Kapitel 1

Sammelsurium (Notation, logische Aussagen, Zahlenräume etc.)

Dieses Kapitel dient im Wesentlichen dazu, eine einheitliche Notation und Sprechweise für die späteren Kapitel zu schaffen. Die Inhalte (nicht unbedingt die Formalisierungen) sollten aus der Schule bekannt sein. Eventuell sind die komplexen Zahl neu. $\boxed{\downarrow 14.10.03}$

1.1 Notation, logische Aussagen

Notation 1.1:

Folgende Standardbezeichnungen und Symbole sollten aus der Schule bekannt sein und werden auch hier verwendet werden:

- **Mengen:** $\{1, 2, 3, x, y, z\}$, $\{x \in \mathbb{R}; x^2 + 25 \cdot x > 29\}$,
- **die leere Meng:** $\emptyset$ *oder auch* $\{\}$,
- **Elemente von Mengen:** $x \in A$ *bedeutet „x ist aus der Menge A“,*
- **Vereinigung:** $A \cup B = \{x;\ x \in A \text{ oder } x \in B\}$,
- **Durchschnitt:** $A \cap B = \{x;\ x \in A \text{ und } x \in B\}$,
- **Teilmengen:** $A \subset B$ *heißt: alle Elemente von A sind auch in B enthalten,*
- **Obermengen:** $B \supset A$ *ist äquivalent zu* $A \subset B$,
- **Differenzmengen:** $A \setminus B = \{x \in A;\ x \notin B\}$,
- $\mathbb{N}$ = *die Menge der natürlichen Zahlen* $\{1, 2, 3, \dots\}$,

- $\mathbb{N}_0 = \{0\} \cup \mathbb{N} = \{0, 1, 2, 3, \dots\}$,
- $\mathbb{Z}$ = *die Menge der ganzen Zahlen* $= \{\dots, -2, -1, 0, 1, 2, \dots\}$,
- $\mathbb{Q}$ = *die Menge der rationalen Zahlen* $= \{\frac{z}{n}; z \in \mathbb{Z}; n \in \mathbb{N}\}$,
- $\mathbb{R}$ = *die Menge der reellen Zahlen,*
  Intervalle:

$$
\begin{array}{lcll}
(a,b) & = & \{x \in \mathbb{R}; a < x < b\}, & \text{„offene" Intervalle,} \\
[a,b] & = & \{x \in \mathbb{R}; a \le x \le b\}, & \text{„geschlossene" Intervalle,} \\
(a,b] & = & \{x \in \mathbb{R}; a < x \le b\}, & \text{„halboffene" Intervalle,} \\
[a,b) & = & \{x \in \mathbb{R}; a \le x < b\}, & \text{„halboffene" Intervalle,}
\end{array}
$$

- $\mathbb{C}$ = *die Menge der komplexen Zahlen (siehe Abschnitt 1.5),*
- **kartesisches Produkt:** $A \times B = \{(a,b); a \in A, b \in B\}$,
- **„Unendlich"**: $\pm\infty$,
- **Wurzel:** $x = \sqrt{y}$ *mit* $y \in [0, \infty)$ *ist die* **positive** *Lösung von* $x^2 = y$.

Elementare Rechenregeln 1.2:
Einige aus der Schule bekannte Rechenregeln für's Potenzieren:

- $x^0 = 1$,
- $x^{-\alpha} = \frac{1}{x^\alpha}$,
- $x^{\alpha+\beta} = x^\alpha \cdot x^\beta$,
- $(x^\alpha)^\beta = x^{\alpha \cdot \beta}$.

Eine logische Aussage ist eine Aussage, der man genau einen der Wahrheitswerte „wahr" oder „falsch" zuordnen kann. Beispiele:

(A1) Die Zahl 5 ist eine Primzahl.

(A2) Es gibt eine größte Primzahl.

(A3) Die Zahl $2^{(2^{(2^{1234567})})} + 123$ ist eine Primzahl.

Hierbei ist A1 wahr und A2 falsch, während der Wahrheitsgehalt von A3 unklar ist (aber entweder 'wahr' oder 'falsch' sein muss).

Logische Aussagen werden über die Symbole

$$\begin{array}{rcl} A_1 \Rightarrow A_2 & \equiv & \text{„aus } A_1 \text{ folgt } A_2\text{“} \\ & \equiv & \text{„}A_2 \text{ ist wahr, wenn } A_1 \text{ wahr ist“} \\ & \equiv & \text{„}A_1 \text{ ist falsch, wenn } A_2 \text{ falsch ist“,} \\ A_1 \Leftrightarrow A_2 & \equiv & \text{„}A_2 \text{ ist genau dann wahr, wenn } A_1 \text{ wahr ist“} \\ & \equiv & \text{„}A_2 \text{ ist genau dann falsch, wenn } A_1 \text{ falsch ist“} \\ & \equiv & \text{„}A_1 \text{ und } A_2 \text{ sind äquivalent“} \end{array}$$

zu neuen logischen Aussagen miteinander verknüpft. Beispiele:

(a) Ist n eine Primzahl, so ist $n+1$ keine Primzahl.

(b) Ist n ein ungerade Primzahl, so ist $n+1$ keine Primzahl.

(c) $x \in \mathbb{R}$ ist genau dann positiv, wenn $(x+1)^2 > (x-1)^2$ gilt.

Mit

$$\begin{array}{ll} a_1: & n \text{ ist eine Primzahl,} \\ a_2: & n+1 \text{ ist keine Primzahl,} \\ a_3: & n \text{ ist eine ungerade Primzahl,} \\ a_4: & x \in \mathbb{R} \text{ ist positiv,} \\ a_5: & \text{für } x \text{ gilt } (x+1)^2 > (x-1)^2 \end{array}$$

werden die obigen Aussagen kürzer als

$$(a): \quad a_1 \;\Rightarrow\; a_2, \qquad (b): \quad a_3 \;\Rightarrow\; a_2, \qquad (c): \quad a_4 \;\Leftrightarrow\; a_5$$

formuliert. Hierbei ist (a) nicht für jedes n wahr (die Ausnahme ist $n = 2$), während b) und c) für jedes n bzw. x wahr sind.

Typische Aussage werden mit Hilfe „logischer Quantoren“ formuliert:

$$\begin{array}{ll} \forall & \text{„für alle“,} \\ \exists & \text{„es gibt (mindestens) ein“,} \\ \exists! & \text{„es gibt genau ein“,} \\ \nexists & \text{„es gibt kein“.} \end{array}$$

Beispiele:

(A) $\forall x \in \mathbb{R}\ \exists y \in \mathbb{R}: \ x = y^2$.

Klartext: für jedes $x \in \mathbb{R}$ gibt es (mindestens) ein $y \in \mathbb{R}$ mit $x = y^2$.

(B) $\forall x \in [0, \infty)\ \exists y \in \mathbb{R}: \ x = y^2$.

Klartext: Für jedes $x \in [0, \infty)$ gibt es (mindestens) ein $y \in \mathbb{R}$ mit $x = y^2$.

(C) $\forall x \in [0, \infty)\ \exists! y \in \mathbb{R} : x = y^2$.

Klartext: Für jedes $x \in [0, \infty)$ gibt es genau ein $y \in \mathbb{R}$ mit $x = y^2$.

(D) $\forall x \in (-\infty, 0) \quad \nexists y \in \mathbb{R} : x = y^2$.

Klartext: Für jedes $x \in (-\infty, 0)$ gilt: es gibt kein reelles y mit $x = y^2$. In besserem Deutsch: Für kein $x \in (-\infty, 0)$ gibt es ein reelles y mit $x = y^2$.

Hierbei ist (A) falsch, (B) richtig, (C) falsch und (D) richtig. Man sollte diese Schreibweisen kennen, um Aussagen in Mathe-Büchern lesen zu können. Ich werde hier versuchen, die Quantorenschreibweise weitmöglichst zu vermeiden und die Aussagen im Klartext ausformulieren.

1.2 Die natürlichen Zahlen, vollständige Induktion

15.10.03↓

Die Menge der natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \dots\}$ bzw. $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$ ist aus der Schule wohlvertraut. Man kann $\mathbb{N}$ mathematisch durch die folgenden Axiome (die „Peano-Axiome“) festlegen:

Axiomatisierung von $\mathbb{N}$ 1.3:

(P1) Es gibt ein ausgezeichnetes („kleinstes“) Element $1 \in \mathbb{N}$.

(P2) Es gibt eine „Nachfolgerfunktion“ $\nu : \mathbb{N} \to \mathbb{N} \setminus \{1\}$ *mit den Eigenschaften:*

(P2a) Für jedes Paar $n_1, n_2 \in \mathbb{N}$ *gilt:* $n_1 \neq n_2 \;\Rightarrow\; \nu(n_1) \neq \nu(n_2)$.

