalent zu (3') $\neg(\Phi \wedge \neg\Psi)$.

Das bedeutet nichts anderes als daß ein Satz der Form (1) bis (3) genau dann gilt, wenn sein äquivalenter Satz (1'), bzw. (2') bzw. (3') gilt.

Wir benutzen für die "logische" Seite der Nonstandardanalysis wie zum Beispiel im folgenden Abschnitt unsere besonders reduzierte Sprache, weil das so bequem ist. In den Anwendungen benutzen wir dann natürlich die erweiterte Sprache.

Die Interpretation der Sprache $\mathcal{L}_W$

1. Ein atomarer Satz $a = b$ bzw. $a \in b$ ist *wahr*, wenn in $W(X)$ $\quad a = b$ bzw. $a \in b$ gilt. Andernfalls ist der Satz *falsch*.
2. Sei Φ ein mathematischer Satz. Der Satz $\neg\Phi$ ist genau dann wahr, wenn Φ falsch ist.
3. Seien Φ und Ψ mathematische Sätze. $\Phi \wedge \Psi$ ist genau dann wahr, wenn sowohl Φ als auch Ψ wahr sind.
4. Sei Φ eine Formel, in der x als einzige Variable vorkommt.
   $\exists(x \in a)\Phi$ ist genau dann wahr, wenn es ein $b \in a$ gibt, so daß der mathematische Satz $\Phi(b)$, den man erhält, indem man in Φ überall x durch b ersetzt, wahr ist.

Damit kann man nach dem Aufbau des mathematischen Satzes in endlichen vielen Schritten feststellen, ob er wahr ist oder nicht.
Wir hatten schon oben erörtert, daß Formeln der Bauart $\Phi \Rightarrow \Psi$, $\quad \forall(x \in a)\Phi(x)$ usw. logisch äquivalent sind zu Formeln unserer "armen" mathematischen Sprache. Deshalb bringen wir jetzt Beispiele von Sätzen, die über die elementare Analysis aus Kapitel 1 hinausgehen, bereits in der Form mit $\forall(x \in a)$, $\Rightarrow$ und V.

Der Raum $C([0,1]$ aller stetigen Funktionen auf dem Einheitsintervall ist aus $V_4C\mathbb{R})$.
Der folgende Satz drückt gleichmäßige Stetigkeit aus:

$$\forall(f \in c([0,1])\,(\forall\varepsilon \in \mathbb{R}((\varepsilon > 0) \Rightarrow \exists(\delta \in \mathbb{R})(\delta > \wedge)$$

$$\forall(x \in [0,1])\forall(y \in [0,1])$$

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon)$$

Um diesen Satz in der Sprache mit $\in$ und $=$ auszudrücken müssen wir statt $\varepsilon > 0$ schreiben $\varepsilon \in \mathbb{R}_+ \wedge \neg(0 = \varepsilon)$. Dasselbe gilt für $\delta > 0$. Statt $|x - y| < \delta$ müssen wir schreiben

$$(x, y, z) \in D \wedge (z, z_1) \in Abs \wedge (\delta, z_1, u) \in D \Rightarrow (u \in \mathbb{R}_+ \wedge \neg(0 = u))$$

Dabei ist D der Graph der Funktion $z = x - y$ und Abs der Graph von $y = |x|$.
Noch komplizierter wird $|f(x) - f(y)| < \varepsilon$.

$$(x, x_1) \in f \wedge (y, y_1) \in f \Rightarrow |x_1 - y_1| < \varepsilon.$$

Die rechte Seite wird ähnlich wie $|x - y| < \delta$ umgeformt.
Das zweite Beispiel gibt die Additivität der Summation über endliche Mengen wieder.
Sei $\sum : P_e(\mathbb{R}) \to \mathbb{R}\,,\ \sum(A) = \sum_{a \in A} a$

$$\forall(x \in P_e(\mathbb{R}))\forall(y \in P_e(\mathbb{R})(x \cap y = \emptyset \Rightarrow \sum(x \cup y) = \sum(x) + \sum(y))$$

