Mathematik 1+2 für Informatik

(Grundlagen, eindimensionale Analysis und Lineare Algebra)

Vorlesungsskript

Thomas Markwig
Fachbereich Mathematik
Universität Tübingen

Studienjahr 2023/24

Inhaltsverzeichnis

Einleitung

Die vorliegende Ausarbeitung zur Vorlesung Mathematik 1 für Informatiker im Wintersemester 2019/20 wird im wesentlichen wiedergeben, was während der Vorlesung an die Tafel geschrieben wird. Einige wenige Abschnitte werden etwas ausführlicher sein. Manche Beweise, die in der Vorlesung nicht besprochen werden, finden sich der Vollständigkeit halber in der Ausarbeitung und sind durch das graue Schriftbild vom Rest unterschieden. Die Ausarbeitung erhebt nicht den Anspruch, ein Lehrbuch zu ersetzen.

In der Vorlesung Mathematik 1 für Informatiker lernen die Teilnehmer die grundlegenden mathematischen Begriffe kennen und werden in die Theorie der eindimensionalen Analysis eingeführt. Sie werden mit der sehr formalen mathematischen Sprache vertraut gemacht und lernen logisch korrekt und sauber zu argumentieren. Sie erwerben damit das grundlegende mathematische Handwerkszeug, das ein Studierender in den Studiengängen der Informatik für sein Fachstudium benötigt. Die Vorlesung wird sich im wesentlichen darauf beschränken, die mathematischen Theorien zu behandeln und nur sehr sporadisch einen Ausblick geben, wo und wie diese in der Informatik benötigt werden. Die Anwendungen der Mathematik in der Informatik werden stattdessen in den Veranstaltungen der Informatik thematisiert werden.

Teil A

Mathematik für Informatik 1

Kapitel I

Grundlegende Begriffsbildungen

Wir beginnen damit, grundlegende Begriffe einzuführen und zu besprechen, die für alle mathematischen Disziplinen gleich wichtig sind.

§ 1 Etwas Logik

Ausblick 1.0 (Aussagenlogik in der Informatik).
Die formale Aussagenlogik spielt in der Informatik in einer ganzen Reihe von Bereichen eine wichtige Rolle, z.B.

- beim Entwurf von Schaltkreisen in der Digitalelektronik (siehe Beispiel 1.7),
- bei der Verifikation,
- beim automatischen Beweisen,
- bei Anfragen an Suchmaschinen im Internet,
- beim Programmieren (z.B. Abbruchbedingungen von Schleifen).

In diesem Zusammenhang sind auch die Begriffe *Logikgatter*, siehe etwa

https://de.wikipedia.org/wiki/Logikgatter,

sowie *Satisfyability*, siehe etwa

http://www.mqasem.net/sat/sat/index.php,

interessant, die Anwendungen der logischen Operationen in der Informatik zeigen.

Nach diesem kurzen Ausblick, weshalb die Begrifflichkeiten der Aussagenlogik für Informatiker interessant sind, wollen wir uns dem Thema selbst zuwenden.

Wie alle Wissenschaftler versuchen auch die Mathematiker *Aussagen* über die Objekte ihrer Forschungsarbeit aufzustellen und *als wahr nachzuweisen*. Anders aber als etwa in den Naturwissenschaften werden die zu untersuchenden Objekte nicht von außen an die Mathematiker herangetragen, vielmehr schaffen sie sie sich selbst durch die Vorgabe

sogenannter *Axiome*. Wie hat man dies zu verstehen? Was ist ein Axiom? Was heißt es, eine Aussage als wahr nachzuweisen? Und was eigentlich ist eine Aussage?

Nun, sobald wir uns auf eine Sprache geeinigt haben, in der wir uns verständigen wollen, sind wir in der Lage, Sätze zu bilden, Sätze, wie etwa (in unserer Alltagssprache)

"Dieser Satz enthält fünf Worte."

oder

"Löse die folgende Aufgabe."

Ein solcher Satz stellt eine *Aussage* in unserem Sinne dar, wenn wir entscheiden können, ob er wahr oder falsch ist. Gemäß dieser Konvention ist der erste der obigen Sätze eine – wahre – Aussage, während beim zweiten Satz, einer Aufforderung, die Frage nach wahr oder falsch wenig Sinn ergibt. Er ist mithin keine Aussage. Wir halten fest:

> Aussagen erkennen wir daran, daß ihnen ein Wahrheitswert zugeordnet ist, **w** für *wahr* oder **f** für *falsch*.

Im folgenden werden wir als Platzhalter für Aussagen meist Großbuchstaben verwenden: $A, B, C, \ldots$.

Eine Aussage als *wahr nachzuweisen*, soll bedeuten, daß wir sie durch logische Schlüsse auf andere, uns als wahr bekannte Aussagen zurückführen. Nehmen wir etwa den folgenden Satz:

A : Der Bundespräsident ist stets mindestens vierzig Jahre alt.

Wir stellen zunächst einmal fest, daß es sich um eine Aussage handelt – und zwar um eine *wahre* Aussage, wie wir aus Artikel 54 des Grundgesetzes ableiten. Dort nämlich finden wir zur Wahl des Bundespräsidenten folgende Aussage:

B : Wählbar ist jeder Deutsche, der das Wahlrecht zum Bundestage besitzt und das vierzigste Lebensjahr vollendet hat.

Weil nun das Grundgesetz gültig ist, ist Aussage A wahr. Wir haben Aussage A also auf eine uns bekannte wahre Aussage zurückgeführt.

Daß die von uns aus dem Grundgesetz zitierte Aussage B ihrerseits wahr ist, läßt sich nicht weiter auf andere Aussagen zurückführen. Vielmehr handelt es sich hierbei um eine Festlegung des Gesetzgebers, der das Gesetz erlassen und damit diese Aussage für wahr erklärt hat.

> Eine Aussage, der der Wahrheitswert **w** schlicht durch Festlegung zugewiesen wurde, nennen wir ein *Axiom*.

Man kann in diesem Sinne das Grundgesetz als eine Sammlung von Axiomen, oder ein Axiomensystem, auffassen – auch wenn der Vergleich in mancher Hinsicht hinken mag.

Eingangs haben wir erklärt, daß die Mathematiker sich die Welt, die sie untersuchen, und ihre Objekte selbst erschaffen. Sie tun dies, indem sie sich einige wenige Aussagen als Axiome vorgeben und sodann studieren, was sich aus diesen durch logisch korrekte Schlüsse ableiten läßt. Freilich, so wie der Gesetzgeber seine Gesetze nicht willkürlich erläßt, so wählen auch die Mathematiker die Axiome, die sie sich vorgeben, mit Bedacht, das heißt, mit dem Ziel, interessante Strukturen zu gewinnen – und die vielfältigen Anwendungen zeigen, daß die Mathematiker bei diesem Vorhaben nicht nur sehr kreativ, sondern auch sehr erfolgreich gewesen sind. Immer wieder haben sie sich von Fragestellungen der Alltagswelt inspirieren lassen, haben die Probleme auf wenige Kernpunkte reduziert und in ein (mathematisches) *Modell* übersetzt. Dabei bedeutet letzteres nichts anderes, als daß man die zu benutzende Sprache und die geltenden Axiome festlegt und daß man die Fragen in dieser neuen Sprache formuliert. Die Stärke dieser *Modellbildung* besteht nun darin, daß man innerhalb des Modells exakt und ohne Wenn und Aber feststellen kann, ob eine Aussage wahr ist oder nicht. Wahr ist sie stets dann, wenn sie durch eine ganze Reihe logisch korrekter Schlüsse aus den vorgegebenen Axiomen hervorgeht. Wann aber ist denn eine Aussage aus einer anderen durch einen *logisch korrekten Schluß* hervorgegangen?

Bevor wir uns dieser Frage erneut zuwenden, wollen wir klären, wie man aus gegebenen Aussagen überhaupt neue Aussagen gewinnen und so das Arsenal an Aussagen erweitern kann.

Eine ganz natürliche Möglichkeit ist die Verneinung oder *Negation* einer Aussage, etwa

$\neg A$: Der Bundespräsident ist *nicht* stets vierzig Jahre alt.

Wir wollen generell die Negation einer Aussage X mit dem Symbol $\neg X$ bezeichnen, und es sollte gelten, wenn X wahr ist, so ist $\neg X$ falsch, und umgekehrt. Das heißt insbesondere, der Wahrheitswert von $\neg X$ hängt nur vom Wahrheitswert von X ab. Dies erlaubt es uns, den Wahrheitswert von $\neg X$ in Abhängigkeit des Wahrheitswertes von X in einer Tabelle festzuhalten:

X	$\neg X$
w	**f**
f	**w**

Aus unserer Alltagssprache sind wir es gewohnt, mehrere Aussagen in auflistender Weise durch das Wort "und" miteinander zu verbinden. Betrachten wir etwa die folgenden Aussagen

C : Wählbar sind nur Deutsche, die das Wahlrecht zum Bundestag besitzen.

sowie

D : Wählbar sind nur Deutsche, die das vierzigste Lebensjahr vollendet haben.

Man erkennt unschwer, daß die Verknüpfung der Aussagen C und D durch "und" inhaltlich mit unserer

obigen Aussage B übereinstimmt, und man spricht von der *Konjunktion* von C und D. Auch hier wollen wir wieder eine symbolische Schreibweise einführen. Sind X und Y zwei Aussagen, so schreiben wir für "X *und* Y" auch $X \wedge Y$. Wenn nun $X \wedge Y$ wieder eine Aussage ist, so muß ihr auch ein Wahrheitswert zugeordnet sein. Dabei sollte wohl $X \wedge Y$ nur dann wahr sein, wenn sowohl X als auch Y wahr sind. Wir können den Wahrheitswert von $X \wedge Y$ also wieder in Abhängigkeit von den Wahrheitswerten von X und Y in einer Tabelle, auch *Wahrheitstafel* genannt, festhalten.

X	Y	$X \wedge Y$
w	**w**	**w**
w	**f**	**f**
f	**w**	**f**
f	**f**	**f**

Ebenso ist uns aus unserem alltäglichen Gebrauch ein weiteres Bindewort bekannt, "oder", welches wir hier instrumentalisieren wollen. Sind X und Y wieder Aussagen, so werden wir gewöhnlich $X \vee Y$ statt "X *oder* Y" schreiben. Die so entstandene neue Aussage nennt man die *Disjunktion* von X und Y, und damit sie wahr ist, soll es uns reichen, daß eine der Aussagen X und Y wahr ist. Dies führt zur folgenden Wahrheitstafel:

X	Y	$X \vee Y$
w	**w**	**w**
w	**f**	**w**
f	**w**	**w**
f	**f**	**f**

Man beachte, daß *oder* hier nicht das ausschließende *entweder oder* ist!

Die Aussage etwa, daß die Kinder unserer Bundestagsabgeordneten stets die deutsche *oder* eine andere Staatsangehörigkeit haben, ist wahr, weil sie nicht ausschließt, daß sie die deutsche und eine andere Staatsangehörigkeit haben.

Im Absatz zur Konjunktion heißt es, daß die Aussage B mit der Konjunktion der Aussagen C und D inhaltlich übereinstimme. Sprachlich sind beide Aussagen aber deutlich verschieden. Anstatt sie *gleich* zu nennen, wollen wir deshalb nur davon sprechen, daß B und $C \wedge D$ *gleichwertig* oder *äquivalent* sind. Dies soll zum Ausdruck bringen, daß sie den gleichen Wahrheitswert besitzen. Gehen wir einen Schritt weiter, so können wir eine neue Verknüpfung zweier Aussagen X und Y einführen, die *Äquivalenz* von X und Y, in Symbolen $X \Leftrightarrow Y$. Sie soll genau dann wahr sein, wenn X und Y den

gleichen Wahrheitswert besitzen. Dies führt zu folgender Wahrheitstafel:

X	Y	$X \Leftrightarrow Y$
w	**w**	**w**
w	**f**	**f**
f	**w**	**f**
f	**f**	**w**

Ein kurzer Blick auf die bislang eingeführten Operationen zur Gewinnung neuer Aussagen aus gegebenen zeigt, daß die Wahrheitswerte der neuen Aussagen stets allein von den Wahrheitswerten der gegebenen Aussagen abhängen, und nicht von deren konkretem Inhalt.

Wir erlauben uns deshalb, eine letzte Verknüpfung von Aussagen, die *Implikation*, dadurch einzuführen, daß wir bei gegebenen Aussagen X und Y den Wahrheitswert der Aussage "X impliziert Y" oder "wenn X, dann Y", in Zeichen $X \Rightarrow Y$, festlegen:

(1)

X	Y	$X \Rightarrow Y$
w	**w**	**w**
w	**f**	**f**
f	**w**	**w**
f	**f**	**w**

Die Wortwahl legt nahe, daß die Aussage $X \Rightarrow Y$ es erlaubt, aus der Wahrheit von X Rückschlüsse auf die Wahrheit von Y zu ziehen. Dies kommt auch in den ersten beiden Zeilen der Wahrheitstafel zum Ausdruck, wird aber noch deutlicher, wenn wir zeigen, daß die Aussagen $X \Rightarrow Y$ und $\neg X \vee Y$ zueinander äquivalent sind. Ist dann nämlich X wahr, so ist $\neg X$ falsch. Damit $\neg X \vee Y$ wahr sein kann, muß mithin Y wahr sein. Dies läßt sich so interpretieren, daß sich bei wahrer Aussage X und korrekter Implikation $X \Rightarrow Y$ für Y nur die Möglichkeit ergibt, ebenfalls wahr zu sein.

In dieser Weise werden wir die Implikation immer wieder anwenden. Wir werden mit einer wahren Aussage starten und mittels einer logisch korrekten Argumentationskette Y aus X ableiten – sprich wir werden $X \Rightarrow Y$ als wahr erweisen. Damit haben wir dann zugleich die Wahrheit von Y bewiesen.

Die Gültigkeit der behaupteten Äquivalenz leiten wir durch eine Betrachtung der Wahrheitstafeln her. Es reicht, festzustellen, daß die Werte in den Spalten von $X \Rightarrow Y$ und von $\neg X \vee Y$ übereinstimmen:

X	Y	$\neg X$	$\neg X \vee Y$	$X \Rightarrow Y$
w	**w**	**f**	**w**	**w**
w	**f**	**f**	**f**	**f**
f	**w**	**w**	**w**	**w**
f	**f**	**w**	**w**	**w**

Die bisherigen Betrachtungen erläutern die ersten beiden Zeilen der Wahrheitstafel der Implikation. Mysteriöser sind auf den ersten Blick zweifellos die beiden letzten, erlauben sie es doch, aus einer falschen Aussage eine beliebige andere Aussage herzuleiten und den vorgenommenen Schluß als korrekt anzusehen. Widerstrebt uns das nicht zutiefst? Wir möchten an einem Beispiel, das auf ein wenig Schulwissen in Mathematik zurückgreift, verdeutlichen, daß die obige Festlegung sehr wohl sinnvoll ist. Will man etwa die Lösungen der Gleichung

$$x^2 - 2x = -1$$

finden, so wird man auf beiden Seiten der Gleichung zunächst die Zahl 1 addieren, um so auf der linken Seite den Ausdruck $(x-1)^2$ zu erhalten, ein Verfahren, welches als *quadratische Ergänzung* bekannt ist. Man leitet aus der Aussage $x^2 - 2x = -1$ die Aussage $x^2 - 2x + 1 = 0$ her. Dieser Schluß läßt sich formulieren als die Implikation

$$\big(x^2 - 2x = -1\big) \quad \Rightarrow \quad \big(x^2 - 2x + 1 = 0\big).$$

Der Schluß, daß die Addition einer Zahl auf beiden Seiten einer Gleichung, die Gleichheit nicht zerstört, ist uns wohl vertraut und wir sehen ihn als korrekt an, unabhängig davon, was auf beiden Seiten der Gleichung steht. Wenden wir diesen Schluß nun auf eine andere Gleichung an, etwa auf die Gleichung $0 = 1$, so erhalten wir die Implikation

$$(0 = 1) \quad \Rightarrow \quad (0 + 1 = 1 + 1).$$

Die beiden Aussagen links und rechts des Implikationspfeiles sind offenbar falsch, der Schluß an sich ist jedoch nach dem eben Gesagten zulässig. Mithin sollte die Implikation den Wahrheitswert **w** tragen.

Ein Beispiel dafür, daß sich aus einer falschen Aussage durch einen korrekten Schluß auch eine wahre Aussage herleiten läßt, erhalten wir in analoger Weise, wenn wir uns vergegenwärtigen, daß die Gleichheit auch durch Multiplikation mit einer Zahl nicht zerstört wird. Dies führt dann zu der wahren Implikation

$$(0 = 1) \quad \Rightarrow \quad (0 \cdot 0 = 1 \cdot 0),$$

bei der die Aussage auf der linken Seite des Implikationspfeiles falsch ist, während die auf der rechten Seite wahr ist.

Wir halten fest:

> Der Wahrheitswert der Implikation $X \Rightarrow Y$ bewertet nur die Korrektheit des Schließens, nicht jedoch die Wahrheit der Aussagen X und Y.

Es sei deshalb jedem ans Herz gelegt, die Voraussetzungen, auf die er seine Aussagen gründet, genauestens auf ihren Wahrheitsgehalt zu prüfen! Sonst nützt auch noch so sauberes Schließen nicht viel.

Wir wollen den eingeführten Begriffsapparat nun an zwei Beispielen testen, die uns einige wichtige Erkenntnisse liefern werden.

Beispiel 1.1.
Es seien X und Y zwei Aussagen.

a. Wir haben bereits bei der Definition der Äquivalenz davon gesprochen, daß $X \Leftrightarrow Y$ bedeuten solle, daß "X *genau dann* wahr ist, *wenn* Y wahr ist". Dies wollte verkürzt ausdrücken, "wenn X, dann Y" und "wenn Y, dann X". Wir behaupten deshalb, daß die Aussagen "$X \Leftrightarrow Y$" und "$(X \Rightarrow Y) \wedge (Y \Rightarrow X)$" äquivalent sind, mit anderen Worten, die Aussagen X und Y sind genau dann äquivalent, wenn Y aus X folgt und umgekehrt.
Diese Tatsache werden wir immer wieder verwenden, wenn wir die Äquivalenz zweier Aussagen beweisen wollen. Ihre Gültigkeit leiten wir wieder durch eine Betrachtung der Wahrheitstafeln her.

X	Y	$X \Rightarrow Y$	$Y \Rightarrow X$	$(X \Rightarrow Y) \wedge (Y \Rightarrow X)$	$X \Leftrightarrow Y$
w	**w**	**w**	**w**	**w**	**w**
w	**f**	**f**	**w**	**f**	**f**
f	**w**	**w**	**f**	**f**	**f**
f	**f**	**w**	**w**	**w**	**w**

b. Die Aussagen "$X \Rightarrow Y$" und "$\neg Y \Rightarrow \neg X$" sind ebenfalls äquivalent, wie die folgende Tabelle zeigt:

X	Y	$\neg X$	$\neg Y$	$X \Rightarrow Y$	$\neg Y \Rightarrow \neg X$
w	**w**	**f**	**f**	**w**	**w**
w	**f**	**f**	**w**	**f**	**f**
f	**w**	**w**	**f**	**w**	**w**
f	**f**	**w**	**w**	**w**	**w**

Man nennt diese Äquivalenz auch *Kontraposition*. Will man also zeigen, daß eine Aussage X eine Aussage Y impliziert, so kann man statt dessen beide Aussagen verneinen und zeigen, daß aus $\neg Y$ die Aussage $\neg X$ folgt.

□

Kehren wir nun zu der Frage zurück, wann eine Aussage Y aus einer Aussage X durch einen *logisch korrekten Schluß* hervorgegangen ist. Bedeutet dies nur, daß $X \Rightarrow Y$ den Wahrheitswert **w** besitzt? Ja ... und nein! Ist X wahr und hat die Implikation $X \Rightarrow Y$ den Wahrheitswert **w**, so folgt unmittelbar, daß Y wahr ist. In diesem Sinne gilt die Antwort *ja*. Aber damit haben wir das Problem nur verlagert, da die Frage bleibt, wie wir prüfen, ob $X \Rightarrow Y$ denn wahr ist, ohne den Wahrheitswert von Y zu kennen. Wir haben bereits weiter oben – sehr vage – angedeutet, daß wir hierzu meist eine Kette von

logisch korrekten und in sich schlüssigen Argumenten verwenden, und viel deutlicher wollen wir hier auch nicht werden. Im Verlauf der folgenden Kapitel werden wir viele Beispiele dafür sehen, wie eine Implikation durch eine Reihe von Argumenten bewiesen – oder besser untermauert – wird; und es wird sicher immer wieder vorkommen, daß Euch diese auf den ersten Blick *nicht* wirklich schlüssig vorkommen, daß es eines genaueren Hinsehens und vielleicht auch der Ergänzung einiger Argumente bedarf, bis Ihr der Kette das Prädikat *logisch korrekt und in sich schlüssig* verleihen wollt. Und das ist eine wichtige Erkenntnis: ob ein Schluß als logisch korrekt erkannt wird, hängt vom Betrachter ab. Und deshalb ist die Frage, ob ein Schluß logisch korrekt ist, weit mehr als nur die Frage, ob $X \Rightarrow Y$ wahr ist.

Beispiel 1.2.
Hier nun einige mathematische Aussagen.

A. Jede gerade Zahl ist Summe zweier ungerader Zahlen.

B. Es gibt unendlich viele Primzahlen.

C. Jede gerade Zahl größer zwei ist Summe zweier Primzahlen.

D. Zu jedem Kreis läßt sich, nur mit Zirkel und Lineal, ein Quadrat konstruieren, das den gleichen Flächeninhalt hat.

E. Die Gleichung $x^n + y^n = z^n$ besitzt für $n > 2$ keine Lösung mit positiven ganzen Zahlen x, y, z.

F. Gegeben sei eine Ansammlung nicht-leerer Mengen. Dann läßt sich aus jeder der Mengen ein Element auswählen.

Die Aussage A ist offensichtlich wahr, und auch die Aussage B ist richtig, allerdings ist dies keine triviale Aussage. Sie muß bewiesen werden. Die Aussage C ist die bekannte *Goldbachsche Vermutung* aus dem Jahre 1742. Sie ist bis heute weder bewiesen noch widerlegt.

Die Aussage D ist unter dem Begriff *Quadratur des Kreises* bekannt. Sie ist falsch, was sich daraus ableiten läßt, daß die Kreiszahl π transzendent ist (Lindemann 1882). Umgangssprachlich sollte man also die Quadratur des Kreises nicht als Synonym für etwas extrem Schwieriges verwenden, sondern für etwas Unmögliches.

Die Aussage E hat jahrhundertelang als *Fermatsche Vermutung* die Mathematiker beschäftigt. Sie wurde erst 1995 von dem englischen Mathematiker Andrew Wiles als wahr nachgewiesen. Für den Beweis wurden modernste und tiefste mathematische Methoden verwendet.

Die Aussage F, möchte man meinen, ist offensichtlich wahr, eher noch als Aussage A. In gewissem Sinne ist diese Aussage jedoch weder beweisbar noch widerlegbar. Sie ist im Axiomensystem der Mengenlehre von Zermelo und Fraenkel unabhängig von den

anderen Axiomen. In der Tat kann man die Aussage F, die als *Auswahlaxiom* bezeichnet wird, als Axiom der Mengenlehre zulassen (was wir, wie die überwiegende Zahl der Mathematiker, tun wollen) oder auch nicht. Da das Auswahlaxiom, wenn überhaupt, so nur für sogenannte überabzählbare Ansammlungen strittig ist, sind Zustimmung oder Ablehnung in dieser Vorlesung kaum von praktischer Relevanz. □

Wir wollen nun der besseren Übersichtlichkeit halber in einer Bemerkung zusammenfassen, was wir bisher gelernt haben.

Bemerkung 1.3.

a. Eine *Aussage* ist eine Äußerung, der eindeutig ein Wahrheitswert wahr (**w**) oder falsch (**f**) zugeordnet ist.

b. Aus Aussagen X und Y können wir durch Anwenden *logischer Operatoren* neue Aussagen bilden:

Symbol	Bedeutung	Bezeichnung	Alternative Beschreibung
$\neg X$	nicht X	*Negation*	
$X \vee Y$	X oder Y	*Disjunktion*	
$X \wedge Y$	X und Y	*Konjunktion*	
$X \Rightarrow Y$	aus X folgt Y	*Implikation*	$(\neg X) \vee Y$
$X \Leftrightarrow Y$	genau dann X, wenn Y	*Äquivalenz*	$(X \Rightarrow Y) \wedge (Y \Rightarrow X)$

Neben Aussagen, die wahr oder falsch sein können, sind *Aussageformen* oder *Prädikate* wichtig.

> Eine *Aussageform* ist eine Äußerung, die eine oder mehrere Variablen enthält und zu einer Aussage (d.h. wahr oder falsch) wird, wenn man zulässige Werte für diese Variablen einsetzt.

So ist etwa

$$a > b$$

eine Aussageform, die von den Variablen a und b abhängt, für die wir die ganzen Zahlen als zulässige Werte ansehen wollen. Setzen wir konkrete Werte ein, so entsteht eine Aussage, die wahr sein kann (z.B. für $a = 42$ und $b = 37$) oder falsch (z.B. für $a = 2$ und $b = 4$).

> Aussageformen werden in der Praxis häufig mit *Quantoren* gebraucht:

$\forall$: "für alle".
$\exists$: "es existiert ein".
$\exists_1$: "es existiert genau ein".
$\nexists$: "es existiert kein".

Ist P eine Aussageform, die von einer Variablen x abhängt, so bedeutet:

$\forall\, x \,:\, P(x)$: "für alle x gilt $P(x)$",
$\exists\, x \,:\, P(x)$: "es gibt ein x, so daß $P(x)$ gilt".

Mit Hilfe der Quantoren haben wir aus den Aussageformen neue Aussagen gebildet.

Beispiel 1.4.

$$\forall\, x, \forall\, y, \forall\, z, \forall\, n \;:\; n > 2 \Rightarrow x^n + y^n \neq z^n.$$

Dies ist für positive natürliche Zahlen x, y, z und n die in Beispiel 1.2 formulierte Fermatsche Vermutung. □

Wichtig ist das richtige Verneinen einer Aussage.

$$\neg\big(\forall\, x \;:\; P(x)\big) \;\Leftrightarrow\; \exists\, x \;:\; \big(\neg P(x)\big).$$

Die Verneinung der Aussage "für alle x gilt die Aussage $P(x)$" ist gleichbedeutend mit "es gibt ein x, für das die Aussage $P(x)$ nicht gilt".

$$\neg\big(\exists\, x \;:\; P(x)\big) \;\Leftrightarrow\; \forall\, x \;:\; \big(\neg P(x)\big).$$

Die Verneinung der Aussage "es gibt ein x, für das die Aussage $P(x)$ gilt" ist gleichbedeutend mit "für alle x gilt die Aussage $P(x)$ nicht" bzw. mit "für kein x gilt die Aussage $P(x)$".

$$(A \Rightarrow B) \;\Leftrightarrow\; (\neg B \Rightarrow \neg A).$$

Die Aussage "aus A folgt B" ist gleichbedeutend mit "aus nicht B folgt nicht A". Letzteres bezeichnet man auch als *Kontraposition* von ersterem.

Proposition 1.5.
Es seien X, Y und Z Aussagen.

a. *Assoziativgesetze*
- $(X \vee Y) \vee Z \iff X \vee (Y \vee Z)$.
- $(X \wedge Y) \wedge Z \iff X \wedge (Y \wedge Z)$.

b. *Kommutativgesetze*
- $X \vee Y \iff Y \vee X$.
- $X \wedge Y \iff Y \wedge X$.

c. *Distributivgesetze*
- $X \wedge (Y \vee Z) \iff (X \wedge Y) \vee (X \wedge Z)$.
- $X \vee (Y \wedge Z) \iff (X \vee Y) \wedge (X \vee Z)$.

Beweis: Den Nachweis der Äquivalenzen überlassen wir dem Leser als Übungsaufgabe. □

Bemerkung 1.6.
Eine Aussageform, die für jeden möglichen Wert der Variablen eine wahre Aussage liefert, nennt man auch eine *Tautologie*. Die in Proposition 1.5 gezeigten Äquivalenzen aufgefaßt als Aussageformen sind in diesem Sinne Tautologien.

Beispiel 1.7 (Nand-Gatter).
Die Digitalelektronik basiert auf den Prinzipien der Aussagenlogik. Die Bauelemente elektronischer Schaltungen realisieren im Grunde Aussageformen, die je nach Wert der boolschen Variablen zwei Zustände annehmen können. Dabei kann man zeigen, daß alle logischen Aussagen, die wir aus den in diesem Abschnitt eingeführten Operationen erstellen können, sich auch mittels einer einzigen Operation gewinnen lassen, dem sogenannten Nand $\uparrow$, die durch folgende Wahrheitstafel definiert wird:

X	Y	$X \uparrow Y$
w	**w**	**f**
w	**f**	**w**
f	**w**	**w**
f	**f**	**w**

Wir überlassen es dem Leser, mittels Wahrheitstafeln die folgenden Äquivalenzen zu zeigen:

$$\begin{aligned} \neg(X \wedge Y) &\iff (X \uparrow Y) \\ \neg X &\iff (X \uparrow X) \\ (X \vee Y) &\iff ((X \uparrow X) \uparrow (Y \uparrow Y)) \\ (X \wedge Y) &\iff ((X \uparrow Y) \uparrow (X \uparrow Y)) \end{aligned}$$

Da sich die anderen Operatoren alle auf die Negation, die Disjunktion und die Konjunktion zurück führen lassen, ist damit auch gezeigt, daß sich alle auch allein durch das Nand ausdrücken lassen.

Bemerkung 1.8 (Griechisches Alphabet).
Es hat sich in der Mathematik eingebürgert, neben den lateinischen auch griechische

Buchstaben zu verwenden, um Objekte und Variablen zu bezeichnen, und das werden wir immer wieder mal tun. Deshalb füge ich hier das griechische Alphabet an:

$A\ \alpha$ Alpha	$B\ \beta$ Beta	$\Gamma\ \gamma$ Gamma	$\Delta\ \delta$ Delta	$E\ \epsilon\ \varepsilon$ Epsilon	$Z\ \zeta$ Zeta	$H\ \eta$ Eta	$\Theta\ \theta\ \vartheta$ Theta
$I\ \iota$ Iota	$K\ \kappa$ Kappa	$\Lambda\ \lambda$ Lambda	$M\ \mu$ My	$N\ \nu$ Ny	$\Xi\ \xi$ Xi	$O\ o$ Omikron	$\Pi\ \pi$ Pi
$P\ \rho$ Rho	$\Sigma\ \sigma$ Sigma	$T\ \tau$ Tau	$Y\ \upsilon$ Ypsilon	$\Phi\ \phi\ \varphi$ Phi	$X\ \chi$ Chi	$\Psi\ \psi$ Psi	$\Omega\ \omega$ Omega

Aufgaben

Aufgabe 1.9.

a. Negiere die folgenden Aussagen:

(i) Jedes Auto, das am Samstag um 9:00 auf dem Parkplatz parkte, war rot.

(ii) Mindestens ein Auto, das am Samstag um 9:00 auf dem Parkplatz parkte, war rot.

(iii) Am Samstag um 9:00 parkten rote Autos auf dem Parkplatz.

(iv) Es gibt keine größte ganze Zahl.

(v) Keine Regel ohne Ausnahme.

Warum ist das Sprichwort *Keine Regel ohne Ausnahme*, in sich widersprüchlich?

b. Beweise oder widerlege Aussage (iv).

Aufgabe 1.10.

Es seien X und Y Aussagen. Zeige die folgenden Äquivalenzen:

a. *De Morgansche Regeln*

- $\neg(X \vee Y) \iff \neg X \wedge \neg Y$.
- $\neg(X \wedge Y) \iff \neg X \vee \neg Y$.

b. $(\neg X \implies f) \iff X$.

Aufgabe 1.11 (Nand).

Es seien X und Y zwei Aussagen.

Aufgabe 1.12.

a. Drücke die folgenden Aussagen in Worten aus und, falls eine Aussage falsch sein sollte, ersetze sie dabei durch ihre Negation.

(i) $\forall\, m \in \mathbb{N},\ \exists\, n \in \mathbb{N}\ :\ m = n + n$,

(ii) $\exists\, m \in \mathbb{N},\ \exists\, n \in \mathbb{N}\ :\ (m \neq n) \wedge (m^n = n^m)$.

b. Drücke die folgenden Aussagen in Symbolen aus:

(i) Zwischen je zwei verschiedenen reellen Zahlen gibt es eine weitere reelle Zahl.

(ii) Es gibt keine größte Primzahl in den natürlichen Zahlen.

Aufgabe 1.13.

Welche der folgenden Schlußfolgerungen ist korrekt?

a. Falls es anfängt zu regnen, wird die Straße naß. Aber, da die Straße nicht naß werden wird, wird es auch nicht regnen.

b. Einige Politiker sind ehrlich. Einige Frauen sind Politiker. Also sind einige weibliche Politiker ehrlich.

Aufgabe 1.14.

Drücke die folgende Aussage in Worten aus:

$$\forall\, m \in \mathbb{N},\ \forall\, n \in \mathbb{N}\ :\ m \geq n \implies \exists\, l \in \mathbb{N}\ :\ m = n + l.$$

Aufgabe 1.15. a. Negiere die folgenden Aussagen:

(i) Zu jedem Vorschlag gibt es jemanden, der den Vorschlag kritisiert.

(ii) In manchen Häusern haben nicht alle Wohnungen fließendes Wasser.

b. Beweise oder widerlege die folgenden Aussagen:

(i) Jede ganze Zahl ist ein Vielfaches von drei.

(ii) Die Summe von je zwei ungeraden Zahlen ist gerade.

Aufgabe 1.16.

Seien X, Y und Z Aussagen. Beweise die folgende Aussage:

$$(X \vee Y) \iff ((X \uparrow X) \uparrow (Y \uparrow Y)).$$

§ 2 Mengen

Definitionsversuch 2.1 (Georg Cantor).
Eine *Menge* ist eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen. Die in einer Menge zusammengefaßten Objekte nennen wir die *Elemente* der Menge.

Notation 2.2.

a. *Mengen angeben durch Auflisten der Elemente*:

$$\text{z.B. } \{1, 2, 5, 3, 4, 0\}$$

b. *Mengen angeben durch Vorschreiben einer Eigenschaft*:

$$\text{z.B. } \{x \mid x \text{ ist eine natürliche Zahl kleiner als } 6\}$$

c. Sei M eine Menge.
- $x \in M$ heißt "x *ist Element von* M"
- $x \notin M$ heißt "x *ist nicht Element von* M"

d. $\{\}$ und $\emptyset$ bezeichnen die *leere Menge*, d.h. die Menge, die kein Element enthält.

Definition 2.3 (Inklusionsrelationen).
Für zwei Mengen M und N definieren wir:

1) $M \subseteq N \;:\Longleftrightarrow\; (x \in M \;\Rightarrow x \in N)$ "M ist *Teilmenge* von N"

2) $M = N \;:\Longleftrightarrow\; (M \subseteq N \;\wedge\; N \subseteq M)$
$\Longleftrightarrow\; (x \in M \;\Leftrightarrow\; x \in N)$

3) $M \neq N \;:\Longleftrightarrow\; \neg(M = N)$
$\Longleftrightarrow\; ((\exists\, x \in M \;:\; x \notin N) \;\vee\; (\exists\, x \in N \;:\; x \notin M))$

4) $M \subsetneq N \;:\Longleftrightarrow\; (M \subseteq N \;\wedge\; M \neq N)$ "M ist *echte Teilmenge* von N"

Beispiel 2.4.

a. $\{1, 2, 5, 3, 4, 0\} = \{x \mid x \text{ ist eine natürliche Zahl kleiner als } 6\}$.

b. $\{1, 3\} \subsetneq \{1, 2, 3\}$.

c. $\{1, 2, 1\} = \{1, 2\} = \{2, 1\}$.

d. $1 \notin \{2, 3\}$, $2 \in \{2, 3\}$.

Bemerkung 2.5 (Die Zahlbereiche).
Wir setzen die folgenden Mengen in unserer Vorlesung als bekannt voraus:

- $\mathbb{N} = \{0, 1, 2, 3, 4, \ldots\}$ die Menge der *natürlichen Zahlen*,
- $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ die Menge der *ganzen Zahlen*,
- $\mathbb{Q} = \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0 \right\}$ die Menge der *rationalen Zahlen*,
- $\mathbb{R}$, die Menge der *reellen Zahlen*, d.h. der Dezimalbrüche.

Beachte:

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}.$$

Im Verlauf der Vorlesung werden wir viele bekannte Eigenschaften dieser Mengen nochmals ausführlich thematisieren.

Definition 2.6 (Operationen von Mengen).
Es seien M, N, P sowie M_i für $i \in I$ Mengen.

a. $M \cap N := \{x \mid x \in M \wedge x \in N\}$ heißt der *Durchschnitt* von M und N.

b. $M \cup N := \{x \mid x \in M \vee x \in N\}$ heißt die *Vereinigung* von M und N.

c. $M \setminus N := \{x \mid x \in M \wedge x \notin N\}$ heißt die *Differenzmenge* von M und N. Wir sagen auch M *ohne* N.

d. $M \times N := \{(x, y) \mid x \in M \wedge y \in N\}$ heißt das *kartesische Produkt* von M und N. Dabei ist (x, y) ein *geordnetes Paar*, und für zwei geordnete Paare $(x, y), (u, v) \in M \times N$ gilt

$$(x, y) = (u, v) \iff (x = u \ \wedge \ y = v).$$

e. M und N heißen genau dann *disjunkt*, wenn $M \cap N = \emptyset$, d.h. wenn sie kein Element gemeinsam besitzen.

f. $P = M \dot\cup N \ :\iff \ (P = M \cup N \ \wedge \ M \cap N = \emptyset)$.
Wir sagen dann, P ist die *disjunkte Vereinigung* von M und N.

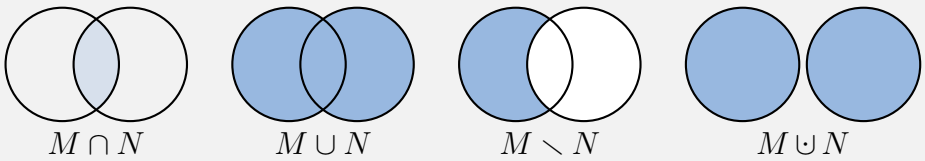


Abbildung 1: Durchschnitt, Vereinigung, Differenzmenge, disjunkte Vereinigung

g. $\bigcap_{i \in I} M_i := \{x \mid x \in M_i \ \forall\, i \in I\}$ heißt der *Durchschnitt* der M_i.

h. $\bigcup_{i \in I} M_i := \{x \mid \exists\, i \in I \ : \ x \in M_i\}$ heißt die *Vereinigung* der M_i.

i. $P = \dot\bigcup_{i \in I} M_i \ :\iff \ (P = \bigcup_{i \in I} M_i \ \wedge \ M_i \cap M_j = \emptyset \ \forall\, i, j \in I \text{ mit } i \neq j)$.
Wir nennen die $(M_i)_{i \in I}$ dann auch eine *disjunkte Zerlegung* von P, und wir sagen, die M_i sind *paarweise disjunkt*.

Beispiel 2.7.
Ist $M = \{1, 2\}$ und $N = \{e, \pi, i\}$, so ist

$$M \times N = \{(1, e), (1, \pi), (1, i), (2, e), (2, \pi), (2, i)\}.$$

Proposition 2.8 (Einfache Rechengesetze für Mengenoperationen).
Es seien M, N, P Mengen.

a. *Assoziativgesetze*
   - $(M \cup N) \cup P = M \cup (N \cup P)$.
   - $(M \cap N) \cap P = M \cap (N \cap P)$.

b. *Kommutativgesetze*
   - $M \cup N = N \cup M$.
   - $M \cap N = N \cap M$.

c. *Distributivgesetze*
   - $M \cap (N \cup P) = (M \cap N) \cup (M \cap P)$.
   - $M \cup (N \cap P) = (M \cup N) \cap (M \cup P)$.

d. *Identitätsgesetze*
   - $M \cup \emptyset = M$.
   - $M \subseteq N \implies M \cap N = M$.

e. *Komplementgesetze*
   - $M \subseteq N \implies M \cup (N \setminus M) = N$.
   - $M \subseteq N \implies M \cap (N \setminus M) = \emptyset$.

Beweis: a., d. und e. überlassen wir dem Leser als Übungsaufgabe.

b. Es gilt:

$$M \cup N = \{x \mid x \in M \vee x \in N\} \overset{1.5}{=} \{x \mid x \in N \vee x \in M\} = N \cup M$$

und

$$M \cap N = \{x \mid x \in M \wedge x \in N\} \overset{1.5}{=} \{x \mid x \in N \wedge x \in M\} = N \cap M.$$

c. Es gilt:

$$\begin{aligned} x \in M \cap (N \cup P) &\iff x \in M \ \wedge \ x \in N \cup P \\ &\iff x \in M \ \wedge \ (x \in N \ \vee \ x \in P) \\ &\overset{1.5}{\iff} (x \in M \ \wedge \ x \in N) \ \vee \ (x \in M \ \wedge \ x \in P) \\ &\iff x \in M \cap N \ \vee \ x \in M \cap P \\ &\iff x \in (M \cap N) \cup (M \cap P) \end{aligned}$$

und

$$\begin{aligned} x \in M \cup (N \cap P) &\iff x \in M \ \vee \ x \in N \cap P \\ &\iff x \in M \ \vee \ (x \in N \ \wedge \ x \in P) \\ &\overset{1.5}{\iff} (x \in M \ \vee \ x \in N) \ \wedge \ (x \in M \ \vee \ x \in P) \\ &\iff x \in M \cup N \ \wedge \ x \in M \cup P \\ &\iff x \in (M \cup N) \cap (M \cup P). \end{aligned}$$

□

Bemerkung 2.9 (Paradoxon von Russel).
Man muß bei der Definition von Mengen mittels Eigenschaften vorsichtig sein!

Betrachte die "*Menge*"

$$M = \{X \mid X \text{ ist Menge } \wedge \ X \notin X\}$$

aller Mengen, die sich nicht selbst als Element enthalten!

Angenommen, M wäre eine Menge. Dann sind zwei Fälle zu unterscheiden.

1. **Fall:** $M \notin M$**:** Dann ist M eine Menge, die sich nicht selbst als Element enthält. Mithin gilt $M \in M$ aufgrund der Definition von M. Dies ist ein Widerspruch.
2. **Fall:** $M \in M$**:** Dann ist M eine Menge, die sich selbst als Element enthält. Mithin gilt $M \notin M$ aufgrund der Definition von M. Dies ist ebenfalls ein Widerspruch.

Also kann keiner der beiden Fälle auftreten, und wir haben insgesamt einen Widerspruch hergeleitet.

Fazit: M ist *keine* Menge! Auch die *Menge aller Mengen* gibt es nicht!

Aufgaben

Aufgabe 2.10 (De Morgansche Regeln).
Es seien M und M_i, $i \in I$, Mengen. Zeige, die de Morganschen Regeln

$$M \setminus \bigcup_{i \in I} M_i = \bigcap_{i \in I} M \setminus M_i$$

und

$$M \setminus \bigcap_{i \in I} M_i = \bigcup_{i \in I} M \setminus M_i.$$

§ 3 Abbildungen

In diesem Abschnitt wollen wir den für die Mathematik zentralen Begriff der Abbildung einführen.

Definition 3.1 (Abbildungen).
Es seien M und N zwei Mengen. Eine *Abbildung* oder *Funktion* f von M nach N ist eine *eindeutige Zuordnung*, die *jedem* Element $x \in M$ *genau ein* Element $f(x) \in N$ zuweist. Wir werden den Begriff *Funktion* nur dann verwenden, wenn $N = \mathbb{R}$ ist.
Wir nennen M den *Definitionsbereich* von f und N den *Ziel-* oder *Wertebereich*.
Notation:

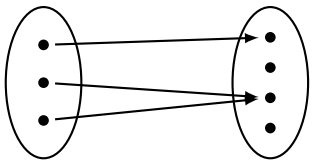
$$f : M \longrightarrow N : x \mapsto f(x).$$

Beachte, aufgrund der Definition einer Abbildung, gilt für zwei Abbildungen $f : M \longrightarrow N$ und $g : X \longrightarrow Y$:

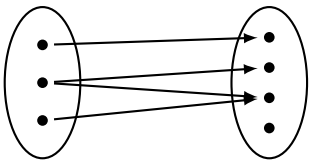
$$f = g \iff \big(M = X \ \wedge \ N = Y \ \wedge \ \forall\, x \in M \ : \ f(x) = g(x)\big).$$

Beispiel 3.2.

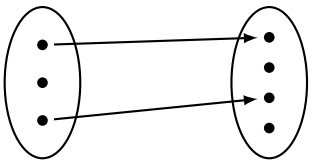
a. Die folgenden Bilder sollen den Begriff der Abbildung graphisch veranschaulichen:



Abbildung



keine Abbildung



keine Abbildung

b. $f : \mathbb{N} \longrightarrow \mathbb{N} : x \mapsto x^2$ und $g : \mathbb{Z} \longrightarrow \mathbb{N} : x \mapsto x^2$. Beachte: $f \neq g$, da ihre Definitionsbereiche nicht übereinstimmen.

c. Sei $f : M \longrightarrow N$ eine Abbildung und $A \subseteq M$. Dann heißt die Abbildung

$$f_{|A} : A \longrightarrow N : x \mapsto f(x)$$

die *Einschränkung* von f auf A.

d. Sei M eine Menge. Dann heißt die Abbildung

$$\mathrm{id}_M : M \longrightarrow M : x \mapsto x$$

die *Identität* auf M.