(P2b) Besitzt eine Teilmenge $A \subset \mathbb{N}$ *die Eigenschaften*
1) $1 \in A$,
2) $n \in A \quad \Rightarrow \quad \nu(n) \in A$,
so folgt $A = \mathbb{N}$.

Die durch diese Eigenschaften bestimmte Menge $\mathbb{N} = \{1, \nu(1), \nu(\nu(1)), \dots\}$ ist uns als $\mathbb{N} = \{1, 2, 3, \dots, \}$ vertraut. Wir benutzen dabei die übliche Bezeichnung

$$n = \underbrace{\nu(\dots(\nu(}_{n-1} 1 \underbrace{))\dots)}_{n-1}.$$

Die Nachfolgerfunktion $\nu(n) = n + 1$ ist eine schrecklich umständliche Notation und wird durch die übliche Addition ausgedrückt. Das Axiom (P2b) ist das „Prinzip der vollständigen Induktion“:

Sei $A(n)$ eine Aussage über eine natürliche Zahl n. Können wir $A(1)$ beweisen und zeigen, dass aus der Aussage $A(n)$ die Aussage $A(n+1)$ folgt, so folgt aus (P2), dass die Menge $A := \{n \in \mathbb{N}; A(n) \text{ ist richtig}\}$ ganz $\mathbb{N}$ ist, also $A(n)$ für alle n in $\mathbb{N}$ gilt.

Beispiel 1.4: Behauptung:

$$A(n): \qquad \sum_{i=1}^{n} i \;=\; \frac{n\cdot(n+1)}{2}.$$

Beweis durch Induktion nach n:

<u>Induktionsstart</u>: $A(1)$ ist die Aussage $\sum_{i=1}^{1} i = \frac{1\cdot(1+1)}{2}$. Diese Aussage $A(1):\ 1=1$ ist trivialerweise richtig.

<u>Schritt $n \to n+1$</u>: Für irgendein gegebenes n gelte $A(n)$, also $\sum_{i=1}^{n} = n\cdot(n+1)/2$. Wir folgern hieraus $A(n+1)$:

$$\sum_{i=1}^{n+1} i = \Big(\sum_{i=1}^{n} i\Big) \;+\; (n+1) \stackrel{A(n)}{=} \frac{n\cdot(n+1)}{2} \;+\; n+1$$

$$= \frac{n^2+n+2\cdot n+2}{2} = \frac{(n+1)\cdot(n+2)}{2}.$$

Unter Benutzung der Gültigkeit von $A(n)$ haben wir die Gültigkeit von $A(n+1)$ bewiesen. Das Prinzip der vollständigen Induktion sagt uns, dass $A(n)$ für alle $n\in\mathbb{N}$ richtig ist.

1.3 Die rationalen Zahlen

Wir haben $\mathbb{N}$ über die Peano-Axiome mathematisch sauber definiert. Angenommen, wir hätten die ganzen Zahlen $\mathbb{Z}$ (aus vielen Schuljahren zumindestens intutitiv bekannt) mathematisch sauber definiert (was wir hier nicht haben), so könnten wird die Menge der rationalen Zahlen als die Menge aller Brüche

$$\mathbb{Q} = \Big\{\frac{z}{n};\ z\in\mathbb{Z};\ n\in\mathbb{N}\Big\}$$

definieren. Eine rationale Zahl wäre somit so etwas wie das Paar

$$\frac{z}{q} \equiv (\text{Zähler } z, \text{ Nenner } n).$$

Hier gibt es ein formales Problem, da z.B. $1/2 \equiv (1,2)$ und $2/4 \equiv (2,4)$ übereinstimmen sollen. Man kann das „Kürzen" von Brüchen mathematisch sauber fassen, indem man „Äquivalenzklassen" von Paaren einführt:

$$(z,n) \sim (\tilde z, \tilde n),\ \text{ wenn } z\cdot\tilde n = \tilde z\cdot n \text{ gilt.}$$

Hier gehen Physikstudent und Mathematikstudentin nun zunächst eine Weile getrennte Wege (können später aber wieder zueinander finden). Wir akzeptieren als Physiker den aus der Schule bekannten „intuitiven“ Umgang mit rationalen Zahlen über die vertrauten Rechenregeln wie z.B.

$$\frac{z_1}{n_1} + \frac{z_2}{n_2} = \frac{z_1 \cdot n_2 + n_1 \cdot z_2}{n_1 \cdot n_2}, \qquad \frac{z_1}{n_1} \cdot \frac{z_2}{n_2} = \frac{(z_1 \cdot z_2)}{(n_1 \cdot n_2)}.$$

Die Mathematikerin würde formal eine Addition und eine Multiplikation auf den Äquivalenzklassen der Paare (Zähler, Nenner) definieren und dann zeigen, dass diese Definitionen sich als die obige Addition/Multiplikation von Brüchen interpretieren lassen. Sind auf diese Weise alle aus der Schule vertrauten Rechenregeln formal bewiesen, trifft man sich wieder und entwickelt die Theorie weiter. Der Physiker-Weg des „es wird schon alles zusammenpassen“ ist dabei deutlich kürzer.

Satz 1.5: (Abzählbarkeit von $\mathbb{Q}$)

Die Menge $\mathbb{Q}$ der rationalen Zahlen ist abzählbar, d.h., es gibt eine invertierbare Abbildung von $\mathbb{N}$ nach $\mathbb{Q}$.

Beweis: Wir betrachten zunächst die positiven rationalen Zahlen

$$\mathbb{Q}_+ = \Big\{ \frac{z}{n};\ z \in \mathbb{N}, n \in \mathbb{N} \Big\}.$$

Jedes $q \in \mathbb{Q}_+$ findet sich nach Kürzen von Nenner und Zähler an einem eindeutig bestimmten Punkt des Schemas

$n\backslash z$	1	2	3	$\ldots$
1	$*$	$*$	$*$	$\ldots$
2	$*$	$*$	$*$	$\ddots$
3	$*$	$*$	$*$	$\ddots$
$\vdots$	$\vdots$	$\ddots$	$\ddots$	$\ddots$

Dieses Schema wird links oben startend wie folgt durchlaufen:

$n\backslash z$	1	2	3	$\ldots$
1	$\rightarrow$	$\swarrow$	$\rightarrow$	$\swarrow$
2	$\downarrow$	$\nearrow$	$\swarrow$	
3	$\nearrow$	$\swarrow$		
$\vdots$	$\downarrow$	$\nearrow$		
$\vdots$	$\nearrow$			

Überspringt man dabei die Stellen, die nach Kürzen von Zähler und Nenner einer früher schon gefundenen rationalen Zahlen entsprechen, hat man die invertierbare Abbildung $\mathbb{N} \to \mathbb{Q}_+$ explizit konstruiert (wenngleich ich mich scheue, die Abbildungsvorschrift explizit hinzuschreiben). Analog ist die Menge $\mathbb{Q}_-$ der negativen rationalen Zahlen abzählbar (denn $\mathbb{Q}_+$ und $\mathbb{Q}_-$ lassen sich durch $q \in \mathbb{Q}_+ \to -q \in \mathbb{Q}_-$ ein-eindeutig aufeinander abbilden). Die Vereinigung $\mathbb{Q}_+ \cup \mathbb{Q}_-$ abzählbarer Mengen ist offensichtlich wieder abzählbar (Aufgabe 5, Blatt 1) und damit auch $\mathbb{Q} = \{0\} \cup \mathbb{Q}_+ \cup \mathbb{Q}_-$.

Q.E.D.

1.4 Die reellen Zahlen

Die reellen Zahlen und die komplexen Zahlen des nächsten Abschnitts sind die Zahlen, mit denen es der Physiker meistens zu tun hat.

Axiomatisierung von $\mathbb{R}$ 1.6:

Die Menge $\mathbb{R}$ der reellen Zahlen mit der Addition

$$+: \quad (x,y) \in \mathbb{R} \times \mathbb{R} \quad \to \quad x + y \in \mathbb{R}$$

und der Multiplikation

$$\cdot: \quad (x,y) \in \mathbb{R} \times \mathbb{R} \quad \to \quad x \cdot y \in \mathbb{R}$$

hat folgende Eigenschaften:

(R0) Addition und Multiplikation sind **kommutativ**, **assoziativ** *und* **distributiv**, *d.h., für alle $x, y, z \in \mathbb{R}$ gilt:*

$$x + y = y + x, \ (x + y) + z = x + (y + z),$$
$$x \cdot y = y \cdot x, \ (x \cdot y) \cdot z = x \cdot (y \cdot z),$$
$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

(R1) Es gibt zwei ausgezeichnete Elemente 0 und 1 in $\mathbb{R}$, für die $x + 0 = x$ und $1 \cdot x = x$ für alle $x \in \mathbb{R}$ gilt. Bezeichnung: 0 ist das **„neutrale Element der Addition“**, *1 ist das* **„neutrale Element der Multiplikation“**.