Hier muß man $x \cap y = \emptyset$ weiter umformulieren: Wir betrachten hierzu die Funktionen
$\mathcal{D} : P_1(\mathbb{R}) \times P_1(\mathbb{R}) \to P_1(\mathbb{R}) \quad \mathcal{D}(x, y) = x \cap y$
und $\mathcal{V} : P_1(\mathbb{R}) \times P_1(\mathbb{R}) \to P_1(\mathbb{R}) \quad \mathcal{V}(x, y) = x \cup y$
Für $x, y \in P_1(\mathbb{R}) \subset V_1(\mathbb{R})$ ist $(x, y, z) \in V_3(\mathbb{R})$
also $\mathcal{D}, \mathcal{V} \in V_5(\mathbb{R})$.
Statt $x \cap y = \emptyset$ schreiben wir $(x, y, \emptyset) \in \mathcal{D}$ und statt $\sum(x \cup y) = \sum(x) + \sum(y)$ schreiben wir $(S : \mathbb{R} \times \mathbb{R} \to \mathbb{R}\,,\ (r, s) \to r + s)$:

$$\forall(u \in \mathbb{R})\forall(v \in \mathbb{R})\forall(w \in \mathbb{R})\forall(z \in P_e(\mathbb{R}))$$

$$(\underbrace{(x, y, z) \in \mathcal{U}}_{z=x\cup y} \forall \underbrace{(x, u) \in \sum}_{u=\sum(x)} \wedge \underbrace{(y, v) \in \sum}_{v=\sum(y)} \wedge \underbrace{(u, v, w) \in S}_{w=u+v} \Rightarrow \underbrace{(z, w) \in \sum}_{w=\sum(z)})$$

Die beiden Beispiele zeigen, daß man alle mathematischen Sachverhalte - $\mathbb{R}$ betreffend - in unserer verarmten Sprache ausdrücken kann.

2.1.7 Monomorphismen zwischen Superstrukturen

Im Folgenden seien $V(X) \cap Y$ und $V(Y)$ die vollen Superstrukturen über X bzw. Y, wo $\mathbb{N} \subset X$ gilt. Wir betrachten eine injektive Abbildung von $V(X)$ **in** $V(Y)$. Das Bild von $a \in V(X)$ sei ${}^*a$. Eine solche induziert eine Abbildung der Formeln der Sprache $\mathcal{L}_{V(X)}$ in die Menge der Formeln in $\mathcal{L}_{V(Y)}$, indem man alle Konstanten $a, b, c, \cdots$ in der Formel Φ durch ihre Bilder ${}^*a, {}^*b, {}^*c \cdots$ ersetzt. So erhält man die *–transformierte Formel ${}^*\Phi$. Zum Beispiel geht $x \in P_e(X)$ über in $x \in {}^*P_2(X)$. Für ein Element $a \in V_{n+1}(X) \setminus V_0(X)$, also für eine Menge, sei $P_e(a)$ die Menge der endlichen Teilmengen von a. Es ist $P_e(a) \in V_{n+1}(X)$.

Zur Würdigung unserer Axiomatik definieren wir:

Deffnition 2.10 *Sei $A \subset V_{n+1}(X)$ eine Menge. A hat die endliche Durchschnittseigenschaft, wenn für jede endliche Teilmenge $B \subset A$ stets $\bigcap_{C \in B} C \neq \emptyset$ gilt.*

Damit formulieren wir:

Deffnition 2.11 *Sei $\mathbb{N} \subset X \cap Y$. Die injektive Abbildung $^* : V(x) \to V(Y)$ heißt Monomorphismus, wenn sie die folgenden Eigenschaften hat:*