Definition 3.3 (Bilder und Urbilder).
Es sei $f : M \longrightarrow N$ eine Abbildung, $A \subseteq M$ und $B \subseteq N$.

a. $\mathrm{Graph}(f) := \{(x, f(x)) \mid x \in M\} \subseteq M \times N$ heißt der *Graph* von f.

b. $f(A) := \{f(x) \mid x \in A\} \subseteq N$ heißt das *Bild* von A unter f.

c. $\mathrm{Im}(f) := f(M) \subseteq N$ heißt das *Bild* von f.

d. $f^{-1}(B) := \{x \in M \mid f(x) \in B\} \subseteq M$ heißt das *Urbild* von B unter f.

Beispiel 3.4.

a. Wir betrachten die Abbildung $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$.

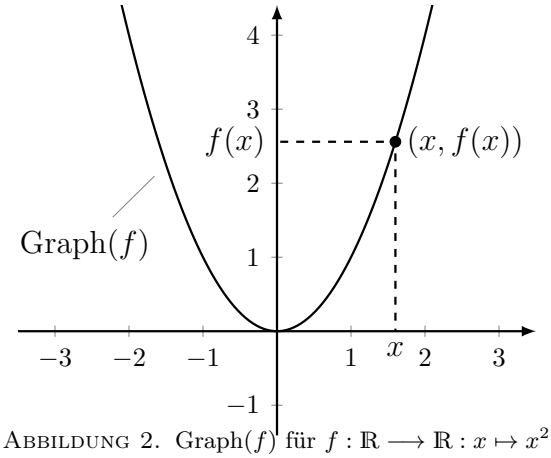


Abbildung 2. Graph(f) für $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$

- Der Graph von f ist in Abbildung 2 zu sehen.
- Für $A = \{-1, 0, 1, 2\}$ ist $f(A) = \{0, 1, 4\}$.
- Für $B = \{0, 1\}$ ist $f^{-1}(B) = \{0, 1, -1\}$.
- Für $B' = \{-1\}$ ist $f^{-1}(B') = \emptyset$.
- $\mathrm{Im}(f) = \{x \in \mathbb{R} \mid x \geq 0\}$.

b. Die Abbildung $\mathrm{nf} : \mathbb{N} \longrightarrow \mathbb{N} : x \mapsto x + 1$ nennen wir die *Nachfolgerfunktion*. Es gelten

$$\mathrm{Im}(\mathrm{nf}) = \mathbb{N} \setminus \{0\}$$

und

$$\forall\, y \in \mathrm{Im}(f) \;:\; \mathrm{nf}^{-1}(\{y\}) = \{y - 1\}.$$

Bemerkung 3.5 (Abbildungen und ihre Graphen).

a. Für zwei Abbildungen $f : M \longrightarrow N$ und $g : P \longrightarrow N$ gilt:

$$f = g \iff \mathrm{Graph}(f) = \mathrm{Graph}(g).$$

b. Ist $\Gamma \subseteq M \times N$ so, daß

$$\forall\, x \in M\ \exists_1\, y \in N \;:\; (x, y) \in \Gamma,$$

dann gibt es eine Abbildung $f : M \longrightarrow N$ mit $\Gamma = \mathrm{Graph}(f)$.

Fazit: Man hätte Abbildungen von M nach N auch als Teilmengen von $M \times N$ definieren können, die die Bedingung in b. erfüllen. So würde man vorgehen, wenn man die Mathematik ausgehend vom Begriff der Menge sukzessive aufbauen möchte.

Mit dieser Beschreibung sieht man übrigens sehr schön, daß es für jede Menge M *genau eine* Abbildung $f : \emptyset \longrightarrow M$ gibt, und daß es für eine nicht-leere Menge M keine Abbildung $f : M \longrightarrow \emptyset$ geben kann.

Definition 3.6 (Injektiv, surjektiv, bijektiv).
Es sei $f : M \longrightarrow N$ eine Abbildung.

a. f heißt genau dann *injektiv*, wenn

$$\forall\, x, x' \in M \;:\; f(x) = f(x') \implies x = x'.$$

b. f heißt genau dann *surjektiv*, wenn

$$\forall\, y \in N\ \exists\, x \in M \;:\; f(x) = y,$$

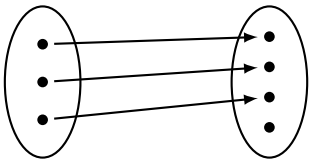
d.h. wenn $\mathrm{Im}(f) = N$.

c. f heißt genau dann *bijektiv*, wenn f injektiv und surjektiv ist.

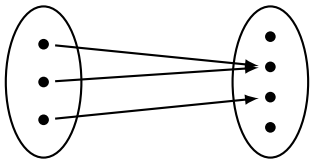
Bemerkung 3.7 (Injektiv, surjektiv, bijektiv).
Es sei $f : M \longrightarrow N$ eine Abbildung.

a. Ist $y \in N$ und $x \in M$ mit $f(x) = y$, so nennen wir x *ein Urbild* von y unter f.

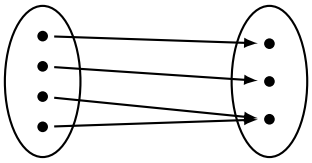
b. Es gelten:
- f ist injektiv $\iff$ jedes $y \in N$ hat *höchstens* ein Urbild.
- f ist surjektiv $\iff$ jedes $y \in N$ hat *mindestens* ein Urbild.
- f ist bijektiv $\iff$ jedes $y \in N$ hat *genau* ein Urbild.



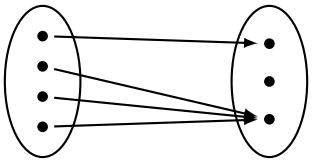
injektiv



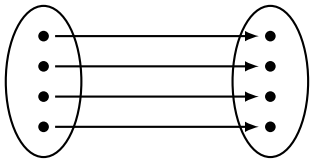
nicht injektiv



surjektiv



nicht surjektiv



bijektiv

Beispiel 3.8.

a. Die Nachfolgerfunktion $\mathrm{nf} : \mathbb{N} \longrightarrow \mathbb{N} : x \mapsto x+1$ ist injektiv, aber nicht surjektiv. Denn, $x+1 = \mathrm{nf}(x) = \mathrm{nf}(y) = y+1$ für $x, y \in \mathbb{N}$ impliziert $x = y$, und $0 \notin \mathrm{Im}(f)$.

b. $g : \mathbb{Z} \longrightarrow \mathbb{N} : x \mapsto x^2$ ist nicht injektiv.
Denn, für $x = 1 \neq -1 = y$ gilt $g(x) = g(1) = 1 = g(-1) = g(y)$.

c. Die Abbildung id_M ist bijektiv für jede Menge M.
Denn, für $y \in M$ gilt $\mathrm{id}_M(y) = y$, so daß id_M surjektiv ist, und für $x, x' \in M$ mit $\mathrm{id}_M(x) = \mathrm{id}_M(x')$ gilt $x = x'$, so daß id_M injektiv ist.

d. Ist $f : M \longrightarrow N$ injektiv, so ist die Abbildung $M \longrightarrow \mathrm{Im}(f) : x \mapsto f(x)$ offenbar bijektiv.

Definition 3.9 (Komposition von Abbildungen).
Seien $f : M \longrightarrow N$ und $g : N \longrightarrow P$ zwei Abbildungen. Die Abbildung

$$g \circ f : M \longrightarrow P : x \mapsto g(f(x))$$

heißt die *Komposition* oder *Verkettung* von f und g.

Beispiel 3.10.
Seien $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$ und $g : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x+1$. Dann gilt

$$(g \circ f)(x) = g(f(x)) = g(x^2) = x^2 + 1$$

und

$$(f \circ g)(x) = f(g(x)) = f(x+1) = (x+1)^2 = x^2 + 2x + 1.$$

Man beachte, daß die Abbildungen $g \circ f$ und $f \circ g$ nicht gleich sind, da $(g \circ f)(1) = 2 \neq 4 = (f \circ g)(1)$.

Proposition 3.11 (Assoziativität der Komposition).
Seien $f : M \longrightarrow N$, $g : N \longrightarrow P$ und $h : P \longrightarrow Q$ Abbildungen. Dann gilt

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Wir schreiben deshalb auch kurz $h \circ g \circ f$.

Beweis: Da die Definitions- und Zielbereiche der beiden Funktionen übereinstimmen, reicht es, die Abbildungsvorschrift zu überprüfen. Sei dazu $x \in M$. Dann gilt

$$\big((h \circ g) \circ f\big)(x) = (h \circ g)\big(f(x)\big) = h\Big(g\big(f(x)\big)\Big) = h\big((g \circ f)(x)\big) = \big(h \circ (g \circ f)\big)(x).$$

Dies zeigt die Behauptung. □

Satz 3.12 (Bijektivität = Existenz einer Umkehrabbildung).
Es sei $f : M \longrightarrow N$ eine Abbildung.

a. f ist genau dann bijektiv, wenn eine Abbildung $g : N \longrightarrow M$ existiert, so daß $g \circ f = \mathrm{id}_M$ und $f \circ g = \mathrm{id}_N$.

b. Die Abbildung g in Teil a. ist dann eindeutig bestimmt und bijektiv. Wir nennen sie die *Inverse* oder *Umkehrabbildung* von f und bezeichnen sie mit f^{-1}.

Beweis:

a. "$\Longleftarrow$": Wir wollen zunächst zeigen, daß f surjektiv ist. Sei dazu $y \in N$ gegeben. Setze $x := g(y) \in M$. Dann gilt

$$f(x) = f\big(g(y)\big) = (f \circ g)(y) = \mathrm{id}_N(y) = y.$$

Also ist f surjektiv.
Dann wollen wir zeigen, daß f injektiv ist. Seien dazu $x, x' \in M$ mit $f(x) = f(x')$ gegeben. Dann gilt

$$x = \mathrm{id}_M(x) = (g \circ f)(x) = g\big(f(x)\big) = g\big(f(x')\big) = (g \circ f)(x') = \mathrm{id}_M(x') = x'.$$

Also ist f injektiv.
"$\Longrightarrow$": Da f bijektiv ist, gibt es für jedes $y \in N$ genau ein Urbild $x_y \in M$ von y unter f, d.h. $f(x_y) = y$. Wir definieren nun eine Abbildung

$$g : N \longrightarrow M : y \mapsto x_y.$$

Dann gilt zunächst für $y \in N$

$$(f \circ g)(y) = f\big(g(y)\big) = f(x_y) = y = \mathrm{id}_N(y).$$

Also ist $f \circ g = \mathrm{id}_N$.
Zudem gilt für $x \in M$ und $y := f(x) \in N$

$$f(x_y) = y = f(x).$$

Da f injektiv ist, folgt daraus $x = x_y$, und wir erhalten

$$(g \circ f)(x) = g\big(f(x)\big) = g(y) = x_y = x = \mathrm{id}_M(x).$$

Damit ist auch $g \circ f = \mathrm{id}_M$ gezeigt.

b. Sei $h : N \longrightarrow M$ eine zweite Abbildung mit $h \circ f = \mathrm{id}_M$ und $f \circ h = \mathrm{id}_N$. Dann gilt für $y \in N$

$$f\big(g(y)\big) = (f \circ g)(y) = \mathrm{id}_N(y) = (f \circ h)(y) = f\big(h(y)\big).$$

Da f injektiv ist, folgt mithin $g(y) = h(y)$, und somit $g = h$. Die Eindeutigkeit von g ist also gezeigt. Außerdem ist g nach Teil a. auch bijektiv.

□

Beispiel 3.13.
Die Abbildung $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto 2x + 1$ ist bijektiv mit $f^{-1} : \mathbb{R} \longrightarrow \mathbb{R} : y \mapsto \frac{1}{2} \cdot y - \frac{1}{2}$.

Denn für $y \in \mathbb{R}$ gilt

$$\big(f \circ f^{-1}\big)(y) = 2 \cdot \left(\frac{1}{2} \cdot y - \frac{1}{2}\right) + 1 = y = \mathrm{id}_{\mathbb{R}}(y)$$

und für $x \in \mathbb{R}$ gilt

$$\big(f^{-1} \circ f\big)(x) = \frac{1}{2} \cdot (2x + 1) - \frac{1}{2} = x = \mathrm{id}_{\mathbb{R}}(x).$$

Die Behauptung folgt also aus Satz 3.12.

Proposition 3.14 (Injektivität, Surjektivität, Bijektivität unter Komposition).
Seien $f : M \longrightarrow N$ und $g : N \longrightarrow P$ zwei Abbildungen.

a. Sind f und g injektiv, so ist $g \circ f$ injektiv.

b. Sind f und g surjektiv, so ist $g \circ f$ surjektiv.

c. Sind f und g bijektiv, so ist $g \circ f$ bijektiv.

Beweis: a. Seien $x, x' \in M$ mit $(g \circ f)(x) = (g \circ f)(x')$. Dann gilt

$$g(f(x)) = (g \circ f)(x) = (g \circ f)(x') = g(f(x')).$$

Da g injektiv ist, ist $f(x) = f(x')$, und da f injektiv ist, ist auch $x = x'$. Also ist $g \circ f$ injektiv.

b. Sei $z \in P$. Da g surjektiv ist, gibt es ein $y \in N$ mit $g(y) = z$, und da f surjektiv ist, gibt es ein $x \in M$ mit $f(x) = y$. Die Surjektivität von $g \circ f$ folgt dann aus

$$(g \circ f)(x) = g(f(x)) = g(y) = z.$$

c. Wegen a. ist $g \circ f$ injektiv und wegen b. ist $g \circ f$ auch surjektiv, also bijektiv.

□

Aufgaben

Aufgabe 3.15.
Ist $f : M \longrightarrow N$ eine surjektive Abbildung und $y \in N$, so ist

$$g : M \setminus f^{-1}(\{y\}) \longrightarrow N \setminus \{y\} : x \mapsto f(x)$$

eine surjektive Abbildung.

Aufgabe 3.16.
Es sei $f : M \longrightarrow N$ eine injektive Abbildung, $x' \in M$ und $y' = f(x') \in N$.

a. Dann ist $g : M \setminus \{x'\} \longrightarrow N \setminus \{y'\} : x \mapsto f(x)$ eine injektive Abbildung.

b. Ist f bijektiv, so ist g auch bijektiv.

Aufgabe 3.17.
Für eine Abbildung $f : M \longrightarrow N$, $M \neq \emptyset$, beweise man die folgenden Aussagen:

a. f ist injektiv $\iff \exists\, g : N \longrightarrow M$, so dass $g \circ f = \mathrm{id}_M$.

b. f ist surjektiv $\iff \exists\, g : N \longrightarrow M$, so dass $f \circ g = \mathrm{id}_N$.

Aufgabe 3.18.
Untersuche ob die folgenden Abbildungen injektiv, surjektiv oder bijektiv sind:

a. $f_1 : \mathbb{R} \longrightarrow \mathbb{R} : \; x \longmapsto 3x + 2$

b. $f_2 : \mathbb{Z} \longrightarrow \mathbb{Z} : \; x \longmapsto 3x + 2$

c. $f_3 : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \times \mathbb{R} : \; (x, y) \longmapsto (xy, x + 1)$

d. $f_4 : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \times \mathbb{R} : \; (x, y) \longmapsto (x - 2y, 2x + y)$

Aufgabe 3.19.
Seien M, N zwei nicht-leere Mengen und $f : M \longrightarrow N$ eine Abbildung. Formuliere die folgende Aussage in Quantorenschreibweise und beweise sie:
f ist genau dann surjektiv, wenn für alle nicht-leeren Mengen X und für alle Abbildungen $g : N \longrightarrow X$ und $h : N \longrightarrow X$ aus $g \circ f = h \circ f$ schon $g = h$ folgt.

Aufgabe 3.20.
Seien L, M, N Mengen und $f : L \longrightarrow M$, $g : M \longrightarrow N$ Abbildungen. Beweise oder widerlege - durch Gegenbeispiel - die folgenden Aussagen:

a. Ist $g \circ f$ injektiv, so ist g injektiv.

b. Ist $g \circ f$ injektiv, so ist f injektiv.

c. Ist $g \circ f$ surjektiv, so ist g surjektiv.

d. Ist $g \circ f$ surjektiv, so ist f surjektiv.

Aufgabe 3.21.
Seien M, N Mengen, $A_1, A_2 \subseteq M$ und $B, B_1, B_2 \subseteq N$ Teilmengen und $f : M \longrightarrow N$ eine Abbildung. Beweise die folgenden Aussagen:

a. $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$.

b. $f(f^{-1}(B)) \subseteq B$.

c. $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$.

d. $f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$.

Gib außerdem konkrete Beispiele dafür an, dass in b. und d. keine Gleichheit gilt.

Aufgabe 3.22.
Gib für die folgenden Abbildungsvorschriften die maximale Teilmenge von $\mathbb{R}$ an, die als Definitionsbereich in Frage kommt, und berechne jeweils auch das Bild der Abbildung:

a. $f(x) = \frac{1}{x}$,

b. $g(x) = x^2 - 4$,

c. $h(x) = \sqrt{(g \circ f)(x)}$,

d. $k(x) = \frac{1}{g(x)}$.

Aufgabe 3.23.
Bestimme die Menge $(h \circ g \circ f)^{-1}(B)$ für $B = \{x \in \mathbb{R} \mid x \geq 3\}$ und

$$f : \mathbb{R} \longrightarrow \mathbb{R} \times \mathbb{R} : x \mapsto (x, x-2),$$

$$g : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{Z} \times \mathbb{R} : (x, y) \mapsto \begin{cases} (1, x \cdot y), & \text{wenn } x > y \\ (-1, x \cdot y), & \text{wenn } x \leq y, \end{cases}$$

$$h : \mathbb{Z} \times \mathbb{R} \longrightarrow \mathbb{R} : (z, x) \mapsto x^z.$$

Aufgabe 3.24.
Überprüfe die folgenden Abbildungen auf Injektivität, Surjektivität und Bijektivität:

a. $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x^2+1}$,

b. $g : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} \setminus \{0\} : x \mapsto \frac{1}{x}$,

c. $h : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} : (x, y) \mapsto 2x + y$,

d. $k : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \times \mathbb{R} : (x, y) \mapsto (x + y, x - y)$.

Zeige oder widerlege jeweils, daß die Abbildung die jeweilige Eigenschaft besitzt, und gib im Falle der Bijektivität auch die Umkehrabbildung an.

§ 4 Vollständige Induktion

Ausblick 4.0 (Beweisen in der Informatik).
Wir lernen in diesem Abschnitt ein wichtiges Beweisverfahren in der Mathematik kennen, die *vollständige Induktion*. Andere Beweistechniken wie den *direkten Beweis* und die *Kontraposition* haben wir schon zuvor kennen gelernt. Auch für Informatiker gehört das Beweisen zum Alltagsgeschäft:

- Liefert das Protokoll das, was es liefern soll?
- Ist der Algorithmus korrekt?
- Terminiert er (in endlicher Zeit)?

Bemerkung 4.1 (Prinzip der vollständigen Induktion).
Die folgende Eigenschaft der natürlichen Zahlen ist uns wohl vertraut:

> Addiert man zur Zahl 0 sukzessive die Zahl 1, so erhält man nach und nach alle natürlichen Zahlen.

Man nennt sie das *Prinzip der vollständigen Induktion*.

Mit Hilfe der Nachfolgerfunktion $\mathrm{nf} : \mathbb{N} \longrightarrow \mathbb{N} : n \mapsto n+1$ können wir die Eigenschaft auch wie folgt formulieren:

> Ist $M \subseteq \mathbb{N}$ mit $0 \in M$ und $\forall\, n \in M \ : \ n+1 = \mathrm{nf}(n) \in M$, so ist $M = \mathbb{N}$.

Daraus leitet sich das im folgenden Satz formulierte Beweisprinzip für Aussagen ab, die von einer natürlichen Zahl abhängen.

Satz 4.2 (Prinzip der vollständigen Induktion).
Sei $\mathcal{A}(n)$ eine Aussageform mit zulässigen Werten $n \in \mathbb{N}$.
Falls $\mathcal{A}(0)$ wahr ist und $\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$ wahr ist, so ist $\mathcal{A}(n)$ wahr für alle $n \in \mathbb{N}$.

Beweis: Setze $M := \{n \in \mathbb{N} \mid \mathcal{A}(n) \text{ wahr}\}$. Nach Voraussetzung gilt dann $0 \in M$ und für $n \in M$ folgt $n+1 \in M$. Aus dem Prinzip der Vollständigen Induktion in Bemerkung 4.1 folgt dann $M = \mathbb{N}$. Also ist $\mathcal{A}(n)$ wahr für alle $n \in \mathbb{N}$. □

Bemerkung 4.3.
Man beachte, um den Schluß $\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$ als wahr zu erweisen, reicht es, den Fall zu betrachten, daß $\mathcal{A}(n)$ wahr ist, da andernfalls der Schluß ohnehin den Wahrheitswert wahr trägt.

Wir nennen:

- "$\mathcal{A}(0)$ wahr" den *Induktionsanfang*,
- "$\mathcal{A}(n)$ wird als wahr vorausgesetzt" die *Induktionsvoraussetzung* und
- "$\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$" den *Induktionsschluß*.

Beispiel 4.4.
Die Zahl $n^3 - n$ ist für jedes $n \in \mathbb{N}$ durch 6 teilbar.

Beweis: Wir führen den Beweis durch vollständige Induktion und formulieren dazu zunächst unsere Aussageform $\mathcal{A}(n)$:

$$\mathcal{A}(n) : \text{Es gibt ein } k \in \mathbb{N} \text{ mit } n^3 - n = 6 \cdot k.$$

Induktionsanfang: $n = 0$**:** $0^3 - 0 = 0 = 6 \cdot 0$. Also ist $\mathcal{A}(0)$ wahr.

Induktionsvoraussetzung: Wir setzen voraus, daß $\mathcal{A}(n)$ wahr ist, d.h. es gibt ein $k \in \mathbb{N}$ mit $n^3 - n = 6 \cdot k$.

Induktionsschritt: $n \mapsto n+1$**:** Man beachte, daß eine der beiden Zahlen n oder $n+1$ gerade sein muß, und daß deshalb die Zahl $n \cdot (n+1)$ gerade ist. Es gibt also eine natürliche Zahl $l \in \mathbb{N}$ mit $n \cdot (n+1) = 2 \cdot l$. Damit erhalten wir dann

$$(n+1)^3 - (n+1) = (n^3 - n) + 3 \cdot n \cdot (n+1) = 6k + 6l = 6 \cdot (k+l).$$

Wir haben also gezeigt, daß $\mathcal{A}(n+1)$ wahr ist.

Also ist $\mathcal{A}(n)$ wahr für alle $n \in \mathbb{N}$, und das heißt, daß $n^3 - n$ stets durch 6 teilbar ist. □

Bemerkung 4.5 (Varianten der vollständigen Induktion).

a. *Alternativer Induktionsanfang*:
Statt $n = 0$ als Induktionsanfang zu wählen, kann eine beliebige ganze Zahl $n_0 \in \mathbb{Z}$ als Induktionsanfang dienen. Man erhält dann, daß $\mathcal{A}(n)$ wahr ist für alle $n \geq n_0$. Denn, man erhält alle ganzen Zahlen $n \geq n_0$, indem man zu n_0 sukzessive 1 addiert.

b. *Alternative Induktionsvoraussetzung*:
Im Induktionsschritt schließen wir von $\mathcal{A}(n)$ auf $\mathcal{A}(n+1)$, d.h. wir setzen nur $\mathcal{A}(n)$ als richtig voraus und schließen daraus die Korrektheit von $\mathcal{A}(n+1)$. Stattdessen können wir auch $\mathcal{A}(k)$ für $k = n_0, \ldots, n$ als richtig voraussetzen und auf $\mathcal{A}(n+1)$ schließen (wobei $\mathcal{A}(n_0)$ der Induktionsanfang sein soll). Das ist manchmal hilfreich.

Bemerkung 4.6 (Vollständige Induktion ist ohne Induktionsanfang wertlos.).
Bei einem Induktionsbeweis ist es meist weit leichter, den Induktionsanfang zu zeigen als den Induktionsschluß. Man sollte diesen aber auf keinen Fall vernachlässigen, da ein Beweis mittels Induktion ohne Induktionsanfang schlicht kein Beweis ist.

Wir wollen dies an folgendem Beispiel demonstrieren. Wir betrachten dazu die Aussageform

$$\mathcal{A}(n) : \exists\, k \in \mathbb{N} \;:\; 10^n = 3 \cdot k,$$

d.h. jede Potenz von 10 ist durch 3 teilbar. Wollte man dies durch vollständige Induktion beweisen, wäre der Induktionsschluß leicht zu führen. Man nimmt dazu an, die Aussage sei für eine natürliche Zahl n schon gezeigt, d.h. es gibt ein $k \in \mathbb{N}$ mit

$$10^n = 3 \cdot k$$

und folgert dann, daß auch

$$10^{n+1} = 10^n \cdot 10 = (3 \cdot k) \cdot 10 = 3 \cdot (10k)$$

ein Vielfaches von 3 ist. Die Korrektheit des Schlusses allein reicht aber nicht aus, um die Aussage auch nur für eine einzige natürliche Zahl n zu zeigen. Solange wir nicht mindestens eine natürliche Zahl gefunden haben, für die die Aussage korrekt ist und die uns dann als Induktionsanfang dient, ist der Beweis unvollständig. Einen solchen Induktionsanfang werden wir im vorliegenden Beispiel aber nicht finden, weil schlicht keine Potenz von 10 durch 3 teilbar ist!

Das Beispiel sollte auch noch mal verdeutlichen, daß der Beweis der Korrektheit des Schlusses "$\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$" *nicht* bedeutet, daß $\mathcal{A}(n+1)$ wahr ist! Dies folgt nur, *wenn* $\mathcal{A}(n)$ wirklich wahr ist, nicht aber, wenn $\mathcal{A}(n)$ falsch ist, was der Korrektheit der Schlußfolgerung an sich keinen Abbruch tut!

Aufgaben

Aufgabe 4.7.
Zeige, daß $3^{n+1} - 3$ für jede natürliche Zahl $n \in \mathbb{N}$ durch 6 teilbar ist.

Aufgabe 4.8.
Es sei $a \in \mathbb{N}$ eine natürliche Zahl. Zeige, daß dann $a^{2n+1} - a$ für jede natürliche Zahl $n \in \mathbb{N}$ durch 6 teilbar ist.

Aufgabe 4.9.
Beweise mittels vollständiger Induktion $\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \cdots + \frac{1}{\sqrt{n-1}} + \frac{1}{\sqrt{n}} > \sqrt{n}$ für alle $n \geq 2$.

Aufgabe 4.10.
Beweise die folgenden Aussagen mittels vollständiger Induktion:

a. Für $n \geq 2$ gilt $\prod_{k=2}^{n} \left(1 - \frac{1}{k}\right) = \frac{1}{n}$.

b. Für $n \in \mathbb{N}$ gilt $\sum_{k=0}^{n} (k+1) \cdot \binom{n}{k} = 2^{n-1} \cdot (n+2)$.

c. Für $n \geq 2$ gilt $\frac{4^n}{n+1} < \frac{(2n)!}{(n!)^2}$.

d. Für $n \in \mathbb{N}$ gilt $\sum_{k=1}^{n} (-1)^{n-k} \cdot k^2 = \binom{n+1}{2}$.

§ 5 Mächtigkeit von Mengen

A) Endliche Mengen

Definition 5.1 (Die Mächtigkeit von Mengen).

a. Wir nennen eine Menge M *endlich*, wenn sie nur endlich viele Elemente enthält. In diesem Fall bezeichnen wir mit $\#M = |M|$ die Anzahl an Elementen in M und nennen die Zahl die *Mächtigkeit* von M. Enthält M unendlich viele Elemente, so nennen wir M *unendlich* und setzen $\#M := |M| := \infty$.

b. Zwei Mengen M und N heißen *gleichmächtig*, wenn es eine bijektive Abbildung $f : M \longrightarrow N$ gibt.

c. Eine Menge heißt *abzählbar unendlich*, wenn sie gleichmächtig zu $\mathbb{N}$ ist.

d. Eine Menge heißt *überabzählbar*, wenn sie weder endlich noch abzählbar unendlich ist.

e. Für zwei ganze Zahlen $m, n \in \mathbb{Z}$ bezeichnen wir mit

$$\{m, \ldots, n\} := \{k \in \mathbb{Z} \mid m \leq k \leq n\}$$

die Menge der ganzen Zahlen zwischen m und n. Man beachte, daß $\{m, \ldots, n\} = \emptyset$, wenn $m > n$.

Bemerkung 5.2 (Einfache Eigenschaften der Mächtigkeit endlicher Mengen).

a. Ist eine Menge endlich und enthält genau n Elemente, so können wir die Elemente in M abzählen, etwa $x_1, x_2, x_3, \ldots, x_n$ und wir erhalten so eine bijektive Abbildung

$$f : \{1, \ldots, n\} \longrightarrow M : i \mapsto x_i.$$

Umgekehrt erlaubt eine solche Abbildung, die Elemente von M abzuzählen und wir erhalten $|M| = n$. Damit sehen wir, daß eine Menge genau dann endlich von Mächtigkeit n ist, wenn es eine Bijektion von $\{1, \ldots, n\}$ nach M gibt.

b. Ist M endlich und $A \subseteq M$, so ist auch A endlich und $|A| \leq |M|$.

c. Ist $M = A \dot\cup B$ eine endliche Menge, so gilt $|M| = |A| + |B|$.

Wir wollen den in Bemerkung 5.2 angedeuteten Zusammenhang zwischen der Mächtigkeit einer endlichen Menge und der Existenz von Abbildungen mit bestimmten Eigenschaften im folgenden Satz vertiefen.

Satz 5.3.
Es seien M und N zwei nicht-leere endliche Mengen.

a. Genau dann gilt $|M| \leq |N|$, wenn es eine injektive Abbildung $f : M \to N$ gibt.

b. Genau dann gilt $|M| \geq |N|$, wenn es eine surjektive Abbildung $f : M \to N$ gibt.

c. Genau dann gilt $|M| = |N|$, wenn es eine bijektive Abbildung $f : M \to N$ gibt.

Beweis: Es seien $M = \{x_1, \ldots, x_m\}$ und $N = \{y_1, \ldots, y_n\}$ mit paarweise verschiedenen Elementen $x_i \neq x_j$ für $i \neq j$ und $y_i \neq y_j$ für $i \neq j$. Es gilt $|M| = m > 0$ und $|N| = n > 0$.

a. Ist $m \leq n$, so definiere $f : M \to N$ durch $f(x_i) = y_i$ für $i = 1, \ldots, m$. Dann gilt für $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$f(x_i) = y_i \neq y_j = f(x_j).$$

Mithin ist f injektiv.
Ist umgekehrt $f : M \to N$ eine injektive Abbildung, so gilt $f(M) = \{f(x_1), \ldots, f(x_m)\} \subseteq N$ eine Teilmenge von paarweise verschiedenen Elementen. Mithin enthält N mindestens m Elemente, und folglich gilt $m \leq n$.

b. Ist $m \geq n$, so definiere $f : M \to N$ durch $f(x_i) = y_i$ für $i = 1, \ldots, n$ und $f(x_i) = y_1$ für $i = n+1, \ldots, m$. Dann gilt offenbar $f(M) = \{y_1, \ldots, y_n\} = N$ und f ist surjektiv.
Ist umgekehrt $f : M \to N$ eine surjektive Abbildung, so gilt $\{y_1, \ldots, y_n\} = N = f(M) = \{f(x_1), \ldots, f(x_m)\}$. Mithin enthält die Menge $\{f(x_1), \ldots, f(x_m)\}$ auch n verschiedene Elemente, und folglich ist $m \geq n$.

c. Die Rückrichtung folgt unmittelbar aus den ersten beiden Teilen. Für die Hinrichtung beachte man, daß die in a. und b. definierten Abbildungen im Fall $|M| = |N|$ übereinstimmen und somit bijektiv sind.

□

Bemerkung 5.4 (Schubfachprinzip).
Die Kontraposition von Teil a. in Satz 5.3 nennt man auch das *Schubfachprinzip*: wenn $f : M \longrightarrow N$ eine Abbildung ist und $|M| > |N|$ gilt, dann ist f nicht injektiv.

Übersetzt, wenn man $m > n$ Gegenstände auf n Schubfächer verteilen möchte, dann muß man in mindestens ein Schubfach zwei legen.

Beispiel 5.5.
In einem Raum mit 13 Personen haben mindestens zwei im selben Monat Geburtstag.

Aus Satz 5.3 leitet sich zudem unmittelbar ab, daß für Selbstabbildungen endlicher Mengen die Begriffe injektiv, surjektiv und bijektiv zusammen fallen.

Korollar 5.6 (Injektiv = surjektiv = bijektiv für gleichmächtige endliche Mengen).
Es seien M und N endliche Mengen mit $|M| = |N|$. Dann sind die folgenden Aussagen für eine Abbildung $f : M \longrightarrow N$ äquivalent:

a. f ist injektiv.

b. f ist surjektiv.

c. f ist bijektiv.

Beweis:

a. $\Longrightarrow$ b.: Angenommen, f wäre nicht surjektiv, dann gibt es ein

$$y \in N \setminus \operatorname{Im}(f)$$

und mithin ist

$$\operatorname{Im}(f) \subseteq N \setminus \{y\}.$$

Da f injektiv ist, ist $g : M \longrightarrow \operatorname{Im}(f) : x \mapsto f(x)$ nach Beispiel 3.8 bijektiv, so daß mit Satz 5.3

$$|M| \overset{5.3}{=} |\operatorname{Im}(f)| \leq |N| - 1 < |N| = |M|$$

folgt, was ein offensichtlicher Widerspruch ist. Mithin muß f surjektiv sein.

b. $\Longrightarrow$ c.: Wir müssen zeigen, daß f injektiv ist. Dazu nehmen wir an, f sei nicht injektiv. Dann gibt es $x, x' \in M$ mit $x \neq x'$ und $y := f(x) = f(x')$. Die Abbildung

$$h : M \setminus f^{-1}(\{y\}) \longrightarrow N \setminus \{y\} : z \mapsto f(z)$$

ist nach Aufgabe 3.15 surjektiv. Mithin gilt nach Satz 5.3

$$|M| - 1 \overset{\text{Vor.}}{=} |N| - 1 = |N \setminus \{y\}| \overset{5.3}{\leq} |M \setminus f^{-1}(\{y\})| \leq |M \setminus \{x, x'\}| = |M| - 2,$$

was offenbar ein Widerspruch ist. Mithin muß f injektiv sein.

c. $\Longrightarrow$ a.: Jede bijektive Abbildung ist auch injektiv, also ist f injektiv.

Damit haben wir die Aussage durch einen *Ringschluß* gezeigt. □

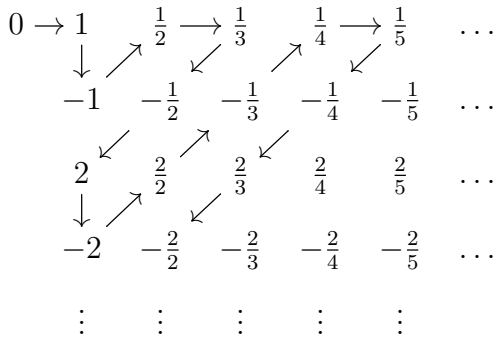
B) Das Cantorsche Diagonalisierungsverfahren

Nachdem wir uns bislang im wesentlichen mit endlichen Mengen beschäftigt haben, wollen wir uns nun unendlichen Mengen zuwenden und dabei zeigen, daß es unterschiedliche Qualitäten der Unendlichkeit gibt.

Proposition 5.7 (Cantorsches Diagonalverfahren).
Die Menge $\mathbb{Q}$ der rationalen Zahlen ist abzählbar unendlich.

Beweis: Wir zeigen, wie man mit Hilfe des Cantorschen Diagonalverfahrens eine bijektive Abbildung von $\mathbb{N}$ nach $\mathbb{Q}$ konstruiert.

Dazu listen wir die rationalen Zahlen zunächst wie folgt auf



und laufen sie dann wie angedeutet entlang der Pfeile ab. Dabei sammeln wir jede rationale Zahl, die mehrfach vorkommt, nur bei ihrem ersten Auftreten auf. Auf dem Weg erhalten wir eine bijektive Abbildung von $\mathbb{N}$ nach $\mathbb{Q}$. □

Proposition 5.8 ($\mathbb{R}$ ist überabzählbar.).
Die Menge $\mathbb{R}$ der reellen Zahlen ist überabzählbar.

Beweis: Auch dies zeigen wir mit Hilfe einer Variante des Cantorschen Diagonalverfahrens.

$\mathbb{R}$ ist sicherlich nicht endlich. Wäre $\mathbb{R}$ abzählbar unendlich, so gäbe es eine bijektive Abbildung $\varphi : \mathbb{N} \to \mathbb{R}$, und wir schreiben dann $\varphi(i)$, $i \in \mathbb{N}$, in Dezimaldarstellung (siehe auch Satz 12.45):

$$\begin{array}{rcccccccc}
\varphi(0) & = & a_{0,-p_0} & a_{0,-p_0+1} & \dots & \underline{a_{0,0}}, & a_{01} & a_{02} & a_{03} & \cdots \\
\varphi(1) & = & a_{1,-p_1} & a_{1,-p_1+1} & \dots & a_{1,0}, & \underline{a_{11}} & a_{12} & a_{13} & \cdots \\
\varphi(2) & = & a_{2,-p_2} & a_{2,-p_2+1} & \dots & a_{2,0}, & a_{21} & \underline{a_{22}} & a_{23} & \cdots \\
\vdots & & & & & & & & \ddots &
\end{array}$$

Dann setzen wir $a := a_{00}, a_{11}a_{22}a_{33}\cdots \in \mathbb{R}$, d. h. a ist diejenige Zahl, die in obiger Aufzählung durch die unterstrichenen Diagonalelemente gegeben ist. Nun ändern wir

jede der Ziffern von a ab (etwa $b_{ii} = 2$, falls $a_{ii} = 0$ und $b_{ii} = 0$ sonst) und erhalten eine Zahl

$$b = b_{00}, b_{11}b_{22}b_{33} \cdots \in \mathbb{R},$$

mit $a_{ii} \neq b_{ii}$ für alle $i \in \mathbb{N}$. Da φ bijektiv ist, gibt es ein $i \in \mathbb{N}$ mit $\varphi(i) = b$, also $a_{ii} = b_{ii}$, im Widerspruch zur Konstruktion von b. (Wir müssen noch berücksichtigen, daß $0,9999\cdots = 1$, was aber die einzige Zweideutigkeit der Dezimaldarstellung ist, und dieser weichen wir durch unsere Wahl der b_{ii} aus.) Also ist $\mathbb{R}$ überabzählbar. □

Bemerkung 5.9 (Kontinuumshypothese).
Da $\mathbb{Q}$ und $\mathbb{R}$ nicht gleichmächtig sind und $\mathbb{Q}$ eine Teilmenge von $\mathbb{R}$ ist, stellt sich ganz natürlich die Frage, ob es eine Menge M mit $\mathbb{Q} \subsetneqq M \subsetneqq \mathbb{R}$ gibt, die weder zu $\mathbb{Q}$ noch zu $\mathbb{R}$ gleichmächtig ist. Es hat lange gedauert, bis man feststellen mußte, daß die Frage auf der Grundlage des allgemein anerkannten Axiomensystems der Mengenlehre von Zermelo-Fränkel nicht entscheidbar ist. Man hat nun also die Wahl, als neues Axiom hinzuzufügen, daß es eine solche Menge gibt, oder auch, daß es keine solche Menge gibt. Die lange bestehende Vermutung, daß man schon mit den übrigen Axiomen beweisen könnte, daß es keine solche Menge gibt, ist als *Kontinuumshypothese* bekannt.

C) Potenzmengen

Definition 5.10 (Potenzmenge).
Es sei M ein eine Menge. Wir nennen die Menge

$$\mathcal{P}(M) := \{A \mid A \subseteq M\}$$

aller Teilmengen von M die *Potenzmenge* von M.

Beispiel 5.11.
$\mathcal{P}(\emptyset) = \{\emptyset\}$, $\mathcal{P}(\{1\}) = \{\emptyset, \{1\}\}$, $\mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$.

Proposition 5.12 (Potenzmengen endlicher Mengen).
Sei M eine endliche Menge mit $n = |M|$, so ist $|\mathcal{P}(M)| = 2^n$.

Beweis: Wir führen den Beweis durch Induktion nach n.

Induktionsanfang: $n = 0$**:** Dann ist $M = \emptyset$ und $\mathcal{P}(M) = \{\emptyset\}$ hat genau $1 = 2^0$ Elemente.

Induktionsschritt: $n \mapsto n+1$**:** Sei also $|M| = n+1$. Wir wählen ein $y \in M$ und setzen $N = M \setminus \{y\}$, so daß $|N| = |M| - 1 = n$. Die Potenzmenge $\mathcal{P}(M)$ läßt sich

nun wie folgt disjunkt aufspalten:

$$\mathcal{P}(M) = \{A \subseteq M \mid y \notin A\} \cup \{A \subseteq M \mid y \in A\}.$$

Dabei ist

$$\{A \subseteq M \mid y \notin A\} = \{A \subseteq M \mid A \subseteq N\} = \mathcal{P}(N)$$

und

$$\{A \subseteq M \mid y \in A\} = \{B \cup \{y\} \subseteq M \mid B \subseteq N\} = \{B \cup \{y\} \subseteq M \mid B \in \mathcal{P}(N)\}.$$

Beide Mengen sind offenbar gleichmächtig zu $\mathcal{P}(N)$, und nach Induktionsvoraussetzung gilt $|\mathcal{P}(N)| = 2^n$. Insgesamt erhalten wir also

$$|\mathcal{P}(M)| = 2^n + 2^n = 2^{n+1}.$$

Damit folgt die Aussage mittels Induktion. □

Aufgaben

Aufgabe 5.13.
Die Menge $\mathbb{Z}$ der ganzen Zahlen ist abzählbar unendlich.

Aufgabe 5.14.
In einer Gruppe mit mindestens zwei Personen gibt es mindestens zwei, die dieselbe Anzahl an Personen in der Gruppe kennen.

Aufgabe 5.15.
Sei M eine Menge und $N := \{g \mid g : M \to \{0,1\}$ Abbildung$\}$. Zeige, die Menge N ist gleichmächtig zur Potenzmenge $\mathcal{P}(M)$.

§ 6 Äquivalenzrelationen

Ausblick 6.0 (Relationen in der Informatik).
In diesem Abschnitt und in Abschnitt 8 beschäftigen wir uns mit speziellen Typen von *Relationen*. Eine Relation stellt eine Beziehung zwischen den Elementen einer Menge mit den Elementen einer anderen Menge her. Dieses Prinzip spielt in der Informatik etwa bei *relationalen Datenbanken* eine wichtige Rolle. Für ein simples Beispiel siehe etwa

http://wikis.gm.fh-koeln.de/Datenbanken/Kartesisches-Produkt

Äquivalenzrelationen stellen ein sehr wichtiges *Ordnungs-* und *Konstruktionsprinzip* innerhalb der Mathematik dar, das im Verlauf der ersten Semester an einigen zentralen Stellen benötigt wird, etwa im Zusammenhang mit Faktorräumen, der Äquivalenz von Matrizen oder der Konjugation von Matrizen und der Jordanschen Normalform.

A) Äquivalenzrelationen und Äquivalenzklassen

Definition 6.1 (Relation).
Seien M und N zwei Mengen, so nennen wir jede Teilmenge $R \subseteq M \times N$ eine *Relation* zwischen M und N.

Bemerkung 6.2.
Ist R eine Relation zwischen M und N, $x \in M$ und $y \in N$, so wollen wir sagen *x steht in Relation zu y bezüglich R*, wenn $(x, y) \in R$. Die Menge R legt also fest, wann zwei Elemente in Relation zueinander stehen. Wir schreiben auch xRy statt $(x, y) \in R$.

Beispiel 6.3 (Abbildungen als Relationen).

a. Der Graph einer Abbildung $f : M \longrightarrow N$ ist ein Beispiel einer Relation, bei der jedes $x \in M$ zu genau einem $y \in N$ in Relation steht.

b. Ist M die Menge der Hörer der Vorlesung und N die Menge der in Tübingen studierbaren Fächer, so ist

$$R = \{(x, y) \in M \times N \mid x \text{ studiert } y\}$$

eine Relation zwischen M und N, die ganz sicher nicht Graph einer Funktion ist.

Bemerkung 6.4 (Motivation des Begriffs Äquivalenzrelation).
Der folgende Begriff der *Äquivalenzrelation* bereitet den Studenten oft extreme Schwierigkeiten. Dabei liegt auch ihm ein ganz einfaches Prinzip zugrunde, das wir zunächst an einem Beispiel erläutern wollen.