(R2) (Inverse bzgl. der Addition) Zu jedem $x \in \mathbb{R}$ gibt es genau ein Element $\tilde{x}$ mit $x + \tilde{x} = 0$. Bezeichnung: $\tilde{x} = -x$.

(R3) (Inverse bzgl. der Multiplikation) Zu jedem $x \in \mathbb{R} \setminus \{0\}$ gibt es genau ein Element $\hat{x} \in \mathbb{R} \setminus \{0\}$ mit $x \cdot \hat{x} = 1$. Bezeichnung: $\hat{x} = 1/x = x^{-1}$.

(R4) (Anordnung) Es gibt eine Teilmenge $\mathbb{R}_+ \subset \mathbb{R}$ (die positiven reellen Zahlen), für die gilt:

$$0 \neq x \in \mathbb{R}_+ \ \Rightarrow \ -x \notin \mathbb{R}_+,$$

$$0 \neq x \notin \mathbb{R}_+ \;\Rightarrow\; -x \in \mathbb{R}_+,$$

$$x, y \in \mathbb{R}_+ \;\Rightarrow\; x + y \in \mathbb{R}_+,\; x \cdot y \in \mathbb{R}_+.$$

Bezeichnung: wir schreiben $x < y$, wenn $y - x \in \mathbb{R}_+$ gilt. Wir schreiben $x \leq y$, wenn $x < y$ oder $x = y$ gilt.

Definition: *Eine Zahl $\beta \in \mathbb{R}$ heißt* **obere Schranke** *einer Teilmenge $A \subset \mathbb{R}$, wenn $a \leq \beta$ gilt für alle $a \in A$. Eine Teilmenge $A \subset \mathbb{R}$ heißt* **nach oben beschränkt**, *wenn es (mindestens) eine obere Schranke von A gibt.*

(R5) (Supremumsaxiom) Jede nach oben beschränkte nicht-leere Teilmenge $A \subset \mathbb{R}$ besitzt eine kleinste obere Schranke in $\mathbb{R}$, die als das **„Supremum von A"**

$$\sup A = \min\{\beta \in \mathbb{R}; a \leq \beta \;\forall a \in A\} \in \mathbb{R}.$$

bezeichnet wird.

Wir sind aus der Schule bereits an viele weitere Vereinbarungen und Notationen gewöhnt. Beispielsweise schreiben wir einfach

$$x + y + z,\; x \cdot y \cdot z,$$

denn die Assoziativität garantiert, dass eine Klammerung wie z.B. $(x + y) + z$ bzw. $x + (y + z)$ nicht nötig ist. Wir benutzen die abkürzenden Notationen

$$x - y := x \cdot (-y), \qquad \frac{x}{y} := x \cdot y^{-1}$$

oder benutzen auch

$$y \geq x \quad \text{statt} \quad x \leq y$$

usw. Die aus der Schule vertrauten Rechenregeln lassen sich durch mühselige Kleinarbeit alle aus den Axiomen (R0) – (R5) herleiten („beweisen"). Hier nur ein paar kleine Beispiele.

Behauptung: Man darf Ungleichungen addieren.
Beweis: Für $i = 1, 2$ gilt: $x_i < y_i$ bedeutet $y_i - x_i \in \mathbb{R}_+$. Nach den Axiomen folgt $y_1 - x_1 + y_2 - x_2 \in \mathbb{R}_+$, was als $x_1 + x_2 < y_1 + y_2$ geschrieben wird, also:

$$x_1 < y_1,\; x_2 < y_2 \quad \Rightarrow \quad x_1 + x_2 < y_1 + y_2.$$

Q.E.D.

Behauptung: $0 \cdot x = 0 \; \forall x \in \mathbb{R}$.
Beweis: Aus den Axiomen folgt:

$$0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x.$$

Addiert man das nach den Axiomen existierende $-(0 \cdot x)$ zu dieser Gleichung, erhält man $0 = 0 \cdot x$.

Q.E.D.

Behauptung: $(-1) \cdot x = -x \; \forall x \in \mathbb{R}$.
Beweis: Mit dem Distributivgesetz:

$$x + (-1) \cdot x = x \cdot (1 + (-1)) = x \cdot 0 = 0.$$

Also ist $(-1) \cdot x$ das Inverse von x bzgl. der Addition, welches als $-x$ bezeichnet wird.

Q.E.D.

Aus der Schule ist folgende "Rechenregel" bekannt:

> „Bei Multiplikation einer Ungleichung mit einer negativen Zahl dreht sich die Ungleichung um": $x < 0, \; y < z \quad \Rightarrow \quad x \cdot y > x \cdot z$.

Beweis: $x < 0$, $y < z$ bedeuten $-x \in \mathbb{R}_+$, $z - y \in \mathbb{R}_+$. Aufgrund der Axiome für $\mathbb{R}_+$ folgt $(-x) \cdot (z - y) \in \mathbb{R}_+$, also $x \cdot y - x \cdot z \in \mathbb{R}_+$, was wiederum als $x \cdot z < x \cdot y$ geschrieben wird.

Q.E.D.

↓16.10.03

Körper-Axiome 1.7:

> *Eine Menge $\mathbb{R}$ mit einer Addition und einer Multiplikation, die die Eigenschaften (R0) – (R3) erfüllt, heißt* **Körper**. *Ein Körper, der zusätzlich die Eigenschaft (R4) hat, heißt* **angeordneter Körper**.

Die Menge der rationalen Zahlen $\mathbb{Q}$ erfüllt offensichtlich die Eigenschaften (R0) – (R4), also: auch $\mathbb{Q}$ ist ein angeordneter Körper. Was also ist der Unterschied zwischen $\mathbb{Q}$ und $\mathbb{R}$? Es ist das Supremumsaxiom (R5)! Historisch gesehen geht es um die Tatsache, dass es Zahlen wie $\sqrt{2} \in \mathbb{R}$ gibt, die sich (zum Entsetzen der alten Griechen) als irrational herausstellten:

Satz 1.8: ($\sqrt{2}$ ist irrational)

Es gibt keine rationale Zahl $q \in \mathbb{Q}$ mit $q^2 = 2$.

Beweis: Wir brauchen nur die Nichtexistenz von positiven $q \in \mathbb{Q}_+$ mit $q^2 = 2$ zu betrachten, denn mit $q^2 = 2$ gilt auch $(-q)^2 = 2$, d.h., mit einer negativen rationalen Lösung von $q^2 = 2$ hätten wir automatisch auch eine positive rationale Lösung.

Die typische Formulierung ist: wir können o.B.d.A. („ohne Beschränkung der Allgemeinheit") $q > 0$ voraussetzen.

Wir führen einen **indirekten Beweis** zur Nichtexistenz von $q \in \mathbb{Q}_+$ mit $q^2 = 2$. Angenommen, es existiert $q = z/n$ mit $z \in \mathbb{Z}, n \in \mathbb{N}$ und $q^2 = 2$. Wir setzen voraus, dass z und n teilerfremde natürliche Zahlen sind (wenn nicht, ersetzen wir z und n durch die teilerfremden Werte, die nach Kürzen entstehen). Aus $q^2 = 2$ folgt

$$z^2 = 2 \cdot n^2 \quad \text{mit} \quad z, n \in \mathbb{N}.$$

Hieraus folgt, dass z gerade sein muss (sonst könnte z^2 nicht den Faktor 2 enthalten). Damit enthält z^2 den Faktor 4. Damit muss n gerade sein, denn sonst wäre $2 \cdot n^2$ nicht durch 4 teilbar.

Sowohl z als auch n müssen also durch 2 teilbar sein, wenn $(z/n)^2 = 2$ gilt. Widerspruch zur Teilerfremdheit von z und n! Das einzig zweifelhafte Glied der Beweiskette ist die Annahme, es gäbe ein $q = z/n \in \mathbb{Q}_+$ mit $q^2 = 2$. Also muss diese Annahme falsch gewesen sein.

Q.E.D.

Beispiel 1.9: Wir wissen, dass $\sqrt{2} \notin \mathbb{Q}$ gilt Wir können nun aber die Existenz von $\sqrt{2}$ als die positive Lösung von $x^2 = 2$ mit $x \in \mathbb{R}$ **beweisen**. Wir setzen ganz konstruktiv

$$\beta := \sup \{x \in \mathbb{R}_+ ; \ x^2 \leq 2\}$$

und beweisen, dass $\beta \in \mathbb{R}$ existiert und $\beta^2 = 2$ gilt.