1. $^*(\emptyset) = \emptyset, \quad ^*(X) = Y, \quad ^*n = n$ *für alle* $n \in \mathbb{N} \subset X$.
2. $a \in X$ *impliziert* $^*a \in Y$.
3. $a \in V_{n+1}(X) \setminus V_n(X)$ *impliziert* $^*a \in V_{n+1}(Y) \setminus V_n(Y)$.
4. $a \in V_{n+1}(X)$ *und* $b \in a$ *impliziert* $b \in {^*V_n(X)}$.
5. *Es gilt das Transferprinzip: Ein mathematischer Satz Φ der Sprache $\mathcal{L}_{V(X)}$ ist genau dann in $V(X)$ wahr, wenn $^*\Phi$ in $V(Y)$ wahr ist.*
6. *(Erweiterungseigenschaft) Für jede Menge $a \subset V_{n+1}(X)$ mit der endlichen Durchschnittseigenschaft gilt $\emptyset \neq \cap_{b \in a} {^*b}$.*
7. *(Saturiertheit) Für jede abzählbare Menge $B \subset {^*V_{n+1}(X)} \setminus {^*V_0(X)}$ mit der endlichen Durchschnittseigenschaft gilt $\emptyset \neq \cap_{b \in B} b$.*

Ob es einen solchen Monomorphismus gibt, beantwortet der folgende Satz:

Theorem 2.12 *(Zakon-Robinson) Zu jeder unendlichen Menge X gibt es eine unendliche Menge Y und einen Monomorphismus * von $V(X)$ in $V(Y)$.*

Was ist der Unterschied zum Kapitel 1? Dort hatten wir den "Monomorphismus" nur auf der Ebene $V_1(X)$ konstruiert und das Transferprinzip auf unsere elementaren Sätze eingeschränkt, die nur Quantoren über Elemente aus $V_0(\mathbb{R}) = \mathbb{R}$ zuließen. Hier haben wir das Transferprinzip in voller Allgemeinheit.
Der Beweis des Satzes läuft im Prinzip wie unser Beweis in Abschnitt 1.6, nur braucht man ein Ultrafunktional auf einer viel größeren Menge als $\mathbb{N}$. Siehe A. Hurd, P.A. Loeb: An introduction to nonstandard real analysis, Kapitel II, II.4. oder P.A. Loeb, M. Wolff, Nonstandard Analysis for the working mathematican, Kluwer 2000, S. 57 ff.

Wir stellen zunächst die einfachsten Eigenschaften eines Monomorphismus zusammen:

Satz 2.13 (Elementare Eigenschaften eines Monomorphismus)

a) Seien $a, b, a_1, \cdots a_n$ fest gewählte Elemente in $V(X)$. Dann gilt

1. $^*\{a_1, \cdots, a_n\} = \{^*a_1, \cdots, ^*a_n\}$.
2. $^*(a_1, \cdots, a_n) = (^*a_1, \cdots, ^*a_n)$.
3. $a \in b$ *gilt genau dann, wenn* $^*a \in {}^*b$ *gilt.*
4. $a = b$ *gilt genau dann, wenn* $^*a = {}^*b$ *gilt.*
5. $a \subset b$ *gilt genau dann, wenn* $^*a = {}^*b$ *gilt.*
6. $^*(\cup_{k=1}^n a_k) = \cup_{k=1}^n {}^*a_k$, $^*(\cap_{k=1}^n a_k) = \cap_{k=1}^n {}^*a_k$.

b) Ist f eine Abbildung von a nach b, so ist $^*f$ eine Abbildung von $^*a$ nach $^*b$ und für alle x aus a gilt $^*(f(x)) = {}^*f(^*x)$.

c) $f : a \to b$ *ist genau dann* $\left\{\begin{array}{c} \textit{injektiv} \\ \textit{surjektiv} \\ \textit{bijektiv} \end{array}\right\}$ *wenn* $^*f$ $\left\{\begin{array}{c} \textit{injektiv} \\ \textit{surjektiv} \\ \textit{bijektiv} \end{array}\right\}$ *ist.*

d) $^*(a_1 \times a_2 \times \cdots \times a_n) = {}^*a_1 \times {}^*a_2 \times \cdots \times {}^*a_n$.