Die Gesamtheit aller Schüler einer Schule werden von der Schulleitung zwecks sinnvoller Organisation des Unterrichts in Schulklassen eingeteilt. Dabei achtet die Schulleitung darauf, daß jeder Schüler zu einer Schulklasse gehört und auch nur zu dieser einen. Etwas mathematischer ausgedrückt, die Schulleitung teilt die *Menge* S der Schüler in *paarweise disjunkte Teilmengen* K_i, $i = 1, \ldots, k$, ein, so daß wir anschließend eine *disjunkte Zerlegung*

$$S = \dot{\bigcup_{i=1}^{k}} K_i$$

der Menge S in die Schulklassen $K_1, \ldots, K_k$ haben. Dabei kann man für die Zugehörigkeit der Schüler Alfred, Ben und Christoph zu einer Schulklasse folgendes feststellen:

1) Alfred gehört zu einer Schulklasse.
2) Wenn Alfred in derselben Schulklasse ist wie Ben, dann ist Ben auch in derselben Schulklasse wie Alfred.
3) Wenn Alfred in derselben Schulklasse ist wie Ben und wenn zugleich Ben in derselben Schulklasse ist wie Christoph, dann ist auch Alfred in derselben Schulklasse wie Christoph.

Diese Aussagen sind so offensichtlich, daß man kaum glauben mag, daß es einen tieferen Sinn hat, sie zu erwähnen. Aber nehmen wir für einen Augenblick an, die Schulleitung hat ihre Einteilung der Schüler vorgenommen und für jede Schulklasse eine Liste mit den Namen der Schüler erstellt, die zu dieser Schulklasse gehören sollen. Nehmen wir ferner an, die Schulleitung hat noch nicht überprüft, ob jeder Schüler in genau einer Schulklasse eingeteilt ist. Dann behaupte ich, wenn man in den drei Aussagen 1)-3) die Schüler Alfred, Ben und Christoph durch beliebige Schüler ersetzt und die Aussagen richtig sind für jede Kombination der Schülernamen, dann ist sichergestellt, daß auch jeder Schüler in genau einer Schulklasse eingeteilt ist.

Als Mathematiker suchen wir nach möglichst einfachen Regeln, denen die Einteilung der Schulklassen genügen muß, um sicherzustellen, daß sie wirklich eine disjunkte Zerlegung von S ist, d.h. daß wirklich jeder Schüler in genau einer Schulklasse ist, und die Regeln 1)-3) sind genau die Regeln, die wir dazu brauchen. Wenn wir nun die Zugehörigkeit zweier Schüler x und y zur selben Klasse verstehen als "x *steht in Relation zu* y", dann definieren uns die drei Regeln 1)-3) zudem eine Teilmenge von $S \times S$, nämlich die Relation

$$R = \{(x, y) \in S \times S \mid x \text{ ist in derselben Schulklasse wie } y\}.$$

Die Regeln 1)-3) lassen sich für Schüler $x, y, z \in S$ dann wie folgt formulieren:

- $(x, x) \in R$.
- Wenn $(x, y) \in R$, dann ist auch $(y, x) \in R$.
- Wenn $(x, y) \in R$ und $(y, z) \in R$, dann ist auch $(x, z) \in R$.

Eine solche Relation nennt man eine *Äquivalenzrelation*, man nennt Schüler derselben Schulklasse *äquivalent* und die Schulklassen nennt man dann auch *Äquivalenzklassen*.

Wir führen den Begriff der *Äquivalenzrelation* nun für beliebige Mengen ein.

Definition 6.5 (Äquivalenzrelation).
Es sei M eine Menge. Eine *Äquivalenzrelation* auf M ist eine Teilmenge $R \subseteq M \times M$, so daß für alle $x, y, z \in M$ gilt:

R1: $(x, x) \in R$, ("Reflexivität")
R2: $(x, y) \in R \implies (y, x) \in R$, ("Symmetrie")
R3: $(x, y), (y, z) \in R \implies (x, z) \in R$. ("Transitivität")

Bei Äquivalenzrelationen hat sich eine alternative Schreibweise zu $(x, y) \in R$ durchgesetzt, die auch wir im folgenden verwenden wollen.

Notation 6.6 (Schreibweise $\sim$ für Äquivalenzrelationen).
Äquivalenzrelationen werden oft eher mit Symbolen wie $\sim$ statt Buchstaben wie R bezeichnet. Mit der Schreibweise aus Bemerkung 6.2 gilt dann

$$x \sim y \iff (x, y) \in \sim$$

und die drei Axiome in Definition 6.5 lassen sich wie folgt formulieren. Für $x, y, z \in M$ soll gelten:

R1: $x \sim x$, ("Reflexivität")
R2: $x \sim y \implies y \sim x$, ("Symmetrie")
R3: $x \sim y, y \sim z \implies x \sim z$. ("Transitivität")

Definition 6.7 (Äquivalenzklassen).
Es sei M eine Menge und $\sim$ eine Äquivalenzrelation auf M. Für $x \in M$ heißt die Menge

$$\overline{x} := \{y \in M \mid y \sim x\}$$

die *Äquivalenzklasse* von x. Jedes $y \in \overline{x}$ heißt ein *Repräsentant* der Klasse $\overline{x}$. Mit

$$M/\sim := \left\{\overline{x} \mid x \in M\right\}$$

bezeichnen wir die Menge der *Äquivalenzklassen modulo der Äquivalenzrelation* $\sim$.

Beispiel 6.8 (Der Abstand vom Ursprung als Äquivalenzrelation).
Wir betrachten die Menge $M = \mathbb{R}^2$ der Punkte in der reellen Zahlenebene und wir bezeichnen mit $|P|$ den Abstand von P zum Ursprung $(0,0)$. Für zwei Punkte $P, Q \in M$ definieren wir

$$P \sim Q \iff |P| = |Q|,$$

d.h. wir nennen die Punkte *äquivalent*, falls ihr Abstand zum Ursprung gleich ist. Dann ist $\sim$ eine Äquivalenzrelation.

R1: Sei $P \in M$, dann ist $|P| = |P|$, also $P \sim P$.
R2: Falls $P, Q \in M$ mit $P \sim Q$, dann ist $|P| = |Q|$ und somit auch $|Q| = |P|$. Damit gilt aber $Q \sim P$.
R3: Falls $P, Q, R \in M$ mit $P \sim Q$ und $Q \sim R$, dann gilt $|P| = |Q|$ und $|Q| = |R|$. Aber damit gilt auch $|P| = |R|$ und somit $P \sim R$.

Die Äquivalenzklasse

$$\overline{P} = \{Q \in M \mid |Q| = |P|\}$$

von $P \in M$ ist der Kreis um den Ursprung vom Radius $|P|$.

B) Äquivalenzrelationen und disjunkte Zerlegungen

Wir haben anfangs behauptet, daß die drei Axiome einer Äquivalenzrelation sicherstellen, daß die zugehörigen Äquivalenzklassen eine disjunkte Zerlegung von M induzieren, und umgekehrt, daß jede disjunkte Zerlegung eine Äquivalenzrelation mit sich bringt. Dies wollen wir im Folgenden beweisen. Dazu sollten wir zunächst den Begriff disjunkt klären.

Proposition 6.9 (Die Äquivalenzrelation zu einer disjunkten Zerlegung).
Ist $(M_i)_{i\in I}$ eine disjunkte Zerlegung von M und definieren wir eine Relation auf M durch

$$x \sim y \iff \exists\, i \in I \,:\, x, y \in M_i,$$

dann ist $\sim$ eine Äquivalenzrelation auf M.

Beweis: Ist $x \in M = \bigcup_{i\in I} M_i$, so gibt es ein $i \in I$ mit $x \in M_i$ und somit gilt $x \sim x$. $\sim$ ist also reflexiv.

Sind $x, y \in M$ mit $x \sim y$, so gibt es ein $i \in I$ mit $x, y \in M_i$. Dann gilt aber auch $y \sim x$. Die Relation ist also symmetrisch.

Sind $x, y, z \in M$ mit $x \sim y$ und $y \sim z$, so gibt es $i, j \in I$ mit $x, y \in M_i$ und $y, z \in M_j$. Da die Zerlegung disjunkt ist und $y \in M_i \cap M_j$, folgt $M_i = M_j$. Also gilt $x, z \in M_i$ und somit $x \sim z$. $\sim$ ist also auch transitiv. □

Proposition 6.10 (Die disjunkte Zerlegung zu einer Äquivalenzrelation).
Es sei M eine Menge. Ist $\sim$ eine Äquivalenzrelation auf M, dann bilden die Äquivalenzklassen eine disjunkte Zerlegung von M, d. h. jedes $x \in M$ liegt in genau einer Äquivalenzklasse.
Insbesondere gilt für Äquivalenzklassen $\overline{x}$ und $\overline{y}$ entweder $\overline{x} = \overline{y}$ oder $\overline{x} \cap \overline{y} = \emptyset$.

Beweis: Sei $x \in M$ beliebig. Aus $x \sim x$ folgt $x \in \overline{x} \subseteq \bigcup_{\overline{y} \in M/\sim} \overline{y}$. Mithin gilt

$$M = \bigcup_{\overline{y} \in M/\sim} \overline{y}.$$

Es bleibt also zu zeigen, daß die Äquivalenzklassen paarweise disjunkt sind.

Seien $\overline{x}, \overline{y} \in M/\sim$ mit $\overline{x} \cap \overline{y} \neq \emptyset$. Dann gibt es ein $z \in \overline{x} \cap \overline{y}$, und es gilt $z \sim x$ und $z \sim y$. Wegen der Symmetrie gilt aber auch $x \sim z$ und mittels der Transitivität dann $x \sim y$. Sei nun $u \in \overline{x}$ beliebig, dann gilt $u \sim x$ und wieder wegen der Transitivität $u \sim y$. Also $u \in \overline{y}$ und damit $\overline{x} \subseteq \overline{y}$. Vertauschung der Rollen von x und y in der Argumentation liefert schließlich $\overline{x} = \overline{y}$. □

Korollar 6.11 (Äquivalenzrelationen auf endlichen Mengen).
Sei M eine endliche Menge, $\sim$ eine Äquivalenzrelation auf M und $M_1, \ldots, M_s$ seien die paarweise verschiedenen Äquivalenzklassen von $\sim$. Dann gilt:

$$|M| = \sum_{i=1}^{s} |M_i|.$$

Beweis: Mit M sind auch alle M_i endlich und die Behauptung folgt aus Proposition 6.10 und Bemerkung 5.2. □

Ein Beispiel aus dem Alltag für eine Äquivalenzrelation haben wir oben bereits gesehen. Ein weiteres wichtiges und wohlbekanntes Beispiel sind die rationalen Zahlen! Ein Bruch ist nichts weiter als die Äquivalenzklasse eines Tupels von ganzen Zahlen, und das Kürzen des Bruches, z.B. $\frac{1}{2} = \frac{2}{4}$, ist nur die Wahl eines möglichst einfachen Repräsentanten.

Beispiel 6.12 (Die rationalen Zahlen).
Man kann die rationalen Zahlen wie folgt als Äquivalenzklassen von Paaren ganzer Zahlen definieren. Für $(p, q), (p', q') \in M := \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ definiere

$$(p, q) \sim (p', q') \quad :\Longleftrightarrow \quad pq' = p'q.$$

Wir wollen nun zeigen, daß hierdurch wirklich eine Äquivalenzrelation auf M definiert wird. Seien dazu $x = (p, q), x' = (p', q'), x'' = (p'', q'') \in M$ gegeben:[1]

R1: Für die Reflexivität müssen wir $x \sim x$ zeigen. Nun gilt aber $pq = pq$, woraus $x = (p, q) \sim (p, q) = x$ folgt.

R2: Für die Symmetrie nehmen wir an, daß $x \sim x'$ gilt und müssen $x' \sim x$ folgern. Wegen $x \sim x'$ gilt aber nach Definition $pq' = p'q$, und folglich auch $p'q = pq'$. Letzteres bedeutet aber, daß $x' = (p', q') \sim (p, q) = x$.

R3: Für die Transitivität nehmen wir schließlich an, daß $x \sim x'$ und $x' \sim x''$ gilt, und müssen daraus schließen, daß $x \sim x''$. Wegen $x \sim x'$ gilt nun aber $pq' = p'q$, und wegen $x' \sim x''$ gilt $p'q'' = p''q'$. Multiplizieren wir die erste der Gleichungen mit q'' und die zweite mit q, so erhalten wir

$$pq'q'' = p'qq'' = p'q''q = p''q'q.$$

Da nach Voraussetzung $q' \neq 0$, können wir beide Seiten der Gleichung durch q' teilen und erhalten:

$$pq'' = p''q.$$

Das wiederum bedeutet, daß $x = (p, q) \sim (p'', q'') = x''$ gilt.

Die drei Axiome einer Äquivalenzrelation sind also erfüllt.

Wir setzen nun $\mathbb{Q} := M/\sim$ und für $(p, q) \in M$ setzen wir $\frac{p}{q} := \overline{(p, q)}$, d. h. die rationale Zahl $\frac{p}{q}$ ist die Äquivalenzklasse des Paares (p, q) unter der obigen Äquivalenzrelation. Dann bedeutet die Definition von $\sim$ soviel wie, daß $\frac{p}{q}$ und $\frac{p'}{q'}$ gleich sind, wenn die kreuzweisen Produkte von Zähler und Nenner, pq' und $p'q$, übereinstimmen, oder in der vielleicht etwas bekannteren Formulierung, wenn die Brüche nach *Erweitern* mit q' bzw. mit q übereinstimmen: $\frac{p}{q} = \frac{pq'}{qq'} \overset{!}{=} \frac{p'q}{q'q} = \frac{p'}{q'}$.

Auch die Rechenregeln für rationale Zahlen lassen sich mit Hilfe der Äquivalenzklassen definieren. Für $(p, q), (r, s) \in M$ definiere:

$$\overline{(p, q)} + \overline{(r, s)} := \overline{(ps + qr, qs)},$$
$$\overline{(p, q)} \cdot \overline{(r, s)} := \overline{(pr, qs)}.$$

In Anlehnung an unser erstes Beispiel, der Einteilung der Schüler in Schulklassen, kann man das obige Rechenprinzip als "Rechnen mit Klassen" bezeichnen. Will man zwei Klassen addieren (bzw. multiplizieren), so nimmt man aus jeder der Klassen ein Element,

[1] Man sollte sich nicht dadurch verwirren lassen, daß die Elemente von M nun selbst schon Zahlenpaare sind! Wollte man die Relation als Teilmenge von $M \times M$ schreiben, so müßte man

$$R = \{((p, q), (p', q')) \in M \times M \mid pq' = p'q\}$$

betrachten. Das erläutert vielleicht auch, weshalb wir die *alternative* Schreibeweise bevorzugen – solche Paare von Paaren werden doch leicht unübersichtlich.

addiert (bzw. multipliziert) diese Elemente und schaut, in welche Klasse das Resultat gehört. Diese Klasse ist dann die Summe (bzw. das Produkt) der beiden Klassen.

Was man sich bei diesem Vorgehen allerdings klar machen muß, ist, daß das Ergebnis nicht von der Wahl der Repräsentanten (d.h. der Elemente aus den Klassen) abhängt. Man spricht davon, daß die Operation *wohldefiniert* ist. Wir führen das für die Addition der rationalen Zahlen vor.

Sind $(p',q') \in \overline{(p,q)}$ und $(r',s') \in \overline{(r,s)}$ andere Repräsentanten, dann gilt $p'q = q'p$ und $r's = s'r$. Es ist zu zeigen, daß $(p's' + q'r', q's') \in \overline{(ps+qr, qs)}$ gilt. Ausmultiplizieren liefert

$$(p's' + q'r')(qs) = p'qs's + q'qr's = q'ps's + q'qs'r = (ps+qr)(q's'),$$

was zu zeigen war. □

Aufgaben

Aufgabe 6.13.
Wir definieren für zwei Punkte $(x,y),(x',y') \in \mathbb{R}^2$

$$(x,y) \sim (x',y') \quad :\Longleftrightarrow \quad |x| + |y| = |x'| + |y'|.$$

Zeige, $\sim$ ist eine Äquivalenzrelation auf $\mathbb{R}^2$. Zeichne die Äquivalenzklassen zu $(1,1)$ und zu $(-2,3)$ in die Zahlenebene $\mathbb{R}^2$ ein.

Aufgabe 6.14 (Die ganzen Zahlen).
Es sei $M = \mathbb{N} \times \mathbb{N}$ und $m = (a,b) \in M$ und $m' = (a',b') \in M$ seien zwei Elemente in M. Wir definieren

$$m \sim m' \quad \Longleftrightarrow \quad a + b' = a' + b.$$

Zeige, daß $\sim$ eine Äquivalenzrelation ist und daß die folgende Abbildung bijektiv ist:

$$\Phi : \mathbb{Z} \longrightarrow M/\sim : z \mapsto \begin{cases} \overline{(z,0)}, & \text{falls } z \geq 0, \\ \overline{(0,-z)}, & \text{falls } z < 0. \end{cases}$$

Aufgabe 6.15 (Die projektive Gerade).
Wir definieren für $v = (v_1,v_2), w = (w_1,w_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$

$$v \sim w \quad \Longleftrightarrow \quad \exists\, \lambda \in \mathbb{R} \setminus \{0\} \ : \ v = \lambda \cdot w$$

wobei $\lambda \cdot w := (\lambda \cdot w_1, \lambda \cdot w_2)$.

a. Zeige, daß $\sim$ eine Äquivalenzrelation auf $M = \mathbb{R}^2 \setminus \{(0,0)\}$ ist. Es ist üblich die Äquivalenzklasse $\overline{(v_1, v_2)}$ von (v_1, v_2) mit $(v_1 : v_2)$ zu bezeichnen, und man nennt die Menge $M/\sim$ der Äquivalenzklassen die *projektive Gerade* über $\mathbb{R}$ und bezeichnet sie mit $\mathbb{P}^1_{\mathbb{R}}$.

b. Die Menge $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ ist Kreis vom Radius Eins um den Mittelpunkt $(0, 0)$. Zeige, daß die Abbilung

$$\Phi : S^1 \longrightarrow \mathbb{P}^1_{\mathbb{R}} : (x, y) \mapsto \overline{(x, y)}$$

surjektiv ist.

c. Wenn wir in der Definition von $\sim$ alle Elemente $v, w \in \mathbb{R}^2$ zulassen, definiert $\sim$ dann eine Äquivalenzrelation auf $\mathbb{R}^2$? Falls ja, was ist die Äquivalenzklasse von $(0, 0)$?

Aufgabe 6.16 (Kongruenz modulo n).
Ist $n \in \mathbb{Z}_{>0}$ eine positive ganze Zahl, so definieren wir für $x, y \in \mathbb{Z}$

$$x \equiv y \iff x - y \text{ ist ein Vielfaches von } n.$$

Zeige, daß $\equiv$ eine Äquivalenzrelation ist mit genau den n paarweise verschiedenen Äquivalenzklassen $\overline{0}, \overline{1}, \ldots, \overline{n-1}$.
Man nennt zwei äquivalente Zahlen x und y dann auch *kongruent modulo* n. Diese Äquivalenzrelation wird in der Vorlesung algebraische Strukturen genauer untersucht.

Aufgabe 6.17.
Für zwei reelle Zahlen $x, y \in \mathbb{R}$ definieren wir

$$x \sim y \iff x^2 - y^2 = 2x - 2y.$$

Zeige, daß $\sim$ eine Äquivalenzrelation auf $\mathbb{R}$ ist und bestimme die Äquivalenzklassen von 0 und 1.

§ 7 Gruppen und Körper

Gruppen und Körper sind Themen, die ausführlich in Vorlesungen zur Algebra behandelt werden. Wir wollen die Begriffe hier nur so weit einführen, wie sie für die Analysis erforderlich sind.

A) Gruppen

Definition 7.1 (Gruppen).

a. Eine *Gruppe* ist ein Paar $(G, *)$ bestehend aus einer *nicht-leeren* Menge G und einer zweistelligen Operation "$*$", d. h. einer Abbildung

$$* : G \times G \to G : (g, h) \mapsto g * h,$$

so daß die folgenden *Gruppenaxiome* gelten:

G1: $(g * h) * k = g * (h * k) \quad \forall\, g, h, k \in G,$ ("Assoziativgesetz")

G2: $\exists\, e \in G \;:\; \forall\, g \in G \;:\; e * g = g,$ ("Existenz eines Neutralen")

G3: $\forall\, g \in G\, \exists\, g^{-1} \in G \;:\; g^{-1} * g = e.$ ("Existenz von Inversen")

Ein Element mit der Eigenschaft von e nennt man *neutrales Element* der Gruppe G. Ein Element mit der Eigenschaft von g^{-1} nennt man ein *Inverses zu* g.

b. Eine Gruppe $(G, *)$ heißt *abelsch* oder *kommutativ*, wenn $(G, *)$ zudem noch dem folgenden Axiom genügt:

G4: $g * h = h * g \quad \forall\, g, h \in G$ ("Kommutativgesetz")

Beispiel 7.2.

a. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$ mit der üblichen Addition als Gruppenoperation sind abelsche Gruppen. Die Zahl Null erfüllt jeweils die Rolle eines neutralen Elements, und zu einer Zahl g existiert mit $-g$ ein inverses Element.

b. $(\mathbb{Q} \setminus \{0\}, \cdot)$ und $(\mathbb{R} \setminus \{0\}, \cdot)$ mit der üblichen Multiplikation als Gruppenoperation sind ebenfalls abelsche Gruppen. Die Zahl 1 ist jeweils ein neutrales Element, und zu einer Zahl g existiert als inverses Element die Zahl $\frac{1}{g}$.

c. Ist M eine nicht-leere Menge, so ist die Menge

$$\mathrm{Sym}(M) := \{f : M \longrightarrow M \mid f \text{ ist bijektiv}\}$$

mit der Komposition von Abbildungen als Gruppenoperation eine Gruppe. Die Assoziativität von "$\circ$" haben wir in Proposition 3.11 gezeigt, die Identität ist das neutrale Element und in Satz 3.12 haben wir gezeigt, daß jede bijektive Abbildung

ein Inverses besitzt. Wir nennen $(\mathrm{Sym}(M), \circ)$ die *symmetrische Gruppe* auf M. Enthält M mehr als zwei Elemente, so ist $\mathrm{Sym}(M)$ nicht abelsch.

Bemerkung 7.3.
Es sei $(G, *)$ eine Gruppe.

a. Das neutrale Element $e \in G$ ist eindeutig bestimmt und hat die Eigenschaft:
$$e * g = g * e = g \quad \forall\, g \in G.$$

b. Sei $g \in G$. Das inverse Element g^{-1} zu g ist eindeutig bestimmt und hat die Eigenschaft:
$$g^{-1} * g = g * g^{-1} = e.$$

c. Für $g, h \in G$ gelten $\left(g^{-1}\right)^{-1} = g$ und $(g * h)^{-1} = h^{-1} * g^{-1}$.

d. Wird die Gruppenoperation als Multiplikation und mit "$\cdot$" bezeichnet, so schreiben wir für das Neutrale Element meist 1 und für das Inverse zu g weiterhin g^{-1} oder $\frac{1}{g}$.
Wird die Gruppenoperation als Addition und mit "$+$" bezeichnet, so schreiben wir für das Neutrale Element meist 0 und für das Inverse zu g meist $-g$. Zudem schreiben wir statt $g + (-h)$ in aller Regel $g - h$.

e. In Ermangelung eines besseren Namens nennen wir auch "$*$" oft einfach die *Gruppenmultiplikation*.

Die Aussagen in der Bemerkung werden in Vorlesungen zur (Linearen) Algebra bewiesen. Für den interessierten Leser fügen wir hier einen Beweis ein.

Beweis von Bemerkung 7.3: Da wir für das Paar $(G, *)$ die Axiome G1-G3 aus Definition 7.1 voraussetzen, gibt es ein neutrales Element $e \in G$, und zu beliebigem, aber fest gegebenem $g \in G$ gibt es ein Inverses $g^{-1} \in G$.

Wir wollen zunächst zeigen, daß für dieses e und dieses g^{-1} die in a. und b. geforderten zusätzlichen Eigenschaften gelten.

Da $(G, *)$ eine Gruppe ist, gibt es ein $(g^{-1})^{-1} \in G$ mit

$$(g^{-1})^{-1} * g^{-1} = e. \tag{2}$$

Also folgt:

$$\begin{aligned} g * g^{-1} &\overset{G2}{=} e * \left(g * g^{-1}\right) \overset{(2)}{=} \left((g^{-1})^{-1} * g^{-1}\right) * \left(g * g^{-1}\right) \overset{G1}{=} (g^{-1})^{-1} * \left(g^{-1} * (g * g^{-1})\right) \\ &\overset{G1}{=} (g^{-1})^{-1} * \left((g^{-1} * g) * g^{-1}\right) \overset{G3}{=} (g^{-1})^{-1} * \left(e * g^{-1}\right) \overset{G2}{=} (g^{-1})^{-1} * g^{-1} \overset{(2)}{=} e. \end{aligned} \tag{3}$$

Damit ist gezeigt, daß g^{-1} die zusätzliche Eigenschaft in b. erfüllt, und wir erhalten:

$$g * e \overset{G3}{=} g * \left(g^{-1} * g\right) \overset{G1}{=} \left(g * g^{-1}\right) * g \overset{(3)}{=} e * g \overset{G2}{=} g. \tag{4}$$

Nun war aber g ein beliebiges Element in G, so daß damit die zusätzliche Eigenschaft von e in a. gezeigt ist.

Sei nun $\tilde{e} \in G$ irgendein Element mit der Eigenschaft des Neutralen, d.h.

$$\tilde{e} * h = h \tag{5}$$

für alle $h \in G$. Wir müssen zeigen, daß $e = \tilde{e}$ gilt. Da wir bereits wissen, daß e die zusätzliche Eigenschaft in a. erfüllt, können wir diese, d.h. (4), mit $\tilde{e}$ in der Rolle von g anwenden, und anschließend (5) mit e in der Rolle von h:

$$\tilde{e} \overset{(4)}{=} \tilde{e} * e \overset{(5)}{=} e.$$

Schließlich müssen wir noch zeigen, wenn $\tilde{g}^{-1} \in G$ ein weiteres inverses Element zu g ist, d.h. wenn

$$\tilde{g}^{-1} * g = e \tag{6}$$

gilt, dann ist schon $g^{-1} = \tilde{g}^{-1}$. Wenden wir das bislang Gezeigte an, so gilt:

$$\tilde{g}^{-1} \overset{(4)}{=} \tilde{g}^{-1} * e \overset{(3)}{=} \tilde{g}^{-1} * \left(g * g^{-1}\right) \overset{G1}{=} \left(\tilde{g}^{-1} * g\right) * g^{-1} \overset{(6)}{=} e * g^{-1} \overset{G2}{=} g^{-1}.$$

Damit sind die Aussagen in Teil a. und b. gezeigt und es bleibt noch, die Aussagen in Teil c. zu zeigen.

Um die erste Gleichheit zu zeigen, reicht es wegen der Eindeutigkeit des Inversen zu g^{-1} zu zeigen, daß g die Eigenschaft *des* Inversen zu g^{-1} besitzt. Beim Beweis können wir die Gruppenaxiome sowie die in a. und b. bewiesenen zusätzlichen Eigenschaften des Inversen anwenden:

$$g * g^{-1} \overset{b.}{=} e.$$

Also ist g ein Inverses zu g^{-1}, und damit gilt wie angedeutet wegen der Eindeutigkeit des Inversen zu g^{-1}:

$$\left(g^{-1}\right)^{-1} = g.$$

Analog ist nach Voraussetzung $(gh)^{-1}$ ein Inverses zu gh, und es reicht wegen der Eindeutigkeit des Inversen zu gh zu zeigen, daß $h^{-1}g^{-1}$ ebenfalls die Eigenschaft eines Inversen zu gh hat:

$$\left(h^{-1} * g^{-1}\right) * (g * h) \overset{G1}{=} h^{-1} * \left(g^{-1} * \left(g * h\right)\right) \overset{G1}{=} h^{-1} * \left(\left(g^{-1} * g\right) * h\right)\right)$$
$$\overset{G3}{=} h^{-1} * (e * h) \overset{G2}{=} h^{-1} * h \overset{G3}{=} e.$$

Mithin ist $h^{-1} * g^{-1}$ ein Inverses zu gh, und somit

$$(g * h)^{-1} = h^{-1} * g^{-1}.$$

Damit sind nun alle Aussagen der Bemerkung bewiesen. □

Lemma 7.4 (Kürzungsregeln).
Sei $(G, *)$ eine Gruppe, $g, a, b \in G$. Dann gelten die *Kürzungsregeln*:

a. $g * a = g * b \implies a = b$, und

b. $a * g = b * g \implies a = b$.

Beweis: Die erste Kürzungsregel folgt durch Multiplikation mit dem Inversen zu g von links:

$$a \overset{G2}{=} e * a \overset{G3}{=} \left(g^{-1} * g\right) * a \overset{G1}{=} g^{-1} * (g * a)$$
$$\overset{Vor.}{=} g^{-1} * (g * b) \overset{G1}{=} \left(g^{-1} * g\right) * b \overset{G3}{=} e * b \overset{G2}{=} b.$$

Entsprechend folgt die zweite Kürzungsregel durch Multiplikation mit g^{-1} von rechts und unter Berücksichtigung der zusätzlichen Eigenschaft des Inversen in Bemerkung 7.3. Die Details überlassen wir dem Leser. □

B) Körper

Definition 7.5 (Körper).
Ein *Körper* ist ein Tripel $(K, +, \cdot)$ bestehend aus einer Menge K zusammen mit zwei zweistelligen Operationen

$$+ : K \times K \to K : (x, y) \mapsto x + y, \qquad (\text{“}\textit{Addition}\text{”})$$

und

$$\cdot : K \times K \to K : (x, y) \mapsto x \cdot y, \qquad (\text{“}\textit{Multiplikation}\text{”})$$

so daß folgende Axiome erfüllt sind:

a. $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0.

b. $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element 1.

c. Es gelten die *Distributivgesetze* $x \cdot (y+z) = x \cdot y + x \cdot z$ und $(y+z) \cdot x = y \cdot x + z \cdot x$ für $x, y, z \in K$.

Ist eine Teilmenge $L \subseteq K$ eines Körpers mit den *gleichen* Operationen wieder selbst ein Körper, so nennen wir L einen *Teilkörper* von K.

Beispiel 7.6 (Die endlichen Körper $\mathbb{F}_p$).

a. Die rationalen Zahlen $(\mathbb{Q}, +, \cdot)$ und die reellen Zahlen $(\mathbb{R}, +, \cdot)$ mit der üblichen Addition und Multiplikation sind Körper. $\mathbb{Q}$ ist ein Teilkörper von $\mathbb{R}$.

b. Die ganzen Zahlen $(\mathbb{Z}, +, \cdot)$ sind kein Körper, da z.B. der Zahl 2 ein multiplikatives Inverses fehlt.

c. Auf der Menge $\mathbb{F}_2 := \{0, 1\}$ definieren wir zwei Operationen durch folgende Additions- und Multiplikationstafeln:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Mit ein wenig Aufwand kann man nachrechnen, daß alle Körperaxiome erfüllt sind und daß mithin $\mathbb{F}_2$ ein Körper ist. $\mathbb{F}_2$ ist der kleinstmögliche Körper, da nach Definition ein Körper stets mindestens zwei Elemente, nämlich ein Neutrales bezüglich der Addition und ein davon verschiedenes Neutrales bezüglich der Multiplikation enthalten muß. Man beachte auch, daß aufgrund von Lemma 7.8 keine andere Möglichkeit für die obigen Verknüpfungstafeln besteht, wenn man einen Körper mit genau zwei Elementen haben möchte. — Beachte auch, daß $\mathbb{F}_2$ kein Teilkörper von $\mathbb{R}$ ist, da das Ergebnis von $1 + 1$ in den beiden Körpern nicht übereinstimmt.

d. Allgemeiner zeigt man in Vorlesungen zur (Linearen) Algebra, daß man für eine Primzahl p die Menge

$$\mathbb{F}_p := \{0, 1, \ldots, p-1\}$$

auf folgende Weise zu einem Körper machen kann. Für eine natürliche Zahl $a \in \mathbb{N}$ können wir Division mit Rest durch die Zahl p durchführen. Wir erhalten dann eindeutig bestimmte Zahlen $q \in \mathbb{N}$ und $0 \leq r < p$ mit

$$a = q \cdot p + r.$$

Die Zahl r heißt der Rest von a bei Division mit Rest durch p, und wir bezeichnen sie $r(a : p)$.
Mit dieser Notation definieren wir für zwei Zahlen $a, b \in \mathbb{F}_p$

$$a + b := r(a + b : p)$$

und

$$a \cdot b := r(a \cdot b : p),$$

wobei das "$+$" bzw. das "$\cdot$" auf der rechten Seite jeweils die Operation in den ganzen Zahlen bezeichnet, während das "$+$" und das "$\cdot$" auf der linken Seite neu definierte Operationen sind. Formal wäre es besser, für diese neuen Operationen neue Symbole zu verwenden, etwa "$\oplus$" und "$\otimes$", aber Mathematiker sind bequeme Menschen und schreiben nur ungerne mehr als nötig. Deshalb bleiben wir bei den bewährten Symbolen und müssen nur drauf achten, wo wir gerade rechnen. Jedenfalls gilt, daß $\mathbb{F}_p$ mit diesen beiden Operationen ein Körper ist.
Man beachte auch, daß in $\mathbb{F}_p$ für jede Primzahl p stets

$$\underbrace{1 + 1 + \ldots + 1}_{p-\text{mal}} = r(p : p) = 0$$

gilt! Damit ist auch das Negative einer Zahl $a \in \mathbb{F}_p$ leicht zu berechnen als $p - a$, hingegen ist das multiplikative Inverse $\frac{1}{a}$ einer Zahl $0 \neq a \in \mathbb{F}_p$ nicht so ohne weiteres anzugeben. Man lernt in der (Linearen) Algebra, wie man dieses mit Hilfe des Euklidischen Algorithmus' berechnen kann.
Z.B., gilt in $\mathbb{F}_5$

$$3 + 4 = r(3 + 4 : 5) = r(7 : 5) = 2$$

und

$$3 \cdot 4 = r(3 \cdot 4 : 5) = r(12 : 5) = 2.$$

In der (Linearen) Algebra schreibt man übrigens oft $\mathbb{Z}/p\mathbb{Z}$ oder $\mathbb{Z}_p$ anstatt $\mathbb{F}_p$, und die Zahl a wird dort meist mit $\overline{a}$ oder $[a]$ bezeichnet. Das liegt daran, daß man den Körper mit der Menge der Äquivalenzklassen der Kongruenz modulo p identifizieren kann (siehe Aufgabe 6.16).

Notation 7.7.
Ist K ein Körper und sind $x, y, z \in K$ mit $z \neq 0$, so schreiben wir statt $x + (-y)$ in aller Regel $x - y$, und statt $x \cdot z^{-1}$ schreiben wir oft $\frac{x}{z}$. Außerdem schreiben wir statt $x \cdot y$ meist nur xy.

Lemma 7.8 (Rechenregeln).
Es sei K ein Körper, $x, y, z \in K$ und $u, v \in K \setminus \{0\}$.

a. $-(-x) = x$,

b. $x + y = z \iff x = z - y$,

c. $-(x + y) = -x - y$,

d. $0 \cdot x = x \cdot 0 = 0$,

e. $(-x) \cdot y = x \cdot (-y) = -(x \cdot y)$,

f. $(-x) \cdot (-y) = x \cdot y$,

g. $x \cdot (y - z) = x \cdot y - x \cdot z$.

h. $\left(x^{-1}\right)^{-1} = x$, für $x \neq 0$,

i. $x \cdot y = 0 \iff x = 0$ oder $y = 0$,

j. $z \cdot x = z \cdot y,\ z \neq 0 \implies x = y$,

k. $\frac{x}{u} \cdot \frac{y}{v} = \frac{x \cdot y}{u \cdot v}$,

l. $\frac{x}{u} + \frac{y}{v} = \frac{x \cdot v + y \cdot u}{u \cdot v}$.

Beweis: Die Aussagen a., b., c. und h. folgen unmittelbar aus Bemerkung 7.3 und Lemma 7.4.

d. Für $x \in K$ gilt $0 \cdot x = (0+0) \cdot x = 0 \cdot x + 0 \cdot x$, also folgt $0 \cdot x = 0$ mittels der Kürzungsregeln in $(K, +)$. Analog sieht man $x \cdot 0 = 0$.

e. Für $x, y \in K$ gilt wegen d.:

$$x \cdot y + (-x) \cdot y = (x - x) \cdot y = 0 \cdot y = 0,$$

also $-(x \cdot y) = (-x) \cdot y$. Die Gleichheit des Ausdrucks zu $x \cdot (-y)$ folgt analog.

f. Für $x, y \in K$ folgt unter Zuhilfenahme von a. und e.:

$$(-x) \cdot (-y) = -\big(x \cdot (-y)\big) = -\big(-(x \cdot y)\big) = x \cdot y.$$

g. Für $x, y, z \in K$ impliziert e.:

$$x \cdot (y - z) = x \cdot y + x \cdot (-z) = x \cdot y + \big(-(x \cdot z)\big) = x \cdot y - x \cdot z.$$

i. Ist $x = 0$ oder $y = 0$, so ist nach d. auch $x \cdot y = 0$. Ist $x \neq 0$ und $y \neq 0$, so ist $x \cdot y \in K \setminus \{0\}$, da $K \setminus \{0\}$ bezüglich der Multiplikation abgeschlossen ist.

j. Die Aussage zeigt man genau wie die Kürzungsregeln für Gruppen (siehe Lemma 7.4).

k. Unter Beachtung der Assoziativität und Kommutativität der Multiplikation sowie der Notation 7.7 gilt

$$\frac{x}{u} \cdot \frac{y}{v} = \big(x \cdot u^{-1}\big) \cdot \big(y \cdot v^{-1}\big) = (x \cdot y) \cdot (u \cdot v)^{-1} = \frac{x \cdot y}{u \cdot v}.$$

l. Dies geht analog zu k. mit etwas mehr Schreibarbeit.

□

C) Summen, Produkte und der Binomische Lehrsatz

Notation 7.9 (Produkte und Summen).
Es sei K ein Körper und $x_0, \ldots, x_n \in K$ seien $n+1$ Elemente in K, $n \in \mathbb{N}$. Wir schreiben

$$\prod_{i=0}^{n} x_i := x_0 \cdot \ldots \cdot x_n$$

für das *Produkt* der Zahlen $x_0, \ldots, x_n$ und

$$\sum_{i=0}^{n} x_i := x_0 + \ldots + x_n$$

für die *Summe* der Zahlen $x_0, \ldots, x_n$. Wir einigen uns dabei darauf, daß das leere Produkt (d.h. ein Produkt, bei dem der obere Index kleiner als der untere ist) den Wert 1 hat und die leere Summe den Wert 0.

Außerdem definieren wir für $x \in K$ und $n \in \mathbb{N}$ die *Potenzen* von x durch

$$x^n := \underbrace{x \cdot \ldots \cdot x}_{n-\text{mal}} = \prod_{i=1}^{n} x$$

falls $n \geq 1$ sowie $x^0 := 1$. Ist zudem $x \neq 0$, so definieren wir

$$x^{-n} := \left(x^{-1}\right)^n = \underbrace{x^{-1} \cdot \ldots \cdot x^{-1}}_{n-\text{mal}} = \frac{1}{x^n}.$$

Analog dazu setzen wir

$$n \cdot x := \underbrace{x + \ldots + x}_{n-\text{mal}} = \sum_{i=1}^{n} x$$

und

$$(-n) \cdot x := n \cdot (-x) = \underbrace{(-x) + \ldots + (-x)}_{n-\text{mal}}$$

für $n \geq 1$, sowie $0 \cdot x = 0$.

Bemerkung 7.10 (Rekursionsprinzip).
Dem Prinzip der vollständigen Induktion ist das *Rekursionsprinzip* eng verwandt. Wollen wir einen Ausdruck für alle natürlichen Zahlen definieren, so definieren wir ihn für die Zahl 0 und führen die Definition für die Zahl n auf die Definition für die Zahl $n-1$ zurück.

Die Notation mit Punkten "..." in Notation 7.9 ist stets eine versteckte Induktion oder Rekursion. Formal korrekt wäre es das Produkt rekursiv zu definieren durch $\prod_{i=0}^{0} x_i := x_0$ und $\prod_{i=0}^{n} x_i := \left(\prod_{i=0}^{n-1} x_i\right) \cdot x_n$. Analog sollte man die Summe rekursiv definieren durch $\sum_{i=0}^{0} x_i := x_0$ und $\sum_{i=0}^{n} x_i := \left(\sum_{i=0}^{n-1} x_i\right) + x_n$. Und für die Definition von x^n und $n \cdot x$ gilt Entsprechendes.

Beispiel 7.11 (Gauß).
Die Summe der natürlichen Zahlen bis zu einer gegebenen Zahl n ist

$$\sum_{k=0}^{n} k = \frac{n \cdot (n+1)}{2}.$$

Man beweist die Aussage durch Induktion nach n, wobei sie für $n = 0$ offenbar richtig ist. Nehmen wir nun an, daß sie für n gilt, so folgt

$$\sum_{k=0}^{n+1} k = \sum_{k=0}^{n} k + (n+1) \overset{Ind}{=} \frac{n \cdot (n+1)}{2} + (n+1) = \frac{(n+1) \cdot (n+2)}{2}.$$

Also gilt die Aussage für alle $n \in \mathbb{N}$ nach dem Prinzip der vollständigen Induktion.

Satz 7.12 (Endliche geometrische Reihe).
Ist K ein Körper, $1 \neq q \in K$ und $n \in \mathbb{N}$, so gilt

$$\sum_{k=0}^{n} q^k = \frac{1 - q^{n+1}}{1 - q}.$$

Beweis: Der Beweis ist eine einfache Anwendung des Prinzips der vollständigen Induktion. □

Definition 7.13 (Fakultät).
Für eine natürliche Zahl $n \in \mathbb{N}$ definieren wir die *Fakultät* durch

$$n! := \prod_{i=1}^{n} i = 1 \cdot \ldots \cdot n,$$

falls $n \geq 1$, und durch $0! := 1$.
Für eine natürliche Zahl $n \in \mathbb{N}$ und $k \in \mathbb{Z}$ erklären wir den *Binomialkoeffizienten* von n über k durch

$$\binom{n}{k} := \frac{n!}{(n-k)! \cdot k!} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{k \cdot (k-1) \cdot \ldots \cdot 1},$$

falls $0 \leq k \leq n$, und durch $\binom{n}{k} := 0$ sonst.

Proposition 7.14 (Binomialkoeffizienten).
Es seien $n, k \in \mathbb{N}$ natürliche Zahlen. Dann gilt

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$

Beweis: Wir unterscheiden mehrere Fälle.

1. **Fall:** $k = 0$**:**

$$\binom{n+1}{k} = \binom{n+1}{0} = 1 = 0 + 1 = \binom{n}{-1} + \binom{n}{0} = \binom{n}{k-1} + \binom{n}{k}.$$

2. **Fall:** $k = n + 1$**:**

$$\binom{n+1}{k} = \binom{n+1}{n+1} = 1 = 1 + 0 = \binom{n}{n} + \binom{n}{n+1} = \binom{n}{k-1} + \binom{n}{k}.$$

3. **Fall:** $k < 0$ **oder** $k > n + 1$**:**

$$\binom{n+1}{k} = 0 = 0 + 0 = \binom{n}{k-1} + \binom{n}{k}.$$

4. Fall: $1 \leq k \leq n$**:**

$$\begin{aligned}\binom{n}{k-1} + \binom{n}{k} =& \frac{n!}{(n+1-k)! \cdot (k-1)!} + \frac{n!}{(n-k)! \cdot k!} \\ =& \frac{n! \cdot k}{(n+1-k)! \cdot k!} + \frac{n! \cdot (n+1-k)}{(n+1-k)! \cdot k!} \\ =& \frac{n! \cdot (k+n+1-k)}{(n+1-k)! \cdot k!} = \frac{(n+1)!}{(n+1-k)! \cdot k!} = \binom{n+1}{k}.\end{aligned}$$

□

Satz 7.15 (Binomischer Lehrsatz).
Es sei K ein Körper, $x, y \in K$ und $n \in \mathbb{N}$, so gilt

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot x^k \cdot y^{n-k}.$$

Beweis: Wir führen den Beweis durch Induktion nach n.