Für die Existenz von β ist nur zu zeigen, dass die betrachtete Menge

$$A = \{x \in \mathbb{R}_+ ; \ x^2 \leq 2\}$$

nach oben beschränkt ist. Dies ist leicht: aus $x^2 \leq 2$ folgt sicherlich $x < 1.5$, denn für $x \geq 1.5$ würde $x^2 \geq 1.5^2 = 2.25 > 2$ folgen, wir betrachten aber nur x mit $x^2 \leq 2$. Also ist 1.5 eine obere Schranke der Menge A.

Als allgemeines Prinzip gilt, dass das Supremum einer nach oben beschränkten Menge eindeutig ist. (Dies ist ein Quickie. Die einzige wirkliche Herausforderung besteht darin, dies auf einem Zeitungsrand zu beweisen).

Damit stellt obiges β wegen des Supremumsaxioms eine eindeutig definierte Zahl in $\mathbb{R}$ dar. Es bleibt zu beweisen, dass $\beta^2 = 2$ gilt. Das ist etwas mühseliger.

Behauptung: $\beta = \sup\,\{x \in \mathbb{R}_+ ;\ x^2 \leq 2\}$ erfüllt $\beta^2 = 2$.

Beweis: Sicherlich gilt $\beta > 0$, also speziell $2 + \beta > 0$.

Angenommen, es gilt $\beta^2 < 2$. Dann kann β keine obere Schranke von A sein, denn z.B. die Zahl

$$a = \beta + \frac{2-\beta^2}{\underbrace{2+\beta}_{>0}} = \frac{2\cdot\beta + \beta^2 + 2 - \beta^2}{2+\beta} = \frac{2+2\cdot\beta}{2+\beta}$$

wäre echt größer als β. Sie liegt aber in A, denn es gilt

$$\begin{aligned} a^2 &= \frac{4+8\cdot\beta+4\cdot\beta^2}{(2+\beta)^2} = \frac{(8+8\cdot\beta+2\cdot\beta^2)-2\cdot(2-\beta^2)}{(2+\beta)^2} \\ &= \frac{2\cdot(2+\beta)^2-2\cdot(2-\beta^2)}{(2+\beta)^2} = 2-2\cdot\frac{2-\beta^2}{\underbrace{(2+\beta)^2}_{>0}} < 2. \end{aligned}$$

Widerspruch zur Annahme, dass β obere Schranke von A ist!

Angenommen, es gilt $\beta^2 > 2$. Dann kann β nicht die kleinste obere Schranke von A sein. Z. B. ist die Zahl

$$\beta' = \beta + \frac{2-\beta^2}{\underbrace{2+\beta}_{<0}} = \frac{2\cdot\beta + \beta^2 + 2 - \beta^2}{2+\beta} = \frac{2+2\cdot\beta}{2+\beta}$$

kleiner als β. Sie ist aber eine obere Schranke von A, denn wegen

$$\begin{aligned} \beta'^2 &= \frac{4+8\cdot\beta+4\cdot\beta^2}{(2+\beta)^2} = \frac{(8+8\cdot\beta+2\cdot\beta^2)+2\cdot(\beta^2-2)}{(2+\beta)^2} \\ &= \frac{2\cdot(2+\beta)^2+2\cdot(\beta^2-2)}{(2+\beta)^2} = 2+2\cdot\frac{\beta^2-2}{\underbrace{(2+\beta)^2}_{>0}} > 2 \end{aligned}$$

gilt:

$$a \in A \quad \Rightarrow \quad a^2 \leq 2 \quad \Rightarrow \quad a^2 < \beta'^2 \quad \Rightarrow \quad a < \beta'.$$

(Im letzten Schritt wird ausgenutzt, dass wir bereits wissen, dass $\beta' = (2{+}2{\cdot}\beta)/(2{+}\beta) > 0$ gilt, da sicherlich $\beta > 0$ gilt.) Widerspruch!

Wir haben gezeigt, dass weder $\beta^2 < 2$ noch $\beta^2 > 2$ gelten kann. Es bleibt nur $\beta^2 = 2$.
Q.E.D.

Analog zu oberen Schranken heißt $\alpha \in \mathbb{R}$ **untere Schranke** einer Teilmenge $A \subset \mathbb{R}$, wenn $\alpha \leq a$ für alle $a \in A$ gilt. Wir führen die größte untere Schranke als das **Infimum** von A ein, indem wir es auf das Supremum zurückführen:

$$\inf A = -\sup\{-a;\ a \in A\}.$$

Man kann leicht nachrechnen, dass dies in der Tat die größte untere Schranke von A ist. Da die Existenz von Suprema/Infima eine so wichtige strukturelle Eigenschaft der reellen Zahlen darstellt, stellen wir das Supremumsaxiom noch einmal heraus:

Das Supremumsaxiom 1.10:

*Jede nach oben beschränkte nichtleere Teilmenge A von $\mathbb{R}$ besitzt eine kleinste obere Schranke (**„das Supremum“ von** A):*

$$\sup A = \min\{\beta \in \mathbb{R};\ a \leq \beta\ \forall a \in A\}.$$

*Jede nach unten beschränkte nichtleere Teilmenge A von $\mathbb{R}$ besitzt eine größte untere Schranke (**„das Infimum“ von** A):*

$$\inf A = \max\{\alpha \in \mathbb{R};\ \alpha \leq a\ \forall a \in A\}.$$

Die Existenz des Minimums/Maximums aller oberen/unteren Schranken, welche das Supremum/Infimum definieren, ist die gewünschte Vollständigkeit der reellen Zahlen, die $\mathbb{R}$ z.B. von $\mathbb{Q}$ unterscheidet und uns die Existenz der über $\mathbb{Q}$ hinausgehenden irrationalen Zahlen wie Wurzeln etc. garantiert. In der Tat ist „$\mathbb{R}$ deutlich größer als $\mathbb{Q}$". Beispielsweise können die reellen Zahlen nicht mehr durch $\mathbb{N}$ indiziert werden (d.h., es gibt keine Zahlenfolge (x_n), die $\mathbb{R}$ vollständig durchläuft):

Satz 1.11: (Überabzählbarkeit von $\mathbb{R}$)

Die Menge $\mathbb{R}$ der reellen Zahlen ist nicht abzählbar, d.h., es gibt keine invertierbare Abbildung von $\mathbb{N}$ nach $\mathbb{R}$.

Beweis: (Nach Chr. Blatter, Analysis I) Betrachte eine *beliebige* Abbildung $n \in \mathbb{N} \to x_n \in \mathbb{R}$. Wir können stets (mindestens) eine Zahl in $\mathbb{R}$ konstruieren, die nicht in der Menge $\{x_n; n \in \mathbb{N}\}$ enthalten ist. Dazu definieren wir die abgeschlossenen Intervalle

$$[a_1, b_1] = [x_1 + 1, x_1 + 2],$$

und

$$[a_{n+1}, b_{n+1}] = \begin{cases} [a_n, \frac{3\cdot a_n + b_n}{4}], & \text{falls } x_{n+1} \geq \frac{a_n + b_n}{2}, \\ [\frac{a_n + 3\cdot b_n}{4}, b_n] & \text{falls } x_{n+1} < \frac{a_n + b_n}{2}. \end{cases}$$

Anschaulich: $[a_{n+1}, b_{n+1}]$ ist das linke Viertel von $[a_n, b_n]$, wenn x_{n+1} rechts von der Mitte sitzt und das rechte Viertel, wenn x_{n+1} links von der Mitte sitzt. Damit steigen die a_n monoton, die b_n fallen monoton, und es gilt

$$a_1 \le a_2 \le a_3 \le \dots \quad \le \quad \dots \le b_3 \le b_2 \le b_1,$$

also

$$[a_1, b_1] \supset [a_2, b_2] \supset [a_3, b_3] \supset \dots \supset [\alpha, \beta],$$

wobei $\alpha = \sup\{\alpha_n; n \in \mathbb{N}\}$, $\beta = \inf\{\beta_n; n \in \mathbb{N}\}$. Ausserdem gilt nach Konstruktion $x_n \notin [a_n, b_n]$, also

$$x_n \notin [a_n, b_n] \supset [a_{n+1}, b_{n+1}] \supset \dots \supset [\alpha, \beta].$$

Damit gibt es die beiden Zahlen $\alpha \in \mathbb{R}$ und $\beta \in \mathbb{R}$ (die evtl. übeinstimmen können), die von allen x_n verschieden sind.