Beweis: a) 1) Sei $b = \{a_1, \cdots, a_n\} \in V_{n+1}(X)$. Dann gilt

$$\forall (x \in V_n(X))(x \in b \Leftrightarrow ((x = a_1) v (x = a_2) v \cdots v (x = a_n))).$$

Der gesternte Satz sagt dann $x \in {}^*b \Leftrightarrow x = {}^*a_j$ für ein $j \leq n$.
2) Wegen $(a, b) = \{\,\{a\}, \{a, b\}\,\}$ folgt mit 2) 2-maliger Anwendung von 1):

$$^*(a, b) = \{^*\{a\}, ^*\{a, b\}\,\} = \{\,\{^*a\}, \{^*a, ^*b\}\,\} = (^*a, ^*b).$$

$(a_1, a_2, \cdots, a_n) = \{(1, a_1), \cdots, (n, a_n)\}$ (das ist der Graph der Abbildung $a : \{1, \cdots, n\} \ni \mapsto a_j$). Also folgt mit 1)

$$\begin{aligned} ^*(a_1, \cdots, a_n) &= \{^*(1, a_1), \cdots, ^*(n, a_n)\} \\ &= \{(^*1, ^*a_1), \cdots, (^*n, ^*a_n)\} = (^*a_1, \cdots, ^*a_n). \end{aligned}$$

3) und 4) $a \in b$ und $a = b$ sind atomare Sätze. Das TFP liefert die Behauptung.
5) $a \subset b \in V_n(X)$. Das bedeutet: $\forall (x \in V_n(X))(x \in a \Rightarrow x \in b)$. Das TFP liefert die Behauptung.
6) Beweis durch Induktion. Für $n = 2$ formuliere man einen entsprechenden Satz und wende das TFP an.

b) Sei f eine Abbildung von a nach b, identifiziert mit ihrem Graphen $\subset V_n(X) \quad (n \geq 1)$. Dann gilt

$$\forall(x \in V_n(X))(x \in f \Leftrightarrow [\exists(y \in a)\exists(z \in b)((y,z) \in f \wedge (y,z) = x))$$
$$\wedge \forall(u \in b)((y,u) \in f \Leftrightarrow u = z))]$$

Das TFP liefert die Behauptung, daß ${}^*f$ eine Abbildung ist. Der Rest ist klar.

c) Verwenden Sie das TFP.

d) Sei $b = \cup_{j=1}^n a_j \subset V_q(X)$ und $c = \{f \in V_p(X) : f(\{1, \cdots, n\}) \subset b\}$ wo $p = q + 2$ ist.

$$\forall(x \in V_p(X))(x \in a_1 \times \cdots \times a_n \Leftrightarrow x \in C \wedge (1, x(1) \in a_1 \wedge \cdots \wedge (n, x(n)) \in a_n)$$

Das Transferprinzip liefert die Behauptung. □

Sei $X = \mathbb{R}$. Nach dem vorstehenden Satz ist ${}^*(\mathbb{R}^p) = ({}^*\mathbb{R})^p$. Wir hatten schon in Kapitel 1 gesehen: Sind f, g zwei Funktionen von $D \subset \mathbb{R}^p \to \mathbb{R}^q$, so ist ${}^*(f+g) = {}^*f + {}^*g$. Wir beweisen dies um für beliebige Definitionsbereiche $d \in V_n(\mathbb{R})$. Wir wissen schon, daß die Menge F aller Abbildungen von d nach $\mathbb{R}^p$ ein Element aus einem geeigneten $V_p(\mathbb{R})$ ist. Wir setzen $S_F = \{(f, g, h) : f, g \in F \text{ und } h = f + g\}$. Dies ist also der Graph der Summenabbildung $+_F$. Sei S_p die Addition auf $\mathbb{R}^p$. Dann gilt

$$\forall(f \in F)\forall(g \in F)\forall(h \in F)((f,g,h) \in S_F \Rightarrow \forall x \in d(f(x), g(x), h(x) \in S_p) \tag{2.2}$$

und daraus folgt:

Satz 2.14 *Sind $f, g \in {}^*F$, so ist*
$(f +_F g)(x) = f(x) +_p g(x)$ für alle $x \in {}^*d$.