Induktionsanfang: $n = 0$**:** Nach Definition gilt

$$(x+y)^0 = 1 = 1 \cdot 1 \cdot 1 = \sum_{k=0}^{0} \binom{0}{k} \cdot x^k \cdot y^{0-k}.$$

Induktionsschluß: $n \mapsto n+1$**:** Es gilt

$$\begin{aligned}(x+y)^{n+1} =& (x+y)^n \cdot (x+y) = (x+y)^n \cdot x + (x+y)^n \cdot y \\ \overset{Ind.}{=}& \sum_{k=0}^{n} \binom{n}{k} \cdot x^{k+1} \cdot y^{n-k} + \sum_{k=0}^{n} \binom{n}{k} \cdot x^k \cdot y^{n+1-k} \\ =& x^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k} \cdot x^{k+1} \cdot y^{n-k} + \sum_{k=1}^{n} \binom{n}{k} \cdot x^k \cdot y^{n+1-k} + y^{n+1} \\ =& x^{n+1} + \sum_{k=1}^{n} \binom{n}{k-1} \cdot x^k \cdot y^{n+1-k} + \sum_{k=1}^{n} \binom{n}{k} \cdot x^k \cdot y^{n+1-k} + y^{n+1} \\ \overset{7.14}{=}& x^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} \cdot x^k \cdot y^{n+1-k} + y^{n+1} \\ =& \sum_{k=0}^{n+1} \binom{n+1}{k} \cdot x^k \cdot y^{n+1-k}\end{aligned}$$

Die Aussage folgt damit aus dem Prinzip der vollständigen Induktion. □

Bemerkung 7.16 (Pascalsches Dreieck).
Man ordnet die Binomialkoeffizienten gerne in der folgenden Form an, die als Pascalsches

Dreieck bekannt ist:

$$\begin{array}{ccccccccc}
 & & & \binom{0}{0} & & & & & \\
 & & \binom{1}{0} & & \binom{1}{1} & & & & \\
 & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & \\
\binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & \\
\binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4}
\end{array}$$

Berechnet man die Werte der Binomialkoeffizienten, erhält man die folgende Gestalt:

0. Zeile:				1				
1. Zeile:			1		1			
2. Zeile:		1		2		1		
3. Zeile:		1		3		3		1
4. Zeile:	1		4		6		4		1

Aufgrund von Proposition 7.14 kann man die Einträge der $n+1$-ten Zeile aus den Einträgen der n-ten Zeile berechnen. Graphisch im Pascalschen Dreieck nimmt die Proposition folgende Gestalt an:

$$\begin{array}{ccccc}
\binom{n}{k-1} & & + & & \binom{n}{k} \\
 & \searrow & & \swarrow & \\
 & & \binom{n+1}{k} & &
\end{array}$$

D.h. die Summe zweier benachbarter Einträge der n-ten Zeile liefert den mittig unter ihnen stehenden Eintrag der $n+1$-ten Zeile.

Aufgrund des binomischen Lehrsatzes sind die Einträge der n-ten Zeile des Pascalschen Dreiecks genau die Koeffizienten, die wir erhalten, wenn wir $(x+y)^n$ ausschreiben. Z.B.

$$(x+y)^3 = 1 \cdot x^3 + 3 \cdot x^2 y + 3 \cdot xy^2 + 1 \cdot y^3.$$

Aufgaben

Aufgabe 7.17.
Sei M eine Menge. Zeige, die Potenzmenge $\mathcal{P}(M)$ wird mittels der symmetrischen Differenz

$$A + B := (A \cup B) \setminus (A \cap B)$$

für $A, B \in \mathcal{P}(M)$ eine abelsche Gruppe.

Aufgabe 7.18.
Es sei K ein Körper und $x \in K$. Zeige, $x^2 = 1$ genau dann, wenn $x \in \{1, -1\}$.

Aufgabe 7.19.

a. Auf der Menge $G := \mathbb{R} \times \mathbb{R}$ definieren wir eine zweistellige Operation

$$+ : G \times G \longrightarrow G : \big((x, y), (u, v)\big) \mapsto (x + u, y + v).$$

Zeige, $(G, +)$ ist eine abelsche Gruppe mit neutralem Element $(0, 0)$.

b. Auf der Menge $H := (\mathbb{R} \times \mathbb{R}) \setminus \{(0, 0)\}$ definieren wir eine zweistellige Operation

$$\cdot : H \times H \longrightarrow H : \big((x, y), (u, v)\big) \mapsto (xu - yv, xv + yu).$$

Zeige, $(H, \cdot)$ ist eine abelsche Gruppe mit neutralem Element $(1, 0)$.

c. Zeige, daß $(\mathbb{R} \times \mathbb{R}, +, \cdot)$ ein Körper ist, wenn die Operationen "+" und "·" wie in a. und b. definiert sind.

Aufgabe 7.20.
Zeige durch vollständige Induktion, daß

$$\sum_{k=0}^{n} (k + 1) \cdot \binom{n}{k} = 2^{n-1} \cdot (n + 2)$$

für $n \in \mathbb{N}$ gilt.

Aufgabe 7.21 (Die projektive Gerade als Gruppe).
Wir haben in Aufgabe 6.15 die Projektive Gerade $\mathbb{P}^1_{\mathbb{R}}$ als Menge von Äquivalenzklassen auf $\mathbb{R}^2 \setminus \{(0, 0)\}$ eingeführt.
Zeige, daß die zweistellige Operation

$$(v_1 : v_2) \cdot (w_1 : w_2) := (v_1 \cdot w_1 - v_2 \cdot w_2 : v_1 \cdot w_2 + v_2 \cdot w_1).$$

wohldefiniert ist, d.h. nicht von der Wahl der Repräsentanten für die Äquivalenzklasse abhängt, und daß $\mathbb{P}^1_{\mathbb{R}}$ mit dieser Operation eine Gruppe ist.

Aufgabe 7.22.
Für zwei natürliche Zahlen $1 \leq k \leq n$ gilt

$$k \cdot \binom{n}{k} = n \cdot \binom{n-1}{k-1}.$$

§ 8 Ordnungsrelationen

In diesem Abschnitt wollen wir Körper mit einer zusätzlichen Struktur einführen und betrachten, sogenannte angeordnete Körper. Dazu führen wir zunächst den Begriff der Ordnungsrelation ein.

A) Ordnungsrelationen

Definition 8.1 (Ordnungsrelation).
Es sei M eine Menge. Eine *Ordnungsrelation* auf M, auch *Halbordnung* oder *partielle Ordnung* genannt, ist eine Relation $R \subseteq M \times M$, so daß für alle $x, y, z \in M$ gilt:

O1: $(x, x) \in R$, ("Reflexivität")
O2: $(x, y), (y, x) \in R \implies x = y$, ("Antisymmetrie")
O3: $(x, y), (y, z) \in R \implies (x, z) \in R$. ("Transitivität")

Wir nennen dann (M, R) auch eine *partiell* oder *teilgeordneten Menge*.

Beispiel 8.2.
Es sei $M = \mathbb{N}$.

a. Die übliche Größerrelation

$$R = \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid x \leq y\}$$

ist eine Ordnungsrelation auf $\mathbb{N}$.

b. Die Relation

$$R = \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid x \text{ teilt } y\}$$

ist eine weitere Ordnungsrelation auf $\mathbb{N}$ (siehe Aufgabe 8.18).

Notation 8.3 (Schreibweise $\leq$ für Ordnungsrelationen).
Ordnungsrelationen bezeichnet man meist eher mit Symbolen wie $\leq$ anstatt mit R. Sei also M eine Menge und $\leq$ eine Ordnungsrelation auf M. Dann gilt in unserer Notation aus Bemerkung 6.2 für $x, y \in M$

$$x \leq y \iff (x, y) \in \leq$$

Mit dieser Schreibweise lassen sich die drei Axiome in Definition 8.1 wie folgt formulieren. Für $x, y, z \in M$ soll gelten:

O1: $x \leq x$, ("Reflexivität")
O2: $x \leq y \wedge y \leq x \implies x = y$, ("Antisymmetrie")
O3: $x \leq y \wedge y \leq z \implies x \leq z$. ("Transitivität")

Gilt für $x, y \in M$, daß $x \leq y$ und $x \neq y$, so schreiben wir auch $x < y$.

Beispiel 8.4.
Ist M eine Menge, so ist die Potenzmenge $\mathcal{P}(M)$ von M durch

$$A \leq B \ :\Longleftrightarrow\ A \subseteq B, \text{ für } A, B \in \mathcal{P}(M),$$

partiell geordnet, aber im allgemeinen sind zwei Elemente von $\mathcal{P}(M)$ nicht unbedingt vergleichbar bezüglich dieser Ordnungsrelation. Z. B. sind im Fall $M = \mathbb{N}$ die Elemente $\{2\}$ und $\{3\}$ in $\mathcal{P}(\mathbb{N})$ nicht vergleichbar.

Allgemeiner gilt, ist N eine Menge, deren Elemente wieder Mengen sind, so wird N mit der analogen Definition von "$\leq$" eine partiell geordnete Menge.

Definition 8.5 (Total- und Wohlordnungen).
Es sei M ein Menge.

a. Eine Ordnungsrelation "$\leq$" auf M heißt *Totalordnung* oder *lineare Ordnung*, falls je zwei Elemente aus M vergleichbar sind, d. h. für je zwei Elemente $x, y \in M$ gilt $x \leq y$ oder $y \leq x$.

b. Ist "$\leq$" eine Ordnungsrelation auf M, $A \subseteq M$ und $x \in A$, so heißt x *minimal* (bzw. *maximal*) in A, falls für alle $y \in A$ mit $y \leq x$ (bzw. $x \leq y$) gilt $x = y$.

c. Eine Totalordnung heißt *Wohlordnung*, falls jede nicht-leere Teilmenge von M ein minimales Element besitzt.

Bemerkung 8.6 (Minimum und Maximum).
Das Minimum bzw. Maximum einer Menge M bezüglich einer Totalordnung ist offenbar eindeutig bestimmt, sofern es existiert. Wir bezeichnen es mit $\min(M)$ bzw. mit $\max(M)$.

Beispiel 8.7.

a. Die reellen Zahlen $(\mathbb{R}, \leq)$ mit der üblichen Kleiner-Gleich-Relation $\leq$ sind total geordnet, aber nicht wohlgeordnet.

b. Gleiches trifft auf $(\mathbb{Z}, \leq)$ mit der üblichen Kleiner-Gleich-Relation

$$\ldots -2 < -1 < 0 < 1 < 2 < \ldots$$

zu. Allerdings definiert die "unübliche" Anordnung

$$0 < -1 < 1 < -2 < 2 < -3 < 3 < \ldots$$

in der Tat eine Wohlordnung auf $\mathbb{Z}$.

Bemerkung 8.8 (Archimedisches Prinzip).
Die natürlichen Zahlen sind bezüglich der üblichen Ordnungsrelation "$\leq$" wohlgeordnet, d.h.:

Jede nicht-leere Menge natürlicher Zahlen enthält eine kleinste Zahl.

Diese wohlbekannte Eigenschaft der natürlichen Zahlen nennen wir auch das *archimedische Prinzip*.

B) Das Supremumsaxiom

Definition 8.9 (Supremum und Infimum).
Es sei "$\leq$" eine Totalordnung auf einer Menge M und $\emptyset \neq A \subseteq M$ eine nicht-leere Teilmenge von M.

a. Wir nennen $s \in M$ eine *obere Schranke* von A, falls $s \geq x$ für alle $x \in A$.

b. Wir nennen A *nach oben beschränkt*, falls A eine obere Schranke besitzt.

c. Wir nennen $s \in M$ das *Supremum* von A, falls s das Minimum der Menge der oberen Schranken von A ist. Dieses Minimum ist eindeutig bestimmt, wenn es existiert, und wir bezeichnen es dann mit $\sup(A)$.

d. Wir nennen $s \in M$ eine *untere Schranke* von A, falls $s \leq x$ für alle $x \in A$.

e. Wir nennen A *nach unten beschränkt*, falls A eine untere Schranke besitzt.

f. Wir nennen $s \in M$ das *Infimum* von A, falls s das Maximum der Menge aller unteren Schranken von A ist. Dieses Maximum ist eindeutig bestimmt, wenn es existiert, und wir bezeichnen es dann mit $\inf(A)$.

g. Wir nennen A *beschränkt*, wenn A nach oben und nach unten beschränkt ist.

Beispiel 8.10.

a. Besitzt eine Teilmenge A einer totalgeordneten Menge M ein Maximum, so ist dieses offenbar auch das Supremum von A. Analog ist das Minimum einer Menge A auch ihr Infimum.

b. Betrachten wir die reellen Zahlen mit ihrer üblichen Ordnung und die Menge $A = \{x \in \mathbb{R} \mid 0 < x \leq 1\}$, so ist $1 = \sup(A) = \max(A)$ das Supremum von A, das zugleich ein Maximum ist, und $0 = \inf(A)$ ist ein Infimum von A, das kein Minimum ist.

c. Betrachten wir die rationalen Zahlen mit ihrer üblichen Ordnungsrelation, so ist

$$\{x \in \mathbb{Q} \mid x > 0 \text{ und } x^2 \leq 2\}$$

nach oben beschränkt, besitzt aber kein Supremum in $\mathbb{Q}$ (siehe Satz 9.10).

Bemerkung 8.11 (Supremumsaxiom).
Die reellen Zahlen sind bezüglich ihrer üblichen Ordnungsrelation nicht wohlgeordnet, d.h. nicht jede nicht-leere Teilmenge besitzt ein kleinstes Element. Selbst, wenn wir voraussetzen, daß die Teilmenge nach unten beschränkt ist, muß sie kein kleinstes Element besitzen, d.h. kein Minimum enthalten, wie wir in Beispiel 8.10 gesehen haben. Es gilt aber, daß zu jeder nicht-leeren, nach unten beschränkten Teilmenge von $\mathbb{R}$ ein Infimum in $\mathbb{R}$ existiert. Äquivalent dazu ist die duale Aussage für das Supremum:

> Jede nicht-leere, nach oben beschränkte Teilmenge von $\mathbb{R}$ besitzt ein Supremum in $\mathbb{R}$.

Diese Eigenschaft ist als *Supremumsaxiom* der reellen Zahlen bekannt. Auch wenn sich die Korrektheit der Aussage nicht unmittelbar aus unserer Alltagserfahrung mit den reellen Zahlen als Dezimalzahlen erschließt, wollen wir sie ohne weiteren Beweis als gegeben voraussetzen.

C) Angeordnete Körper

Definition 8.12 (Angeordnete Körper).
Es sei K ein Körper und "$\leq$" eine Totalordnung auf K. Wir nennen das Quadrupel $(K, +, \cdot, \leq)$ einen *angeordneten Körper*, wenn die Totalordnung mit der Addition und der Multiplikation verträglich ist, d.h. wenn für alle $x, y, z \in K$

$$x < y \quad \Longrightarrow \quad x + z < y + z$$

und

$$x < y, \;\; 0 < z \quad \Longrightarrow \quad x \cdot z < y \cdot z$$

gilt. Ist $x \in K$ und $x > 0$, so nennen wir x *positiv*, ist $x < 0$, so nennen wir x *negativ*.

Beispiel 8.13.

a. Die rationalen Zahlen $\mathbb{Q}$ und die reellen Zahlen $\mathbb{R}$ mit der üblichen Ordnungsrelation sind Beispiele für angeordnete Körper. $\mathbb{Q}$ erfüllt das Supremumsaxiom nicht (siehe Beispiel 8.10), $\mathbb{R}$ erfüllt es.

b. Es gibt keine Totalordnung auf $\mathbb{F}_2$, durch die $\mathbb{F}_2$ ein angeordneter Körper würde. Denn würde es eine solche Totalordnung "$\leq$" geben, so wäre entweder $0 < 1$, was zum Widerspruch $1 = 0 + 1 < 1 + 1 = 0$ führt, oder es wäre $1 < 0$, was zum Widerspruch $0 = 1 + 1 < 0 + 1 = 1$ führt.

Lemma 8.14 (Rechenregeln in angeordneten Körpern).
Es sei $(K, +, \cdot, \leq)$ ein angeordneter Körper und $x, y, u, v \in K$.

a. $x > 0 \iff -x < 0$.

b. Ist $x \neq 0$, so ist $x^2 > 0$.

c. $1 > 0$.

d. Ist $0 < x < y$, so ist $0 < \frac{1}{y} < \frac{1}{x}$.

e. Ist $x < y$ und $u < v$, so ist $x + u < y + v$.

f. Ist $0 < x$ und $n \in \mathbb{N}$, so ist $0 < x^n$.

g. Ist $0 \leq x, y$ und $n \in \mathbb{N}$ mit $n \geq 1$, so gilt

$$x < y \iff x^n < y^n.$$

Beweis:

a. Aus $0 < x$ folgt durch Addition von $-x$

$$-x = 0 + (-x) < x + (-x) = 0.$$

Umgekehrt folgt aus $-x < 0$ durch Addition von x

$$0 = -x + x < 0 + x = x.$$

b. Ist $x > 0$, so folgt unmittelbar

$$0 = 0 \cdot x < x \cdot x = x^2.$$

Ist $x < 0$, so ist $0 < -x$ und es gilt

$$0 = 0 \cdot (-x) < (-x) \cdot (-x) = x \cdot x = x^2.$$

c. $1 = 1^2 > 0$.

d. Nach Voraussetzung ist $y > 0$. Nehmen wir an, $\frac{1}{y} < 0$, so folgt

$$1 = \frac{1}{y} \cdot y < 0 \cdot y = 0$$

im Widerspruch zu Teil c., also ist $0 < \frac{1}{y}$. Entsprechend gilt $0 < \frac{1}{x}$, so daß auch

$$0 = 0 \cdot \frac{1}{y} < \frac{1}{x} \cdot \frac{1}{y} = \frac{1}{xy}$$

und somit wegen $x < y$ auch

$$\frac{1}{y} = x \cdot \frac{1}{xy} < y \cdot \frac{1}{xy} = \frac{1}{x}.$$

e. Wir wenden die Verträglichkeit der Totalordnung mit der Addition mehrfach an:

$$x + u < y + u < y + v.$$

f./g. Den Beweis überlassen wir dem Leser als Übungsaufgabe.

□

Proposition 8.15 (Charakterisierung des Supremums und Infimums).
Ist $(K,+,\cdot,\leq)$ ein angeordneter Körper, $A \subseteq K$ und $s \in K$, dann gelten

$$s = \sup(A) \iff \begin{array}{ll} 1) & \forall\, x \in A \,:\, x \leq s \text{ und} \\ 2) & \forall\, 0 < \varepsilon \in K \,:\, \exists\, x \in A \,:\, s - \varepsilon < x \end{array}$$

sowie

$$s = \inf(A) \iff \begin{array}{ll} 1) & \forall\, x \in A \,:\, x \geq s \text{ und} \\ 2) & \forall\, 0 < \varepsilon \in K \,:\, \exists\, x \in A \,:\, s + \varepsilon > x. \end{array}$$

Beweis: Ist $s = \sup(A)$, so ist s ein obere Schranke von A und somit gilt Bedingung 1). Sei also $0 < \varepsilon \in K$, so ist $s - \varepsilon < s$ und mithin ist $s - \varepsilon$ keine obere Schranke von A. Also gibt es ein $x \in A$ mit $x > s - \varepsilon$ und Bedingung 2) ist erfüllt.

Nehmen wir nun umgekehrt an, daß die Bedingungen 1) und 2) gelten. Wegen 1) ist s dann eine obere Schranke von A, und wir müssen nur noch zeigen, daß es keine kleinere obere Schranke geben kann. Dazu betrachten wir eine beliebige kleinere Zahl $t \in K$ mit $t < s$. Für $\varepsilon := s - t \in K$ gilt $\varepsilon > 0$ und wegen 2) gibt es dann ein $x \in A$ mit $x > s - \varepsilon = t$. Also ist t keine obere Schranke von A.

Die Aussage für das Infimum zeigt man analog. □

Das folgende Lemma ist interessant bei der Definition des Riemann-Integrals einer Funktion (siehe Definition 19.7).

Lemma 8.16.
Seien $A, B \subseteq \mathbb{R}$ zwei nicht-leere Teilmengen von $\mathbb{R}$ mit $a \leq b$ für alle $a \in A$, $b \in B$. Dann gilt

$$\sup(A) \leq \inf(B).$$

Beweis: Aus der Voraussetzung folgt unmittelbar, daß A nach oben und B nach unten beschränkt ist, so daß $\sup(A) \in \mathbb{R}$ und $\inf(B) \in \mathbb{R}$ existieren.

Angenommen, $\sup(A) > \inf(B)$, so ist $\varepsilon := \frac{\sup(A)-\inf(B)}{2} > 0$. Somit ist $\sup(A) - \varepsilon$ keine obere Schranke von A und $\inf(B) + \varepsilon$ keine untere Schranke von B. Es gibt also ein $a \in A$ und ein $b \in B$ mit

$$a > \sup(A) - \varepsilon = \frac{\sup(A) + \inf(B)}{2} = \inf(B) + \varepsilon > b,$$

was im Widerspruch zur Voraussetzung steht. □

Aufgaben

Aufgabe 8.17.
Ist M eine endliche Menge, so gilt

$$|M| = \min\{n \in \mathbb{N} \mid \exists\, f : M \longrightarrow \{1, \ldots, n\} \text{ injektiv}\},$$

und jede injektive Abbildung $f : M \longrightarrow \{1, \ldots, |M|\}$ ist bijektiv.

Aufgabe 8.18.
Zeige, daß durch

$$R := \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid x \text{ teilt } y\}$$

eine Ordnungsrelation auf $\mathbb{N}$ definiert wird. Ist R eine Totalordnung?

Aufgabe 8.19.
Definiere auf $M = \mathbb{N} \times \mathbb{N}$ eine Relation durch

$$\begin{aligned}(m,n) \leq (k,l) \quad \Longleftrightarrow \quad & 1.\ \max\{m,n\} < \max\{k,l\} \ \text{ oder} \\ & 2.\ (\max\{m,n\} = \max\{k,l\} \ \text{ und } m < k) \ \text{ oder} \\ & 3.\ (\max\{m,n\} = \max\{k,l\} \ \text{ und } m = k \ \text{ und } n > l) \ \text{ oder} \\ & 4.\ (m,n) = (k,l).\end{aligned}$$

Zeige, daß "$\leq$" eine Totalordnung auf M definiert. Stelle graphisch in der Zahlenebene $\mathbb{R}^2$ dar, wie die Elemente (m, n) in M mit $\max\{m, n\} \leq 4$ angeordnet sind.

Aufgabe 8.20.
Sei K ein angeordneter Körper und $A, B \subseteq K$ Teilmengen, so daß $\sup(A)$ und $\sup(B)$ existieren. Wir setzen $A + B := \{a + b \mid a \in A, b \in B\}$. Beweise, daß auch $\sup(A + B)$ existiert und $\sup(A + B) = \sup(A) + \sup(B)$ gilt.

Aufgabe 8.21.
Bestimme Supremum, Infimum, Maximum und Minimum (sofern sie existieren) der Mengen:

$$A = \left\{ \frac{m+n}{m \cdot n} \;\middle|\; m, n \in \mathbb{N}_{>0} \right\} \subseteq \mathbb{R}$$

und

$$B = \left\{ n + \frac{(-1)^n}{n} \;\middle|\; n \in \mathbb{N}_{>0} \right\} \subseteq \mathbb{R}.$$

Aufgabe 8.22 (Jeder angeordnete Körper enthält $\mathbb{N}$.).
Sei $(K, +, \cdot, \leq)$ ein angeordneter Körper mit Eins 1_K. Zeige, die Abbildung

$$\mathbb{N} \longrightarrow K : n \mapsto \sum_{i=1}^{n} 1_K$$

ist injektiv und verträglich mit den Operationen $+$ und $\cdot$.

Aufgabe 8.23.
Gib alle möglichen Ordnungsrelationen auf der 2-elementigen Menge $M = \{a, b\}$ als Teilmengen von $M \times M$ an. Begründe, weshalb dies alle sind.

Aufgabe 8.24.
Für zwei Tupel $(a, b), (c, d) \in \mathbb{N} \times \mathbb{N}$ definieren wir

$$(a, b) \ \leq \ (c, d) \quad :\Longleftrightarrow \quad a + b < c + d \ \text{ oder } \ (a + b = c + d \ \text{ und } \ a \leq c).$$

Zeige, $\leq$ ist eine Wohlordnung auf der Menge $M = \mathbb{N} \times \mathbb{N}$ und stelle graphisch (in der Ebene) dar, wie die Tupel (a, b) mit $a + b \leq 4$ angeordnet sind.

§ 9 Eigenschaften der reellen Zahlen ℝ

In diesem Abschnitt wollen wir einige wichtige Eigenschaften der reellen Zahlen betrachten.

A) Axiomatische Charakterisierung der reellen Zahlen

Theorem 9.1 (Charakterisierung der reellen Zahlen).
Der Körper ℝ der reellen Zahlen mit der üblichen Ordnungsrelation ist der einzige angeordnete Körper, in dem jede nicht-leere, nach oben beschränkte Menge ein Supremum besitzt.

Bemerkung 9.2.
Die Aussage in Theorem 9.1 besagt zweierlei. Zum einen wird festgestellt, daß ℝ ein angeordneter Körper ist und dem Supremumsaxiom genügt. Zum anderen wird festgestellt, daß dies für keinen *anderen* angeordneten Körper gilt. Das soll heißen, wenn es einen anderen angeordneten Körper $(K, +, \cdot, \leq)$ mit diesen Eigenschaften gibt, dann gibt es eine *bijektive* Abbildung

$$f : \mathbb{R} \longrightarrow K,$$

so daß $f(x+y) = f(x) + f(y)$, $f(x \cdot y) = f(x) \cdot f(y)$ und

$$x \leq y \quad \Longleftrightarrow \quad f(x) \leq f(y)$$

für alle $x, y \in \mathbb{R}$ gilt. In dem Fall kann man die beiden Körper nicht mehr unterscheiden. Man sagt deshalb auch, daß die reellen Zahlen durch die Eigenschaften in Theorem 9.1 charakterisiert sind, und man könnte die reellen Zahlen axiomatisch durch Angabe der Eigenschaften einführen.

Wir wollen Theorem 9.1 in dieser Vorlesung *nicht* beweisen. Stattdessen werden wir von den reellen Zahlen von nun an nur noch die im Satz angegebenen Eigenschaften wirklich verwenden. Wenn wir uns also ℝ als einen beliebigen angeordneten Körper mit Supremumsaxiom denken, dann wird alles, was wir von nun an beweisen, dort genauso gelten. Wir müßten die reellen Zahlen also noch gar nicht kennen, um die weitere Theorie betreiben zu können. Die wenigen oben gegebenen Axiome reichen uns aus. Insofern befinden wir uns von jetzt an auf wesentlich sichererem Grund und müssen nicht mehr immer wieder Bezug auf unser Vorwissen zu den Zahlsystemen nehmen.

Satz 9.3 (ℝ ist archimedisch angeordnet.).
Für $x, y \in \mathbb{R}$ mit $0 < x < y$ gibt es eine natürliche Zahl $n \in \mathbb{N}$, so daß $y < n \cdot x$.

Beweis: Wir betrachten die nicht-leere Teilmenge

$$A := \{n \cdot x \mid n \in \mathbb{N}\} \subsetneq \mathbb{R}$$

der reellen Zahlen und müssen zeigen, daß y keine obere Schranke dieser Menge ist.

Nehmen wir an, dies wäre doch der Fall, dann ist A nach oben beschränkt und somit existiert das Supremum

$$s := \sup(A).$$

Da $x > 0$ ist, ist $s - x < s$ und somit ist $s - x$ keine obere Schranke von A, d.h. es gibt eine natürliche Zahl $n \in \mathbb{N}$ mit

$$s - x < n \cdot x.$$

Dann ist aber auch

$$s = (s - x) + x < n \cdot x + x = (n + 1) \cdot x,$$

im Widerspruch dazu, daß s eine obere Schranke von A ist.

Damit haben wir gezeigt, daß A keine obere Schranke besitzt und insbesondere, daß y keine solche ist, d.h. es gibt eine natürliche Zahl $n \in \mathbb{N}$ mit $y < n \cdot x$. □

Korollar 9.4 (Konsequenzen der archimedischen Anordnung).

a. Für alle $x \in \mathbb{R}$ gibt es eine ganze Zahl n, so daß $n \leq x < n + 1$.

b. Für alle $\varepsilon \in \mathbb{R}$ mit $\varepsilon > 0$ gibt es eine natürliche Zahl n, so daß $0 < \frac{1}{n} < \varepsilon$.

Beweis:

a. Ist $0 \leq x < 1$, so ist $n = 0$. Ist $1 \leq x$, so gibt es nach Satz 9.3 eine Zahl $m \in \mathbb{N}$ mit $x < m \cdot 1 = m$. Nach dem Archimedischen Prinzip 8.8 besitzt dann die nicht-leere Menge

$$M := \{k \in \mathbb{N} \mid x < k\}$$

ein Minimum $m_0 = \min(M)$, und für $n := m_0 - 1 < m_0$ gilt mithin

$$n \leq x < m_0 = n + 1.$$

Ist $x < 0$, so ist $-x > 0$ und wir haben schon gezeigt, daß es eine natürliche Zahl $m \in \mathbb{N}$ mit $m \leq -x < m + 1$ gibt. Dann ist aber

$$-m - 1 < x \leq -m.$$

Falls $x = -m$, so setzen wir $n := -m$, und sonst setzen wir $n := -m - 1$.

b. Wegen $\varepsilon > 0$ ist nach Lemma 8.14 auch $\frac{1}{\varepsilon} > 0$, und nach a. gibt es dann eine natürliche Zahl $n \in \mathbb{N}$ so, daß

$$0 < \frac{1}{\varepsilon} < n.$$

Mit Lemma 8.14 folgt dann

$$0 < \frac{1}{n} < \frac{1}{\frac{1}{\varepsilon}} = \varepsilon.$$

□

Definition 9.5 (Intervalle).
Es seien $a, b \in \mathbb{R}$. Wir nennen eine Menge der Form

$$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$$

ein *abgeschlossenes Intervall*, eine Menge der Form

$$(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$$

ein *offenes Intervall* und Mengen der Form

$$[a, b) := \{x \in \mathbb{R} \mid a \leq x < b\}$$

bzw.

$$(a, b] := \{x \in \mathbb{R} \mid a < x \leq b\}$$

halboffene Intervalle. Mengen der Form

$$[a, \infty) := \{x \in \mathbb{R} \mid a \leq x\},$$
$$(a, \infty) := \{x \in \mathbb{R} \mid a < x\},$$
$$(-\infty, a] := \{x \in \mathbb{R} \mid x \leq a\},$$
$$(-\infty, a) := \{x \in \mathbb{R} \mid x < a\}$$
$$(-\infty, \infty) := \mathbb{R}$$

heißen *uneigentliche Intervalle*.

Satz 9.6 (ℚ liegt dicht in ℝ.).
Sind $a, b \in \mathbb{R}$ mit $a < b$, so gibt es eine rationale Zahl im Intervall (a, b).

Beweis: Wegen $b - a > 0$ gibt es nach Korollar 9.4 eine natürliche Zahl $n \in \mathbb{N}$ mit

$$0 < \frac{1}{n} < b - a. \tag{7}$$

Zudem gibt es nach Korollar 9.4 eine ganze Zahl $m \in \mathbb{Z}$ mit

$$m \leq n \cdot a < m + 1. \tag{8}$$

Damit gilt dann

$$a \overset{(8)}{<} \frac{m+1}{n} = \frac{m}{n} + \frac{1}{n} \overset{(8)}{\leq} a + \frac{1}{n} \overset{(7)}{<} b$$

und $\frac{m+1}{n} \in \mathbb{Q}$ ist eine rationale Zahl. □

B) Weitere Eigenschaften der reellen Zahlen

Satz 9.7 (Bernoullische Ungleichung).
Es sei $x \in \mathbb{R}$ mit $x \geq -1$ und $n \in \mathbb{N}$, dann gilt

$$(1+x)^n \geq 1 + n \cdot x.$$

Beweis: Wir führen den Beweis durch Induktion nach n.

Induktionsanfang: $n = 0$**:** $(1+x)^0 = 1 = 1 + 0 \cdot x$.

Induktionsschluß: $n \mapsto n+1$**:** Nach Lemma 8.14 b. ist $x^2 \geq 0$ und nach Voraussetzung gilt zudem $1 + x \geq 0$. Damit erhalten wir dann:

$$(1+x)^{n+1} = (1+x)^n \cdot (1+x) \overset{\text{Ind}}{\geq}$$

$$(1 + n \cdot x) \cdot (1+x) = 1 + (n+1) \cdot x + n \cdot x^2 \overset{8.14b.}{\geq} 1 + (n+1) \cdot x.$$

Die Aussage ist damit also mittels Induktion gezeigt. □

Mit Hilfe der Bernoulli-Ungleichung kann man die Existenz von n-ten Wurzeln nicht-negativer reeller Zahlen für alle $n \geq 2$ herleiten. Wir führen den Beweis hier auf, werden in der Vorlesung in Beispiel 14.23 aber mit Hilfe des Umkehrsatzes für streng monotone stetige Funktionen einen viel kürzeren und eleganteren Beweis geben.

Satz 9.8 (Existenz von n-ten Wurzeln in $\mathbb{R}$).
Zu jeder reellen Zahl $x \in \mathbb{R}$ mit $x \geq 0$ und jeder natürlichen Zahl $n \in \mathbb{N}$ mit $n \geq 2$ gibt es genau eine reelle Zahl $a \in \mathbb{R}$ mit $a \geq 0$ und $a^n = x$.
Wir nennen diese Zahl die n-te *Wurzel* aus x und bezeichnen sie mit $\sqrt[n]{x}$ oder $x^{\frac{1}{n}}$.

Beweis: Wir wollen uns zunächst der Eindeutigkeit der Lösung zuwenden, sofern sie existiert. Nehmen wir also an, es würde zwei verschiedene nicht-negative reelle Zahlen $a, b \in \mathbb{R}$ mit $a^n = b^n = x$ geben. Dann ist eine der beiden echt kleiner als die andere und wir können ohne Einschränkung annehmen, daß dies a ist, d.h. $0 \leq a < b$. Aus Lemma 8.14 g. folgt dann $x = a^n < b^n = x$, was ein offensichtlicher Widerspruch ist. Mithin haben wir gezeigt, daß es höchstens eine nicht-negative Zahl $a \in \mathbb{R}$ mit $a^n = x$ geben kann.

Es bleibt noch zu zeigen, daß es auch wirklich eine solche nicht-negative Zahl a gibt. Ist $x = 0$, so ist $a = 0$ eine Lösung für $a^n = 0$. Wir können im weiteren Verlauf des Beweises also voraussetzen, daß $x > 0$.

Wir betrachten dann die Teilmenge

$$A := \{y \in \mathbb{R} \mid y \geq 0, y^n \leq x\}$$

der reellen Zahlen, und wir behaupten, daß $1+x$ eine obere Schranke für A ist. Dazu betrachten wir eine reelle Zahl $y \in \mathbb{R}$ mit $y \geq 1+x > 0$. Aus der Bernoullischen Ungleichung folgt dann

$$y^n \overset{8.14g.}{\geq} (1+x)^n \overset{9.7}{\geq} 1+n\cdot x > x,$$

und somit ist $y \notin A$. Also ist A nach oben beschränkt durch $x+1$. Wegen $0 \in A$ ist A zudem nicht-leer und deshalb existiert das Supremum

$$a := \sup(A) \geq 0.$$

Wir wollen nun zeigen, daß $a^n = x$ gilt.

Zeige: $a^n \geq x$**:** Nehmen wir an, es gelte $a^n < x$.

Idee: *Finde eine reelle Zahl* $\varepsilon > 0$*, so daß* $a+\varepsilon \in A$*.* $\lightning$

Wegen $a \geq 0$ ist

$$c := \sum_{k=1}^{n} \binom{n}{k} \cdot a^{n-k} \geq \binom{n}{n} = 1 > 0$$

und somit auch $\frac{1}{c} > 0$ nach Lemma 8.14. Aus unserer Annahme folgt dann

$$\frac{x-a^n}{c} > 0.$$

Somit ist auch

$$\varepsilon := \min\left\{\frac{x-a^n}{c},\ 1\right\} > 0$$

und es folgt

$$a^n + c\cdot\varepsilon \leq x. \tag{9}$$

Wegen $0 < \varepsilon \leq 1$ ist $\varepsilon^k \leq \varepsilon$ für alle $k \geq 1$, und aus dem Binomischen Lehrsatz 7.15 folgt dann

$$\begin{aligned}(a+\varepsilon)^n =& a^n + \sum_{k=1}^{n} \binom{n}{k} \cdot a^{n-k}\cdot\varepsilon^k \\ \leq& a^n + \sum_{k=1}^{n} \binom{n}{k} \cdot a^{n-k}\cdot\varepsilon \ = \ a^n + c\cdot\varepsilon \overset{(9)}{\leq} x.\end{aligned}$$

Somit ist $a+\varepsilon \in A$ und $a+\varepsilon > a$ im Widerspruch dazu, daß a das Supremum von A ist. Mithin muß $a^n \geq x$ sein.

Zeige: $a^n \leq x$**:** Nehmen wir an, es gelte $a^n > x$.

Idee: *Finde ein* $\varepsilon > 0$ *und ein* $y \in A$*, so daß* $y^n > (a-\varepsilon)^n \geq x$*.* $\lightning$

Wegen $a^n > x$ ist $a > 0$ und dann ist auch die Zahl

$$\frac{a\cdot(a^n-x)}{n\cdot a^n} > 0$$

positiv. Wir setzen nun

$$\varepsilon := \min\left\{\frac{a\cdot(a^n-x)}{n\cdot a^n}\ ,\ \ a\right\} > 0.$$

Aus der Definition von ε folgt zum einen

$$-\frac{\varepsilon}{a} \geq -1 \tag{10}$$

und zum anderen unter Anwendung der Bernoullischen Ungleichung

$$x \leq a^n\cdot\left(1+n\cdot\frac{-\varepsilon}{a}\right) \overset{9.7}{\leq} a^n\cdot\left(1-\frac{\varepsilon}{a}\right)^n = (a-\varepsilon)^n\,; \tag{11}$$

dabei beachten, daß wir die Bernoullische Ungleichung wegen (10) anwenden können.

Da a das Supremum von A ist und $a-\varepsilon < a$ ist, muß es eine Zahl $y\in A$ geben mit

$$y > a-\varepsilon > 0.$$

Dann gilt nach Lemma 8.14 auch

$$y^n > (a-\varepsilon)^n \overset{(11)}{\geq} x,$$

im Widerspruch dazu, daß $y\in A$. Also muß auch $a^n\leq x$ gelten.

Da sowohl $a^n\geq x$, als auch $a^n\leq x$ gilt, folgt aus der Antisymmetrie der Ordnungsrelation, daß $a^n = x$, und wir haben die n-te Wurzel von x gefunden. □

Bemerkung 9.9.

In $\mathbb{R}$ besitzt also insbesondere jede nicht-negative Zahl eine Quadratwurzel. Dies gilt in den rationalen Zahlen nicht (siehe Satz 9.10), und man kann die reellen Zahlen als eine Erweiterung des Zahlbereichs der rationalen Zahlen ansehen, die unter anderem deshalb notwendig war. Negative Zahlen besitzen aber auch in $\mathbb{R}$ noch keine Quadratwurzeln, und wir werden im folgenden Kapitel deshalb unseren Zahlbereich noch einmal erweitern zu den sogenannten komplexen Zahlen, die dieses Manko dann beheben.

Satz 9.10 ($\sqrt{2}$ ist irrational.).
Es gibt keine rationale Zahl $a\in\mathbb{Q}$ mit $a^2 = 2$.

Beweis: Nehmen wir an, es wäre $a = \frac{p}{q}\in\mathbb{Q}$ eine solche Zahl. Wir können ohne weiteres annehmen, daß der Bruch in gekürzter Form vorliegt. Aus

$$\frac{p^2}{q^2} = a^2 = 2$$

folgt dann

$$p^2 = q^2\cdot 2.$$

Also ist p^2 eine gerade Zahl, und dann muß notwendigerweise auch p eine gerade Zahl sein. D.h. es gibt ein $b \in \mathbb{Z}$ mit $p = 2 \cdot b$. Also ist

$$4 \cdot b^2 = p^2 = 2 \cdot q^2,$$

und somit

$$2 \cdot b^2 = q^2.$$

Mit dem gleichen Argument sind dann auch q^2 und q gerade Zahlen, und somit ist q von der Form $q = 2 \cdot c$. Aber das widerspricht der Voraussetzung, daß der Bruch $\frac{p}{q}$ in gekürzter Form vorgelegen hat. □

Aufgaben

Aufgabe 9.11.
Zeige durch vollständige Induktion, daß

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \ldots + \frac{1}{\sqrt{n-1}} + \frac{1}{\sqrt{n}} > \sqrt{n}$$

für alle $n \geq 2$ gilt.

Aufgabe 9.12.
Zeige, für je drei reelle Zahlen $a, b, c \in \mathbb{R}$ gilt

$$\frac{|a-c|}{1+|a-c|} \leq \frac{|a-b|}{1+|a-b|} + \frac{|b-c|}{1+|b-c|}$$

Aufgabe 9.13.

a. Schreibe die folgenden Mengen als Vereinigung von Intervallen:
   (i) $A_1 = \{x \in \mathbb{R} \mid |x-1| < 1 \ \vee \ |x+2| \leq 2\}$.
   (ii) $A_2 = \{x \in \mathbb{R} \mid |x-1| > 1 \ \wedge \ |x+2| \leq 2\}$.
   (iii) $A_3 = \{x \in \mathbb{R} \mid x^2 - 2 \leq 2 \ \vee \ \sqrt{|x-10|} > 5\}$.

b. Überprüfe zudem jede der Mengen in Teil (a) sowie die folgende Menge, ob sie nach oben oder unten beschränkt ist, gib ggf. das Infimum oder Supremum der Menge an und überprüfe auch, ob es ein Minimim oder Maximum ist:

$$M = \left\{ \frac{m-n}{m+n} \mid 0 \neq m, n \in \mathbb{N} \right\}.$$

Aufgabe 9.14.
Für je zwei reelle Zahlen $x, y \in \mathbb{R}$ gilt

$$\max\{x, y\} = \frac{1}{2} \cdot (x + y + |x - y|).$$

§ 10 Der Körper der komplexen Zahlen

Ausblick 10.0 (Komplexe Zahlen in der Informatik).
Mit Hilfe der komplexen Zahlen lassen sich *Schwingungsvorgänge* in elektrischen Schaltkreisen auf sehr kompakte Art und Weise beschreiben, weshalb die komplexen Zahlen in der Elektrotechnik und damit in *Technischen Informatik* eine besondere Rolle spielen.

Wir betachten in diesem Abschnitt den Körper $\mathbb{C}$ der komplexen Zahlen, dem neben $\mathbb{R}$ wichtigsten Körper. Warum reichen eigentlich die reellen Zahlen nicht aus, wozu braucht man die komplexen Zahlen? Ja, man kann sogar fragen, warum wir überhaupt die reellen Zahlen benötigen, wenn wir doch ohnehin nur mit endlichen Dezimalbrüchen, also rationalen Zahlen, rechnen können? Die Antwort auf die zweite Frage ist schnell gegeben. Wir wissen alle, daß etwa ganz natürlich auftretende Größen wie die Länge der Diagonalen eines Quadrates mit Seitenlänge eins, sprich die Zahl $\sqrt{2}$, oder das Verhältnis von Umfang zum Durchmesser eines Kreises, sprich die Kreiszahl π, keine rationalen Zahlen sind. Sie sind aber reelle Zahlen und die reellen Zahlen sind in gewissem Sinne, eine Vervollständigung der rationalen Zahlen. Wir brauchen also die reellen Zahlen, da die rationalen Zahlen Lücken aufweisen. Die komplexen Zahlen werden nun deshalb eingeführt, um einen Mangel, den die reellen Zahlen immer noch haben, zu beheben. Hierbei geht es um das Lösen von Gleichungen, aber nicht mehr linearen, sondern quadratischen. Es ist bekannt, daß das Quadrat einer reellen Zahl stets nicht-negativ ist. Also kann es keine reelle Zahl x geben, die die Gleichung $x^2 = -1$ löst.

Als Lösung genau dieser Gleichung wird nun eine neue Größe eingeführt, die *imaginäre Einheit* i. Definitionsgemäß ist sie diejenige Zahl, für die $i^2 = -1$ gilt. Wenn man nun eine solche Größe i einführt, dann ist damit alleine gar nichts gewonnen. Man will ja mit i auch rechnen können, und zwar will man möglichst alle Rechenregeln von $\mathbb{R}$ übertragen. Man will nicht nur $i^2 = i \cdot i$, sondern auch $i + i$ oder Ausdrücke wie $37 + 42i$ bilden können. Dabei sollen die so zu konstruierenden *komplexen Zahlen* die reellen Zahlen als Teilmenge enthalten.

Daß es wirklich ein solches Zahlsystem komplexer Zahlen, in unserer Sprache den Körper der komplexen Zahlen, gibt, ist überhaupt nicht klar und wurde historisch erst spät realisiert und auch akzeptiert.[2] Gauß hat die Zahlen geometrisch, als Punkte in der Ebene, eingeführt, weshalb die komplexen Zahlen heute noch *Gaußsche Zahlenebene* heißen. Wir führen die komplexen Zahlen ebenfalls als reelle Zahlenpaare ein, definieren

[2] Erstmals tauchte $\sqrt{-1}$ wohl um 1540 bei Cardano auf. Wirklich als Zahlsystem wurden die komplexen Zahlen aber erst durch Gauß, 1777-1855, etabliert. Hierzu und zu vielen weiteren interessanten Tatsachen um die komplexen Zahlen vgl. [Ebb92] § 3.

die Addition und die Multiplikation aber algebraisch und werden die Definitionen erst im Anschluß daran geometrisch interpretieren.