Q.E.D.

Bemerkung 1.12: *Existiert in $A \subset \mathbb{R}$ ein maximales Element, so ist dieses Maximum auch das Supremum:*

$$\sup A = \max A.$$

Aber nicht jede beschränkte Menge hat ein maximales Element. Z.B. hat $A = [0, 1)$ kein Maximum, denn das Supremum $\sup A = 1$ (der einzige Kandidat für das Maximum) liegt nicht in A.

Merke: *Im Gegensatz zu Maxima/Minima existieren Suprema/Infima beschränkter Mengen in $\mathbb{R}$ immer!*

Zuletzt noch der für die Analysis unersetzliche Begriff des **Betrags** einer reellen Zahl

$$|x| = \begin{cases} x & \text{falls } x \ge 0, \\ -x & \text{falls } x < 0. \end{cases}$$

Über den Betrag erhält man den Begriff des **Abstands** $|x - y|$ zweier Zahlen. Die fundamentale Eigenschaft des Betrags- bzw. Abstandsbegriffs ist die:

Dreieckungleichung 1.13:

Für alle $x, y \in \mathbb{R}$ gilt die **„Dreiecksungleichung“**

$$|x + y| \le |x| + |y|$$

und die **„umgekehrte Dreiecksungleichung“***:*

$$\Big| |x| - |y| \Big| \le |x - y|.$$

Beweis: Die Dreiecksungleichung ist durch einige Fallunterscheidungen leicht zu überprüfen (Aufgabe 6, Blatt 1). Die umgekehrte Dreiecksungleichung folgt unmittelbar aus der Dreiecksungleichung. Ersetzt man x durch $x - y$, erhält man

$$|(x-y)+y| \leq |x-y| + |y|, \text{ also } |x| \leq |x-y| + |y|.$$

Ersetzt man y durch $y - x$, erhält man

$$|x+(y-x)| \leq |x| + |y-x|, \text{ also } |y| \leq |x| + |x-y|.$$

Die beiden resultierenden Ungleichungen

$$|x| - |y| \leq |x-y|, \quad |y| - |x| \leq |x-y|$$

lassen sich zur umgekehrten Dreiecksungleichung

$$\Big| |x| - |y| \Big| \leq |x-y|$$

zusammenfassen.

Q.E.D.

1.5 Die komplexen Zahlen

21.10.03↓

Motivation: die Gleichung $x^2 = -1$ hat offensichtlich keine reellen Lösungen, da $x^2 \geq 0$ für jedes reelle x gilt. Um auch diese Gleichung lösen zu können, muß man neue Zahlen einführen: die **komplexen Zahlen**. Die grundsätzliche Idee ist ganz einfach: man führt ein neues Symbol i ein, das $\sqrt{-1}$ repräsentieren soll. Es wird einzig und allein durch die Rechenregel $i^2 = -1$ festgelegt. Ansonsten behält man alle aus dem Reellen bekannten Rechenregeln einfach bei.

1.5.1 Definitionen

Definition 1.14: (Die komplexen Zahlen)

Die **Menge der komplexen Zahlen** $\mathbb{C}$ *ist die Menge aller formalen Summen der Form*

$$\mathbb{C} = \{x + i \cdot y;\ x, y \in \mathbb{R}\}.$$

Für $z = x + i \cdot y \in \mathbb{C}$ nennt man x den **Realteil** *und y den* **Imaginärteil** *von z.*
Zahlen $z = x + i \cdot 0$ mit $y = 0$ nennt man **reell**, *schreibt auch kurz $z = x$ und identifiziert z mit $x \in \mathbb{R}$.*
Zahlen $z = 0 + i \cdot y$ mit $x = 0$ nennt man **imaginär** *und schreibt auch kurz $z = i \cdot y$.*
Der **Nullpunkt** *$z = 0 + i \cdot 0$ wird auch kurz als $z = 0$ geschrieben.*
Auf $\mathbb{C}$ definieren wir die Addition

$$z_1 + z_2 = (x_1 + i \cdot y_1) + (x_2 + i \cdot y_2) = \underbrace{(x_1 + x_2)}_{\in\mathbb{R}} + i \cdot \underbrace{(y_1 + y_2)}_{\in\mathbb{R}}$$

sowie die Multiplikation

$$z_1 \cdot z_2 = (x_1 + i \cdot y_1) \cdot (x_2 + i \cdot y_2) = \underbrace{(x_1 \cdot x_2 - y_1 \cdot y_2)}_{\in\mathbb{R}} + i \cdot \underbrace{(x_1 \cdot y_2 + x_2 \cdot y_1)}_{\in\mathbb{R}}.$$

Interpretation 1.15:

Hinter dieser Definition der Multiplikation steckt $i^2 = -1$:

$$i \cdot i = (0 + i \cdot 1) \cdot (0 + i \cdot 1) = (0 \cdot 0 - 1 \cdot 1) + i \cdot (0 \cdot 1 + 1 \cdot 0) = -1.$$

Man braucht sich die formale Definition der Multiplikation nicht zu merken: man benutze einfach die üblichen aus $\mathbb{R}$ bekannten Rechenregeln (Kommutativität $a \cdot b = b \cdot a$, Assoziativität $(a \cdot b) \cdot c = a \cdot (b \cdot c)$, das Distributivgesetz $a \cdot (b + c) = a \cdot b + a \cdot c$ etc.), und setze beim Rechnen

$$i^2 = -1, \quad i^3 = (i^2) \cdot i = -i, \quad i^4 = (i^3) \cdot i = (-i) \cdot i = -(i^2) = 1$$

usw. ein:

$$(x_1 + i \cdot y_1) \cdot (x_2 + i \cdot y_2) = x_1 \cdot x_2 + i \cdot x_1 \cdot y_2 + i \cdot x_2 \cdot y_1 + \underbrace{i^2}_{-1} \cdot y_1 \cdot y_2$$

$$= x_1 \cdot x_2 + i \cdot x_1 \cdot y_2 + i \cdot x_2 \cdot y_1 - y_1 \cdot y_2 = (x_1 \cdot x_2 - y_1 \cdot y_2) + i \cdot (x_1 \cdot y_2 + x_2 \cdot y_1).$$

Folgerung 1.16:

Wir konstruieren eine Division für $z = x + i \cdot y \neq 0 + i \cdot 0 \equiv 0$:

$$\frac{1}{z} = \frac{1}{x + i \cdot y} = \frac{1}{x + i \cdot y} \cdot \frac{x - i \cdot y}{x - i \cdot y}$$

$$= \frac{x - i \cdot y}{(x + i \cdot y) \cdot (x - i \cdot y)} = \frac{x - i \cdot y}{x^2 - (i \cdot y)^2} = \frac{x}{x^2 + y^2} - i \cdot \frac{y}{x^2 + y^2}.$$

Allgemein:

$$\frac{z_1}{z_2} = \frac{x_1 + i \cdot y_1}{x_2 + i \cdot y_2} = \frac{(x_1 + i \cdot y_1) \cdot (x_2 - i \cdot y_2)}{(x_2 + i \cdot y_2) \cdot (x_2 - i \cdot y_2)}$$

$$= \frac{(x_1 \cdot x_2 + y_1 \cdot y_2) + i \cdot (-x_1 \cdot y_2 + x_2 \cdot y_1)}{x_2^2 + y_2^2 + i \cdot (x_2 \cdot y_2 - x_2 \cdot y_2)}$$

$$= \frac{x_1 \cdot x_2 + y_1 \cdot y_2}{x_2^2 + y_2^2} + i \cdot \frac{x_2 \cdot y_1 - x_1 \cdot y_2}{x_2^2 + y_2^2}.$$

Definition 1.17: (komplexe Konjugation etc.)
Es werden folgende speziellen Operationen auf den komplexen Zahlen eingeführt:

$$\begin{array}{rcccl} \Re(z) & = & \Re(x + i \cdot y) & = & x \\ \Im(z) & = & \Im(x + i \cdot y) & = & y \\ |z| & = & |x + i \cdot y| & = & \sqrt{x^2 + y^2} \\ \overline{z} & = & \overline{x + i \cdot y} & = & x - i \cdot y \end{array}$$

(der **Realteil** von z),
(der **Imaginärteil** von z),
(der **Betrag** von z),
(das **komplex Konjugierte** von z).