Beweis: Wende das TFP auf den mathematischen Satz (2.2) an. □

Auf ähnliche Weise erhalten wir nicht nur, daß ${}^*W$ ein ${}^*\mathbb{R}$–Vektorraum ist, falls W ein $\mathbb{R}$–Vektorraum ist, sondern auch: Ist U eine Menge von $\mathbb{R}$–Vektorräumen, $U \subset V_n(X)$ für ein n, so besteht ${}^*U$ nur aus ${}^*\mathbb{R}$–Vektorräumen, also auch die eventuell hinzukommenden Elemente sind ${}^*\mathbb{R}$–Vektorräume. Aber wann kommen Elemente hinzu?

Satz 2.15 $d \in V_n(X)$ *ist genau dann endlich, wenn* ${}^*d = \{{}^*u : u \in d\}$ *ist.*

Anders ausgedrückt: ist d unendlich, so ist ${}^*d \supset_{\neq} \{{}^*u : u \in d\}$.

Beweis: Ist $d = \{u_1, \cdots, u_n\}$, so ist ${}^*d = \{{}^*u_1, \cdots, {}^*u_n\}$ nach Satz? Sei d unendlich und $c = \{a \subset d : d \setminus a \text{ ist endlich }\}$. c hat die endliche Durchschnittseigenschaft, also ist $\emptyset \neq \cap_{a \in c} {}^*a =: A$. Wegen $a \subset d$ ist ${}^*a \subset {}^*d$, also $A \subset {}^*d$. Ist $x \in d$, so ist $x \notin d \setminus \{x\} \in c$, also ${}^*x \notin {}^*(d \setminus \{x\}) \supset A$. Damit ist $\{{}^*x : x \in d\} \cap A = \emptyset$. Also kommen neue Elemente hinzu. □

2.1.8 Interne und externe Mengen

Sei ${}^*V(X) \to V(Y)$ ein Monomorphismus. Dann ist $\{{}^*V_n(X) : n \in \mathbb{N}\}$ eine Unterstruktur der vollen Superstruktur $V(Y)$. Für $n \geq 1$ heißen die Elemente aus ${}^*V_n(X) \setminus {}^*V_0(X)$ **interne Mengen**, diejenigen aus $V_n(Y) \setminus ({}^*V_n(X) \cup V_0(Y)$ heißen **externe Mengen**.

Beispiel: ${}^*\mathbb{N} \setminus \mathbb{N} = A$ ist extern. Denn $A \neq \emptyset$ (s.o). Nun gilt der folgende Satz in $V(X)$:

$$\forall(a \in V_1(X) \setminus V_0(X))[(a \neq \emptyset \wedge \forall(y \in V_0(X))(y \in a \Rightarrow y \in \mathbb{N}))$$
$$\Rightarrow \exists(x \in \mathbb{N})(x \in a \wedge \forall(y \in \mathbb{N})(y \in a \Rightarrow y \geq x))]$$

der besagt, daß jede nicht leere Teilmenge von $\mathbb{N}$ ein kleinstes Element hat. Wäre A intern, also $A \in {}^*V_1(X) \setminus {}^*V_0(X)$ so würde dies nach dem TFP auch für A gelten. Angenommen x sei ein kleinstes Element. Dann wäre $x-1 \notin A$, also $x-1 \in \mathbb{N}$ und damit $x = x-1+1 \in \mathbb{N}$, also $x \notin A$, ein Widerspruch.