A) Die Arithmetik der komplexen Zahlen

Bemerkung 10.1 (Konstruktion der komplexen Zahlen).
Es ist unser erklärtes Ziel, auf der reellen Zahlenebene $\mathbb{R}^2$ mit der Vektoraddtion

$$(x, y) + (u, v) := (x + u, y + v)$$

eine *Multiplikation* zu definieren, so daß einerseits die üblichen Rechenregeln (Assoziativgesetze, Kommutativgesetze und Distributivgesetze) gelten und daß außerdem der Vektor

$$i := (0, 1)$$

eine Lösung der Gleichung

$$z^2 = -1$$

ist. Um letzteres richtig zu interpretieren, denken wir uns die reelle Zahlengerade $\mathbb{R}$ als Teilmenge von $\mathbb{R}^2$, indem wir sie mit der x-Achse identifizieren, d.h.

$$\mathbb{R} \mathrel{\hat{=}} \{(a, 0) \mid a \in \mathbb{R}\} = x\text{-Achse}.$$

Die Multiplikation soll also der Bedingung

$$i^2 = (0, 1) \cdot (0, 1) = (-1, 0) \mathrel{\hat{=}} -1$$

genügen. Außerdem würden wir uns sicher wünschen, daß die Multiplikation eines Vektors mit der reellen Zahl

$$a \mathrel{\hat{=}} (a, 0)$$

wie die Streckung des Vektors um den Faktor a funktioniert, d.h.

$$(a, 0) \cdot (x, y) \mathrel{\hat{=}} a \cdot (x, y) = (ax, ay).$$

Wenn eine Multiplikation diese Wunschliste erfüllt, so gilt offenbar:

$$\begin{aligned}
(x, y) \cdot (u, v) &= \big((x, 0) + (0, y)\big) \cdot \big((u, 0) + (0, v)\big) \\
&= \big((x, 0) + (y, 0) \cdot (0, 1)\big) \cdot \big((u, 0) + (v, 0) \cdot (0, 1)\big) \\
&= (x, 0) \cdot (u, 0) + (y, 0) \cdot (0, 1) \cdot (u, 0) + (x, 0) \cdot (v, 0) \cdot (0, 1) \\
&\qquad + (y, 0) \cdot (0, 1) \cdot (v, 0) \cdot (0, 1) \\
&= (xu, 0) + (yu, 0) \cdot (0, 1) + (xv, 0) \cdot (0, 1) + (yv, 0) \cdot (0, 1) \cdot (0, 1) \\
&= (xu, 0) + (yu, 0) \cdot (0, 1) + (xv, 0) \cdot (0, 1) + (yv, 0) \cdot (-1, 0) \\
&= (xu, 0) + (0, yu) + (0, xv) + (-yv, 0) \\
&= (xu - yv, xv + yu).
\end{aligned}$$

Wir haben für die Definition der Multiplikation also nur *eine einzige* Möglichkeit, und die funktioniert zum Glück auch.

Satz 10.2 (Der Körper der komplexen Zahlen).
Die Menge $\mathbb{C} := \{(x, y) \mid x, y \in \mathbb{R}\}$ zusammen mit der durch

$$(x, y) + (u, v) := (x + u, y + v), \quad \text{für } (x, y), (u, v) \in \mathbb{C},$$

und

$$(x, y) \cdot (u, v) := (xu - yv, xv + yu), \quad \text{für } (x, y), (u, v) \in \mathbb{C},$$

definierten Addition und Multiplikation ist ein Körper, den wir den Körper der *komplexen Zahlen* nennen. .

Beweis: Dies folgt aus Aufgabe 7.19. □

Bemerkung 10.3.

a. Daß $\mathbb{C}$ mit den beiden Operationen ein *Körper* ist, bedeutet, daß die oben erwähnten üblichen Rechenregeln bezüglich der Addition, Subtraktion, Multiplikation und Division gelten, so wie wir sie von den reellen Zahlen her kennen. Man beachte dabei, daß die reelle Zahl $0 \hat{=} (0, 0)$ bei der Addtion nichts tut und die reelle Zahl $1 \hat{=} (1, 0)$ bei der Multiplikation ohne Wirkung ist:

$$(x, y) + (0, 0) = (x + 0, y + 0) = (x, y)$$

und

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

Das multiplikative Inverse der Zahl $(0, 0) \neq (x, y) \in \mathbb{C}$ ist

$$(x, y)^{-1} = \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right).$$

b. Die Abbildung

$$\iota : \mathbb{R} \longrightarrow \mathbb{C} : x \mapsto (x, 0)$$

ist mit der Addition und der Multiplikation verträglich und identifiziert den Körper der reellen Zahlen $\mathbb{R}$ mit dem Teilkörper $\mathbb{R} \times \{0\}$ von $\mathbb{C}$. Wir fassen $\mathbb{R}$ in diesem Sinne als Teilmenge von $\mathbb{C}$ auf.

c. Praktischer als das Rechnen mit Paaren von Zahlen ist die folgende Notation für komplexe Zahlen. Wir setzen $x := (x, 0)$ für $x \in \mathbb{R}$ und $i := (0, 1)$. Dann gilt für $z = (x, y) \in \mathbb{C}$

$$z = (x, y) = (x, 0) + (0, y) = (x, 0) + (0, 1) \cdot (y, 0) \hat{=} x + iy.$$

Diese Schreibweise wollen wir künftig für komplexe Zahlen verwenden. Damit gilt dann:

$$i^2 = (0,1) \cdot (0,1) = -1.$$

Ferner ergibt sich die etwas willkürlich anmutende Definition der Multiplikation ganz "natürlich" aus

$$(x+iy)(u+iv) = \big(xu + i^2yv\big) + i(xv+yu) = (xu-yv) + i(xv+yu).$$

Lemma 10.4 ($\mathbb{C}$ ist nicht angeordnet.).
Es gibt keine Totalordnung "$\leq$" auf $\mathbb{C}$, die $\mathbb{C}$ zu einem angeordneten Körper macht.

Beweis: Angenommen, es gäbe eine Totalordnung "$\leq$", die $\mathbb{C}$ zu einem angeordneten Körper macht. Nach Lemma 8.14 muß dann $0 < i^2 = -1$ gelten, was im Widerspruch zu $0 < 1$ steht. □

Definition 10.5 (Der Betrag und die komplexe Konjugation).

a. Wir definieren die *Betragsfunktion* auf $\mathbb{C}$ durch

$$|\cdot| : \mathbb{C} \longrightarrow \mathbb{R}_{\geq 0} : x+iy \mapsto \sqrt{x^2+y^2}$$

und nennen $|x|$ auch den *Absolutbetrag* von x. Wegen Satz 9.8 ist der Betrag einer komplexen Zahl definiert und ist stets eine nicht-negative reelle Zahl. Beachte zudem, für $x \in \mathbb{R}$ gilt

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

b. Wir definieren die *komplexe Konjugation* als

$$\overline{\cdot} : \mathbb{C} \longrightarrow \mathbb{C} : z = x+iy \mapsto \overline{z} := x - iy.$$

Für $z \in \mathbb{C}$ heißt $\overline{z}$ die zu z *konjugiert komplexe Zahl*.

c. Wir definieren die Abbildungen *Realteil*

$$\mathrm{Re} : \mathbb{C} \longrightarrow \mathbb{R} : x+iy \mapsto x$$

und *Imaginärteil*

$$\mathrm{Im} : \mathbb{C} \longrightarrow \mathbb{R} : x+iy \mapsto y$$

und nennen $\mathrm{Re}(x+iy) = x$ den *Realteil* von z und $\mathrm{Im}(x+iy) = y$ den *Imaginärteil* von z.

Beispiel 10.6.
Wir betrachten die komplexe Zahl

$$z = i - 1 = -1 + i.$$

Dann gilt $\mathrm{Re}(z) = -1$, $\mathrm{Im}(z) = 1$ und

$$\overline{z} = -1 - i = -(1 + i).$$

Für den Betrag von z rechnen wir

$$|z| = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2} = \sqrt{1 + 1} = \sqrt{2}$$

und damit erhalten wir die Gleichung

$$z \cdot \overline{z} = (-1 + i) \cdot (-1 - i) = 2 = |z|^2.$$

Lemma 10.7 (Einfache Rechenregeln in $\mathbb{C}$).
Es seien $z, w \in \mathbb{C}$.

a. Der Betrag ist multiplikativ, d.h.
$$|z| \cdot |w| = |zw|.$$

b. Der Betrag erfüllt die *Dreiecksungleichung*, d.h.
$$|z + w| \leq |z| + |w|,$$
und es gilt stets
$$\big||z| - |w|\big| \leq |z - w|.$$

c. $z = 0 \iff |z| = 0$.

d. $z \cdot \overline{z} = |z|^2$.

e. Wenn $z \neq 0$, dann ist $z^{-1} = \frac{1}{z} = \frac{\overline{z}}{|z|^2}$.

f. Die komplexe Konjugation ist additiv, d.h.
$$\overline{z} + \overline{w} = \overline{z + w}.$$

g. Die komplexe Konjugation ist multiplikativ, d.h.
$$\overline{z} \cdot \overline{w} = \overline{z \cdot w}.$$

h. $\overline{\overline{z}} = z$.

i. $\overline{z} = z \iff z \in \mathbb{R}$.

j. $\mathrm{Re}(z) = \frac{z + \overline{z}}{2} \leq |z|$.

k. $\mathrm{Im}(z) = \frac{z - \overline{z}}{2i} \leq |z|$.

l. $|z| = \big|\overline{z}\big| = |-z|$.

Beweis: Die Aussagen in den Teilen c.-l. überlassen wir dem Leser als Übungsaufgabe.

a. Seien $z = x + iy$ und $w = u + iv$ mit $x, y, u, v \in \mathbb{R}$. Dann gilt

$$\begin{aligned}|zw|^2 =& |(xu - yv) + i \cdot (xv + yu)|^2 = (xu - yv)^2 + (xv + yu)^2\\ =& x^2u^2 - 2xuyv + y^2v^2 + x^2v^2 + 2xvyu + y^2u^2\\ =& x^2u^2 + y^2v^2 + x^2v^2 + y^2u^2 = (x^2 + y^2) \cdot (u^2 + v^2)\\ =& |z|^2 \cdot |w|^2 = \big(|z| \cdot |w|\big)^2.\end{aligned}$$

Aus der Eindeutigkeit der nicht-negativen Quadratwurzel (Satz 9.8) folgt dann

$$|zw| = |z| \cdot |w|.$$

b. Wir wollen nun die Dreiecksungleichung unter Verwendung der übrigen Aussagen zeigen. Es gilt

$$\begin{aligned}|z + w|^2 &\overset{d.}{=} (z + w) \cdot \overline{(z + w)}\\ &\overset{f.}{=} z \cdot \overline{z} + (z \cdot \overline{w} + \overline{z} \cdot w) + w \cdot \overline{w}\\ &\overset{d.,h.}{=} |z|^2 + \big(z \cdot \overline{w} + \overline{z \cdot \overline{w}}\big) + |w|^2\\ &\overset{j.}{=} |z|^2 + 2 \cdot \mathrm{Re}(z \cdot \overline{w}) + |w|^2\\ &\overset{j.}{\leq} |z|^2 + 2 \cdot |z \cdot \overline{w}| + |w|^2\\ &\overset{a.}{=} |z|^2 + 2 \cdot |z| \cdot |\overline{w}| + |w|^2\\ &\overset{l.}{=} |z|^2 + 2 \cdot |z| \cdot |w| + |w|^2\\ &= \big(|z| + |w|\big)^2.\end{aligned}$$

Da dies eine Ungleichung von nicht-negativen Zahlen in dem angeordneten Körper $\mathbb{R}$ ist, folgt aus Lemma 8.14, daß

$$|z + w| \leq |z| + |w|.$$

Es bleibt, die zweite Aussage in Teil b. zu zeigen. Aus der Dreiecksungleichung erhalten wir

$$|z| = |(z - w) + w| \leq |z - w| + |w|,$$

und somit

$$|z| - |w| \leq |z - w|.$$

Analog folgt

$$-(|z| - |w|) = |w| - |z| \leq |w - z| = |-(w - z)| = |z - w|.$$

Wegen

$$\big||z| - |w|\big| = \begin{cases} |z| - |w|, & \text{falls } |z| - |w| \geq 0,\\ -(|z| - |w|) & \text{falls } |z| - |w| < 0,\end{cases}$$

folgt dann $\big||z| - |w|\big| \leq |z - w|$.

□

Beispiel 10.8.

a. Gegeben seien $z = 3 + 2i$ und $w = 5 - i$. Dann gelten

$$z \cdot w = (3 \cdot 5 - 2 \cdot (-1)) + (3 \cdot (-1) + 2 \cdot 5) \cdot i = 17 + 7i$$

sowie

$$|w| = \sqrt{5^2 + (-1)^2} = \sqrt{26}$$

und

$$\begin{aligned} \frac{z}{w} =& z \cdot \frac{\overline{w}}{|w|^2} = (3 + 2i) \cdot \left(\frac{5}{26} + \frac{1}{26} \cdot i\right) \\ =& \left(3 \cdot \frac{5}{26} - 2 \cdot \frac{1}{26}\right) + \left(3 \cdot \frac{1}{26} + 2 \cdot \frac{5}{26}\right) \cdot i \\ =& \frac{13}{26} + \frac{13}{26} \cdot i = \frac{1}{2} + \frac{1}{2} \cdot i. \end{aligned}$$

b. Für die komplexen Zahlen $z = 3 + 4i$ und $w = 5 - 12i$ gilt

$$z + w = (3 + 5) + (4 - 12) \cdot i = 8 - 8i$$

und somit

$$\begin{aligned} |z + w| =& \sqrt{8^2 + 8^2} = \sqrt{2} \cdot 8 < 16 < 18 = 5 + 13 \\ =& \sqrt{25} + \sqrt{169} = \sqrt{3^2 + 4^2} + \sqrt{5^2 + 12^2} = |z| + |w|. \end{aligned}$$

Außerdem gilt

$$\frac{z + \overline{z}}{2} = \frac{(3 + 4i) + (3 - 4i)}{2} = \frac{6}{2} = 3 = \mathrm{Re}(z).$$

B) Geometrische Interpretation der Arithmetik der komplexen Zahlen

Bemerkung 10.9 (Geometrische Deutung und Polarkoordinaten).
Wir wollen hier einige der bisher eingeführten Operationen auf den komplexen Zahlen und der angeführten Eigenschaften derselben geometrisch interpretieren.

- Die Addition ist einfach die komponentenweise Addition, also die Addition der Vektoren (siehe Abbildung 3).

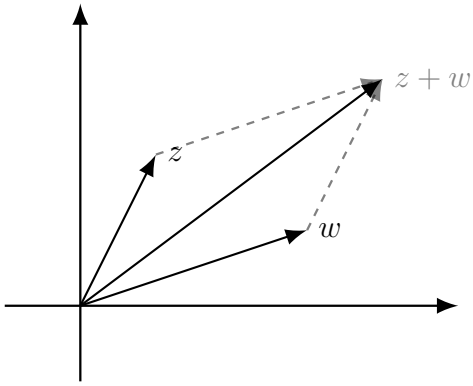


Abbildung 3: Addition in $\mathbb{C}$ als Vektoraddition

- Die komplexe Konjugation ist die Spiegelung an der x-Achse.
- Der Realteil ist die orthogonale Projektion auf die x-Achse und der Imaginärteil die orthogonale Projektion auf die y-Achse.
- Der Betrag $|z| = \sqrt{x^2 + y^2}$ einer komplexen Zahl $z = x + iy$ ist die euklidische Länge des Vektors z, d.h. der Abstand von z zum Ursprung. Dies ergibt sich unmittelbar aus dem Satz von Pythagoras (siehe Abbildung 4).

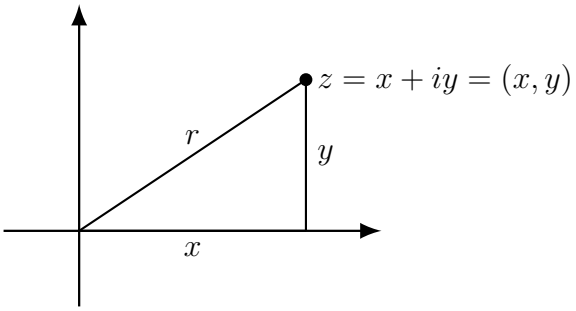


Abbildung 4: Pythagoras: $x^2 + y^2 = r^2$

- Die Dreiecksungleichung besagt deshalb im wesentlichen, daß in einem Dreieck die Summe der Seitenlängen von zwei Seiten stets eine obere Schranke für die Seitenlänge der dritten Seite ist.
- Die Menge

$$K := \{z \in \mathbb{C} \mid |z| = 1\}$$

der Punkte in der Ebene, deren Abstand zum Ursprung genau 1 ist, ist der Einheitskreis um den Ursprung. Man beachte, daß bei einem Punkt

$$z = x + iy,$$

der auf dem Einheitskreis liegt, die kartesichen Koordinaten x und y schon vollständig durch den Winkel $\alpha \in [0, 2\pi)$ bestimmt sind, den der Vektor z mit der x-Achse einschließt. Es gilt nämlich (siehe Abbildung 5)

$$x = \cos(\alpha)$$

und

$$y = \sin(\alpha)$$

und somit

$$z = \cos(\alpha) + i \cdot \sin(\alpha).$$

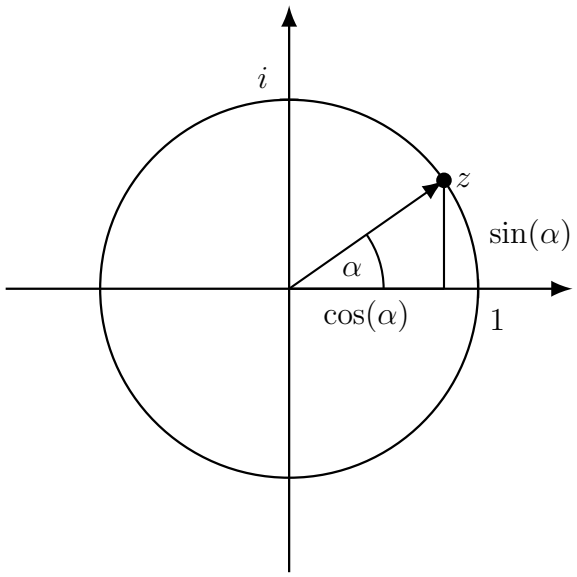


Abbildung 5: Koordinaten eines Punktes $z = \cos(\alpha) + i \cdot \sin(\alpha)$ auf dem Einheitskreis

- Es bleibt, die Multiplikation zweier komplexer Zahlen $0 \neq z, w \in \mathbb{C}$ geometrisch zu deuten. Dazu schreiben wir die Zahl z als

$$z = |z| \cdot \frac{z}{|z|} = r \cdot z'$$

mit $r = |z|$ und $z' = \frac{z}{|z|}$. Man beachte, daß die Zahl z' den Betrag 1 hat, so daß es genau einen Winkel $\alpha \in [0, 2\pi)$ gibt mit

$$z' = (\cos(\alpha), \sin(\alpha)) = \cos(\alpha) + i \cdot \sin(\alpha).$$

Die komplexe Zahl $z \neq 0$ ist also eindeutig durch ihren Betrag und den Winkel α bestimmt. Wir nennen

$$\arg(z) := \alpha$$

das *Argument* von z und das Paar

$$(r, \alpha) = \big(|z|, \arg(z)\big)$$

die *Polarkoordinaten* von z.

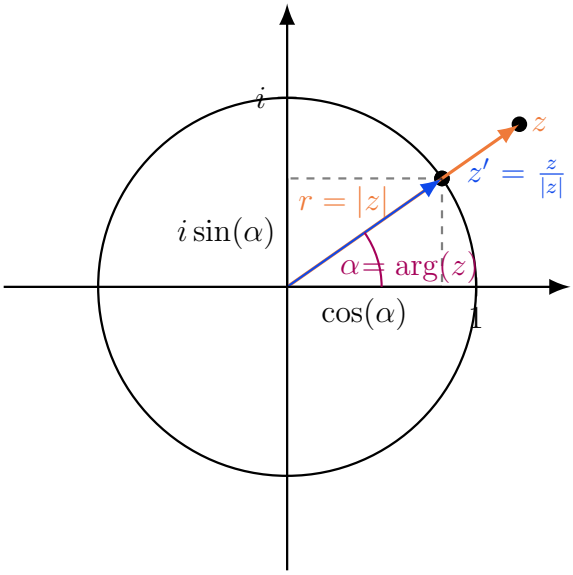


Abbildung 6: Polarkoordinaten von $z = r \cdot (\cos(\alpha) + i \cdot \sin(\alpha))$

Wir erinnern hier an die beiden Additionstheoreme für den Sinus und den Cosinus (siehe auch Satz 12.38):

$$\cos(\alpha + \beta) = \cos(\alpha) \cdot \cos(\beta) - \sin(\alpha) \cdot \sin(\beta) \tag{12}$$

und

$$\sin(\alpha + \beta) = \cos(\alpha) \cdot \sin(\beta) + \sin(\alpha) \cdot \cos(\beta). \tag{13}$$

Betrachten wir zunächst die Multiplikation von zwei komplexen Zahlen $z = |z| \cdot \big(\cos(\alpha) + i \cdot \sin(\alpha)\big)$ und $w = |w| \cdot \big(\cos(\beta) + i \cdot \sin(\beta)\big)$:

$$\begin{aligned} z \cdot w &= |z| \cdot |w| \cdot \big(\cos(\alpha) + i \cdot \sin(\alpha)\big) \cdot \big(\cos(\beta) + i \cdot \sin(\beta)\big) \\ &= |z| \cdot |w| \cdot \big(\cos(\alpha) \cdot \cos(\beta) - \sin(\alpha) \cdot \sin(\beta)\big) + i \cdot \big(\cos(\alpha) \cdot \sin(\beta) + \sin(\alpha) \cdot \cos(\beta)\big) \\ &\overset{(12),(13)}{=} |z| \cdot |w| \cdot \big(\cos(\alpha + \beta) + i \cdot \sin(\alpha + \beta)\big). \end{aligned}$$

Die beiden Zahlen werden also multipliziert, indem man die Argumente addiert und die Beträge multipliziert (siehe Abbildung 7).

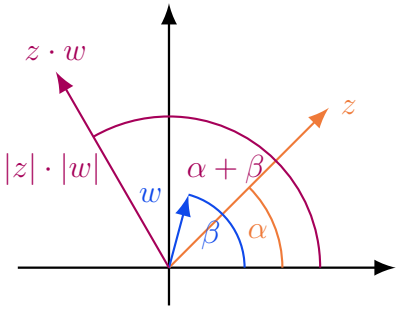


Abbildung 7: Multiplikation zweier komplexer Zahlen

In Polarkoordinaten könnte man dies schreiben als

$$(r, \alpha) \cdot (s, \beta) = (r \cdot s, \alpha + \beta).$$

Beispiel 10.10.
Zur Ermittlung von $\alpha = \arg(z)$ für $z = i - 1$ betrachten wir die Zahl

$$\frac{z}{|z|} = -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i$$

vom Betrag 1, für die gilt

$$\frac{z}{|z|} = \cos(\alpha) + i\sin(\alpha),$$

d.h. $\cos(\alpha) = -\frac{\sqrt{2}}{2}$ und $\sin(\alpha) = \frac{\sqrt{2}}{2}$, also $\alpha = \frac{3}{4}\pi$.

Bemerkung 10.11 (n-te Wurzeln).
Aus der Polarkoordinatendarstellung einer komplexen Zahl

$$w = r \cdot \big(\cos(\alpha) + i \cdot \sin(\alpha)\big)$$

läßt sich leicht ableiten, daß die Zahl

$$a = \sqrt[n]{r} \cdot \left(\cos\left(\tfrac{\alpha}{n}\right) + i \cdot \sin\left(\tfrac{\alpha}{n}\right)\right)$$

eine n-te Wurzel aus w ist, d.h.

$$a^n = w.$$

Dabei ist $\sqrt[n]{r}$ die eindeutig bestimmte nicht-negative n-te Wurzel der nicht-negativen Zahl r.

Die obige Zahl a ist aber nicht die einzige Lösung der Gleichung

$$z^n = w$$

in $\mathbb{C}$. Denn addiert man zum Argument einen der folgenden Winkel

$$\frac{2\pi k}{n}, \quad \text{mit } k = 1, \ldots, n-1,$$

so erhalten wir

$$\begin{aligned}\left(\sqrt[n]{r} \cdot \left(\cos\left(\tfrac{\alpha+2\pi k}{n}\right) + i \cdot \sin\left(\tfrac{\alpha+2\pi k}{n}\right)\right)\right)^n &= \sqrt[n]{r}^n \cdot \left(\cos\left(\alpha + 2\pi k\right) + i \cdot \sin\left(\alpha + 2\pi k\right)\right) \\ &= \sqrt[n]{r}^n \cdot \left(\cos\left(\alpha\right) + i \cdot \sin\left(\alpha\right)\right) = w.\end{aligned}$$

Wir haben also in der Tat n verschiedene n-te Wurzeln von w gefunden:

$$a_k = \sqrt[n]{r} \cdot \left(\cos\left(\tfrac{\alpha+2\pi\cdot k}{n}\right) + i \cdot \sin\left(\tfrac{\alpha+2\pi\cdot k}{n}\right)\right), \quad k = 0, \ldots, n-1.$$

Damit sehen wir, daß die Polynomgleichung

$$z^n = 1$$

in $\mathbb{C}$ genau n Lösungen hat, wobei n der Grad der Gleichung ist. Das ist ein Spezialfall des Fundamentalsatzes der Algebra.

Aufgaben

Aufgabe 10.12.
Bestimme für die folgenden komplexen Zahlen $\operatorname{Re} z$, $\operatorname{Im} z$, $\arg z$, $|z|$, $\overline{z}$ und z^{-1}:

a. $z = i - 1$.

b. $z = \frac{4i}{1+i}$.

c. $z = \frac{(2+2i)^7}{(1-i)^3}$.

d. $z = \frac{4+2i}{2-2i}$.

Aufgabe 10.13.
Berechne die komplexe Zahl $\left(\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right)^{4n}$ für jede natürliche Zahl $n \in \mathbb{N}$.

Aufgabe 10.14.
Es sei $z \in \mathbb{C}$ eine komplexe Zahl mit $\operatorname{Im}(z) \neq 0$ und

$$\frac{2 + 3z + 4z^2}{2 - 3z + 4z^2} \in \mathbb{R}.$$

Bestimme $|z|^2$.

Aufgabe 10.15.
Wie viele Lösungen $z \in \mathbb{C}$ hat die Gleichung

$$\overline{z} - z^2 = i \cdot \left(\overline{z} + z^2\right)?$$

Bestimme alle Lösungen.

Kapitel II

Eindimensionale Analysis

Im folgenden wollen wir die eindimensionale Analysis entwickeln, teilweise nur über den reellen Zahlen, teilweise parallel über den reellen und den komplexen Zahlen. Deshalb führen wir folgende Notation ein.

> Im folgenden sei stets $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ einer der beiden Körper $\mathbb{R}$ oder $\mathbb{C}$.

Ausblick (Analysis in der Informatik)
Die Analysis spielt für die Beschreibung vieler Phänomene in den Natur- und Ingenieurswissenschaften eine wichtige Rolle. In der Informatik sind die Methoden vor allem für die physiknahe Bereiche wie die Computergraphik, Bildverarbeitung oder das Maschinelle Lernen sowie für Komplexitätsabschätzungen wichtig.

§ 11 Folgen und ihre Grenzwerte

A) Konvergente Folgen

Ausblick 11.0 (Folgen in der Informatik).
Lösungen zu vielen Problemen (etwa in der Optimierung) sind formal die Grenzwerte von Folgen und können deshalb durch iterative Verfahren näherungsweise bestimmt werden, für die dann *iterative Algorithmen* zur konkreten Berechnung entwickelt, untersucht und implementiert werden müssen. Zudem spielen Folgen und ihr asymptotisches Verhalten bei der Aufwandsabschätzung von Algorithmen eine Rolle, wie wir in Abschnitt 11.H) sehen werden.

Definition 11.1 (Folgen).
Eine *Folge* in $\mathbb{K}$ ist eine Abbildung

$$\alpha : \mathbb{N} \longrightarrow \mathbb{K}$$

von den natürlichen Zahlen $\mathbb{N}$ nach $\mathbb{K}$.

Notation 11.2 (Familienschreibweise für Folgen).
Eine Folge $\alpha : \mathbb{N} \longrightarrow \mathbb{K}$ ist eindeutig festgelegt durch ihre Funktionswerte $a_n := \alpha(n)$ mit $n \in \mathbb{N}$. Wir schreiben deshalb statt $\alpha : \mathbb{N} \longrightarrow \mathbb{K}$ gemeinhin nur $(a_n)_{n\in\mathbb{N}}$ oder $(a_0, a_1, a_2, \ldots)$.

Manchmal ist es angenehmer, eine Folge nicht bei 0 starten zu lassen, sondern bei einer anderen natürlichen Zahl k. Dann schreiben wir für die Folge schlicht $(a_n)_{n\geq k}$. Formal würde dem dann die Abbildung

$$\mathbb{N} \longrightarrow \mathbb{K} : n \mapsto a_{n+k}$$

entsprechen.

Beispiel 11.3.

a. Ist $c \in \mathbb{K}$, so heißt $\alpha : \mathbb{N} \longrightarrow \mathbb{K} : n \mapsto c$ eine *konstante Folge*. Es gilt $a_n = c$ für $n \in \mathbb{N}$, und mithin $(a_n)_{n\in\mathbb{N}} = (c)_{n\in\mathbb{N}}$. Z.B. $(a_n)_{n\in\mathbb{N}} = (2, 2, 2, 2, 2, \ldots)$.

b. Für $q \in \mathbb{K}$ ist auch $\alpha : \mathbb{N} \longrightarrow \mathbb{K} : n \mapsto q^n$ eine Folge mit $a_n = q^n$, also $(a_n)_{n\in\mathbb{N}} = (q^n)_{n\in\mathbb{N}}$. Z.B. $(a_n)_{n\in\mathbb{N}} = (1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots)$.

c. $\left(\frac{1}{n}\right)_{n\geq 1}$ ist ein Beispiel für eine Folge in $\mathbb{K}$, bei der der Folgenindex nicht bei 0 startet.

Definition 11.4 (Konvergenz und Grenzwert).
Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $a \in \mathbb{K}$.

a. Wir nennen a genau dann einen *Grenzwert* von $(a_n)_{n\in\mathbb{N}}$, wenn

$$\forall\, 0 < \varepsilon \in \mathbb{R} \quad \exists\, n_\varepsilon \in \mathbb{N} \;:\; \forall\, n \geq n_\varepsilon \;:\; |a_n - a| < \varepsilon.$$

In diesem Fall sagen wir auch, daß $(a_n)_{n\in\mathbb{N}}$ *gegen a konvergiert* und schreiben

$$\lim_{n\to\infty} a_n = a$$

oder

$$a_n \longrightarrow a.$$

b. Wir nennen $(a_n)_{n\in\mathbb{N}}$ genau dann *konvergent*, wenn es ein $a \in \mathbb{K}$ gibt, so daß $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert. Andernfalls nennen wir $(a_n)_{n\in\mathbb{N}}$ *divergent*.

c. Wir nennen eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ eine *Nullfolge*, wenn $(a_n)_{n\in\mathbb{N}}$ gegen 0 konvergiert, d.h. $a_n \longrightarrow 0$.

Beispiel 11.5.

a. Die konstante Folge $(a_n)_{n\in\mathbb{N}} = (c)_{n\in\mathbb{N}}$ konvergiert gegen c, d.h. $\lim\limits_{n\to\infty} c = c$.
Um das zu sehen, wählen wir für eine reelle Zahl $\varepsilon > 0$ die natürliche Zahl $n_\varepsilon = 0$, so daß für jedes $n \geq n_\varepsilon = 0$ gilt

$$|a_n - c| = |c - c| = 0 < \varepsilon.$$

b. Die Folge $\left(\frac{1}{n}\right)_{n\geq 1}$ konvergiert gegen 0, d.h. $\lim\limits_{n\to\infty} \frac{1}{n} = 0$.
Denn: sei $0 < \varepsilon \in \mathbb{R}$ gegeben, so gibt es nach Korollar 9.4 eine natürliche Zahl n_ε, so daß $0 < \frac{1}{n_\varepsilon} < \varepsilon$. Ist nun $n \geq n_\varepsilon$, so folgt

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} \leq \frac{1}{n_\varepsilon} < \varepsilon.$$

c. Die Folge $(a_n)_{n\in\mathbb{N}} = \left((-1)^n\right)_{n\in\mathbb{N}}$ ist divergent.
Denn: nehmen wir an, $(a_n)_{n\in\mathbb{N}}$ konvergiert gegen a. Dann gibt es zu $\varepsilon = \frac{1}{2}$ ein n_ε, so daß $|a_n - a| < \varepsilon$ für $n \geq n_\varepsilon$. Insbesondere gilt dann wegen der Dreiecksungleichung

$$2 = |(-1)^{n_\varepsilon} - (-1)^{n_\varepsilon+1}| = |a_{n_\varepsilon} - a_{n_\varepsilon+1}| \leq |a_{n_\varepsilon} - a| + |a - a_{n_\varepsilon+1}| < \varepsilon + \varepsilon = 1,$$

was ein offensichtlicher Widerspruch ist.

Lemma 11.6 (Geometrische Folge).
Es sei $q \in \mathbb{K}$ mit $|q| < 1$, so ist $(q^n)_{n\in\mathbb{N}}$ eine Nullfolge.

Beweis: Wir können ohne Einschränkung annehmen, daß $q \neq 0$, da die Folge sonst sicher eine Nullfolge ist.

Sei $\varepsilon > 0$ gegeben. Wir betrachten die reelle Zahl

$$x := \frac{1}{|q|} - 1 > 0,$$

die positiv ist, da nach Voraussetzung $0 < |q| < 1$. Nach Korollar 9.4 gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$\frac{1}{n_\varepsilon} < x \cdot \varepsilon \tag{14}$$

Ist nun $n \geq n_\varepsilon$, so gilt wegen $|q| = \frac{1}{1+x}$ und der Bernoullischen Ungleichung auch

$$|q^n - 0| = |q|^n = \frac{1}{(1+x)^n} \overset{9.7}{\leq} \frac{1}{1+n\cdot x} < \frac{1}{n\cdot x} \leq \frac{1}{n_\varepsilon \cdot x} \overset{(14)}{<} \frac{x\cdot\varepsilon}{x} = \varepsilon.$$

□

Bemerkung 11.7.
Für eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ gilt offenbar:

$$a_n \longrightarrow a \iff a_n - a \longrightarrow 0 \iff |a_n - a| \longrightarrow 0.$$

Diese Feststellung ist in mancher Anwendung von Nutzen, um Argumente abzukürzen.

Proposition 11.8 (Eindeutigkeit des Grenzwertes von Folgen).
Der Grenzwert einer konvergenten Folge in $\mathbb{K}$ ist eindeutig bestimmt.

Beweis: Nehmen wir an, eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ besitze zwei verschiedene Grenzwerte $a, b \in \mathbb{K}$. Dann ist die reelle Zahl

$$\varepsilon := \frac{|a-b|}{2} > 0$$

positiv. Mithin gibt es zwei natürliche Zahlen $n_\varepsilon, n'_\varepsilon \in \mathbb{N}$, so daß

$$|a_n - a| < \varepsilon$$

für $n \geq n_\varepsilon$ und

$$|a_n - b| < \varepsilon$$

für $n \geq n'_\varepsilon$. Setzen wir nun $N := \max\{n_\varepsilon, n'_\varepsilon\}$, so gilt

$$|a-b| \leq |a-a_N| + |a_N - b| < \varepsilon + \varepsilon = |a-b|,$$

was ein offensichtlicher Widerspruch ist. □

B) Beschränkte Folgen

Definition 11.9 (Beschränkte Folgen).
Eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ heißt *beschränkt*, wenn die Menge

$$\big\{|a_n| \in \mathbb{R} \;\big|\; n \in \mathbb{N}\big\}$$

beschränkt ist, d.h. wenn es eine Zahl $s \in \mathbb{R}$ gibt, so daß $|a_n| \leq s$ für alle $n \in \mathbb{N}$. Man beachte dabei, daß die Menge stets durch 0 nach unten beschränkt ist, und wir nennen eine Zahl s wie oben eine Schranke für $(a_n)_{n\in\mathbb{N}}$.

Beispiel 11.10.

a. Die konvergente Folge $\left(\frac{1}{n}\right)_{n\geq 1}$ ist beschränkt, da $\left|\frac{1}{n}\right| \leq 1$ für alle $n \geq 1$.

b. Die divergente Folge $(a_n)_{n\in\mathbb{N}} = \big((-1)^n\big)_{n\in\mathbb{N}}$ ist ebenfalls beschränkt, da $\big\{|a_n| \in \mathbb{R} \;\big|\; n \in \mathbb{N}\big\} = \{1\}$.

Satz 11.11 (Konvergente Folgen sind beschränkt.).
Jede konvergente Folge in $\mathbb{K}$ ist beschränkt.

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $\lim\limits_{n\to\infty} a_n = a$. Dann gibt es zu $\varepsilon = 1$ eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$|a_n - a| < \varepsilon = 1$$

für $n \geq n_\varepsilon$. Setze

$$s := \max\{1 + |a|, |a_0|, |a_1|, \ldots, |a_{n_\varepsilon - 1}|\},$$

wobei man beachte, daß das Maximum existiert, weil die Menge endlich ist.

Damit erhalten wir dann

$$|a_n| \leq \begin{cases} s, & \text{falls } n < n_\varepsilon, \\ |a_n - a| + |a| < 1 + |a| \leq s, & \text{falls } n \geq n_\varepsilon. \end{cases}$$

Mithin ist s eine Schranke für $(a_n)_{n\in\mathbb{N}}$. □

Beispiel 11.12.

a. Beispiel 11.10 zeigt, daß die Umkehrung von Satz 11.11 nicht gilt.

b. Für $k \in \mathbb{N}$ mit $k \geq 1$ ist die Folge $(a_n)_{n\in\mathbb{N}} = (n^k)_{n\in\mathbb{N}}$ nicht beschränkt, also auch nicht konvergent.

c. Für $q \in \mathbb{K}$ mit $|q| > 1$ ist die Folge $(a_n)_{n\in\mathbb{N}} = (q^n)_{n\in\mathbb{N}}$ nicht beschränkt und somit divergent.
Um dies zu sehen, nehmen wir an, $s > 0$ sei eine Schranke für die Folge $(a_n)_{n\in\mathbb{N}}$ und setzen $x := |q| - 1 > 0$. Da $\mathbb{R}$ archimedisch angeordnet ist (siehe Satz 9.3), gibt es eine natürliche Zahl $n \in \mathbb{N}$, so daß

$$s < n \cdot x.$$

Aus der Bernoullischen Ungleichung erhalten wir damit

$$|q|^n = (1 + x)^n \geq 1 + n \cdot x > s,$$

was ein Widerspruch zur Wahl von s als Schranke von $(a_n)_{n\in\mathbb{N}}$ ist. Dies zeigt, daß $(a_n)_{n\in\mathbb{N}}$ nicht beschränkt ist.

Lemma 11.13.
Ist $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge in $\mathbb{K}$ und $(b_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{K}$, so ist $(a_n \cdot b_n)_{n\in\mathbb{N}}$ eine Nullfolge.

Beweis: Da $(b_n)_{n\in\mathbb{N}}$ beschränkt ist, gibt es eine positive reelle Zahl $s \in \mathbb{R}_{>0}$ mit

$$|b_n| \leq s$$

für alle $n \in \mathbb{N}$.

Sei nun $\varepsilon > 0$ gegeben. Wegen $a_n \longrightarrow 0$ gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$|a_n - 0| < \frac{\varepsilon}{s}$$

für $n \geq n_\varepsilon$. Für $n \geq n_\varepsilon$ erhalten wir damit

$$|a_n \cdot b_n - 0| = |a_n \cdot b_n - 0 \cdot b_n| = |a_n - 0| \cdot |b_n| \leq |a_n - 0| \cdot s < \frac{\varepsilon}{s} \cdot s = \varepsilon.$$

Mithin konvergiert $(a_n \cdot b_n)_{n\in\mathbb{N}}$ gegen 0. □

Beispiel 11.14.
Da die Folge $\left(\frac{1}{n}\right)_{n\geq 1}$ eine Nullfolge ist und da zudem die Folge $\left((-1)^n\right)_{n\geq 1}$ beschränkt ist, gilt

$$\frac{(-1)^n}{n} \longrightarrow 0.$$

C) Grenzwertsätze

Proposition 11.15 (Grenzwertsätze).
Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen in $\mathbb{K}$ mit $a_n \longrightarrow a$ und $b_n \longrightarrow b$.

a. $a_n + b_n \longrightarrow a + b$ und $a_n - b_n \longrightarrow a - b$.

b. $a_n \cdot b_n \longrightarrow a \cdot b$.

c. $|a_n| \longrightarrow |a|$.

d. Ist zudem $b \neq 0$, so gibt es ein $n_0 \in \mathbb{N}$ mit $b_n \neq 0$ für alle $n \geq n_0$ und die Folge $\left(\frac{a_n}{b_n}\right)_{n\geq n_0}$ ist konvergent mit

$$\frac{a_n}{b_n} \longrightarrow \frac{a}{b}.$$

Beweis:

a. Sei $\varepsilon > 0$ gegeben. Dann gibt es natürliche Zahlen $n'_\varepsilon, n''_\varepsilon \in \mathbb{N}$, so daß

$$|a_n - a| < \frac{\varepsilon}{2}$$

für $n \geq n'_\varepsilon$ und

$$|b_n - b| < \frac{\varepsilon}{2}$$

für $n \geq n''_\varepsilon$. Mit $n_\varepsilon := \max\{n'_\varepsilon, n''_\varepsilon\}$ gilt dann

$$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

für $n \geq n_\varepsilon$. Mithin konvergiert $(a_n + b_n)_{n\in\mathbb{N}}$ gegen $a + b$. Analog sieht man $a_n - b_n \longrightarrow a - b$.

b. Nach Satz 11.11 ist $(a_n)_{n\in\mathbb{N}}$ als konvergente Folge beschränkt und nach Voraussetzung ist $(b_n - b)_{n\in\mathbb{N}}$ eine Nullfolge, so daß

$$a_n \cdot (b_n - b) \longrightarrow 0$$

nach Lemma 11.13. Analog ist nach Voraussetzung $(a_n - a)_{n\in\mathbb{N}}$ eine Nullfolge und die konstante Folge $(b)_{n\in\mathbb{N}}$ ist als konvergente Folge beschränkt, so daß

$$(a_n - a) \cdot b \longrightarrow 0.$$

Aus a. folgt dann, daß die Summe der beiden Nullfolgen eine Nullfolge ist, d.h.

$$a_n \cdot b_n - a \cdot b = a_n \cdot (b_n - b) + (a_n - a) \cdot b \longrightarrow 0 + 0 = 0.$$

Also gilt auch $a_n \cdot b_n \longrightarrow a \cdot b$.

c. Ist $\varepsilon > 0$ gegeben, so gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$ mit

$$|a_n - a| < \varepsilon$$

für alle $n \geq n_\varepsilon$. Aber dann gilt nach Lemma 10.7 auch

$$\big||a_n| - |a|\big| \leq |a_n - a| < \varepsilon$$

für alle $n \geq n_\varepsilon$. Es folgt die Behauptung.

d. Wegen $b \neq 0$ ist

$$\epsilon := \frac{|b|}{2} > 0$$

und es gibt ein $n_0 \in \mathbb{N}$ mit

$$|b_n - b| < \epsilon = \frac{|b|}{2}$$

für alle $n \geq n_0$. Mithin ist

$$|b| = |b_n + (b - b_n)| \leq |b_n| + |b - b_n| = |b_n| + |b_n - b| < |b_n| + \frac{|b|}{2},$$

so daß $0 < \frac{|b|}{2} \leq |b_n|$ für $n \geq n_0$. Insbesondere ist $b_n \neq 0$ in diesen Fällen.
Aus Lemma 8.14 folgt zudem

$$0 < \frac{1}{|b_n|} \leq \frac{2}{|b|} \tag{15}$$

für $n \geq n_0$.
Ist nun $\varepsilon > 0$ beliebig gegeben, so gibt es eine natürliche Zahl $n_\varepsilon \geq n_0$ mit

$$|b_n - b| < \frac{\varepsilon \cdot |b|^2}{2} \tag{16}$$

für alle $n \geq n_\varepsilon$. Für diese n erhalten wir damit

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b - b_n|}{|b_n \cdot b|} = \frac{1}{|b_n|} \cdot \frac{1}{|b|} \cdot |b_n - b| \overset{(15),(16)}{<} \frac{2}{|b|^2} \cdot \frac{\varepsilon \cdot |b|^2}{2} = \varepsilon.$$

Also gilt

$$\frac{1}{b_n} \longrightarrow \frac{1}{b},$$

und mit Teil b. folgt dann die Behauptung $\frac{a_n}{b_n} \longrightarrow \frac{a}{b}$.