Merkregel 1.18:
Die Division komplexer Zahlen läuft auf den Standardtrick „Erweitern mit dem komplex konjugierten Nenner hinaus“:

$$\frac{z_1}{z_2} = \frac{x_1 + i \cdot y_1}{x_2 + i \cdot y_2} = \frac{(x_1 + i \cdot y_1) \cdot (x_2 - i \cdot y_2)}{(x_2 + i \cdot y_2) \cdot (x_2 - i \cdot y_2)} = \frac{z_1 \cdot \overline{z}_2}{z_2 \cdot \overline{z}_2}.$$

Satz 1.19: (Rechenregeln)
Für alle $z, z_1, z_2 \in \mathbb{C}$ gilt: „Kommutativität“ und „Assoziativität“ von Multiplikation und Division

$$z_1 \cdot z_2 = z_2 \cdot z_1, \quad (z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3),$$

$$\frac{1}{z_1 \cdot z_2} = \frac{1}{z_2} \cdot \frac{1}{z_1}, \quad \Big(\frac{1}{z_1} \cdot \frac{1}{z_2}\Big) \cdot \frac{1}{z_3} = \frac{1}{z_1} \cdot \Big(\frac{1}{z_2} \cdot \frac{1}{z_3}\Big),$$

„Linearität“ *von $\Re$, $\Im$ und Konjugation*

$$\Re(z_1 + z_2) = \Re(z_1) + \Re(z_2), \quad \Im(z_1 + z_2) = \Im(z_1) + \Im(z_2), \quad \overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2,$$

„Multiplikativität“ *des Betrags und der Konjugation*

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2|, \quad \left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}, \quad \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}, \quad \overline{\left(\frac{z_1}{z_2}\right)} = \frac{\overline{z_1}}{\overline{z_2}}$$

sowie die Beziehungen

$$|z|^2 = |\overline{z}|^2 = z \cdot \overline{z} = \Re(z)^2 + \Im(z)^2, \quad \frac{z_1}{z_2} = \frac{z_1 \cdot \overline{z}_2}{|z_2|^2}$$

und die **„Dreiecksungleichung“***:*

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

Beweis: Alles ist direkt nachzurechnen, z.B.

$$z \cdot \overline{z} = (x + i \cdot y) \cdot (x - i \cdot y) = x^2 + y^2 = |z|^2 = |\overline{z}|^2$$

oder (wie schon oben durchgeführt):

$$\frac{z_1}{z_2} = \frac{z_1 \cdot \overline{z_2}}{z_2 \cdot \overline{z_2}} = \frac{z_1 \cdot \overline{z_2}}{|z_2|^2}.$$

Q.E.D.

Beispiel 1.20: In MuPAD wird $i = \sqrt{-1}$ durch `I` dargestellt:

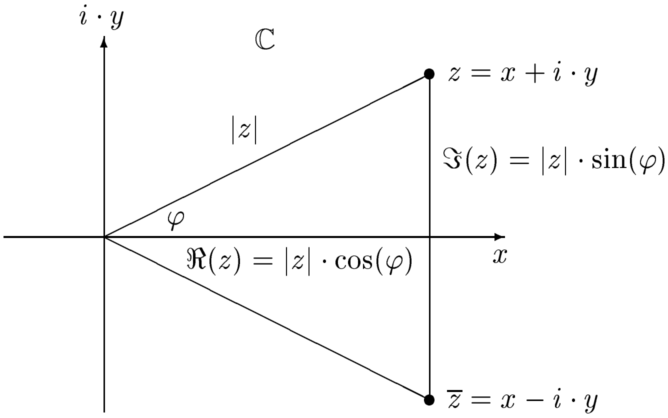
```
>> sqrt(-1)
                                 I
>> I^2
                                -1
```

Die MuPAD-Funktionen `Re`, `Im`, `conjugate` und `abs` berechnen Real- und Imaginärteil, komplexe Konjugation und den Absolutbetrag:

```
>> z:= 2 + 3*I:
>> Re(z), Im(z), conjugate(z), abs(z)
                                          1/2
                      2, 3, 2 - 3 I, 13
```

Geometrische Interpretation 1.21:

*Man stellt sich üblicherweise die Menge der komplexen Zahlen als 2-dimensionale Ebene („**die komplexe Ebene**") vor:*



*Der Betrag von z ist der Abstand zum Ursprung, komplexe Konjugation entspricht der Spiegelung an der x-Achse („**die reelle Achse**"). Die y-Achse wird auch als „**imaginäre Achse**" bezeichnet.*

Geometrisch ist $\mathbb{C}$ nichts anderes als $\mathbb{R}^2$: $x + i \cdot y \in \mathbb{C} \mathrel{\widehat{=}} (x, y) \in \mathbb{R}^2$. Die komplexe Addition entspricht genau der Addition von Vektoren im $\mathbb{R}^2$. Algebraisch besteht der Unterschied zwischen $\mathbb{C}$ und $\mathbb{R}^2$ darin, dass man auf $\mathbb{C}$ neben der Addition noch eine Multiplikation $\mathbb{C} \times \mathbb{C} \mapsto \mathbb{C}$ hat, wohingegen es auf $\mathbb{R}^2$ keine interessante Multiplikation zweier Vektoren gibt, die wieder einen Vektor liefert (außer derjenigen, die $\mathbb{R}^2$ zu $\mathbb{C}$ macht).

Bemerkung 1.22: *Führt man den eingezeichneten Winkel φ zwischen dem „Vektor" z und der reellen Achse ein, so gilt mit den aus der Schule bekannten Winkelfunktionen* sin *und* cos*:*

$$z = x + i \cdot y = |z| \cdot \cos(\varphi) + i \cdot |z| \cdot \sin(\varphi).$$

Die Darstellung $z = x + i \cdot y$ nennt man die **Kartesische Darstellung** *der komplexen Zahl z durch Real- und Imaginärteil. Die Darstellung*

$$z = |z| \cdot \Big(\cos(\varphi) + i \cdot \sin(\varphi) \Big)$$

durch den Betrag $|z|$ und den **„Polarwinkel"** *$\varphi \in [0, 2\pi)$ heißt* **Polardarstellung** *von z (φ heißt auch* **„das Argument von z"***). Wir werden später in Abschnitt 5.3 auf die Polardarstellung komplexer Zahlen zurückkommen, nachdem wir die komplexe Exponentialfunktion eingeführt haben.*

Bemerkung 1.23: *Im Gegensatz zu $\mathbb{R}$ ist $\mathbb{C}$* **prinzipiell nicht orientierbar***! Damit ist gemeint, dass es auf $\mathbb{C}$ keine Ordnung $z_1 < z_2$ im Sinne von (R4) der Axiomatisierung 1.6 geben kann, die mit den Körperaxiomen (R0) – (R3) verträglich ist (diese gelten auch für $\mathbb{C}$: mit der komplexen Addition und Multiplikation ist $\mathbb{C}$ in der Tat ein Körper). Dass* prinzipiell *keine Anordnung mit (R0) – (R4) auf $\mathbb{C}$ existieren kann, ist leicht gezeigt:*

Angenommen, es gäbe $\mathbb{C}_+$ (die Menge der positiven komplexen Zahlen). Das neutrale Element 1 der Multiplikation (dies ist auch auf $\mathbb{C}$ die reelle Zahl 1) muss in $\mathbb{C}_+$ liegen, damit liegt -1 nicht in $\mathbb{C}_+$. Liegt i in $\mathbb{C}_+$? Nein, denn sonst müsste nach (R4) die Zahl $i \cdot i = -1$ wieder in $\mathbb{C}_+$ liegen. Mit demselben Argument kann $-i$ wegen $(-i) \cdot (-i) = -1$ auch nicht in $\mathbb{C}_+$ liegen. Damit haben wir ein Element $z = i \neq 0$ mit der Eigenschaft, dass weder z noch $-z$ in $\mathbb{C}_+$ liegt. Dies verletzt (R4).