□

Beispiel 11.16.

a. Die Folge $\left(\frac{1}{n^2}\right)_{n\geq 1}$ ist eine Nullfolge, da $\lim\limits_{n\to\infty} \frac{1}{n^2} = \lim\limits_{n\to\infty} \frac{1}{n} \cdot \lim\limits_{n\to\infty} \frac{1}{n} = 0 \cdot 0 = 0$.

b. Die Folge $(a_n)_{n\in\mathbb{N}}$ mit

$$a_n = \frac{7n^2+3}{4n^2+n+1}$$

ist wegen der Grenzwertsätze konvergent, denn es gilt

$$a_n = \frac{7+\frac{3}{n^2}}{4+\frac{1}{n}+\frac{1}{n^2}} \longrightarrow \frac{7+0}{4+0+0} = \frac{7}{4}.$$

D) Konvergenzkriterien für reelle Zahlenfolgen

Proposition 11.17 (Einschachtelungssatz).
Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen in $\mathbb{R}$ mit $a_n \longrightarrow a$ und $b_n \longrightarrow b$, und sei $(c_n)_{n\in\mathbb{N}}$ eine weitere Folge reeller Zahlen.

a. Ist $a_n \leq b_n$ für alle $n \geq n_0$, so ist $a \leq b$.

b. Ist $a_n \leq c_n \leq b_n$ für alle $n \geq n_0$ und ist $a = b$, so gilt $c_n \longrightarrow a$.

Beweis:

a. Nehmen wir $b < a$ an, so gibt es für

$$\varepsilon := \frac{a-b}{2} > 0$$

natürliche Zahlen $n'_\varepsilon, n''_\varepsilon \in \mathbb{N}$ mit

$$a_n \in (a-\varepsilon, a+\varepsilon)$$

für alle $n \geq n'_\varepsilon$ und

$$b_n \in (b-\varepsilon, b+\varepsilon)$$

für alle $n \geq n''_\varepsilon$. Mithin gilt für $n = \max\{n_0, n'_\varepsilon, n''_\varepsilon\}$

$$a-\varepsilon < a_n \leq b_n < b+\varepsilon,$$

so daß

$$a-b < 2\cdot\varepsilon = a-b. \quad \lightning$$

b. Sei $\varepsilon > 0$ gegeben, dann gibt es natürliche Zahlen $n'_\varepsilon, n''_\varepsilon \in \mathbb{N}$ mit

$$|a_n - a| < \varepsilon$$

für alle $n \geq n'_\varepsilon$ und

$$|b_n - a| < \varepsilon$$

für alle $n \geq n''_\varepsilon$. Mithin gilt für $n \geq n_\varepsilon := \max\{n_0, n'_\varepsilon, n''_\varepsilon\}$ sicher

$$a - \varepsilon < a_n \leq c_n \leq b_n < a + \varepsilon,$$

d.h.

$$|c_n - a| < \varepsilon.$$

Also konvergiert $(c_n)_{n\in\mathbb{N}}$ gegen a.

□

Beispiel 11.18.
Wegen $0 \leq \frac{1}{n^k} \leq \frac{1}{n}$ für alle $n \geq 1$ und $k \geq 1$ folgt aus $0 \longrightarrow 0$ und $\frac{1}{n} \longrightarrow 0$ auch

$$\frac{1}{n^k} \longrightarrow 0.$$

Definition 11.19 (Monotone Folgen).
Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen. Wir nennen $(a_n)_{n\in\mathbb{N}}$ *monoton wachsend*, falls

$$a_n \leq a_{n+1}$$

für alle $n \in \mathbb{N}$.
Analog nennen wir $(a_n)_{n\in\mathbb{N}}$ *monoton fallend*, falls

$$a_n \geq a_{n+1}$$

für alle $n \in \mathbb{N}$.

Beispiel 11.20.
Die Folge $(n)_{n\in\mathbb{N}}$ ist monoton wachsend und divergent, die Folge $\left(\frac{1}{n}\right)_{n\geq 1}$ ist monoton fallend und konvergent.

Satz 11.21 (Monotoniekriterium).
Jede monoton wachsende oder fallende, beschränkte Folge in $\mathbb{R}$ ist konvergent.

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ eine monoton wachsende, beschränkte Folge reeller Zahlen und sei $s > 0$ eine Schranke. Dann ist die Menge

$$A := \{a_n \mid n \in \mathbb{N}\}$$

nach oben beschränkt durch s, und somit existiert das Supremum

$$a := \sup(A).$$

Wir wollen zeigen, daß $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert.

Dazu sei $\varepsilon > 0$ gegeben. Dann ist $a - \varepsilon$ keine obere Schranke von a, so daß ein $n_\varepsilon \in \mathbb{N}$ existiert mit

$$a - \varepsilon < a_{n_\varepsilon}.$$

Da die Folge monoton wachsend ist, gilt dann aber für alle $n \geq n_\varepsilon$ auch

$$a - \varepsilon < a_{n_\varepsilon} \leq a_n \leq a < a + \varepsilon,$$

oder anders formuliert

$$|a - a_n| = |a_n - a| < \varepsilon.$$

Mithin haben wir $a_n \longrightarrow a$ gezeigt. Der Fall einer monoton fallenden Folge wird analog mit Hilfe des Infimums bewiesen. □

Bemerkung 11.22 (Supremum und Infimum sind Grenzwerte von Folgen.).
Es sei $\emptyset \neq A \subseteq \mathbb{R}$ eine nicht-leere Menge reeller Zahlen.

a. Ist A nach oben beschränkt, so gibt es eine monoton wachsende Folge $(a_n)_{n\in\mathbb{N}}$ in A, die gegen $\sup(A)$ konvergiert.

b. Ist A nach unten beschränkt, so gibt es eine monoton fallende Folge $(a_n)_{n\in\mathbb{N}}$ in A, die gegen $\inf(A)$ konvergiert.

Beweis: Sei zunächst A nach oben beschränkt. Wir wählen $a_0 \in A$ beliebig und setzen $a := \sup(A)$. Für $n \geq 1$ und $\varepsilon = \frac{1}{n}$ gibt es ein $b_n \in A$ mit $a - \varepsilon < b_n \leq a$. Setzen wir nun $a_n := \max\{b_n, a_{n-1}\} \in A$, so definieren wir auf diese Weise rekursiv eine offenbar monoton steigende Folge in A. Für diese gilt zudem

$$0 \leq |a_n - a| = a - a_n \leq a - b_n \leq \frac{1}{n} \longrightarrow 0,$$

woraus mit dem Einschachtelungssatz folgt, daß $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert.

Ist A nach unten beschränkt, so zeigt man die Aussage analog. □

Beispiel 11.23 (Rekursive Folgen — das Heron-Verfahren).
Es sei $c \in \mathbb{R}_{>0}$ eine positive reelle Zahl. Wir setzen $a_0 := 1$ und für $n \in \mathbb{N}$ definieren wir a_{n+1} durch die Rekursionsvorschrift

$$a_{n+1} := \frac{1}{2} \cdot \left(a_n + \frac{c}{a_n} \right) > 0.$$

Wir wollen zeigen, daß die Folge $(a_n)_{n\in\mathbb{N}}$ gegen $\sqrt{c}$ konvergiert.

1. Schritt: $a_{n+1}^2 \geq c$ für alle $n \in \mathbb{N}$: Für $n \in \mathbb{N}$ gilt

$$0 \leq \left(a_n - \frac{c}{a_n}\right)^2 = a_n^2 - 2c + \frac{c^2}{a_n^2}.$$

Addieren wir auf beiden Seiten $4c$, so erhalten wir

$$0 \leq 4c \leq a_n^2 + 2c + \frac{c^2}{a_n^2} = \left(a_n + \frac{c}{a_n}\right)^2 = 4 \cdot a_{n+1}^2.$$

2. Schritt: $(a_n)_{n\geq 1}$ ist monoton fallend: Aus dem 1. Schritt wissen wir, daß $a_n^2 \geq c$ für $n \geq 1$ und mithin auch

$$a_n \geq \frac{c}{a_n}$$

für $n \geq 1$ gilt. Wir erhalten damit

$$a_{n+1} = \frac{1}{2} \cdot \left(a_n + \frac{c}{a_n}\right) \leq \frac{1}{2} \cdot (a_n + a_n) = a_n$$

für alle $n \geq 1$, so daß die Folge monoton fallend ist.

3. Schritt: $(a_n)_{n\geq 1}$ ist beschränkt: Denn $0 < a_n \leq a_1$ für alle $n \geq 1$.

4. Schritt: $a_n \longrightarrow \sqrt{c}$: Da die Folge $(a_n)_{n\geq 1}$ monoton fallend und beschränkt ist, folgt aus Satz 11.21 dann, daß sie konvergent ist, d.h. es gibt ein $a \in \mathbb{R}$ mit $a_n \longrightarrow a$. Den Grenzwert können wir nun mit Hilfe der Grenzwertsätze und der Eindeutigkeit des Grenzwertes bestimmen; es gilt nämlich

$$a \longleftarrow a_{n+1} = \frac{1}{2} \cdot \left(a_n + \frac{c}{a_n}\right) \longrightarrow \frac{1}{2} \cdot \left(a + \frac{c}{a}\right),$$

d.h.

$$a = \frac{1}{2} \cdot \left(a + \frac{c}{a}\right).$$

Lösen wir die Gleichung nach a auf, so erhalten wir

$$a^2 = c,$$

und da die Folgenglieder nie negativ sind, kann auch der Grenzwert nicht negativ sein (siehe Proposition 11.17). Damit ist also $a = \sqrt{c}$ nach Satz 9.8.

Beachte, daß man die Folge $(a_n)_{n\in\mathbb{N}}$ nutzen kann, um die Wurzel $\sqrt{c}$ näherungsweise zu berechnen — man nennt dieses rekursive Verfahren auch das *Heron-Verfahren*. Versucht dies einmal für $c = 2$ oder $c = 4$.

E) Der Satz von Bolzano-Weierstraß

Definition 11.24 (Teilfolge).
Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und ist zudem

$$n_0 < n_1 < n_2 < n_3 < \dots$$

eine aufsteigende Folge natürlicher Zahlen, so nennen wir die Folge

$$(a_{n_k})_{k\in\mathbb{N}} = (a_{n_0}, a_{n_1}, a_{n_2}, a_{n_3}, \ldots)$$

eine *Teilfolge* von $(a_n)_{n\in\mathbb{N}}$.

Beispiel 11.25.
Die Folge $\left(\frac{1}{n^2}\right)_{n\geq 1}$ ist eine Teilfolge von $\left(\frac{1}{n}\right)_{n\geq 1}$.

Satz 11.26 (Bolzano-Weierstraß).
Jede beschränkte Folge in $\mathbb{K}$ besitzt eine konvergente Teilfolge.

Beweis: Wir unterscheiden im Beweis die Fälle $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$.

1. **Fall:** $\mathbb{K} = \mathbb{R}$**:** Da $(a_n)_{n\in\mathbb{N}}$ beschränkt ist, gibt es eine Zahl $s > 0$, so daß

$$-s \leq a_n \leq s$$

für alle $n \in \mathbb{N}$, d.h. das Intervall

$$[b_0, c_0] := [-s, s]$$

enthält unendlich viele Folgenglieder der Folge $(a_n)_{n\in\mathbb{N}}$ und wir wählen eines davon, a_{n_0}. Teilen wir das Intervall in zwei gleichgroße Hälften $[-s, 0]$ und $[0, s]$, so enthält mindestens eines der beiden neuen Intervalle wieder unendlich viele Folgenglieder. Wir wählen ein solches und nennen es $[b_1, c_1]$. Da es unendlich viele Folgenglieder enthält, enthält es auch ein a_{n_1} mit $n_1 > n_0$. Mit dem Intervall $[b_1, c_1]$ verfahren wir in der gleichen Weise und konstruieren so rekursiv eine Folge von Intervallen

$$[b_0, c_0] \supsetneq [b_1, c_1] \supsetneq [b_2, c_2] \supsetneq [b_3, c_3] \supsetneq \ldots,$$

so daß jedes $[b_j, c_j]$ unendlich viele Folgenglieder von $(a_n)_{n\in\mathbb{N}}$ enthält. Zugleich konstruieren wir dabei eine Teilfolge

$$a_{n_0}, a_{n_1}, a_{n_2}, a_{n_3}, \ldots$$

mit

$$b_j \leq a_{n_j} \leq c_j$$

und $n_0 < n_1 < n_2 < \ldots$.
Aufgrund der Konstruktion ist die Folge $(b_j)_{j\in\mathbb{N}}$ eine monoton wachsende beschränkte Folge und besitzt deshalb einen Grenzwert nach dem Monotoniekriterium 11.21, d.h.

$$b_j \longrightarrow b.$$

Analog besitzt $(c_j)_{j\in\mathbb{N}}$ als monoton fallende beschränkte Folge einen Grenzwert c. Da das Intervall $[b_n, c_n]$ aufgrund seiner Definition die Länge $\frac{2s}{2^n}$ hat, folgt dann

$$c - b \longleftarrow c_n - b_n = \frac{2s}{2^n} = 2s \cdot \left(\frac{1}{2}\right)^n \longrightarrow 0,$$

wobei wir für die Konvergenz der rechten Seite die Eigenschaften der geometrischen Folge berücksichtigen (siehe Lemma 11.6). Wegen der Eindeutigkeit des Grenzwertes einer Folge gilt dann $b = c$, und aus dem Einschachtelungssatz folgt dann auch

$$\lim_{j\to\infty} a_{n_j} = b.$$

2. Fall: $\mathbb{K} = \mathbb{C}$**:** Aus Lemma 10.7 wissen wir, daß

$$|\operatorname{Re}(a_n)| \leq |a_n|,$$

so daß die Folge $\big(\operatorname{Re}(a_n)\big)_{n\in\mathbb{N}}$ ebenfalls beschränkt ist. Da wir den Satz von Bolzano-Weierstraß für $\mathbb{K} = \mathbb{R}$ bereits bewiesen haben, gibt es also eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ und eine reelle Zahl b, so daß

$$\operatorname{Re}(a_{n_k}) \longrightarrow b.$$

Ebenfalls aus Lemma 10.7 folgt

$$|\operatorname{Im}(a_{n_k})| \leq |a_{n_k}|,$$

so daß auch die Folge $\big(\operatorname{Im}(a_{n_k})\big)_{k\in\mathbb{N}}$ beschränkt ist, und wieder folgt mittels des Satzes von Bolzano-Weierstraß für $\mathbb{K} = \mathbb{R}$, daß $(a_{n_k})_{k\in\mathbb{N}}$ eine Teilfolge $(a_{n_{k_j}})_{j\in\mathbb{N}}$ besitzt und daß es eine reelle Zahl c gibt, so daß

$$\operatorname{Im}(a_{n_{k_j}}) \longrightarrow c.$$

Aus Aufgabe 11.43 wissen wir, daß die Teilfolge $\big(\operatorname{Re}(a_{n_{k_j}})\big)_{j\in\mathbb{N}}$ von $\big(\operatorname{Re}(a_{n_k})\big)_{k\in\mathbb{N}}$ ebenfalls gegen b konvergiert, und aus Aufgabe 11.42 ergibt sich dann, daß auch die Folge $(a_{n_{k_j}})_{j\in\mathbb{N}}$ konvergent ist mit

$$a_{n_{k_j}} = \operatorname{Re}(a_{n_{k_j}}) + i \cdot \operatorname{Im}(a_{n_{k_j}}) \longrightarrow b + i \cdot c.$$

□

Beispiel 11.27.
Die divergente beschränkte Folge $\big((-1)^n\big)_{n\in\mathbb{N}}$ besitzt als konvergente Teilfolge die konstante Folge $\big((-1)^{2k}\big)_{k\in\mathbb{N}} = (1)_{k\in\mathbb{N}}$.

Satz 11.28 (Abgeschlossene Intervalle sind abgeschlossen.).
Ist $(a_n)_{n\in\mathbb{N}}$ eine konvergente Folge im abgeschlossenen Intervall $[a,b]$, so gilt

$$\lim_{n\to\infty} a_n \in [a,b].$$

Beweis: Nehmen wir an, $c := \lim_{n\to\infty} a_n \in \mathbb{R}\setminus[a,b]$. Ist $c > b$, so gibt es zu $\varepsilon := \frac{c-b}{2} > 0$ ein n_ε für das unter anderem $0 < c - b \le c - a_{n_\varepsilon} = |a_{n_\varepsilon} - c| < \varepsilon = \frac{c-b}{2}$ gilt, was ein offensichtlicher Widerspruch ist. Analog sieht man auch, daß $c < a$ nicht möglich ist. □

F) Das Cauchy-Kriterium

Definition 11.29 (Cauchy-Folge).
Eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ heißt *Cauchy-Folge*, falls

$$\forall\, 0<\varepsilon\in\mathbb{R}\ \ \exists\, n_\varepsilon\in\mathbb{N}\ :\ \forall\, m>n\ge n_\varepsilon\ :\ |a_m - a_n| < \varepsilon.$$

Satz 11.30 (Cauchy-Kriterium: $\mathbb{K}$ ist vollständig.).
Eine Folge in $\mathbb{K}$ ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.

Beweis:

$\Longrightarrow$**:** Wir setzen voraus, daß $(a_n)_{n\in\mathbb{N}}$ eine konvergente Folge ist mit Grenzwert a. Sei nun $\varepsilon > 0$ gegeben, dann gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$|a_n - a| < \frac{\varepsilon}{2}$$

für alle $n \ge n_\varepsilon$. Für zwei natürliche Zahlen $m > n \ge n_\varepsilon$ folgt dann mit der Dreiecksungleichung

$$|a_m - a_n| \le |a_m - a| + |a - a_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Also ist $(a_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge.

$\Longleftarrow$**:** Sei nun umgekehrt $(a_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge. Wir wollen zeigen, daß $(a_n)_{n\in\mathbb{N}}$ konvergent ist und müssen dazu einen Grenzwert von $(a_n)_{n\in\mathbb{N}}$ finden, was nicht ganz leicht ist. Unsere Idee hierzu ist, daß wir eine konvergente Teilfolge von $(a_n)_{n\in\mathbb{N}}$ mit Hilfe des Satzes von Bolzano-Weierstraß finden und dann zeigen, daß deren Grenzwert auch ein Grenzwert von $(a_n)_{n\in\mathbb{N}}$ ist.

1. Schritt: Zeige, daß $(a_n)_{n\in\mathbb{N}}$ beschränkt ist.[1]

[1] Der Beweis geht wie der Beweis von Satz 11.11, wenn man dort den Grenzwert a durch a_{n_ε} ersetzt.

Zu $\varepsilon = 1$ gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$|a_m - a_n| < \varepsilon = 1$$

für alle $m > n \geq n_\varepsilon$. Setze

$$s := \max\{1 + |a_{n_\varepsilon}|, |a_0|, |a_1|, \ldots, |a_{n_\varepsilon - 1}|\}.$$

Damit erhalten wir dann

$$|a_n| \leq \begin{cases} s, & \text{falls } n < n_\varepsilon, \\ |a_n - a_{n_\varepsilon}| + |a_{n_\varepsilon}| < 1 + |a_{n_\varepsilon}| \leq s, & \text{falls } n \geq n_\varepsilon. \end{cases}$$

Mithin ist s eine Schranke für $(a_n)_{n\in\mathbb{N}}$.

2. Schritt: Aufgrund des Satzes von Bolzano-Weierstraß 11.26 besitzt $(a_n)_{n\in\mathbb{N}}$ also eine konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$, und wir setzen

$$a := \lim_{k\to\infty} a_{n_k}.$$

3. Schritt: Zeige, $a_n \longrightarrow a$.

Sei dazu $\varepsilon > 0$ gegeben. Da $(a_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist, gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$ mit

$$|a_m - a_n| < \frac{\varepsilon}{2}$$

für alle $m > n \geq n_\varepsilon$. Da zudem $a_{n_k} \longrightarrow a$ existiert auch ein $k_\varepsilon \in \mathbb{N}$ mit

$$|a_{n_k} - a| < \frac{\varepsilon}{2}$$

für alle $k \geq k_\varepsilon$. Wir wählen nun eine Zahl $k \geq k_\varepsilon$ so, daß $n_k \geq n_\varepsilon$. Dann gilt für jedes $n \geq n_\varepsilon$ auch

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Also konvergiert $(a_n)_{n\in\mathbb{N}}$ gegen a.

□

Beispiel 11.31.

Ist $1 \neq q \in \mathbb{K}$ mit $|q| = 1$, so ist die Folge $(q^n)_{n\in\mathbb{N}}$ keine Cauchy-Folge und mithin auch nicht konvergent.

Um dies zu sehen, betrachten wir $\varepsilon = |q - 1| > 0$ und $n_\varepsilon \in \mathbb{N}$ beliebig. Für $m = n_\varepsilon + 1$ und $n = n_\varepsilon$ gilt dann

$$|q^m - q^n| = |q|^n \cdot |q - 1| = 1^n \cdot \varepsilon = \varepsilon.$$

Wäre die Folge eine Cauchy-Folge, so müßte der Ausdruck für ein geeignetes n_ε echt kleiner als ε werden.

Bemerkung 11.32 ($\mathbb{Q}$ ist nicht vollständig.).
Eine Cauchy-Folge rationaler Zahlen muß in $\mathbb{Q}$ nicht konvergent sein, d.h. ihr Grenzwert in $\mathbb{R}$ muß keine rationale Zahl sein. Zum Beispiel ist $\sqrt{2}$ keine rationale Zahl (siehe Satz 9.10) und ist $\sqrt{2} = \sum_{i=-1}^{\infty} c_i \cdot 10^{-i}$ ihre Dezimalzahldarstellung, so wird durch

$$a_n = \sum_{i=-1}^{n} c_i \cdot 10^{-i}$$

eine Folge $(a_n)_{n\in\mathbb{N}}$ rationaler Zahlen definiert, die in $\mathbb{R}$ gegen $\sqrt{2}$ konvergiert und mithin eine Cauchy-Folge ist, deren Grenzwert $\sqrt{2}$ aber nicht in $\mathbb{Q}$ liegt.

Man sagt auch, die rationalen Zahlen sind nicht vollständig. Dieses Manko der rationalen Zahlen erfordert den Übergang zu den reellen Zahlen. Mit dem gleichen Argument wie für $\sqrt{2}$ sieht man übrigens, daß jede reelle Zahl Grenzwert einer Folge rationaler Zahlen ist. Dies liegt daran, daß $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt (siehe Satz 9.6).

G) Bestimmt divergente Folgen

Definition 11.33 (Bestimmte Divergenz).
Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$.

a. Wir sagen, daß $(a_n)_{n\in\mathbb{N}}$ *bestimmt divergiert gegen* ∞, falls

$$\forall\, s > 0\ \exists\, n_s \in \mathbb{N}\ :\ \forall\, n \geq n_s\ :\ a_n > s.$$

In diesem Fall schreiben wir $a_n \longrightarrow \infty$ oder $\lim\limits_{n\to\infty} a_n = \infty$, und nennen ∞ auch den uneigentlichen *Grenzwert* von $(a_n)_{n\in\mathbb{N}}$.

b. Analog sagen wir, daß $(a_n)_{n\in\mathbb{N}}$ *bestimmt divergiert gegen* $-\infty$, falls

$$\forall\, s < 0\ \exists\, n_s \in \mathbb{N}\ :\ \forall\, n \geq n_s\ :\ a_n < s.$$

In diesem Fall schreiben wir $a_n \longrightarrow -\infty$ oder $\lim\limits_{n\to\infty} a_n = -\infty$, und nennen $-\infty$ auch den uneigentlichen *Grenzwert* von $(a_n)_{n\in\mathbb{N}}$.

Eine Folge, die bestimmt divergiert nennen wir *bestimmt divergent*.

Beispiel 11.34.
Die Folge $(n)_{n\in\mathbb{N}}$ ist bestimmt divergent mit Grenzwert ∞, die Folge $\big((-1)^n \cdot n\big)_{n\in\mathbb{N}}$ ist divergent, aber nicht bestimmt divergent.

Bemerkung 11.35 (Grenzwertsätze für uneigentliche Grenzwerte).
Wir einigen uns für $a \in \mathbb{R}$ auf die folgenden Rechenregeln:

- $a + \infty := \infty$ und $a - \infty := -\infty$.
- $a \cdot \infty := \infty$ und $a \cdot -\infty := -\infty$, falls $a > 0$.

- $a \cdot \infty := -\infty$ und $a \cdot -\infty := \infty$, falls $a < 0$.
- $\frac{a}{\infty} := 0$ und $\frac{a}{-\infty} := 0$.

Damit lassen sich die Grenzwertsätze für Folgen 11.15 verallgemeinern auf Fälle unter Einbeziehung von bestimmt divergenten Folgen. Wann immer man bei der Anwendung der Grenzwertsätze als Grenzwert einen der obigen Ausdrücke erhält, kann man den Grenzwert auf dem Weg berechnen. Die Beweise sind einfach, aber es gilt viele Fälle zu unterscheiden. Z.B. gelten:

a. Wenn $a_n \longrightarrow a$ und $b_n \longrightarrow \infty$, so gilt $a_n + b_n \longrightarrow a + \infty = \infty$.

b. Wenn $a_n \longrightarrow a$ und $b_n \longrightarrow \infty$, so gilt $\frac{a_n}{b_n} \longrightarrow \frac{a}{\infty} = 0$.

Zudem kann man die Grenzwertsätze auch für Brüche von Folgen formulieren, wenn im Nenner eine Nullfolge steht. Allerdings ist dabei etwas Vorsicht geboten:

a. Wenn $a_n \longrightarrow a \neq 0$ und $b_n \longrightarrow 0$ mit $b_n > 0$ für $n \geq n_0$, so gilt $\frac{a_n}{b_n} \longrightarrow \infty \cdot a$.

b. Wenn $a_n \longrightarrow a \neq 0$ und $b_n \longrightarrow 0$ mit $b_n < 0$ für $n \geq n_0$, so gilt $\frac{a_n}{b_n} \longrightarrow -\infty \cdot a$.

c. Ist das Vorzeichen der b_n nicht ab einer gewissen Stelle fest, so existiert $\lim\limits_{n\to\infty} \frac{a_n}{b_n}$ nicht.

H) Landau-Symbole und das asymptotische Verhalten von Folgen

Ausblick 11.36 (Landau-Symbole in der Informatik).
Bei der *Abschätzung des Aufwands eines Algorithmus'* gibt es oft bestimmte Problemgrößen, die den Aufwand dominieren. Bei festem Wert $n \in \mathbb{N}$ für die Problemgröße, z.B. die Größe der Eingabematrix, ergibt sich dann ein Aufwand $a_n \in \mathbb{R}$, z.B. die Anzahl der benötigten Multiplikationen in Abhängigkeit von n. Die sich so ergebende Folge $(a_n)_{n\in\mathbb{N}}$ ist in aller Regel monoton wachsend und unbeschränkt, also bestimmt divergent. Um die Komplexität des Verfahrens zu verstehen, muß man verstehen, wie schnell die Folge gegen ∞ divergiert. Dazu vergleicht man sie mit gewissen Prototypen bestimmt divergenter Folgen und untersucht, ob sie asymptotisch dasselbe Divergenzverhalten aufweisen (siehe auch Bemerkung 11.41). Dies wird mit Hilfe der Landausymbole zum Ausdruck gebracht.

Definition 11.37 (Landau-Symbole $\mathcal{O}$ und o).
Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ Folgen in $\mathbb{R}$ und es gelte $b_n > 0$ für alle $n \in \mathbb{N}$.

a. Wir nennen $(b_n)_{n\in\mathbb{N}}$ eine *asymptotische obere Schranke* für $(a_n)_{n\in\mathbb{N}}$, wenn die Folge $\left(\frac{a_n}{b_n}\right)_{n\in\mathbb{N}}$ beschränkt ist. Wir schreiben dann kurz $a_n \in \mathcal{O}(b_n)$ und $\mathcal{O}(a_n) \leq \mathcal{O}(b_n)$.

b. Wir nennen $(b_n)_{n\in\mathbb{N}}$ eine *scharfe asymptotische obere Schranke* für $(a_n)_{n\in\mathbb{N}}$, wenn $a_n \in \mathcal{O}(b_n)$ und $b_n \in \mathcal{O}(a_n)$ gilt. Wir schreiben dann kurz $a_n \in \Theta(b_n)$ oder auch $\mathcal{O}(a_n) = \mathcal{O}(b_n)$.

c. Wir nennen $(a_n)_{n\in\mathbb{N}}$ *asymptotisch vernachlässigbar gegenüber* $(b_n)_{n\in\mathbb{N}}$, wenn die Folge $\left(\frac{a_n}{b_n}\right)_{n\in\mathbb{N}}$ eine Nullfolge ist. Wir schreiben dann kurz $a_n \in o(b_n)$.

Dabei werden o, $\mathcal{O}$ und Θ auch Landau-Symbole genannt. Natürlich kann die Indizierung der Folgen auch wieder mit einer anderen ganzen Zahl als 0 beginnen.

Bemerkung 11.38.
Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ Folgen in $\mathbb{R}$ und es gelte $b_n > 0$ für alle $n \in \mathbb{N}$.

a. Genau dann gilt $a_n \in \mathcal{O}(b_n)$, wenn

$$\exists\, S > 0 \;:\; \forall\, n \in \mathbb{N} \;:\; |a_n| \leq S \cdot b_n.$$

b. Genau dann gilt $a_n \in \Theta(b_n)$ oder $\mathcal{O}(a_n) = \mathcal{O}(b_n)$, wenn

$$\exists\, s, S > 0 \;:\; \forall\, n \in \mathbb{N} \;:\; s \cdot b_n \leq a_n \leq S \cdot b_n.$$

c. Genau dann gilt $a_n \in o(b_n)$, wenn

$$\forall\, \varepsilon > 0\; \exists\, n_\varepsilon \;:\; \forall\, n \geq n_\varepsilon \;:\; |a_n| < \varepsilon \cdot b_n.$$

d. Aus $a_n \in o(b_n)$ folgt offenbar $a_n \in \mathcal{O}(b_n)$, da jede Nullfolge beschränkt ist.

Beispiel 11.39.
Die Folge $(a_n)_{n\geq 1}$ mit

$$a_n = n^2 + 5n + 2$$

erfüllt $a_n \in \mathcal{O}(n^2)$, da

$$\frac{a_n}{n^2} = 1 + \frac{5}{n} + \frac{2}{n^2} \longrightarrow 1 + 0 + 0 = 1$$

konvergent und damit beschränkt ist. Wegen $n^2 \leq a_n$ für alle n gilt zudem $n^2 \in \mathcal{O}(a_n)$ und somit $a_n \in \Theta(n^2)$ oder $\mathcal{O}(a_n) = \mathcal{O}(n^2)$.

Da die Folge $\left(\frac{a_n}{n^2}\right)_{n\geq 1}$ keine Nullfolge ist, gilt aber $a_n \notin o(n^2)$. Hingegen folgt aus

$$\frac{a_n}{n^3} = \frac{1}{n} + \frac{5}{n^2} + \frac{2}{n^3} \longrightarrow 0 + 0 + 0 = 0$$

dann $a_n \in o(n^3)$.

Proposition 11.40 (Rechenregeln für Landau-Symbole).
Seien $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, $(c_n)_{n\in\mathbb{N}}$ und $(d_n)_{n\in\mathbb{N}}$ strikt positive Folgen in $\mathbb{R}$. Dann gelten:

a. $a_n \in \mathcal{O}(a_n)$.

b. Wenn $a_n \in \mathcal{O}(b_n)$, dann gilt auch $c \cdot a_n \in \mathcal{O}(b_n)$ für alle $c \in \mathbb{R}$.

c. Wenn $a_n \in o(b_n)$, dann gilt auch $c \cdot a_n \in o(b_n)$ für alle $c \in \mathbb{R}$.

d. Wenn $a_n \in \mathcal{O}(c_n)$ und $b_n \in \mathcal{O}(c_n)$, dann gilt $a_n + b_n \in \mathcal{O}(c_n)$.

e. Wenn $a_n \in o(c_n)$ und $b_n \in o(c_n)$, dann gilt $a_n + b_n \in o(c_n)$.

f. Wenn $a_n \in \mathcal{O}(b_n)$ und $b_n \in \mathcal{O}(c_n)$, dann gilt $a_n \in \mathcal{O}(c_n)$.

g. Wenn $a_n \in o(b_n)$ und $b_n \in o(c_n)$, dann gilt $a_n \in o(c_n)$.

h. Wenn $a_n \in \mathcal{O}(c_n)$ und $b_n \in \mathcal{O}(d_n)$, dann gilt $a_n \cdot b_n \in \mathcal{O}(c_n \cdot d_n)$.

i. Wenn $a_n \in o(c_n)$ und $b_n \in o(d_n)$, dann gilt $a_n \cdot b_n \in o(c_n \cdot d_n)$.

Beweis: Die Aussagen folgen unmittelbar aus den Grenwertsätzen für Folgen 11.15.

a. Die Folge der $\frac{a_n}{a_n} = 1$ ist konstant und somit beschränkt.

b. Wenn die Folge der $\frac{a_n}{b_n}$ beschränkt ist, dann ist das auch die Folge der $\frac{c\cdot a_n}{b_n}$.

c. Wenn die Folge der $\frac{a_n}{b_n}$ eine Nullfolge ist, dann ist das auch die Folge der $\frac{c\cdot a_n}{b_n}$.

d. Wenn die Folgen der $\frac{a_n}{c_n}$ und der $\frac{b_n}{c_n}$ beschränkt sind, dann ist auch die Folge der $\frac{a_n+b_n}{c_n} = \frac{a_n}{c_n} + \frac{b_n}{c_n}$ beschränkt.

e. Wenn die Folgen der $\frac{a_n}{c_n}$ und der $\frac{b_n}{c_n}$ Nullfolgen sind, dann ist auch die Folge der $\frac{a_n+b_n}{c_n} = \frac{a_n}{c_n} + \frac{b_n}{c_n}$ eine Nullfolge.

f. Wenn die Folgen der $\frac{a_n}{c_n}$ und der $\frac{b_n}{d_n}$ beschränkt sind, dann ist auch die Folge der $\frac{a_n\cdot b_n}{c_n\cdot d_n} = \frac{a_n}{c_n} \cdot \frac{b_n}{d_n}$ beschränkt.

g. Wenn die Folgen der $\frac{a_n}{c_n}$ und der $\frac{b_n}{d_n}$ Nullfolgen sind, dann ist auch die Folge der $\frac{a_n\cdot b_n}{c_n\cdot d_n} = \frac{a_n}{c_n} \cdot \frac{b_n}{d_n}$ eine Nullfolge.

□

Bemerkung 11.41 (Prototypen für asymptotisches Verhalten).
Die folgende Tabelle enthält die am häufigsten verwendeten Pototypen für den Vergleich von asymptotischem Laufzeitverhalten bei Algorithmen:

$\mathcal{O}(a_n)$	Laufzeitverhalten
$\mathcal{O}(1)$	konstant
$\mathcal{O}(\log(n))$	logarithmisch
$\mathcal{O}(\log^c(n))$	poly-logarithmisch
$\mathcal{O}(n)$	linear
$\mathcal{O}(n\log(n))$	"$n \log n$"
$\mathcal{O}(n^2)$	quadratisch
$\mathcal{O}(n^3)$	kubisch
$\mathcal{O}(n^k)$	polynomial
$\mathcal{O}(2^n)$	exponentiell
$\mathcal{O}(n!)$	faktoriell

Beim Logarithmus haben wir keine Basis angegeben, da aus der Formel (siehe auch Korollar 16.9)

$$\log_a(n) = \frac{\log_2(n)}{\log_2(a)}$$

folgt, daß

$$\mathcal{O}(\log_a(n)) = \mathcal{O}(\log_2(n))$$

für jede Basis $a > 0$ gilt und die Basis somit keine Rolle spielt. In der Praxis wird meist die Basis $a = 2$ verwendet.

Die Folgen sind in obiger Tabelle in der Reihenfolge aufsteigender Komplexität angegeben. Wir geben für einige der Folgen und einige Werte von n in folgender Tabelle Näherungswerte für die Folgenglieder an, um dies zu verdeutlichen:

n	$\log(n)$	$n\log(n)$	n^2	n^3	2^n
10	$\sim 2,3$	~ 23	100	1000	1024
100	$\sim 4,6$	~ 461	10^4	10^6	$\sim 1,27 \cdot 10^{30}$
1000	$\sim 6,9$	~ 6908	10^6	10^9	$\sim 10^{301}$
10000	$\sim 9,2$	~ 92103	10^8	10^{12}	$\sim 10^{3010}$

Um zu sehen, wie gravierend sich dies auf die Laufzeit eines Algorithmus' niederschlägt, nehmen wir vereinfachend an, daß die Problemgröße n bei der Berechnung mit a_n Sekunden zu Buche schlägt. Wenn die Problemgröße nun den Wert $n = 100$ besitzt, so wird ein Algorithmus mit logarithmischer Laufzeit etwa 5 Sekunden benötigen, bei kubischer Laufzeit fast 12 Tage und bei exponentieller Laufzeit mehr als eine Billion mal das Alter des Universums:

1 Stunde	$=$	$3,6 \cdot 10^3$ Sekunden
1 Tag	$\sim$	$8,6 \cdot 10^4$ Sekunden
1 Jahr	$\sim$	$3,2 \cdot 10^7$ Sekunden
100 Jahre	$\sim$	$3,2 \cdot 10^9$ Sekunden
Alter des Universums	$\sim$	$4,3 \cdot 10^{17}$ Sekunden

Aufgaben

Aufgabe 11.42.
Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge komplexer Zahlen. Zeige, daß die folgenden Aussagen äquivalent sind:

a. $a_n \longrightarrow a$.

b. $\mathrm{Re}(a_n) \longrightarrow \mathrm{Re}(a)$ und $\mathrm{Im}(a_n) \longrightarrow \mathrm{Im}(a)$.

Aufgabe 11.43.
Ist $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ bijektiv, so nennen wir die Folge

$$(a_{\sigma(n)})_{n\in\mathbb{N}} = (a_{\sigma(0)}, a_{\sigma(1)}, a_{\sigma(2)}, a_{\sigma(3)}, \ldots)$$

eine *Umordnung* von $(a_n)_{n\in\mathbb{N}}$. Beweise die folgenden beiden Aussagen.

a. Wenn $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert, so konvergiert jede Teilfolge von $(a_n)_{n\in\mathbb{N}}$ gegen a.

b. Wenn $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert, so konvergiert jede Umordnung von $(a_n)_{n\in\mathbb{N}}$ gegen a.

Aufgabe 11.44.

a. Zeige, daß die Folge $(s_n)_{n\in\mathbb{N}}$ mit

$$s_n := \sum_{k=0}^{n} \frac{1}{k!}$$

konvergent ist.

b. Zeige, daß die Folge $(t_n)_{n\in\mathbb{N}}$ mit

$$t_n := \left(1 + \frac{1}{n}\right)^n$$

konvergent ist.

c. Zeige, daß die Grenzwerte von $(s_n)_{n\in\mathbb{N}}$ und $(t_n)_{n\in\mathbb{N}}$ übereinstimmen. Wir nennen den Grenzwert die *Eulersche Zahl* e.

Hinweis zu Teil c., zeige hierfür, daß der Grenzwert von $(t_n)_{n\in\mathbb{N}}$ nach unten durch s_m beschränkt ist.

Aufgabe 11.45.
Untersuche die folgenden Folgen $(a_n)_{n\geq 1}$ auf Konvergenz und berechne gegebenenfalls den Grenzwert:

a. $a_n = \frac{n^4-3n+5}{3n^5+6n^3+11}$.

b. $a_n = \frac{(3n+1)\cdot(n+1)^2-5(n-1)}{1+n+n^2}$.

c. $a_n = \sum_{i=1}^{n} \frac{i}{n^2}$.

d. $a_n = \frac{n}{n^2+1} + \ldots + \frac{n}{n^2+n}$.

e. $a_n = \frac{n}{2^n}$.

Aufgabe 11.46.

Zeige, dass die rekursiv definierte Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_0 = 1$ und

$$a_{n+1} := \sqrt{1 + a_n}$$

konvergiert und bestimme ihren Grenzwert.

Hinweis: Prüfe die Folge (bzw. eine geeignete Teilfolge) auf Monotonie und Beschränktheit. Für die Berechnung des Grenzwertes können dann die Grenzwertsätze geeignet angewandt werden.

Aufgabe 11.47.

Zeige, dass die rekursiv definierte Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_0 = 1$ und

$$a_{n+1} := 1 + \frac{1}{a_n}$$

konvergiert und bestimme ihren Grenzwert.

Hinweis: Prüfe die Folge (bzw. eine geeignete Teilfolge) auf Monotonie und Beschränktheit. Für die Berechnung des Grenzwertes können dann die Grenzwertsätze geeignet angewandt werden.

Aufgabe 11.48.

a. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen, so daß $|a_{n+1} - a_n| < \frac{1}{2^n}$ für alle $n \in \mathbb{N}$. Zeige, daß es sich um eine konvergente Folge handelt.

b. Bleibt die Behauptung aus Aufgabenteil a. korrekt, wenn wir die Bedingung $|a_{n+1} - a_n| < \frac{1}{n}$ voraussetzen? Gib einen Beweis oder ein Gegenbeispiel an.

Aufgabe 11.49.

Bestimme für der in der folgenden Tabelle angegebenen Nullfolgen $(a_n)_{n\geq 1}$ und für jedes der angegebenen $\varepsilon > 0$ eine natürliche Zahl n_ε, so daß $|a_n - 0| < \varepsilon$ für alle $n \geq n_\varepsilon$:

	$\varepsilon = \frac{1}{4}$	$\varepsilon = \frac{1}{16}$	$\varepsilon = \frac{1}{32}$
$a_n = \frac{1}{n^2}$			
$a_n = \frac{n^2}{2^n}$			
$a_n = \frac{n+1}{n} - 1$			

Aufgabe 11.50.
Zeige mit Hilfe der Definition des Grenzwertes, daß $\lim\limits_{n\to\infty} \frac{n^2+1}{4n^2} = \frac{1}{4}$.

Aufgabe 11.51.
Gib zwei divergente Folgen an, deren Produkt konvergiert. Begründe Deine Antwort.

Aufgabe 11.52.
Untersuche die fogenden Folgen $(a_n)_{n\geq 1}$, ob sie konvergent oder divergent sind und bestimme ggf. ihren Grenzwert:

a. $a_n = \frac{n+(-1)^n}{n+1}$
b. $a_n = \frac{(n-1)^3}{n^3+1}$
c. $a_n = \frac{3^{n+1}+2^{n+1}}{3^n+2^n}$
d. $a_n = (-1)^n \cdot \frac{1-n}{2+n}$
e. $a_n = \sqrt{n^2+3n} - \sqrt{n^2-n}$
f. $a_n = \frac{n^3-2n+1}{n^2+1}$
g. $a_n = \frac{1+\sqrt{n-1}}{n+1}$

Aufgabe 11.53.
Welche der folgenden Folgen $(a_n)_{n\geq 1}$ sind monoton, beschränkt, konvergent oder bestimmt divergent? Es ist keine Begründung erforderlich, es darf aber eine Begründung gegeben werden.

a. $a_n = 5$
b. $a_n = \frac{1}{(-1)^n}$
c. $a_n = \frac{(-1)^n}{n^2}$
d. $a_n = \frac{3}{n}$
e. $a_n = \frac{n}{n+1}$
f. $a_n = \frac{n+1}{n}$
g. $a_n = -n$
h. $a_n = n^2 - n$
i. $a_n = 2^n$
j. $a_n = (-2)^n$

k. $a_1 = 1$ und $a_{n+1} = a_n + \frac{1}{n}$ für $n \geq 1$.

l. $a_1 = 1$ und $a_{n+1} = \frac{n}{n+1} \cdot a_n$ für $n \geq 1$.

Aufgabe 11.54.

Die Folge $(a_n)_{n\geq 1}$ sei rekursiv definiert durch $a_1 = 5$ und

$$a_{n+1} = \frac{a_n}{2} + \frac{2}{a_n}.$$

a. Zeige mit Induktion nach n, daß $a_n \geq 2$ für alle $n \geq 1$ gilt.

b. Zeige, die Folge ist monoton fallend.

c. Zeige, die Folge ist konvergent und bestimme den Grenzwert.

Aufgabe 11.55.

Bei Algorithmen wendet man oft das Verfahren *divide and conquer* an. Im einfachsten Fall *teilt* man dabei ein Problem mit Kenngröße $n \geq 1$ in zwei gleich große Teilprobleme der Größe $n-1$ auf. Diese löst man dann rekursiv durch weiteres Aufteilen und setzt die Teillösungen anschließend zusammen.
Nehmen wir an, daß für die Lösung des Problems für $n = 1$ eine Zeiteinheit benötigt wird und daß für das Zerlegen des Problems der Größe n in zwei Teilprobleme der Größe $n-1$ sowie für das Zusammensetzen der Lösungen der zwei Teilprobleme zur Lösung des Problems der Größe n jeweils wieder eine Zeiteinheit benötigt wird. Zudem beschreibe a_n den Aufwand in Zeiteinheiten, zur Lösung des Problems der Größe n.

a. Leite eine Rekursionsvorschrift für die Folge a_n her, d.h. beschreibe a_{n+1} in Abhängigkeit von a_n.

b. Bestimme eine explizite Formel für a_n.

c. Zeige, $a_n \in \mathcal{O}(2^n)$.

Aufgabe 11.56.

Ordne die folgenden Folgen dergestalt, daß für zwei aufeinander folgende Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ stets $a_n \in o(b_n)$ gilt und begründe die Korrektheit der Anordnung:

$$(7n^3)_{n\in\mathbb{N}}, \quad \left(\sqrt{n}\right)_{n\in\mathbb{N}}, \quad (5^n)_{n\in\mathbb{N}}, \quad \left(10^{4^{10}}\right)_{n\in\mathbb{N}}, \quad \left(\left(\frac{1}{10}\right)^{n^2}\right)_{n\in\mathbb{N}}, \quad (10n+3)_{n\in\mathbb{N}}.$$

§ 12 Unendliche Reihen

Ausblick 12.0.
Unendliche Reihen spielen in der Informatik vor allem im Zusammenhang mit der *Zahldarstellung* eine wichtige Rolle, wie wir in Abschnitt 12.H) sehen werden.

A) Konvergenz unendlicher Reihen

Definition 12.1 (Unendliche Reihen).
Ist $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$, so nennen wir die Folge $(s_n)_{n\in\mathbb{N}}$ mit

$$s_n := \sum_{k=0}^{n} a_k$$

der *Partialsummen* von $(a_n)_{n\in\mathbb{N}}$ auch die durch $(a_n)_{n\in\mathbb{N}}$ definierte *Reihe*.
Die Reihe heißt *konvergent*, wenn $(s_n)_{n\in\mathbb{N}}$ eine konvergente Folge ist, und andernfalls heißt sie *divergent*.
Wir bezeichnen sowohl die Reihe $(s_n)_{n\in\mathbb{N}}$ selbst, als auch ihren Grenzwert, sofern er existiert, mit

$$\sum_{n=0}^{\infty} a_n.$$

Beachte, wie stets bei Folgen müssen weder $(a_n)_{n\geq n_0}$ noch $(s_m)_{m\geq n_0} = \sum_{n=n_0}^{\infty} a_n$ (wobei $s_m = \sum_{k=n_0}^{m} a_k$) mit dem Index 0 starten!

Beispiel 12.2 (Teleskopsumme).
Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n\cdot(n+1)}$ ist konvergent mit Grenzwert $\sum_{n=1}^{\infty} \frac{1}{n\cdot(n+1)} = 1$.
Dazu beachten wir, daß $\frac{1}{k\cdot(k+1)} = \frac{1}{k} - \frac{1}{k+1}$ gilt, so daß

$$\begin{aligned} s_n &= \sum_{k=1}^{n} \frac{1}{k\cdot(k+1)} = \sum_{k=1}^{n} \frac{1}{k} - \frac{1}{k+1} \\ &= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \ldots + \left(\frac{1}{n-1} - \frac{1}{n}\right) + \left(\frac{1}{n} - \frac{1}{n+1}\right) \\ &= 1 - \frac{1}{n+1} \longrightarrow 1. \end{aligned}$$

Summen, die sich wie s_n auf zwei Summanden reduzieren, weil sich die übrigen Teile der Summe sukzessive auslöschen, nennt man *Teleskopsummen*.

Beispiel 12.3 (Harmonische Reihe).

Die *harmonische Reihe* $\sum_{n=1}^{\infty} \frac{1}{n}$ ist divergent.

Denn für $n_k = 2^k$ mit $k \in \mathbb{N}$ gilt

$$\begin{aligned} s_{n_k} &= \sum_{i=1}^{2^k} \frac{1}{i} \\ &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \ldots + \left(\frac{1}{2^{k-1}+1} + \ldots + \frac{1}{2^k}\right) \\ &\geq 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \ldots + \left(\frac{1}{2^k} + \ldots + \frac{1}{2^k}\right) \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \ldots + \frac{1}{2} \\ &= 1 + \frac{k}{2} \longrightarrow \infty, \end{aligned}$$

so daß $(s_{n_k})_{k\in\mathbb{N}}$ eine divergente Teilfolge der Folge der Partialsummen ist, weshalb letztere nicht konvergent sein kann.

Lemma 12.4 (Grenzwertsätze für konvergente Reihen).

Seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ zwei konvergente Reihen in $\mathbb{K}$ und $a \in \mathbb{K}$.

a. $\sum_{n=0}^{\infty} (a_n + b_n) = \sum_{n=0}^{\infty} a_n + \sum_{n=0}^{\infty} b_n.$

b. $\sum_{n=0}^{\infty} (a_n - b_n) = \sum_{n=0}^{\infty} a_n - \sum_{n=0}^{\infty} b_n.$

c. $\sum_{n=0}^{\infty} a \cdot a_n = a \cdot \sum_{n=0}^{\infty} a_n.$

d. $\mathbb{K} = \mathbb{R}$ und $a_n \leq b_n$ für alle $n \in \mathbb{N}$, so ist $\sum_{n=0}^{\infty} a_n \leq \sum_{n=0}^{\infty} b_n$.

Insbesondere, sind die Reihen in a.-c. konvergent.

Beweis: Die Aussagen folgen unmittelbar aus den Grenzwertsätzen für Folgen 11.15 sowie aus Proposition 11.17 angewendet auf die Folgen der Partialsummen. □

B) Konvergenzkriterien für unendliche Reihen

Proposition 12.5 (Cauchy-Kriterium für Reihen).
Sei $\sum_{n=0}^{\infty} a_n$ eine Reihe in $\mathbb{K}$. Genau dann ist $\sum_{n=0}^{\infty} a_n$ konvergent, wenn

$$\forall\, \varepsilon > 0\ \exists\, n_\varepsilon \in \mathbb{N}\ :\ \forall\, m > n \geq n_\varepsilon\ :\ \left| \sum_{k=n+1}^{m} a_k \right| < \varepsilon.$$

Beweis: Die Aussage folgt unmittelbar aus dem Cauchy-Kriterium für Folgen 11.30, da

$$s_m - s_n = \sum_{k=n+1}^{m} a_k.$$

□

Lemma 12.6 (Restglieder einer konvergenten Reihe).
Ist die Reihe $\sum_{k=0}^{\infty} a_k$ konvergent, so ist die Folge der Restglieder eine Nullfolge, d.h.

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} a_k = 0.$$

Beweis: Zu $\varepsilon > 0$ gibt es wegen des Cauchy-Kriteriums für Reihen ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $m > n \geq n_\varepsilon$

$$\left| \sum_{k=n+1}^{m} a_k \right| < \varepsilon.$$

Halten wir n fest und betrachten die linke Seite als eine Folge mit Index m, so erhalten wir aus dem Einschachtelungssatz 11.17

$$\left| \sum_{k=n+1}^{\infty} a_k \right| = \lim_{m\to\infty} \left| \sum_{k=n+1}^{m} a_k \right| \leq \varepsilon$$

für alle $n \geq n_\varepsilon$. Also ist die Folge der Restglieder eine Nullfolge. □

Lemma 12.7 (Nullfolgekriterium).
Ist $\sum_{n=0}^{\infty} a_n$ eine konvergente Reihe in $\mathbb{K}$, so ist $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge.

Beweis: Man beachte, daß die Partialsummen s_n der Reihe folgende Eigenschaft erfüllen:

$$a_n = s_n - s_{n-1}.$$

Aus den Grenzwertsätzen für Folgen 11.15 folgt deshalb, daß $(a_n)_{n\in\mathbb{N}}$ als Differenz zweier konvergenter Folgen konvergent ist und daß

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} = \sum_{n=0}^{\infty} a_n - \sum_{n=0}^{\infty} a_n = 0.$$

□

Beispiel 12.8.

a. Die Reihe $\sum_{n=0}^{\infty} \left(1+\frac{1}{n}\right)^n$ ist divergent, da $\left(\left(1+\frac{1}{n}\right)^n\right)_{n\in\mathbb{N}}$ keine Nullfolge ist.

b. Die Umkehrung von Lemma 12.7 gilt nicht, wie das Beispiel der harmonischen Reihe zeigt.

Satz 12.9 (Geometrische Reihe).
Es sei $q \in \mathbb{K}$.

a. Ist $|q| < 1$, so ist die *geometrische Reihe* $\sum_{n=0}^{\infty} q^n$ konvergent mit Grenzwert

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}.$$

b. Ist $|q| \geq 1$, so ist die *geometrische Reihe* $\sum_{n=0}^{\infty} q^n$ divergent.

Beweis:

a. Aus Satz 7.12 wissen wir

$$s_n = \sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q},$$

und da $|q| < 1$ gilt dann wegen Lemma 11.6

$$s_n = \frac{1-q^{n+1}}{1-q} \longrightarrow \frac{1-0}{1-q} = \frac{1}{1-q}.$$

b. Für $|q| \geq 1$ ist die Folge $(q^n)_{n\in\mathbb{N}}$ keine Nullfolge (siehe Beispiele 11.12 und 11.31), und somit ist die Reihe $\sum_{n=0}^{\infty} q^n$ aufgrund des Nullfolgenkriteriums 12.7 divergent.

□

Satz 12.10 (Leibniz-Kriterium für alternierende Reihen in $\mathbb{R}$).
Ist $(a_n)_{n\in\mathbb{N}}$ eine monoton fallende Nullfolge in $\mathbb{R}$, so konvergiert die Reihe

$$\sum_{n=0}^{\infty} (-1)^n \cdot a_n.$$

Beweis: Es sei wieder $s_n = \sum_{k=0}^{n}(-1)^k \cdot a_k$ die n-te Partialsumme der Reihe. Wir betrachten nun zunächst die Teilfolge $(s_{2n})_{n\in\mathbb{N}}$ der *geraden* Partialsummen. Für $n \in \mathbb{N}$ gilt dann

$$s_{2\cdot(n+1)} = s_{2n} - (a_{2n+1} - a_{2n+2}) \leq s_{2n},$$

da nach Voraussetzung $a_{2n+1} \geq a_{2n+2}$. Die Folge $(s_{2n})_{n\in\mathbb{N}}$ ist also monoton fallend.

Analog sieht man, daß die Folge $(s_{2n+1})_{n\in\mathbb{N}}$ der *ungeraden* Partialsummen monoton steigend ist, denn

$$s_{2\cdot(n+1)+1} = s_{2n+1} + (a_{2n+2} - a_{2n+3}) \geq s_{2n+1}.$$

Damit sind beide Folgen dann aber auch beschränkt, denn

$$s_1 \leq s_{2n+1} = s_{2n} - a_{2n+1} \leq s_{2n} \leq s_0$$

für $n \in \mathbb{N}$. Aufgrund des Monotoniekriteriums 11.21 sind also beide Folgen konvergent, d.h. es gibt reelle Zahlen $s, t \in \mathbb{R}$ mit

$$s_{2n} \longrightarrow s \quad \text{und} \quad s_{2n+1} \longrightarrow t.$$

Aus den Grenzwertsätzen für Folgen erhalten wir dann

$$s - t \longleftarrow s_{2n} - s_{2n+1} = a_{2n+1} \longrightarrow 0,$$

so daß $s = t$ gilt.

Sei nun $\varepsilon > 0$ gegeben. Dann gibt es natürliche Zahlen $n'_\varepsilon, n''_\varepsilon \in \mathbb{N}$, so daß

$$|s_{2n} - s| < \varepsilon$$

für alle $n \geq n'_\varepsilon$ und

$$|s_{2n+1} - s| < \varepsilon$$

für alle $n \geq n''_\varepsilon$. Setzen wir nun $n_\varepsilon = \max\{2 \cdot n'_\varepsilon, 2 \cdot n''_\varepsilon + 1\}$, so gilt für $n \geq n_\varepsilon$ offenbar

$$|s_n - s| < \varepsilon.$$

Also ist die Folge $(s_n)_{n\in\mathbb{N}}$ konvergent. □

Beispiel 12.11 (Alternierende harmonische Reihe).
Die alternierende harmonische Reihe $\sum\limits_{n=1}^{\infty} \frac{(-1)^n}{n}$ ist konvergent. Aus dem Beweis des Leibnizkriteriums wissen wir zudem, daß

$$-1 = s_1 \leq \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \leq s_2 = -\frac{1}{2}.$$

Mehr können wir im Augenblick über den Grenzwert der Reihe nicht sagen (siehe dazu Beispiel 18.37).

Lemma 12.12 (Umklammern in Reihen).
Es sei $\sum_{n=0}^{\infty} a_n$ eine konvergente Reihe in $\mathbb{K}$ und $0 = k_0 < k_1 < k_2 < \ldots$ eine aufsteigende Folge natürlicher Zahlen. Setzen wir

$$b_n := \sum_{k=k_n}^{k_{n+1}-1} a_k = a_{k_n} + a_{k_n+1} + \ldots + a_{k_{n+1}-1}$$

so ist die Reihe $\sum_{n=0}^{\infty} b_n$ konvergent und es gilt

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} b_n.$$

Beweis: Ist $(s_n)_{n\in\mathbb{N}}$ die Folge der Partialsummen zu $\sum_{n=0}^{\infty} a_n$ und $(t_n)_{n\in\mathbb{N}}$ die Folge der Partialsummen zu $\sum_{n=0}^{\infty} b_n$, so gilt

$$t_n = s_{k_{n+1}-1}$$

für $n \in \mathbb{N}$. Also ist $(t_n)_{n\in\mathbb{N}}$ eine Teilfolge von $(s_n)_{n\in\mathbb{N}}$ und konvergiert wegen Aufgabe 11.43 gegen den gleichen Grenzwert. □

C) Absolut konvergente Reihen

Definition 12.13 (Umordnung).
Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ bijektiv. Wir nennen die Folge

$$(a_{\sigma(n)})_{n\in\mathbb{N}} = (a_{\sigma(0)}, a_{\sigma(1)}, a_{\sigma(2)}, a_{\sigma(3)}, \ldots)$$

eine *Umordnung* von $(a_n)_{n\in\mathbb{N}}$ und die Reihe

$$\sum_{n=0}^{\infty} a_{\sigma(n)} = a_{\sigma(0)} + a_{\sigma(1)} + a_{\sigma(2)} + a_{\sigma(3)} + \ldots$$

eine *Umordnung* der Reihe $\sum_{n=0}^{\infty} a_n$.

Beispiel 12.14.
Betrachten wir folgende Umordnung der alternierenden harmonischen Reihe

$$\left(-\frac{1}{1} + \frac{1}{2}\right) + \frac{1}{4} + \left(-\frac{1}{3} + \frac{1}{6}\right) + \frac{1}{8} + \left(-\frac{1}{5} + \frac{1}{10}\right) + \frac{1}{12} + \left(-\frac{1}{7} + \frac{1}{14}\right) + \frac{1}{16} + \ldots,$$

d.h. in den Klammern sind sukzessive die ungeraden Folgenglieder a_n zusammen jeweils mit dem zugehörigen Folgenglied a_{2n} aufgeführt, und zwischen den Klammerausdrücken stehen der Reihe nach die Folgenglieder, deren Index durch 4 teilbar ist. Es ist klar, daß man auf dem Weg alle Glieder der harmonischen Reihe auflistet. Wenn diese Umordnung der harmonischen Reihe wieder konvergent ist, so können wir wegen Lemma 12.12 zur

Berechnung des Grenzwertes auch die Klammern wie angegeben setzen. Der Grenzwert der Reihe ist dann

$$-\frac{1}{2}+\frac{1}{4}-\frac{1}{6}+\frac{1}{8}-\frac{1}{10}+\frac{1}{12}-\frac{1}{14}+\frac{1}{16}-\ldots=\frac{1}{2}\cdot\sum_{n=1}^{\infty}\frac{(-1)^n}{n}$$

genau die Hälfte des Grenzwertes der harmonischen Reihe. Daraus ergibt sich folgende Erkenntnis:

Durch Umordnung einer konvergenten Reihe kann sich der Grenzwert ändern.

Definition 12.15.
Eine Reihe $\sum_{n=0}^{\infty} a_n$ in $\mathbb{K}$ heißt *absolut konvergent*, wenn die Reihe ihrer *Absolut*beträge $\sum_{n=0}^{\infty} |a_n|$ konvergiert. Da die Folge der Parialsummen $t_n := \sum_{k=0}^{n} |a_k|$ monoton wächst, ist dies gleichwertig dazu, daß die Folge $(t_n)_{n\in\mathbb{N}}$ beschränkt ist (siehe Monotoniekriterium 11.21 und Satz 11.11).

Beispiel 12.16.
Die alternierende harmonische Reihe ist konvergent, aber nicht absolut konvergent.

Lemma 12.17.
Ist $\sum_{n=0}^{\infty} a_n$ in $\mathbb{K}$ absolut konvergent, so ist sie auch konvergent.

Beweis: Sei $\varepsilon > 0$ gegeben. Dann gibt es wegen des Cauchy-Kriteriums für Reihen eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$\left|\sum_{k=n+1}^{m} |a_k|\right| < \varepsilon$$

für alle $m > n \geq n_\varepsilon$, da die Reihe $\sum_{n=0}^{\infty} |a_n|$ konvergiert. Aus der Dreiecksungleichung wissen wir nun aber, daß dann auch

$$\left|\sum_{k=n+1}^{m} a_k\right| \leq \sum_{k=n+1}^{m} |a_k| = \left|\sum_{k=n+1}^{m} |a_k|\right| < \varepsilon$$

für alle $m > n \geq n_\varepsilon$ gilt. Mithin ist die Reihe $\sum_{n=0}^{\infty} a_n$ nach dem Cauchy-Kriterium für Reihen konvergent. □

Man beachte, daß Beispiel 12.14 zeigt, daß im folgenden Satz die Voraussetzung *absolut konvergent* nicht durch die Bedingung *konvergent* ersetzt werden kann.

Satz 12.18 (Umordnungssatz).
Jede Umordnung einer absolut konvergenten Reihe ist absolut konvergent und konvergiert gegen den gleichen Grenzwert.

Beweis: Sei $\sum_{n=0}^{\infty} a_n$ eine absolut konvergente Reihe und $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ sei bijektiv.

Wir wollen zunächst zeigen, daß die Reihe $\sum_{n=0}^{\infty} \left(a_n - a_{\sigma(n)}\right)$ gegen Null konvergiert. Sei dazu $\varepsilon > 0$ gegeben. Da die Reihe $\sum_{n=0}^{\infty} a_n$ absolut konvergent ist, gibt es ein $n'_\varepsilon \in \mathbb{N}$, so daß

$$\sum_{k=n+1}^{m} |a_k| = \left| \sum_{k=n+1}^{m} |a_k| \right| < \varepsilon$$

für alle $m > n \geq n'_\varepsilon$ gilt. Da die Abbildung σ surjektiv und die Menge $\{0, 1, \ldots, n'_\varepsilon\}$ endlich ist, gibt es eine Zahl $n_\varepsilon \geq n'_\varepsilon$ mit

$$\{0, 1, \ldots, n'_\varepsilon\} \subseteq \{\sigma(0), \sigma(1), \ldots, \sigma(n_\varepsilon)\}.$$

Für $n \geq n_\varepsilon$ heben sich deshalb in der Partialsumme

$$\sum_{k=0}^{n} \left(a_k - a_{\sigma(k)}\right)$$

die a_i mit $i \leq n'_\varepsilon$ heraus, da sie einmal mit positivem und einmal mit negativem Vorzeichen auftreten. Die übrigen a_i können sich herausheben oder auch nicht; in letzterem Fall kommen sie in der Summe genau einmal (entweder mit positivem oder mit negativem Vorzeichen) vor. Setzen wir nun

$$m := \max\{\sigma(0), \sigma(1), \ldots, \sigma(n), 0, 1, \ldots, n\} + 1 > n'_\varepsilon,$$

so erhalten wir insgesamt

$$\left| \sum_{k=0}^{n} \left(a_k - a_{\sigma(k)}\right) - 0 \right| \leq \sum_{k=n'_\varepsilon+1}^{m} |a_k| < \varepsilon.$$

Also konvergiert die Reihe $\sum_{n=0}^{\infty} \left(a_n - a_{\sigma(n)}\right)$ gegen Null.

Aus den Grenzwertsätzen für Reihen 12.4 erhalten wir deshalb, daß die Reihe

$$\sum_{n=0}^{\infty} a_{\sigma(n)} = \sum_{n=0}^{\infty} a_n - \sum_{n=0}^{\infty} \left(a_n - a_{\sigma(n)}\right) = \sum_{n=0}^{\infty} a_n$$

konvergent ist mit dem Grenzwert $\sum_{n=0}^{\infty} a_n$.

Wenden wir das Ergebnis auf die konvergente Reihe $\sum_{n=0}^{\infty} |a_n|$ und ihre Umordnung $\sum_{n=0}^{\infty} \left|a_{\sigma(n)}\right|$ an, so folgt auch, daß die Umordnung absolut konvergent ist. □

D) Konvergenzkriterien für absolute Konvergenz

Satz 12.19 (Majorantenkriterium).
Es seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ zwei Reihen in $\mathbb{K}$. Ist $\sum_{n=0}^{\infty} b_n$ absolut konvergent und $|a_n| \leq |b_n|$ für alle $n \geq n_0$, so ist auch $\sum_{n=0}^{\infty} a_n$ absolut konvergent.
Wir nennen $\sum_{n=0}^{\infty} b_n$ dann eine konvergente *Majorante* von $\sum_{n=0}^{\infty} a_n$.

Beweis: Die Folge von Partialsummen $(s_n)_{n\in\mathbb{N}}$ mit

$$s_n := \sum_{k=0}^{n} |a_k|$$

ist beschränkt durch $s_{n_0} + \sum_{n=0}^{\infty} |b_n|$. Also ist die Reihe absolut konvergent. □

Proposition 12.20 (Minorantenkriterium).
Es seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ zwei Reihen in $\mathbb{R}$.
Ist $\sum_{n=0}^{\infty} b_n$ divergent und $a_n \geq b_n \geq 0$ für alle $n \in \mathbb{N}$, so ist $\sum_{n=0}^{\infty} a_n$ divergent.
Wir nennen $\sum_{n=0}^{\infty} b_n$ dann eine divergente *Minorante* von $\sum_{n=0}^{\infty} a_n$.

Beweis: Wegen $b_n \geq 0$ ist die Folge der Partialsummen $(t_n)_{n\in\mathbb{N}}$ mit

$$t_n := \sum_{k=0}^{n} b_k$$

monoton wachsend. Da die Folge $(t_n)_{n\in\mathbb{N}}$ nach Voraussetzung divergent ist, ist sie wegen des Monotoniekriteriums für Folgen 11.21 nicht beschränkt. Aber dann ist auch die Folge der Partialsummen $(s_n)_{n\in\mathbb{N}}$ mit

$$s_n := \sum_{k=0}^{n} a_k$$

unbeschränkt, wegen $s_n \geq t_n$, und mithin ist sie divergent nach Satz 11.11. □

Beispiel 12.21.
Für $k \geq 2$ ist die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^k}$ konvergent.

Dazu betrachten wir zunächst den Fall $k = 2$. Wegen

$$a_n := \frac{1}{(n+1)^2} \leq \frac{1}{n \cdot (n+1)} =: b_n$$

ist wegen Beispiel 12.2 die Reihe $\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} \frac{1}{n\cdot(n+1)}$ eine konvergente Majorante von $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} \frac{1}{(n+1)^2}$. Nehmen wir nun noch eine Indexverschiebung vor, so sehen wir, daß die Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \sum_{n=1}^{\infty} \frac{1}{(n+1)^2}$$

ebenfalls konvergent ist. Für den Fall $k > 2$ gilt nun wegen $0 \le \frac{1}{n^k} \le \frac{1}{n^2}$, daß die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$ ihrerseits eine konvergente Majorante von $\sum_{n=1}^{\infty} \frac{1}{n^k}$ ist.

Satz 12.22 (Wurzelkriterium).
Es sei $\sum_{n=0}^{\infty} a_n$ eine Reihe in $\mathbb{K}$.

a. Existiert ein $q < 1$ mit $\sqrt[n]{|a_n|} \le q$ für $n \ge n_0$, so ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent.

b. Ist $\sqrt[n]{|a_n|} \ge 1$ für alle $n \ge n_0$, so ist $\sum_{n=0}^{\infty} a_n$ divergent.

Beweis: Ist $q < 1$ mit $\sqrt[n]{|a_n|} \le q$ für $n \ge n_0$, d.h. $|a_n| \le q^n$, so ist die geometrische Reihe $\sum_{n=0}^{\infty} q^n$ nach Satz 12.9 eine konvergente Majorante von $\sum_{n=0}^{\infty} a_n$.

Falls $\sqrt[n]{|a_n|} \ge 1$ für alle $n \ge n_0$, d.h. $|a_n| \ge 1^n = 1$, so ist $(a_n)_{n\in\mathbb{N}}$ keine Nullfolge und mithin ist $\sum_{n=0}^{\infty} a_n$ wegen des Nullfolgekriteriums divergent. □

Satz 12.23 (Quotientenkriterium).
Es sei $\sum_{n=0}^{\infty} a_n$ eine Reihe in $\mathbb{K}$ mit $a_n \neq 0$ für alle $n \ge n_0$.

a. Existiert ein $q < 1$ mit $\left|\frac{a_{n+1}}{a_n}\right| \le q$ für $n \ge n_0$, so ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent.

b. Ist $\left|\frac{a_{n+1}}{a_n}\right| \ge 1$ für alle $n \ge n_0$, so ist $\sum_{n=0}^{\infty} a_n$ divergent.

Beweis: Wenn eine reelle Zahl $0 < q < 1$ existiert mit $\left|\frac{a_{n+1}}{a_n}\right| \le q$ für $n \ge n_0$, so gilt

$$|a_{n+1}| \le q \cdot |a_n|$$

für alle $n \ge n_0$, und mit Induktion sieht man dann, daß

$$|a_n| \le q \cdot |a_{n-1}| \le q^2 \cdot |a_{n-2}| \le \ldots \le q^{n-n_0} \cdot |a_{n_0}|.$$

Also ist die geometrische Reihe

$$\frac{|a_{n_0}|}{q^{n_0}} \cdot \sum_{n=0}^{\infty} q^n = \sum_{n=0}^{\infty} q^{n-n_0} \cdot |a_{n_0}|$$

nach Satz 12.9 eine konvergente Majorante von $\sum_{n=0}^{\infty} a_n$.

Ist $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ für alle $n \geq n_0$, so ist $|a_{n+1}| \geq |a_n| \neq 0$ für alle $n \geq n_0$. Mithin ist $(a_n)_{n\in\mathbb{N}}$ keine Nullfolge und die Reihe $\sum_{n=0}^{\infty} a_n$ ist wegen des Nullfolgekriteriums dann divergent. □

Korollar 12.24 (Praktikables Quotienten-/Wurzelkriterium).
Sei $\sum_{n=0}^{\infty} a_n$ eine Reihe in $\mathbb{K}$ mit $a_n \neq 0$ für alle $n \geq n_0$.

a. Falls $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$ oder $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} < 1$, so ist $\sum\limits_{n=0}^{\infty} a_n$ absolut konvergent.

b. Falls $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| > 1$ oder $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} > 1$, so ist $\sum\limits_{n=0}^{\infty} a_n$ divergent.

c. Im Fall $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 1$ oder $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} = 1$ wird keine Aussage getroffen!

Beweis: Falls $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$ bzw. $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} < 1$, so kann man Satz 12.22 bzw. Satz 12.23 mit

$$q := \frac{1 + \lim\limits_{n\to\infty} \sqrt[n]{|a_n|}}{2} \quad \text{bzw.} \quad q := \frac{1 + \lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|}{2}$$

anwenden.

Falls $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} > 1$ bzw. $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| > 1$, so ist sicher $\sqrt[n]{|a_n|} > 1$ bzw. $\left|\frac{a_{n+1}}{a_n}\right| > 1$ für n hinreichend groß, so daß die Aussage ebenfalls aus Satz 12.22 bzw. Satz 12.23 folgt. □

Bemerkung 12.25.
Man beachte, daß die harmonische Reihe $\sum_{n\geq 1} a_n = \sum_{n\geq 1} \frac{1}{n}$ divergent ist, obwohl stets

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{\frac{1}{n+1}}{\frac{1}{n}} = 1 - \frac{1}{n+1} < 1$$

gilt. Aber, es gibt kein $q < 1$ mit

$$\left|\frac{a_{n+1}}{a_n}\right| = 1 - \frac{1}{n+1} < q$$

für alle hinreichend großen n. Das Quotientenkriterium ist deshalb nicht anwendbar. Beachte auch, daß in diesem Fall

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty} 1 - \frac{1}{n+1} = 1$$

gilt.

Beispiel 12.26.
Die Reihe $\sum_{n=1}^{\infty} \frac{n^2}{n!}$ ist absolut konvergent, da

$$\left|\frac{\frac{(n+1)^2}{(n+1)!}}{\frac{n^2}{n!}}\right| = \frac{(n+1)^2 \cdot n!}{n^2 \cdot (n+1)!} = \frac{n+1}{n^2} = \frac{1}{n} + \frac{1}{n^2} \longrightarrow 0 + 0 = 0.$$

E) Das Cauchy-Produkt absolut konvergenter Reihen

Satz 12.27 (Cauchy-Produkt).
Es seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ zwei absolut konvergente Reihen in $\mathbb{K}$. Für $n \in \mathbb{N}$ setzen wir

$$c_n := \sum_{k=0}^{n} a_k \cdot b_{n-k} = \sum_{i+j=n} a_i \cdot b_j.$$

Dann ist die Reihe $\sum_{n=0}^{\infty} c_n$ absolut konvergent und es gilt

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} a_n \cdot \sum_{n=0}^{\infty} b_n.$$

Beweis: Wir konstruieren zunächst eine bijektive Abbildung $\sigma : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$, indem wir die Elemente in $\mathbb{N} \times \mathbb{N}$ in der in Abbildung 1 angegebenen Weise durchlaufen.

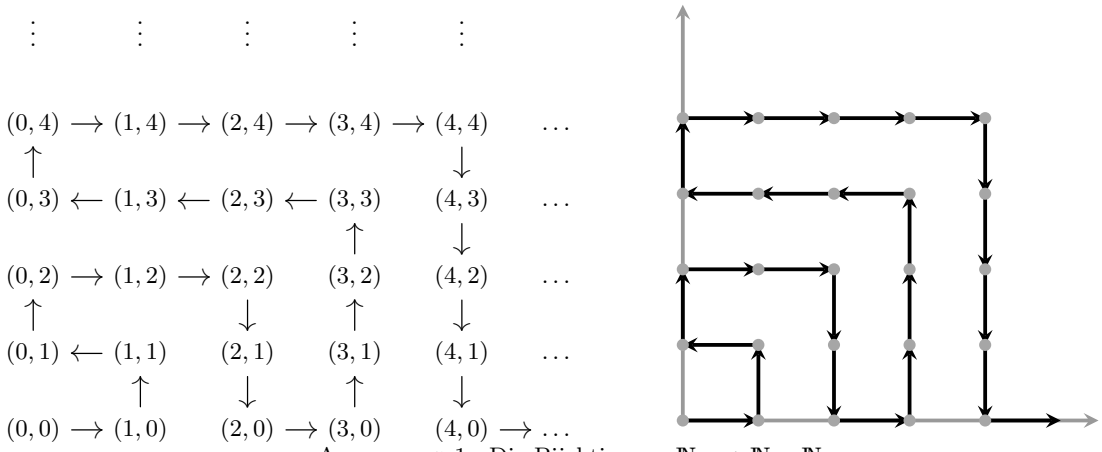


Abbildung 1. Die Bijektion $\sigma : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$

Aufgrund der Definition von σ gilt für $m \in \mathbb{N}$ offenbar

$$\{\sigma(0), \sigma(1), \ldots, \sigma((m+1)^2 - 1)\} = \{(k,l) \mid 0 \le k, l \le m\} \tag{17}$$

Dann definieren wir uns eine Folge $(d_n)_{n\in\mathbb{N}}$ durch

$$d_n := a_k \cdot b_l, \qquad \text{wenn } (k,l) = \sigma(n).$$

Wir wollen nun zunächst zeigen, daß die Reihe $\sum_{n=0}^{\infty} d_n$ absolut konvergent ist. Dazu beachten wir, daß für $m \in \mathbb{N}$ die Ungleichung

$$\sum_{n=0}^{m} |d_n| \le \sum_{n=0}^{(m+1)^2-1} |d_n| \overset{(17)}{=} \sum_{k=0}^{m} \sum_{l=0}^{m} |a_k \cdot b_l| = \sum_{k=0}^{m} |a_k| \cdot \sum_{l=0}^{m} |b_l| \le \sum_{k=0}^{\infty} |a_k| \cdot \sum_{l=0}^{\infty} |b_l|$$

erfüllt ist. Mithin ist die Folge der Partialsummen von $\sum_{n=0}^{\infty} |d_n|$ nach oben beschränkt. Damit ist $\sum_{n=0}^{\infty} d_n$ absolut konvergent und deshalb auch konvergent.

Zudem folgt aus den Grenzwertsätzen

$$\sum_{i=0}^{\infty} d_i \longleftarrow \sum_{i=0}^{(n+1)^2-1} d_i \overset{(17)}{=} \sum_{k=0}^{n}\sum_{l=0}^{n} a_k \cdot b_l = \sum_{k=0}^{n} a_k \cdot \sum_{l=0}^{n} b_l \longrightarrow \sum_{k=0}^{\infty} a_k \cdot \sum_{l=0}^{\infty} b_l,$$

und wegen der Eindeutigkeit des Grenzwertes gilt dann zudem

$$\sum_{n=0}^{\infty} d_n = \sum_{n=0}^{\infty} a_n \cdot \sum_{n=0}^{\infty} b_n.$$

Diese absolut konvergente Reihe $\sum_{n=0}^{\infty} d_n$ werden wir nun umordnen. Dazu konstruieren wir uns nach dem Cantorschen Diagonalverfahren eine weitere bijektive Abbildung $\pi : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$ wie in Abbildung 2 angedeutet.

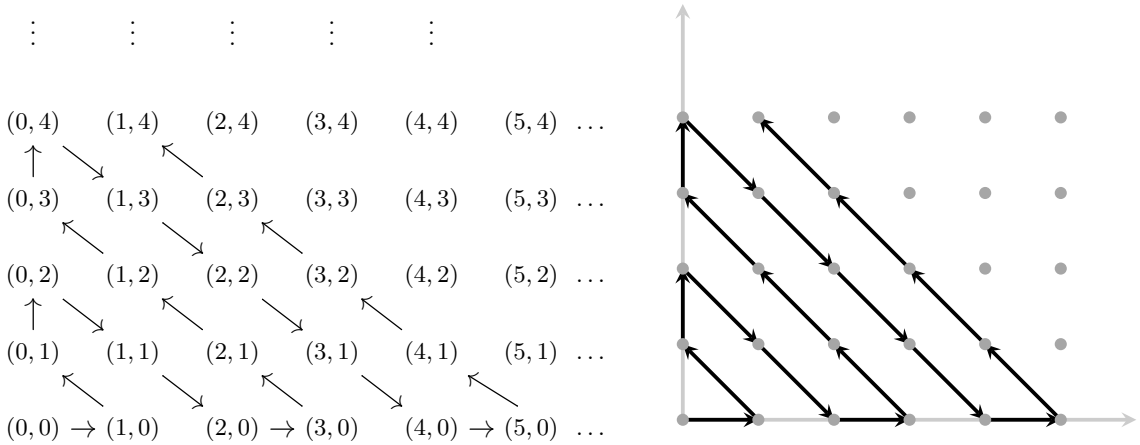


Abbildung 2. Die Bijektion $\pi : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$

Wir setzen nun

$$e_n := d_{\sigma^{-1}(\pi(n))} = a_k \cdot b_l, \quad \text{wenn } \pi(n) = (k,l)$$

für $n \in \mathbb{N}$ und erhalten so eine Umordnung $(e_n)_{n\in\mathbb{N}} = (d_{\sigma^{-1}(\pi(n))})_{n\in\mathbb{N}}$ der Folge $(d_n)_{n\in\mathbb{N}}$. Wegen des Umordnungssatzes 12.18 ist dann auch die Reihe $\sum_{n=0}^{\infty} e_n$ absolut konvergent mit dem gleichen Grenzwert

$$\sum_{n=0}^{\infty} e_n = \sum_{n=0}^{\infty} d_n.$$

Nun entsteht die Reihe $\sum_{n=0}^{\infty} c_n$ offenbar aus der Reihe $\sum_{n=0}^{\infty} e_n$ durch Einfügen von Klammern[2] im Sinne von Lemma 12.12. Mithin ist die Reihe nach eben diesem Lemma ebenfalls konvergent mit dem gleichen Grenzwert, d.h.

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} e_n = \sum_{n=0}^{\infty} d_n = \sum_{n=0}^{\infty} a_n \cdot \sum_{n=0}^{\infty} b_n.$$

[2]Wir wollen dies in der Fußnote etwas ausführen. Aufgrund der Definition von π und unter Verwendung der Formel in Beispiel 7.11 zur Berechnung der Summe der ersten n Zahlen sieht man, daß für $n \in \mathbb{N}$ folgende Gleichheit gilt

$$\{(k,l) \mid k+l=n\} = \left\{\pi(i) \;\middle|\; \frac{n\cdot(n+1)}{2} \le i \le \frac{(n+2)\cdot(n+1)}{2} - 1\right\}. \tag{18}$$

Aufgrund der Dreiecksungleichung erhalten wir für $m \in \mathbb{N}$ zudem

$$\sum_{n=0}^{m} |c_n| = \sum_{n=0}^{m} \left| \sum_{k+l=n} a_k \cdot b_l \right| \leq \sum_{n=0}^{m} \sum_{k+l=n} |a_k \cdot b_l| \leq \sum_{k=0}^{m} |a_k| \cdot \sum_{l=0}^{m} |b_l| \leq \sum_{k=0}^{\infty} |a_k| \cdot \sum_{l=0}^{\infty} |b_l|,$$

so daß auch die Reihe $\sum_{n=0}^{\infty} |c_n|$ beschränkt und monoton wachsend, also konvergent ist, d.h. $\sum_{n=0}^{\infty} c_n$ ist absolut konvergent. □

F) Potenzreihen

Definition 12.28.
Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$, $a \in \mathbb{K}$ und t eine Veränderliche.
Wir nennen einen Ausdruck der Form $\sum_{n=0}^{\infty} a_n \cdot (t-a)^n$ eine *Potenzreihe* über $\mathbb{K}$ in der Veränderlichen t mit *Entwicklungspunkt* a. Im folgenden beschränken wir uns im wesentlichen auf den Fall $a = 0$ und schreiben dann einfach $\sum_{n=0}^{\infty} a_n \cdot t^n$. Unser Ziel ist es, für die Veränderliche t Werte $x \in \mathbb{K}$ einzusetzen und so eine Reihe zu erhalten, die konvergiert oder auch nicht.

Lemma 12.29.
Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $y \in \mathbb{K}$, so daß die Reihe $\sum_{n=0}^{\infty} a_n \cdot y^n$ konvergiert. Dann ist die Reihe $\sum\limits_{n=0}^{\infty} a_n \cdot x^n$ absolut konvergent für alle $x \in \mathbb{K}$ mit $|x| < |y|$.

Beweis: Da die Reihe $\sum_{n=0}^{\infty} a_n \cdot y^n$ konvergent ist, ist die Folge $(a_n \cdot y^n)_{n\in\mathbb{N}}$ nach dem Nullfolgekriterium eine Nullfolge, und mithin ist sie auch beschränkt. D.h. es gibt ein $s > 0$ mit

$$|a_n \cdot y^n| \leq s$$

für alle $n \in \mathbb{N}$. Für $x \in \mathbb{K}$ mit $|x| < |y|$ setzen wir $q := \frac{|x|}{|y|} < 1$ und erhalten dann

$$|a_n \cdot x^n| = |a_n \cdot y^n| \cdot \frac{|x|^n}{|y|^n} \leq s \cdot q^n.$$

Also ist die geometrische Reihe $\sum_{n=0}^{\infty} s \cdot q^n = s \cdot \sum_{n=0}^{\infty} q^n$ eine konvergente Majorante von $\sum_{n=0}^{\infty} a_n \cdot x^n$, so daß diese nach dem Majorantenkriterium absolut konvergiert. □

Für c_n ergibt sich daraus

$$c_n = \sum_{k+l=n} a_k \cdot b_l = \sum_{i=\frac{n\cdot(n+1)}{2}}^{\frac{(n+2)\cdot(n+1)}{2}-1} e_i = e_{\frac{n\cdot(n+1)}{2}} + e_{\frac{n\cdot(n+1)}{2}+1} + \ldots + e_{\frac{(n+2)\cdot(n+1)}{2}-1},$$

d.h. $\sum_{n=0}^{\infty} c_n$ entsteht aus $\sum_{n=0}^{\infty} e_n$ durch Zusammenfassung von Summanden mittels Einfügen von Klammern.

Notation 12.30.
Wir wollen den Begriff des Supremums etwas erweitern, indem wir $\sup(\emptyset) := -\infty$ setzen und $\sup(A) := \infty$, falls $A \subseteq \mathbb{R}$ nicht nach oben beschränkt ist. Damit gilt für jede Teilmenge $A \subseteq \mathbb{R}$

$$\sup(A) \in \mathbb{R} \cup \{\infty, -\infty\}.$$

Wir erinnern uns, daß wir in Bemerkung 11.35 für $x \in \mathbb{R}$ bereits die Konvention $\frac{x}{\infty} := \frac{x}{-\infty} := 0$ eingeführt haben. Wir vereinbaren nun zudem $\frac{x}{0} := \infty$ für $x > 0$ sowie $\frac{x}{0} := -\infty$ für $x < 0$.

Definition 12.31.
Für eine Potenzreihe $\sum_{n=0}^{\infty} a_n \cdot t^n$ über $\mathbb{K}$ nennen wir

$$r := \sup\left\{ |y| \;\middle|\; y \in \mathbb{K}, \sum_{n=0}^{\infty} a_n \cdot y^n \text{ ist konvergent} \right\} \in \mathbb{R}_{\geq 0} \cup \{\infty\}$$

den *Konvergenzradius* der Potenzreihe.

Man beachte, daß die Potenzreihe zumindest für $y = 0$ konvergiert, so daß die angegebene Menge nicht-leer ist!

Satz 12.32 (Konvergenzradius).
Es sei $\sum_{n=0}^{\infty} a_n \cdot t^n$ eine Potenzreihe über $\mathbb{K}$ mit Konvergenzradius r.

a. Ist $x \in \mathbb{K}$ mit $|x| < r$, so ist $\sum_{n=0}^{\infty} a_n \cdot x^n$ absolut konvergent.

b. Ist $x \in \mathbb{K}$ mit $|x| > r$, so ist $\sum_{n=0}^{\infty} a_n \cdot x^n$ divergent.

Setzen wir $U_r(0) := \{x \in \mathbb{K} \mid |x| < r\}$, so definiert die Potenzreihe also eine Abbildung

$$U_r(0) \longrightarrow \mathbb{K} : x \mapsto \sum_{n=0}^{\infty} a_n \cdot x^n.$$

Wir nennen $U_r(0)$ den *Konvergenzbereich* der Potenzreihe.

Bemerkung 12.33.

a. Über den Fall $|x| = r$ wird in Satz 12.32 keine Aussage getroffen! Wir nennen die Menge $\{x \mid |x| = r\}$ den *Rand* des Konvergenzbereiches.

b. Konvergenzradius $r = \infty$ heißt, daß $\sum_{n=0}^{\infty} a_n \cdot x^n$ für alle $x \in \mathbb{K}$ absolut konvergent ist.

c. Ist $\mathbb{K} = \mathbb{R}$, so ist die Menge $U_r(0) = (-r, r)$ ein offenes Intervall; ist $\mathbb{K} = \mathbb{C}$, so ist die Menge $U_r(0)$ ein Kreis mit Radius r um den Ursprung. In Abbildung 3 stellen wir den Konvergenzbereich der Reihe graphisch dar.

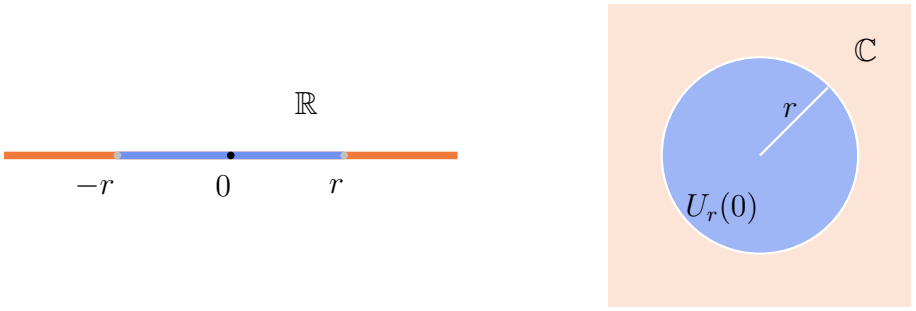


Abbildung 3: Konvergenzbereich $U_r(0)$

Beweis von Satz 12.32: a. Wir betrachten die Menge

$$A := \left\{ |y| \;\middle|\; y \in \mathbb{K}, \sum_{n=0}^{\infty} a_n \cdot y^n \text{ ist konvergent} \right\},$$

so daß $r = \sup(A)$. Ist $r = \infty$, so ist A unbeschränkt und zu jedem $x \in \mathbb{K}$ gibt es ein $y \in A$ mit $|y| > |x|$. Ist $|x| < r < \infty$, so ist $\varepsilon := \frac{r-|x|}{2} > 0$ und $r - \varepsilon = \sup(A) - \varepsilon$ ist keine obere Schranke von A. Es gibt also ein $y \in A$ mit $|y| > r - \varepsilon = \frac{r+|x|}{2} > |x|$. In beiden Fällen $\sum_{n=0}^{\infty} a_n \cdot y^n$ ist konvergent und $|x| < |y|$, und nach Lemma 12.29 ist $\sum_{n=0}^{\infty} a_n \cdot x^n$ mithin absolut konvergent.

b. Ist $|x| > r$, so ist $|x| \notin A$ und mithin muß $\sum_{n=0}^{\infty} a_n \cdot x^n$ divergent sein.