1.5.2 Nullstellen von Polynomen, der Fundamentalsatz der Algebra

$\boxed{\downarrow 22.10.03}$

Die Motivation zur Einführung der komplexen Zahlen war, Polynomgleichungen wie z.B. $x^2 + 1 = 0$ lösen zu können. In der Tat stellt sich nun heraus, dass Polynome vom Grad n immer genau n (evtl. „entartete") komplexe Nullstellen haben. Wir definieren zunächst „Entartung" von Nullstellen, wobei wir auf die aus der Schule bekannte Differentiation zurückgreifen:

Definition 1.24: (Vielfachheit von Nullstellen)

Sei $f : \mathbb{R} \mapsto \mathbb{R}$ eine mehrfach differenzierbare Funktion (siehe Kapitel 6). Man nennt x^* eine Nullstelle der **„Vielfachheit"** *k (oder auch* **„k-fache Nullstelle"***), wenn*

$$f(x^*) = f'(x^*) = \cdots = f^{(k-1)}(x^*) = 0, \quad f^{(k)}(x^*) \neq 0.$$

Beispiel 1.25: (Mehrfache Polynomwurzeln)
Für das Polynom $p(x) = (x - x^*)^n$ mit $n > 0$ ist x^* eine n-fache Nullstelle:

$$p(x) = (x - x^*)^n, \quad p'(x) = n \cdot (x - x^*)^{n-1}, \quad p''(x) = n \cdot (n-1) \cdot (x - x^*)^{n-2},$$

$$\ldots\ , \quad p^{(n-1)}(x) = n \cdot (n-1) \cdot \cdots \cdot 2 \cdot (x - x^*), \quad p^{(n)}(x) = n!,$$

also

$$\begin{aligned} p(x^*) &= (x^* - x^*)^n = 0, \\ p'(x^*) &= n \cdot (x^* - x^*)^{n-1} = 0, \\ p''(x^*) &= n \cdot (n-1) \cdot (x^* - x^*)^{n-2} = 0, \\ &\ldots \end{aligned}$$

$$p^{(n-1)}(x^*) = n \cdot (n-1) \cdot \dots \cdot 2 \cdot (x^* - x^*)^1 = 0,$$
$$p^{(n)}(x^*) = n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1 \cdot (x^* - x^*)^0 = n! \neq 0.$$

Beispiel 1.26: (Mehrfache Polynomwurzeln)
Seien $x_1, \dots, x_k$ verschieden. Für das Polynom

$$p(x) = (x - x_1)^{n_1} \cdot \dots \cdot (x - x_k)^{n_k}$$

vom Grad $n = n_1 + \dots + n_k$ sind $x_1, \dots, x_k$ Nullstellen der Vielfachheit $n_1, \dots, n_k \in \mathbb{N}$. Der Nachweis geht analog zum letzten Beispiel: Betrachte eine der Nullstellen x_i und schreibe

$$p(x) = (x - x_i)^{n_i} \cdot f(x), \quad f(x) = (x - x_1)^{n_1} \cdot \dots \cdot \cancel{(x - x_i)^{n_i}} \cdot \dots \cdot (x - x_k)^{n_k}.$$

Über die aus der Schule bekannte Produktregel $(g \cdot f)' = g' \cdot f + g \cdot f'$ der Differentiation folgt

$$p(x) = (x - x_i)^{n_i} \cdot f(x),$$
$$p'(x) = n_i \cdot (x - x_i)^{n_i - 1} \cdot f(x) + (x - x_i)^{n_i} \cdot f'(x),$$
$$p''(x) = n_i \cdot (n_i - 1) \cdot (x - x_i)^{n_i - 2} \cdot f(x) + 2 \cdot n_i \cdot (x - x_i)^{n_i - 1} \cdot f'(x) + (x - x_i)^{n_i} \cdot f''(x)$$

usw., wobei f, f', f'' etc. Polynome sind. Die ersten $n_i - 1$ Ableitungen verschwinden an der Stelle $x = x_i$:

$$p(x_i) = 0^{n_i} \cdot f(x_i) = 0\ ,$$
$$p'(x_i) = 0^{n_i - 1} \cdot f(x_i) + 0^{n_i} \cdot f'(x_i) = 0$$
$$p''(x_i) = (..) \cdot 0^{n_i - 2} \cdot f(x_i) + (..) \cdot 0^{n_i - 1} \cdot f'(x_i) + 0^{n_i} \cdot f''(x_i) = 0$$

usw. Die n_i-te Ableitung verschwindet nicht:

$$p^{(n_i)}(x) = n_i! \cdot f(x) + (..) \cdot (x - x_i) \cdot f'(x) + \dots + (x - x_i)^{n_i} \cdot f^{(n_i)}(x),$$

also $p^{(n_i)}(x_i) = n_i! \cdot f(x_i)$, wobei

$$f(x_i) = (x_i - x_1)^{n_1} \cdot \dots \cdot \cancel{(x_i - x_i)^{n_i}} \cdot \dots \cdot (x_i - x_k)^{n_k} \neq 0$$

gilt, da $x_i \neq x_1, \dots, x_i \neq x_k$ vorausgesetzt ist. Damit ist x_i eine Nullstelle der Vielfachheit n_i.

Ein Polynom und seine Ableitungen

$$p(x) = a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \dots + a_1 \cdot x + a_0,$$
$$p'(x) = a_n \cdot n \cdot x^{n-1} + a_{n-1} \cdot (n-1) \cdot x^{n-2} + \dots + a_1$$

etc. ist natürlich auch für komplexe Zahlen x wohldefiniert und man kann daher nach (mehrfachen) komplexen Nullstellen fragen. Die Definition 1.24 der Vielfachheit wird dabei auch für komplexe Nullstellen beibehalten.

Wir stellen zunächst das Horner-Schema zur Polynomauswertung und Polynomdivision vor:

Satz 1.27: (Polynomauswertung und -division per Horner-Schema)
Für das Polynom $p(x) = a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \cdots + a_0$ mit Koeffizienten $a_k \in \mathbb{C}$ gilt für jedes $x^* \in \mathbb{C}$:

$$\frac{p(x) - p(x^*)}{x - x^*} = b_0 \cdot x^{n-1} + b_1 \cdot x^{n-2} + \cdots + b_{n-2} \cdot x + b_{n-1},$$

*wobei b_0, b_1 etc. durch die Rekursion (**„Horner-Schema"**)*

$$\boxed{\begin{array}{l} b_0 := a_n \ ; \\ \textit{for } k := 1 \textit{ to } n \textit{ do } \ b_k := b_{k-1} \cdot x^* + a_{n-k} \ ; \end{array}}$$

gegeben sind. Es gilt $b_n = p(x^*)$.

Beweis: Mit $b_0 = a_n$, $b_k - b_{k-1} \cdot x^* = a_{n-k}$ und $-b_{n-1} \cdot x^* = a_0 - b_n$ folgt

$$(x - x^*) \cdot \Big(b_0 \cdot x^{n-1} + b_1 \cdot x^{n-2} + \cdots + b_{n-2} \cdot x + b_{n-1}\Big)$$

$$\begin{aligned}
= {} & b_0 \cdot x^n + b_1 \cdot x^{n-1} + b_2 \cdot x^{n-2} + \cdots + b_{n-1} \cdot x \\
& - \underbrace{b_0 \cdot x^*}\cdot x^{n-1} - \underbrace{b_1 \cdot x^*} \cdot x^{n-2} - \cdots - \underbrace{b_{n-2} \cdot x^* \cdot x} - \underbrace{b_{n-1} \cdot x^*} \\
= {} & a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + a_{n-2} \cdot x^{n-2} + \cdots + a_1 \cdot x + a_0 - b_n \\
= {} & p(x) - b_n .
\end{aligned}$$

Für $x = x^*$ folgt $0 = p(x^*) - b_n$ und damit

$$(x - x^*) \cdot \Big(b_0 \cdot x^{n-1} + b_1 \cdot x^{n-2} + \cdots + b_{n-2} \cdot x + b_{n-1}\Big) = p(x) - p(x^*).$$

Q.E.D.

Das Horner-Schema liefert mittels $b_n = p(x^*)$ die Auswertung des Polynoms an einer Stelle x^* mit n Multiplikationen und $n - 1$ Additionen. In der Tat ist es (für „dicht besetzte" Polynome) das Standardschema, mit dem auf dem Rechner Polynomauswertungen implementiert werden. Bei der Auswertung werden gleichzeitig die Koeffizienten $b_0, \ldots, b_{n-1}$ des **„Faktorpolynoms"** $(p(x) - p(x^*))/(x - x^*)$ mitgeliefert. Das Horner-Schema läuft auf die folgende Darstellung des Polynoms hinaus:

$$p(x^*) \;=\; \underbrace{((\ldots(\underbrace{(\underbrace{\underset{\substack{\uparrow\\ b_0 = a_n}}{a_n} \cdot x^* + a_{n-1}}_{b_1 = a_n \cdot x^* + a_{n-1}}) \cdot x^* + a_{n-2}}_{b_2 = a_n \cdot x^{*2} + a_{n-1} \cdot x^* + a_{n-2}}) \cdot x^* + \cdots) \ldots) \cdot x^* + a_1) \cdot x^* + a_0}_{\substack{\ddots \\ b_n = a_n \cdot x^{*n} + a_{n-1} \cdot x^{*n-1} + \cdots + a_1 \cdot x^* + a_0 = p(x^*)}}.$$

Ist x^* eine Nullstelle („Wurzel") des Polynoms, so folgt $p(x)/(x - x^*) =$ Polynom(x). Es ergibt sich das Grundprinzip, dass man bei einer gegebenen Nullstelle einen „Linearfaktor" $x - x^*$ vom Polynom abspalten kann:

Folgerung 1.28:

Ist x^* eine Wurzel des Polynoms p vom Grad $n > 0$, so gilt

$$p(x) = (x - x^*) \cdot q(x)$$

mit einem „Faktorpolynom" $q(x) = b_0 \cdot x^{n-1} + b_1 \cdot x^{n-2} + \cdots + b_{n-1}$ vom Grad $n - 1$, dessen Koeffizienten z.B. durch das Horner-Schema berechenbar sind.

Merke: *x^* ist dann und genau dann eine Wurzel, wenn sich der Linearfaktor $x - x^*$ vom Polynom abspalten läßt.*

Zwar hat nicht jedes Polynom reelle Nullstellen, aber es gilt das (zu beweisende) wichtige Prinzip:

Jedes Polynom vom Grad > 0 hat (mindestens) eine komplexe Nullstelle.

Setzen wir zur Motivation des kommenden Fundamentalsatzes 1.29 mal dieses Prinzip voraus. Es gilt $p(x) = (x - x^*) \cdot q(x)$ mit einer (garantiert existierenden) Nullstelle x^* von p. Die Nullstellen des Faktorpolynoms q sind offensichtlich wieder Nullstellen des Ausgangspolyms p. Hat man nun eine Nullstelle x^{**} von q, so kann man nach Folgerung 1.28 angewendet auf q einen weiteren Linearfaktor abspalten:

$$q(x) = (x - x^{**}) \cdot \tilde{q}(x), \quad \text{also} \quad p(x) = (x - x^*) \cdot (x - x^{**}) \cdot \tilde{q}(x)$$

mit einem Restpolynom $\tilde{q}$ vom Grad $n - 2$. Dies setzt man fort, bis man nach n Schritten auf ein konstantes Polynom stößt, das keine Nullstellen mehr besitzt. Es folgt eine Faktordarstellung

$$p(x) = (x - x^*) \cdot (x - x^{**}) \cdot (x - x^{***}) \cdot \cdots \cdot (x - x^{*\cdots*}) \cdot a_n,$$

wobei a_n das zuletzt verbleibende konstante Restpolynom (vom Grad 0) ist. Vergleicht man die führenden Koeffizienten auf der linken und rechten Seite dieser Gleichung, so sieht man sofort, dass das verbleibende konstante Restpolynom nichts anderes als der führende Koeffizient von $p(x) = a_n \cdot x^n + \cdots + a_0$ ist. Es folgt das Grundprinzip:

Jedes Polynom vom Grad $n > 0$ hat genau n komplexe Nullstellen.

Hierbei müssen wir aber etwas vorsichtig zählen, da in der Konstruktion die Nullstellen x^*, x^{**}, x^{***} etc. eventuell übereinstimmen können. Eine saubere Formulierung liefert der folgende fundamentale Satz:

Satz 1.29: (Fundamentalsatz der Algebra, Gauß 1799)

Zu jedem Polynom $p(x) = a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \cdots + a_0$ vom Grad $n > 0$ mit $a_0, \ldots, a_n \in \mathbb{C}$ und $a_n \neq 0$ gibt es komplexe Zahlen $z_1, \ldots, z_k$ (die Wurzeln) und $n_1, \ldots, n_k \in \mathbb{N}$ (die Vielfachheiten) mit $n_1 + \cdots + n_k = n$, so dass

$$p(x) = a_n \cdot (x - z_1)^{n_1} \cdot \cdots \cdot (x - z_k)^{n_k}.$$

Für die Anzahl k der *unterschiedlichen* Wurzeln gilt hierbei $k \leq n$ wegen $n_1 + \cdots + n_k = n$.

zum Beweis: Wie in der Motivation gezeigt, braucht man nur zu beweisen, dass jedes Polynom vom Grad > 0 mindestens eine komplexe Nullstelle bezitzt. Dies ist je nach den zur Verfügung stehenden Hilfsmitteln aber gar nicht so einfach und sprengt unseren Rahmen hier. Typischerweise wird der Satz in Lehrbüchern über komplexe Funktionentheorie bewiesen. Ein „elementarer“ Beweis findet sich z.B. unter:

`http://helios.mathematik.uni-kl.de/~luene/kleinodien/laplace.html`

Interpretation 1.30:

Nach Beispiel 1.26 sind $z_1, \ldots, z_k$ die Nullstellen von p mit den Vielfachheiten $n_1, \ldots, n_k$. Zählt man z_1 als n_1 Wurzeln, z_2 als n_2 Wurzeln etc., so ergeben sich insgesamt $n_1 + \cdots + n_k = n$ komplexe Wurzeln des Polynoms, und in der Tat erhalten wir in dieser Zählweise:

$$\boxed{\textit{Jedes Polynom vom Grad } n > 0 \textit{ hat genau } n \textit{ komplexe Nullstellen!}}$$

Über die Anzahl der **reellen** *Nullstellen hingegen kann man i.A. wenig aussagen (z.B. hat $p(x) = x^2 + 1$ überhaupt keine reelle Nullstellen).*

Merkregel 1.31:

> *Der Punkt x^* ist dann und genau dann eine Nullstelle eines Polynoms $p(x)$, wenn sich gemäß Folgerung 1.28 der Linearfaktor $x - x^*$ vom Polynom abfaktorisieren läßt. Die Vielfachheit von x^* gibt an, wie oft sich dieser Linearfaktor abspalten läßt. Man sollte eine Polynomwurzel besser als einen Linearfaktor ansehen. Der Fundamentalsatz besagt, dass sich ein Polynom vom Grad n immer in genau n Linearfaktoren aufspalten läßt.*

Die Existenz der komplexen Wurzeln sagt nichts darüber aus, ob man diese Wurzeln in irgendeiner Weise explizit darstellen kann. In der Tat gibt es z.B. für Polynome vom Grad ≥ 5 keine allgemeingültige Lösungsformel mit Hilfe von (verschachtelten) Wurzeln. Numerisch kann man jedoch stets Gleitpunktapproximationen der Wurzeln finden.

Beispiel 1.32: In MuPAD ist `solve` für exakte Lösungen und `numeric::solve` für numerische Lösungen zuständig. Das folgende Polynom hat 9 Wurzeln, die sich (zufälligerweise) alle explizit darstellen lassen:

```
>> p:= x^9 + 2*x^7 - x^3 - 2*x:
>> solve(p = 0, x)

                   1/2      1/2              1/2
{0, -1, 1, - I 2    , I 2    , - 1/2 I 3    - 1/2,

                   1/2           1/2                  1/2
   1/2 - 1/2 I 3    , 1/2 I 3    - 1/2, 1/2 I 3    + 1/2}
```

Der numerische Gleichungslöser liefert Gleitpunktnäherungen der Wurzeln:

```
>> numeric::solve(p = 0, x)

{0.0, - 0.5 - 0.8660254038 I, - 0.5 + 0.8660254038 I,

   0.5 - 0.8660254038 I, 1.0, -1.414213562 I, 1.414213562 I,

   0.5 + 0.8660254038 I, -1.0}
```

Die zurückgegebenen Objekte {...} sind jeweils Mengen, deren Elementanordnung willkürlich vom System nach internen Kriterien bestimmt wird. Diese sehen für exakte Werte anders aus als für Gleitpunktnäherungen, so dass sich die Reihenfolge der Elemente beim exakten und beim numerischen Lösen unterscheiden kann (was im obigen Beispiel auch in der Tat der Fall ist).

Es fällt hierbei auf, dass die komplexen Wurzeln als komplex konjugierte Paare $x_k \pm i \cdot y_k$ auftauchen. Das ist kein Zufall und liegt daran, dass das eben betrachtete Polynom „reell" ist (damit ist gemeint, dass die Koeffizienten reell sind).

Satz 1.33: (konjugierte Wurzelpaare reeller Polyome)

Ist z eine k-fache Nullstelle des Polynoms $p(x) = a_n \cdot x^n + \cdots + a_0$ mit **reellen** *Koeffizienten $a_0, \ldots, a_n$, so ist auch $\overline{z}$ eine k-fache Nullstelle des Polynoms. Bei reellen Polynomen tauchen nicht-reelle Wurzeln also immer in komplex-konjugierten Paaren auf.*

Beweis: Für ein reelles Polynom gilt wegen $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$ offensichtlich

$$\overline{p(z)} = p(\overline{z}).$$

Also gilt $p(\overline{z}) = 0$ dann und genau dann, wenn $p(z) = 0$ gilt. Da mit p auch alle Ableitungen von p wieder reelle Polynome sind, stimmen auch die Vielfachheiten der Nullstellen z und $\bar{z}$ überein.

Q.E.D.