□

Satz 12.34 (Cauchy-Hadamard).
Sei $\sum_{n=0}^{\infty} a_n \cdot t^n$ eine Potenzreihe über $\mathbb{K}$.

a. Falls der eigentliche oder uneigentliche Grenzwert $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ existiert, so ist der Konvergenzradius von $\sum_{n=0}^{\infty} a_n \cdot t^n$ gegeben durch

$$r = \frac{1}{\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|}.$$

b. Falls der eigentliche oder uneigentliche Grenzwert $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ existiert, so ist der Konvergenzradius von $\sum_{n=0}^{\infty} a_n \cdot t^n$ gegeben durch

$$r = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{|a_n|}}.$$

Beweis:

a. Es sei $r = \frac{1}{\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|}$ und $x \in \mathbb{K}$. Ist $|x| < r$, so ist die Reihe $\sum_{n=0}^{\infty} a_n \cdot x^n$ nach dem Quotientenkriterium in Korollar 12.24 absolut konvergent, da

$$\lim_{n\to\infty} \left|\frac{a_{n+1} \cdot x^{n+1}}{a_n \cdot x^n}\right| = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \cdot |x| < 1.$$

Analog ist die Reihe divergent, wenn $|x| > r$.

b. Der Beweis geht analog zu a., wobei wir das Quotientenkriterium in Korollar 12.24 durch das dortige Wurzelkriterium ersetzen.

□

Beispiel 12.35.

a. Die *geometrische Reihe* $\sum_{n=0}^{\infty} t^n$ hat den Konvergenzradius $r = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{1}} = 1$.
Damit wissen wir, daß die Reihe $\sum_{n=0}^{\infty} x^n$ absolut konvergiert für $|x| < 1$ und daß sie divergiert für $|x| > 1$. Wir haben aber in Beispiel 11.31 schon gesehen, daß sie zudem auch für alle $|x| = 1$ divergiert, d.h. sie divergiert für alle Punkte im Rand des Konvergenzbereiches.

b. Die Potenzreihe $\sum_{n=1}^{\infty} \frac{t^n}{n}$ hat ebenfalls den Konvergenzradius

$$r = \frac{1}{\lim\limits_{n\to\infty} \frac{n}{n+1}} = \frac{1}{\lim\limits_{n\to\infty} \left(1 - \frac{1}{n+1}\right)} = 1.$$

Aber für $x = -1$ erhalten wir die alternierende harmonische Reihe $\sum_{n=1}^{\infty} \frac{x^n}{n} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n}$, die konvergiert, so daß die Potenzreihe nicht für alle x im Rand des Konvergenzbereiches divergiert.

G) Exponentialfunktion, Sinus und Cosinus als Potenzreihen

Satz 12.36 (Exponentialfunktion).
Die Potenzreihe $\sum_{n=0}^{\infty} \frac{t^n}{n!}$ über $\mathbb{K}$ hat Konvergenzradius $r = \infty$.
Die dadurch definierte Abbildung

$$\exp : \mathbb{K} \longrightarrow \mathbb{K} : x \mapsto \exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

nennen wir die *Exponentialfunktion*. Sie genügt der *Funktionalgleichung*

$$\exp(x + y) = \exp(x) \cdot \exp(y)$$

für $x, y \in \mathbb{K}$.

Beweis: Der Konvergenzradius ergibt sich als

$$r = \frac{1}{\lim\limits_{n\to\infty} \frac{n!}{(n+1)!}} = \frac{1}{\lim\limits_{n\to\infty} \frac{1}{n+1}} = \frac{1}{0} = \infty.$$

Zudem folgt aus dem Cauchy-Produkt für Reihen 12.27 und dem Binomischen Lehrsatz 7.15

$$\begin{aligned}
\exp(x)\cdot\exp(y) &= \sum_{n=0}^{\infty}\frac{x^n}{n!}\cdot\sum_{n=0}^{\infty}\frac{y^n}{n!}\\
&\overset{12.27}{=} \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{x^k}{k!}\cdot\frac{y^{n-k}}{(n-k)!}\\
&= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}\cdot x^k\cdot y^{n-k}\\
&\overset{7.15}{=} \sum_{n=0}^{\infty}\frac{(x+y)^n}{n!}\\
&= \exp(x+y).
\end{aligned}$$

□

Bemerkung 12.37.
Nach Aufgabe 11.44 gilt $\exp(1)=e$ und mit Induktion folgt aus der Funktionalgleichung leicht, daß $\exp(n)=e^n$ für $n\in\mathbb{Z}$ und $\exp(\frac{1}{n})=e^{\frac{1}{n}}=\sqrt[n]{e}$ für $n\geq 2$. Wir setzen für $x\in\mathbb{K}$ deshalb allgemein

$$e^x := \exp(x),$$

so daß die neue Notation mit der üblichen Potenzschreibweise und mit der Notation in Satz 9.8 übereinstimmt, und das Potenzgesetz $e^{x+y}=e^x\cdot e^y$ gilt.

Satz 12.38 (Sinus und Cosinus).

a. Die Potenzreihe $\sum_{n=0}^{\infty}(-1)^n\cdot\frac{t^{2n}}{(2n)!}$ über $\mathbb{K}$ hat Konvergenzradius $r=\infty$. Die dadurch definierte Abbildung

$$\cos:\mathbb{K}\longrightarrow\mathbb{K}:x\mapsto\cos(x):=\sum_{n=0}^{\infty}(-1)^n\cdot\frac{x^{2n}}{(2n)!}$$

nennen wir den *Cosinus*.

b. Die Potenzreihe $\sum_{n=0}^{\infty}(-1)^n\cdot\frac{t^{2n+1}}{(2n+1)!}$ über $\mathbb{K}$ hat Konvergenzradius $r=\infty$. Die dadurch definierte Abbildung

$$\sin:\mathbb{K}\longrightarrow\mathbb{K}:x\mapsto\sin(x):=\sum_{n=0}^{\infty}(-1)^n\cdot\frac{x^{2n+1}}{(2n+1)!}$$

nennen wir den *Sinus*.

c. Für $x\in\mathbb{K}$ gelten

$$\sin(-x)=-\sin(x)$$

und

$$\cos(-x) = \cos(x).$$

Wir nennen den Sinus eine *ungerade* Funktion und den Cosinus eine *gerade*.

d. Für $x \in \mathbb{K}$ gilt

$$e^{i\cdot x} = \exp(i \cdot x) = \cos(x) + i \cdot \sin(x).$$

e. Für $x \in \mathbb{K}$ gilt

$$\cos(x)^2 + \sin(x)^2 = 1.$$

f. Für $x \in \mathbb{K}$ gilt

$$\cos(x) = \frac{1}{2} \cdot \left(e^{ix} + e^{-ix}\right)$$

und

$$\sin(x) = \frac{1}{2i} \cdot \left(e^{ix} - e^{-ix}\right).$$

g. Für zwei reelle Zahlen $x, y \in \mathbb{R}$ gelten die *Additionstheoreme*

$$\cos(x+y) = \cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y)$$

und

$$\sin(x+y) = \cos(x) \cdot \sin(y) + \sin(x) \cdot \cos(y).$$

h. Für eine reelle Zahl $x \in \mathbb{R}$ gilt $|e^{ix}| = 1$.

Beweis:

a. Ist $x \in \mathbb{K}$, so setzen wir

$$a_n := \begin{cases} (-1)^m \cdot \frac{x^{2m}}{(2m)!}, & \text{falls } n = 2m \text{ gerade}, \\ 0, & \text{falls } n \text{ ungerade.} \end{cases}$$

Dann ist $\cos(x) = \sum_{n=0}^{\infty} a_n$ und $|a_n| \leq \left|\frac{x^n}{n!}\right|$. Mithin ist $\exp(x)$ eine konvergente Majorante von $\cos(x)$, und $\cos(x)$ ist absolut konvergent für alle $x \in \mathbb{K}$. Insbesondere ist der Konvergenzradius also

$$r = \sup\left\{|x| \;\middle|\; x \in \mathbb{K}\right\} = \sup(\mathbb{R}_{\geq 0}) = \infty.$$

b. Ist $x \in \mathbb{K}$, so setzen wir

$$a_n := \begin{cases} (-1)^m \cdot \frac{x^{2m+1}}{(2m+1)!}, & \text{falls } n = 2m+1 \text{ ungerade}, \\ 0, & \text{falls } n \text{ gerade.} \end{cases}$$

Dann ist $\sin(x) = \sum_{n=0}^{\infty} a_n$ und $|a_n| \leq \left|\frac{x^n}{n!}\right|$. Mithin ist $\exp(x)$ eine konvergente Majorante von $\sin(x)$, und $\sin(x)$ ist absolut konvergent für alle $x \in \mathbb{K}$. Insbesondere ist der Konvergenzradius also

$$r = \sup\left\{|x| \;\middle|\; x \in \mathbb{K}\right\} = \sup(\mathbb{R}_{\geq 0}) = \infty.$$

c. Für $x \in \mathbb{K}$ gilt

$$\sin(-x) = \sum_{n=0}^{\infty}(-1)^n \frac{(-x)^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty}(-1)^n \frac{(-1)\cdot(x^{2n+1})}{(2n+1)!} = -\sin(x)$$

und

$$\cos(-x) = \sum_{n=0}^{\infty}(-1)^n \frac{(-x)^{2n}}{(2n)!} = \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n}}{(2n)!} = \cos(x).$$

d. Wir beachten, daß für die imaginäre Einheit i stets $i^{2n} = (-1)^n$ und $i^{2n+1} = (-1)^n \cdot i$ gilt. Dadurch erhalten wir

$$\begin{aligned}\cos(x) + i\cdot\sin(x) &= \sum_{n=0}^{\infty}(-1)^n \cdot \frac{x^{2n}}{(2n)!} + i \cdot \sum_{n=0}^{\infty}(-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} i^{2n} \cdot \frac{x^{2n}}{(2n)!} + \sum_{n=0}^{\infty} i^{2n+1} \cdot \frac{x^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} \frac{(i\cdot x)^n}{n!} = \exp(i\cdot x)\end{aligned}$$

unter Berücksichtigung der Grenzwertsätze für konvergente Reihen 12.4 und des Lemmas 12.12 zum Umklammern konvergenter Reihen.

e. Für $x \in \mathbb{K}$ gilt

$$\begin{aligned}\cos(x)^2 + \sin(x)^2 &= \big(\cos(x) + i\cdot\sin(x)\big)\cdot\big(\cos(x) - i\cdot\sin(x)\big) \\ &\overset{c.}{=} \big(\cos(x) + i\cdot\sin(x)\big)\cdot\big(\cos(-x) + i\cdot\sin(-x)\big) \\ &\overset{d.}{=} \exp(ix)\cdot\exp(-ix) \\ &\overset{12.36}{=} \exp(ix - ix) = \exp(0) = 1.\end{aligned}$$

f. Für $x \in \mathbb{K}$ gilt

$$e^{ix} + e^{-ix} \overset{d.}{=} \cos(x) + i\cdot\sin(x) + \cos(-x) + i\cdot\sin(-x) \overset{c.}{=} 2\cdot\cos(x)$$

und

$$e^{ix} - e^{-ix} \overset{d.}{=} \cos(x) + i\cdot\sin(x) - \cos(-x) - i\cdot\sin(-x) \overset{c.}{=} 2\cdot i\cdot\sin(x).$$

g. Werten wir den Sinus oder den Cosinus an einer reellen Zahl aus, so erhalten wir eine reelle Zahl. Für $x, y \in \mathbb{R}$ betrachten wir nun die komplexe Zahl

$$\begin{aligned}\cos(x+y) + i\cdot\sin(x+y) &\overset{d.}{=} \exp(i\cdot(x+y)) \overset{12.36}{=} \exp(ix)\cdot\exp(iy) \\ &\overset{d.}{=} \big(\cos(x) + i\cdot\sin(x)\big)\cdot\big(\cos(y) + i\cdot\sin(y)\big) \\ &= \big(\cos(x)\cdot\cos(y) - \sin(x)\cdot\sin(y)\big) \\ &\quad + i\cdot\big(\cos(x)\cdot\sin(y) + \sin(x)\cdot\cos(y)\big).\end{aligned}$$

Durch einen Vergleich des Realteils bzw. des Imaginärteils der beiden Seiten der Gleichung, erhalten wir die gewünschten Formeln.

h. Für $x \in \mathbb{R}$ gilt $|\exp(ix)| = |\cos(x) + i\sin(x)| = \sqrt{\cos(x)^2 + \sin(x)^2} = 1$.

□

Bemerkung 12.39 (Alternativer Beweis für $\cos(x)^2 + \sin(x)^2 = 1$).
Wir wollen zunächst $\cos(x)^2$ mit Hilfe des Cauchy-Produktes 12.27 ausrechnen.

$$\begin{aligned}
\cos(x)^2 &= \sum_{n=0}^{\infty}(-1)^n \cdot \frac{x^{2n}}{(2n)!} \cdot \sum_{n=0}^{\infty}(-1)^n \cdot \frac{x^{2n}}{(2n)!} \\
&= \sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k \cdot \frac{x^{2k}}{(2k)!} \cdot (-1)^{n-k} \cdot \frac{x^{2n-2k}}{(2n-2k)!} \\
&= \sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^n \cdot \frac{x^{2n}}{(2k)! \cdot (2n-2k)!} \\
&= \sum_{n=0}^{\infty}(-1)^n \cdot \frac{x^{2n}}{(2n)!} \cdot \sum_{k=0}^{n}\binom{2n}{2k} \\
&= 1 + \sum_{n=0}^{\infty}(-1)^{n+1} \cdot \frac{x^{2n+2}}{(2n+2)!} \cdot \sum_{k=0}^{n+1}\binom{2n+2}{2k},
\end{aligned}$$

wobei die letzte Gleichheit durch eine Indexverschiebung zustande kommt.

Dann wenden wir uns $\sin(x)^2$ zu und erhalten analog

$$\begin{aligned}
\sin(x)^2 &= \sum_{n=0}^{\infty}(-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!} \cdot \sum_{n=0}^{\infty}(-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!} \\
&= \sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k \cdot \frac{x^{2k+1}}{(2k+1)!} \cdot (-1)^{n-k} \cdot \frac{x^{2n-2k+1}}{(2n-2k+1)!} \\
&= \sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^n \cdot \frac{x^{2n+2}}{(2k+1)! \cdot (2n-2k+1)!} \\
&= \sum_{n=0}^{\infty}(-1)^n \cdot \frac{x^{2n+2}}{(2n+2)!} \cdot \sum_{k=0}^{n}\binom{2n+2}{2k+1}.
\end{aligned}$$

Addieren wir die Gleichungen, so erhalten wir mit dem Binomischen Lehrsatz 7.15

$$\begin{aligned}
\cos(x)^2+\sin(x)^2 &= 1+\sum_{n=0}^{\infty}(-1)^{n+1}\cdot\frac{x^{2n+2}}{(2n+2)!}\cdot\left(\sum_{k=0}^{n+1}\binom{2n+2}{2k}-\sum_{k=0}^{n}\binom{2n+2}{2k+1}\right)\\
&= 1+\sum_{n=0}^{\infty}(-1)^{n+1}\cdot\frac{x^{2n+2}}{(2n+2)!}\cdot\left(\sum_{l=0}^{2n+2}\binom{2n+2}{l}\cdot(-1)^l\cdot 1^{2n-l}\right)\\
&\overset{7.15}{=} 1+\sum_{n=0}^{\infty}(-1)^{n+1}\cdot\frac{x^{2n+2}}{(2n+2)!}\cdot(1-1)^{2n}\\
&= 1
\end{aligned}$$

unter Berücksichtigung der Grenzwertsätze für Reihen 12.4 sowie des Umordnungssatzes für absolut konvergente Reihen 12.18.

Bemerkung 12.40.
Die Additionstheoreme

$$\cos(x+y)=\cos(x)\cdot\cos(y)-\sin(x)\cdot\sin(y)$$

und

$$\sin(x+y)=\cos(x)\cdot\sin(y)+\sin(x)\cdot\cos(y).$$

gelten in der Tat nicht nur für reelle Zahlen x und y, sondern auch für beliebige komplexe Zahlen $x,y\in\mathbb{C}$. Allerdings funktioniert dann der oben geführte Beweis nicht, da dann die Aufteilung in den Real- und Imaginärteil in der angegebenen Form nicht möglich ist. Stattdessen kann man die Formeln direkt aus der Definition von Sinus und Cosinus mittels Potenzreihen herleiten, wie wir das in der obigen Bemerkung für die Gleichung $\cos(x)^2+\sin(x)^2=1$ getan haben. Das zu tun, überlassen wir dem Leser als Übungsaufgabe.

Bemerkung 12.41 (Potenzreihen mit beliebigem Entwicklungspunkt a).
Für eine Potenzreihe $\sum_{n=0}^{\infty}a_n\cdot(t-a)^n$ über $\mathbb{K}$ mit Entwicklungspunkt $a\in\mathbb{K}$ ist der Konvergenzradius immer noch definiert als

$$r:=\sup\left\{|y|\;\middle|\;y\in\mathbb{K},\sum_{n=0}^{\infty}a_n\cdot y^n\text{ ist konvergent}\right\}\in\mathbb{R}_{\geq 0}\cup\{\infty\},$$

und wir erhalten dann aus Satz 12.32

- $\forall\, x\in\mathbb{K}$ mit $|x-a|<r$ ist $\sum_{n=0}^{\infty}a_n\cdot(x-a)^n$ absolut konvergent,
- $\forall\, x\in\mathbb{K}$ mit $|x-a|>r$ ist $\sum_{n=0}^{\infty}a_n\cdot(x-a)^n$ divergent,

und Satz 12.34 impliziert, daß sich der Konvergenzradius ggf. berechnen läßt als

$$r=\frac{1}{\lim\limits_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|}$$

bzw. als

$$r = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{|a_n|}}.$$

H) Darstellung von Zahlen in Zahlsystemen

Ausblick 12.42 (Zahldarstellung in der Informatik).
Wir sind es gewohnt, reelle Zahlen als Dezimalzahlen darzustellen, d.h. eine Zahldarstellung in einem Stellenwertsystem zur Basis $b = 10$ zu verwenden. Die Wahl der Basis ist dabei eigentlich willkürlich. In der Antike waren auch andere Basen üblich, bei den Babyloniern etwa die Basis $b = 60$. In der Informatik nutzt man vornehmlich die Basen $b = 2$, das Dualsystem, und $b = 16$, das Hexadezimalsystem.

Definition und Satz 12.43 (b-adische Zahldarstellung).
Es sei $b \in \mathbb{N}$ eine natürliche Zahl mit $b \geq 2$ und es sei $z \in \mathbb{Z}$. Dann ist die Reihe

$$s = \sum_{n=z}^{\infty} a_n \cdot b^{-n}$$

konvergent in $\mathbb{R}$ für jede Wahl einer Folge $(a_n)_{n\geq z}$ mit

$$a_n \in \{0, 1, \ldots, b-1\}$$

für alle $n \geq z$. Wir nennen die Reihe dann eine Darstellung der reellen Zahl s als *b-adischen Bruch* und schreiben diesen auch in der Form

$$s = (a_z, a_{z+1}a_{z+2}a_{z+3}\ldots \; E - z)_b.$$

Wenn $a_n = 0$ für $n \geq N$ gilt, dann schreiben wir auch

$$s = (a_z, a_{z+1}a_{z+2}a_{z+3}\ldots a_N \; E - z)_b$$

und sprechen von einem *endlichen b-adischen Bruch*. Wenn sich die Ziffernfolge ab einer Stelle N periodisch wiederholt, d.h.

$$a_{N+i} = a_{N+i+k\cdot m}$$

für alle $i = 1, \ldots, m$ und alle $k \in \mathbb{N}$, dann schreiben wir

$$s = (a_z, a_{z+1}\ldots a_N \; \overline{a_{N+1}\ldots a_{N+m}} \; E - z)_b$$

und sprechen von einem *periodischen b-adischen Bruch*.

Beweis: Da die a_n alle nicht negativ und durch $b-1$ beschränkt sind, gilt

$$|a_n \cdot b^{-n}| = a_n \cdot b^{-n} \leq (b-1) \cdot b^{-n},$$

und wegen $\frac{1}{b} < 1$ ist dann die geometrische Reihe

$$\sum_{n=z}^{\infty}(b-1)\cdot b^{-n} = (b-1)\cdot b^{-z}\cdot\sum_{n=0}^{\infty}\frac{1}{b^n} = \frac{(b-1)\cdot b^{-z}}{1-\frac{1}{b}} = b^{1-z} \tag{19}$$

ist eine konvergente Majorante von s. Mithin ist die Reihe absolut konvergent und ihr Grenzwert s ist eine reelle Zahl. □

Beispiel 12.44.

a. Die natürliche Zahl 1274 hat die 10-adische Darstellung

$$1274 = 1\cdot 10^3 + 2\cdot 10^2 + 7\cdot 10^1 + 4\cdot 10^0 = (1,274\ E\ 3)_{10}.$$

b. Die rationale Zahl $\frac{1}{3}$ hat die 10-adische Darstellung

$$\frac{1}{3} = \sum_{n=1}^{\infty} 3\cdot 10^{-n} = (3,\overline{3}\ E\ -1)_{10} = (0,\overline{3}\ E\ 0)_{10}.$$

c. Die rationale Zahl $\frac{1685}{49}$ hat die 7-adische Darstellung

$$\frac{1685}{49} = 4\cdot 7^1 + 6\cdot 7^0 + \frac{2}{7} + \frac{5}{7^2} = (4,625\ E\ 1)_7.$$

d. Die natürliche Zahl 2 besitzt die 2-adische Darstellungen $2 = (1,0\ E1)_2$ und

$$2 = \frac{1}{1-\frac{1}{2}} = \sum_{n=0}^{\infty} 2^{-n} = (0,\overline{1}\ E\ 1)_2.$$

Satz 12.45 (b-adische Zahldarstellung).
Sei $b \in \mathbb{N}$ eine natürliche Zahl mit $b \geq 2$.

a. Jede nicht-negative reelle Zahl besitzt eine Darstellung als b-adischer Bruch.

b. Die Darstellung einer reellen Zahl als b-adischer Bruch ist bis auf die folgende Uneindeutigkeit eindeutig: Ist $N \geq z$ eine feste ganze Zahl mit $a_n = c_n$ für $z \leq n \leq N-1$, $a_N < c_N$ und

$$(a_z, a_{z+1}a_{z+2}a_{z+3}\ldots\ E - z)_b = (c_z, c_{z+1}c_{z+2}c_{z+3}\ldots\ E - z)_b,$$

so gilt $a_N + 1 = c_N$ sowie $a_n = b-1$ und $c_n = 0$ für alle $n > N$.

c. Jede natürliche Zahl $s \in \mathbb{N}$ besitzt eine Darstellung als endlicher b-adischer Bruch der Form

$$s = (a_z, a_{z+1}\ldots a_0\ E - z)_b.$$

d. Ein b-adischer Bruch ist genau dann in $\mathbb{Q}$, wenn er endlich oder periodisch ist.

Beweis:

a. Es reicht, die Aussage für positive reelle Zahlen $s \in \mathbb{R}_{>0}$ zu zeigen, da sie für $s = 0$ offenbar gilt. Weil $(b^{-n})_{n\in\mathbb{N}}$ streng monoton fallend mit

$$\lim_{n\to\infty} b^{-n} = 0$$

und $(b^n)_{n\in\mathbb{N}}$ streng monoton wachsend mit

$$\lim_{n\to\infty} b^n = \infty$$

ist, können wir eine ganze Zahl $z \in \mathbb{Z}$ wählen, so daß

$$b^{-z} \leq s < b^{-z+1}.$$

Wir wollen die Koeffizienten a_n für $n \geq z$ nun rekursiv so definieren, daß stets

$$a_n := \max(A_n)$$

für

$$A_n := \left\{ a \in \mathbb{N} \;\middle|\; s - \sum_{i=z}^{n-1} a_i \cdot b^{-i} - a \cdot b^{-n} \geq 0 \right\}$$

gilt. Wir nehmen dazu an, daß wir für ein festes $n \geq z$ solche Koeffizienten $a_z, \ldots, a_{n-1}$ bereits gefunden haben. Dann ist die Menge A_n nicht leer, da $a = 0$ die gesuchte Bedingung sicher erfüllt. Wir zeigen nun, daß

$$A_n \subseteq \{0, 1, \ldots, b-1\} \tag{20}$$

gilt. Wäre dies nicht der Fall, dann wäre $b \in A_n$ und es würde

$$0 \leq s - \sum_{i=z}^{n-1} a_i \cdot b^{-i} - b \cdot b^{-n} = s - \sum_{i=z}^{n-2} a_i \cdot b^{-i} - (a_{n-1} + 1) \cdot b^{-(n-1)}$$

gelten und somit wäre $a_{n-1} + 1 \in A_{n-1}$ im Widerspruch dazu, daß a_{n-1} das Maximum dieser Menge ist. Also ist (20) erfüllt und es folgt

$$a_n \in \{0, 1, \ldots, b-1\}.$$

Wir erhalten auf diese Weise also einen b-adischen Bruch und für die Folge der zugehörigen Partialsummen gilt nach Konstruktion

$$s - \sum_{i=z}^{n} a_i \cdot b^{-i} - b^{-n} = s - \sum_{i=z}^{n-1} a_i \cdot b^{-i} - (a_n + 1) \cdot b^{-n} < 0$$

und somit

$$0 \leq s - \sum_{i=z}^{n} a_i \cdot b^{-i} \leq b^{-n}.$$

Die Folge in der Mitte ist nun durch zwei Nullfolgen eingeschachtelt und ist somit auch eine Nullfolge, woraus

$$s = \sum_{n=z}^{\infty} a_n \cdot b^{-n}$$

folgt.

b. Wegen $\sum_{n=z}^{\infty} a_n \cdot b^{-n} = \sum_{n=z}^{\infty} c_n \cdot b^{-n}$ und $a_n = c_n$ für $n = z, \ldots, N-1$ gilt auch

$$\sum_{n=N}^{\infty} a_n \cdot b^{-n} = \sum_{n=N}^{\infty} c_n \cdot b^{-n}. \tag{21}$$

Zudem folgt aus $a_N < c_N$ auch

$$c_N - a_N \in \{1, \ldots, b-1\}.$$

Insgesamt erhalten wir damit

$$0 < (c_N - a_N) \cdot b^{-N} = \sum_{N+1}^{\infty} a_n \cdot b^{-n} - \sum_{n=N+1}^{\infty} c_n \cdot b^{-n} = \sum_{N+1}^{\infty} (a_n - c_n) \cdot b^{-n}$$
$$\leq \sum_{n=N+1}^{\infty} (b-1) \cdot b^{-n} \overset{(19)}{=} b^{-N}.$$

Dies bedingt aber $c_N - a_N = 1$ und mithin

$$a_N + 1 = c_N.$$

Setzen wir dies in (21) ein und ziehen auf beiden Seiten $a_N \cdot b^{-N}$ ab, so erhalten wir

$$0 \leq b^{-N} + \sum_{n=N+1}^{\infty} c_n \cdot b^{-n} = \sum_{n=N+1}^{\infty} a_n \cdot b^{-n} \leq \sum_{n=N+1}^{\infty} (b-1) \cdot b^{-n} \overset{(19)}{=} b^{-N} \tag{22}$$

und damit

$$\sum_{n=N+1}^{\infty} c_n \cdot b^{-n} = 0.$$

Wäre einer der Koeffizienten c_n positiv, so wäre der Grenzwert der Reihe durch dieses c_n nach unten beschränkt und nicht null. Also gilt $c_n = 0$ für alle $n \geq N+1$. Damit folgt aus (22) dann

$$\sum_{n=N+1}^{\infty} a_n \cdot b^{-n} = b^{-N} \overset{(19)}{=} \sum_{n=N+1}^{\infty} (b-1) \cdot b^{-n}$$

und somit

$$\sum_{n=N+1} (b - 1 - a_n) \cdot b^{-n} = 0.$$

Da die Koeffizienten der Reihe alle nicht-negativ sind, folgt wie oben auch in diesem Fall, daß sie alle null sind, d.h.

$$a_n = b - 1$$

für alle $n \geq N+1$.

c. Wir zeigen mittels Induktion nach s, daß s eine Darstellung als b-adischer Bruch der Form

$$s = \sum_{n=z}^{0} a_n \cdot b^{-n} = (a_z, a_{z+1} \ldots a_0 \; E - z)_b$$

besitzt. Für $s = 0$ ist dies offensichtlich der Fall. Sei also $s > 0$. Teilen wir s durch b mit Rest, so erhalten wir natürliche Zahlen $q, r \in \mathbb{N}$, so daß

$$s = q \cdot b + r$$

und $0 \leq r < b$. Die Zahl $q = \frac{s-r}{b}$ ist kleiner als s und besitzt mittels Induktion eine Darstellung als b-adischer Bruch der Form

$$q = \sum_{n=z}^{0} a_n \cdot b^{-n}.$$

Dann ist aber

$$s = q \cdot b + r = \sum_{n=z+1}^{1} a_{n-1} \cdot b^{-n} + r = (a_z, a_{z+1} \ldots a_0 r \; E - (z+1))_b$$

ebenfalls eine solche.

d. Ist $s = (a_z, a_{z+1} a_{z+2} \ldots a_N \; E - z)_b$ ein endlicher b-adischer Bruch, dann ist

$$s = \sum_{n=z}^{N} a_n \cdot b^{-n} \in \mathbb{Q}$$

als Summe endlich vieler rationaler Zahlen eine rationale Zahl.
Ist $s = (a_z, a_{z+1} \ldots a_N \; \overline{a_{N+1} \ldots a_{N+m}} \; E - z)_b$ ein periodischer b-adischer Bruch, dann gilt

$$\begin{aligned} s - b^m \cdot s &= \sum_{n=z}^{\infty} a_n \cdot b^{-n} - \sum_{n=z-m}^{\infty} a_{n+m} \cdot b^{-n} \\ &= \sum_{n=z}^{N} a_n \cdot b^{-n} - \sum_{n=z-m}^{N} a_{n+m} \cdot b^{-n} =: q, \end{aligned}$$

wobei dann q als Differenz zweier endlicher b-adischer Brüche eine rationale Zahl ist. Aber damit gilt dann auch

$$s = \frac{q}{1 - b^m} \in \mathbb{Q}.$$

Es bleibt noch zu zeigen, daß eine b-adische Darstellung einer rationalen Zahl

$$s = \sum_{n=z}^{\infty} a_n \cdot b^{-n} \in \mathbb{Q}$$

stets endlich oder periodisch ist. Es reicht dabei, dies für

$$s - \sum_{n=z}^{0} a_n \cdot b^{-n} = \sum_{n=1}^{\infty} a_n \cdot b^n$$

zu zeigen, da der endliche Anfangsteil des b-adischen Bruches nichts an der Eigenschaft ändert endlich oder periodisch zu sein. Wir können also ohne Einschränkung $z > 0$ annehmen.

Zunächst schreiben wir $s = \frac{p}{q}$ als Bruch mit $p, q \in \mathbb{N}$ und setzen

$$r_n := q \cdot b^n \cdot \left(s - \sum_{i=z}^{n} a_i \cdot b^{-i} \right) \in \mathbb{N}$$

für $n \geq z - 1 \geq 0$. Wir stellen zunächst fest, daß

$$r_n = q \cdot b^n \cdot \sum_{i=n+1}^{\infty} a_i \cdot b^{-i} \leq q \cdot b^n \sum_{i=n+1}^{\infty} (b-1) \cdot b^{-i} \overset{(19)}{=} q$$

gilt und daß r_n nicht negativ ist.
Gilt $r_n = 0$ für ein n, dann folgt $a_i = 0$ für alle $i \geq n+1$ und s ist ein endlicher b-adischer Bruch.
Gilt $r_n = q$ für ein n, dann gilt $a_i = b-1$ für alle $i \geq n+1$ und s ist ein periodischer Bruch.
Wir können also $0 < r_n < q$ für alle $n \geq z$ annehmen. Aus der Definition von r_n folgt dann unmittelbar

$$\begin{aligned} b \cdot r_{n-1} &= q \cdot b^n \cdot \left(s - \sum_{i=z}^{n-1} a_i \cdot b^{-i} \right) \\ &= q \cdot b^n \cdot \left(s - \sum_{i=z}^{n} a_i \cdot b^{-i} \right) + q \cdot a_n \\ &= a_n \cdot q + r_n, \end{aligned}$$

d.h. r_n ist der Rest bei Division mit Rest von $b \cdot r_{n-1}$ durch q und a_n ist der ganzzahlige Anteil dabei. Beide sind durch den Wert $b \cdot r_{n-1}$ eindeutig bestimmt. Wegen $0 < r_n < q$ kommen für den Rest nur endlich viele Werte in Frage und somit kommen auch für die Zahlen $b \cdot r_{n-1}$, die geteilt werden sollen, nur endlich viele Werte in Frage, so daß einer dieser Werte irgendwann ein zweites Mal auftaucht. Damit wiederholen sich die Reste und die ganzzahligen Anteile von diesem Index an aber periodisch und s ist ein periodischer b-adischer Bruch.

□

Bemerkung 12.46 (Algorithmus zur Berechnung einer b-adischen Darstellung).
Aus dem Beweis von Satz 12.45 läßt sich ein Algorithmus zur Berechnung der b-adischen Darstellung einer rationalen Zahl s herleiten.

Zunächst zerlegen wir

$$s = x + y$$

in seinen ganzzahligen Anteil $x \in \mathbb{N}$ und seinen gebrochenen Anteil $y = \frac{p}{q} \in [0, 1)$.

Für x berechnen wir eine b-adische Darstellung wie im Beweis von Beweis von Satz 12.45 c.: Wir setzen $x_0 = x$ und berechnen dann rekursiv mittels Division mit Rest natürliche

Zahlen $x_n, a_{1-n} \in \mathbb{N}$ mit

$$x_{n-1} = x_n \cdot b + a_{1-n}$$

mit

$$0 \leq a_{1-n} < b$$

für $n \geq 1$. Wegen $0 \leq x_n < x_{n-1}$ bricht der Prozess nach endlich vielen Schritten mit einem $x_N = 0$ ab. Wir erhalten damit die b-adische Darstellung

$$x = \sum_{n=-N}^{0} a_n \cdot b^{-n}.$$

Für $y = \frac{p}{q}$ setzen wir zunächst $r_0 = q \cdot y = p$ und bestimmen rekursiv mittels Division mit Rest ganze Zahlen a_n, r_n mit

$$b \cdot r_{n-1} = a_n \cdot q + r_n$$

und

$$0 \leq r_n < q$$

für $n \geq z$.[3]Wir stoppen den Prozess, wenn der Rest $r_N = 0$ null wird oder erstmals ein Rest r_{N+1} zum zweiten Mal als r_{N+m+1} auftaucht. Im ersten Fall ist

$$y = \sum_{n=1}^{N} a_n \cdot b^{-n} = (a_1, a_{z+1} \ldots a_N \; E - 1)_b$$

ein endlicher b-adischer Bruch, im zweiten Fall ist

$$y = (a_1, a_{z+1} \ldots a_N \; \overline{a_{N+1} \ldots a_{N+m}} \; E - 1)_b$$

ein periodischer b-adischer Bruch. Um die Darstellung für s zu erhalten, brauchen wir dann nur die Darstellungen für x und y zusammenzusetzen.

Beispiel 12.47 (Berechnung einer 2-adischen Darstellung).
Wir wollen mit dem in Bemerkung 12.46 beschriebenen Algorithmus für die rationale Zahl

$$s = \frac{37}{5} = 7 + \frac{2}{5}$$

die Darstellung als 2-adischer Bruch berechnen.

[3]Woher wissen wir, daß $a_n \in \{0, \ldots, b-1\}$ liegt? Einerseits folgt dies aus dem Beweis von Satz 12.45 d., wo gezeigt wurde, daß die Koeffizienten der b-adischen Darstellung dem ganzzahligen Anteil in obiger Division mit Rest entsprechen, und dieser ist eindeutig bestimmt. Also muß der ganzzahlige Anteil, den wir hier bei der Division mit Rest ausrechnen auch kleiner als b sein. Andererseits kann man dies auch direkt nachrechnen. Wir wissen, daß stets $r_{n-1} < q$ gilt und damit erhalten wir

$$b = \frac{b \cdot q}{q} > \frac{b \cdot r_{n-1}}{q} = \frac{a_n \cdot q + r_n}{q} \geq \frac{a_n \cdot q}{q} = a_n.$$

Für den ganzzahligen Anteil $x = x_0 = 7$ erhalten wir durch wiederholte Division mit Rest durch $b = 2$:

$$\begin{aligned} x_0 &= 7 = 3 \cdot 2 + 1 = x_1 \cdot b + a_0 \\ x_1 &= 3 = 1 \cdot 2 + 1 = x_2 \cdot b + a_{-1} \\ x_2 &= 1 = 0 \cdot 2 + 1 = x_3 \cdot b + a_{-2} \end{aligned}$$

Der Algorithmus für x bricht hier wegen $x_3 = 0$ ab und wir erhalten die 2-adische Darstellung

$$x = 7 = 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = (1,11\ E2)_2.$$

Für $y = \frac{p}{q} = \frac{2}{5}$ und $r_0 = p = 2$ erhalten wir durch wiederholte Division mit Rest durch $q = 5$:

$$\begin{aligned} b \cdot r_0 &= 2 \cdot 2 = 0 \cdot 5 + 4 = a_1 \cdot 5 + r_1 \\ b \cdot r_1 &= 2 \cdot 4 = 1 \cdot 5 + 3 = a_2 \cdot 5 + r_2 \\ b \cdot r_2 &= 2 \cdot 3 = 1 \cdot 5 + 1 = a_3 \cdot 5 + r_3 \\ b \cdot r_3 &= 2 \cdot 1 = 0 \cdot 5 + 2 = a_4 \cdot 5 + r_4 \end{aligned}$$

Der Algorithmus bricht wegen $r_4 = r_0$ hier ab und wir erhalten die periodische b-adische Darstellung

$$y = (\overline{0,110}\ E\ -1)_2.$$

Setzen wir diese zusammen, so erhalten wir

$$s = \frac{37}{5} = (1,11\overline{0110}\ E2)_2.$$

Bemerkung 12.48 (Binärzahlen).
Für einige Basen b haben die b-adischen Zahlsysteme eigene Namen, die sich aus dem Wert der Basis ableiten. Bei $b = 10$ spricht man von *Dezimalzahlen*, bei $b = 2$ von *Binärzahlen* und bei $b = 16$ von *Hexadezimalzahlen*.

Bemerkung 12.49 (Binärzahlen in der Informatik).
Reelle Zahlen werden in Rechnern als Binärzahlen gespeichert. Im Standard *IEEE754*, double-precision binary floating point, stehen für eine reelle Zahl 8 Byte (64 Bit) als Speicher zur Verfügung. Dabei wird ein Bit für das Vorzeichen verwendet, damit auch negative Zahlen gespeichert werden können. Für den Exponenten (den Wert nach dem E) stehen 11 Bit zur Verfügung, von denen wiederum eines für das Vorzeichen des Exponenten verwendet wird. Die Zahlenfolge vor dem E in der Binärdarstellung nennt man die *Mantisse* und für diese stehen 52 Bit zur Verfügung womit 53 Stellen gespeichert werden können, indem man die führende Zahl weg läßt, da sie stets eine 1 sein muß.

Eine solche sieht dann in etwa wie folgt aus:

$$s = \underbrace{-}_{1\text{ Bit}} \cdot 1, \underbrace{01100011\ldots001}_{52\text{ Bit}} \cdot 2^{\underbrace{-(1,001111001\ E\ 10)_2}_{11\text{ Bit}}}.$$

Der maximale positve Wert, der so dargestellt werden kann, ist

$$\left(2 - 2^{-52}\right) \cdot 2^{1023} \sim 1,8 \cdot 10^{308},$$

und der minimale positive Wert ist

$$2^{-1022} \sim 2,3 \cdot 10^{-308}.$$

Man beachte aber, daß auch bei weitem nicht alle positiven reellen Zahlen zwischen diesen beiden Werten im Rechner dargestellt werden können, so daß bei Rechnungen of schon beim Abspeichern der Eingabedaten Fehler entstehen (müssen), weil die binäre Darstellung der eingegebenen Zahlen nicht in das Raster paßt. Wie sich solche Fehler auf nachfolgende Rechnungen auswirken und was man tun sollte, um sie im Griff zu behalten, untersucht die Numerische Mathematik.

Aufgaben

Aufgabe 12.50.
Untersuche die folgenden Reihen auf Konvergenz. Die Berechnung der Grenzwerte im Falle der Konvergenz ist nicht erforderlich.

a. $\sum_{n=1}^{\infty} \frac{1-n^4}{100n^4}$.

b. $\sum_{n=1}^{\infty} \frac{n!}{n^{n+1}}$.

c. $\sum_{n=1}^{\infty} \frac{(2n)^n}{(-3)^{n+1}}$.

Aufgabe 12.51 (Nicht-konvergentes Cauchy-Produkt konvergenter Reihen).
Betrachte die Folgen $(a_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ mit

$$a_n = (-1)^n \cdot \frac{1}{\sqrt{n+1}}$$

und

$$c_n = \sum_{k=0}^{n} a_k \cdot a_{n-k} = \sum_{k=0}^{n} \frac{(-1)^k}{\sqrt{k+1}} \cdot \frac{(-1)^{n-k}}{\sqrt{n-k+1}}.$$

Zeige, die Reihe $\sum_{n=0}^{\infty} a_n$ ist konvergent, aber das Cauchy-Produkt $\sum_{n=0}^{\infty} c_n$ ist divergent.

Aufgabe 12.52.
Es sei $q \in \mathbb{K}$ mit $|q| < 1$.

a. Berechne das Cauchy-Produkt $\left(\sum_{n=0}^{\infty} q^n\right)^2$.

b. Berechne den Wert der Reihe $\sum_{n=0}^{\infty} nq^n$.

Aufgabe 12.53.
Seien $\sum_{n=0}^{\infty} a_n t^n$ und $\sum_{n=1}^{\infty} n a_n t^{n-1}$ Potenzreihen in $\mathbb{K}$. Zeige die folgenden Aussagen:

a. Konvergiert $\sum_{n=0}^{\infty} a_n y^n$ für ein $y \in \mathbb{K}$, so konvergiert $\sum_{n=1}^{\infty} n a_n x^{n-1}$ absolut für alle $x \in \mathbb{K}$ mit $|x| < |y|$.

b. Die gegebenen Potenzreihen haben denselben Konvergenzradius.

c. Konvergieren die Potenzreihen vielleicht sogar stets für dieselben $x \in \mathbb{K}$?

Hinweis: Schaut euch hilfestellend den Beweis von Lemma 12.29 an und verwendet Aufgabe 12.52.

Aufgabe 12.54.
Bestimme die Konvergenzradien folgender Potenzreihen.

a. $\sum_{n=0}^{\infty} n^k \cdot t^n$ für $k \in \mathbb{N}$.

b. $\sum_{n=1}^{\infty} \frac{n!}{(2n)^n} \cdot t^n$.

Aufgabe 12.55.
Beweise für $x, y \in \mathbb{K}$ die Gleichung $\sin(x+y) = \sin(x) \cdot \cos(y) + \cos(x) \cdot \sin(y)$.

Aufgabe 12.56.
Zeige, daß die beiden Reihen

$$\sum_{n=0}^{\infty} \frac{\sin(n)}{n!}$$

und

$$\sum_{n=0}^{\infty} \frac{\cos(n)}{n!}$$

absolut konvergent sind und berechne ihren Grenzwert.

Aufgabe 12.57.
Zeige für alle $x, y \in \mathbb{R}$ die Gleichung

$$\cos(x) - \cos(y) = -2 \cdot \sin\left(\frac{x+y}{2}\right) \cdot \sin\left(\frac{x-y}{2}\right).$$

Aufgabe 12.58.
Für $x \in \mathbb{R}$ mit $e^{ix} \neq 1$ und $n \in \mathbb{N}$ zeige

$$\sum_{k=1}^{n} \cos(k \cdot x) = \frac{\sin\left(\frac{2n+1}{2} \cdot x\right)}{2 \cdot \sin\left(\frac{x}{2}\right)} - \frac{1}{2}.$$

Aufgabe 12.59 (Binomialreihe).
Für eine beliebige reelle Zahl $a \in \mathbb{R}$ und eine natürliche Zahl $k \in \mathbb{N}$ definieren wir den *Binomialkoeffizienten*

$$\binom{a}{k} = \frac{a \cdot (a-1) \cdot \ldots (a-k+1)}{k!} \in \mathbb{R}.$$

a. Zeige, für $a \in \mathbb{R} \setminus \mathbb{N}$ hat die Potenzreihe $\sum_{n=0}^{\infty} \binom{a}{n} \cdot t^n$ den Konvergenzradius 1, (siehe auch Aufgabe 18.50).

b. Berechne die Binomialkoeffizienten $\binom{a}{k}$ für den Fall $a = 3$ für alle k und berechne damit auch die Potenzreihe in (a) exakt. Welchen Konvergenzradius hat die Binomialreihe für $a \in \mathbb{N}$?

Aufgabe 12.60.

a. Zeige, daß die geometrische Reihe $\sum_{n=0}^{\infty} q^n$ mit $q = \frac{2}{3}$ eine konvergente Majorante der Reihe $\sum_{n=1}^{\infty} \frac{1}{n!}$ ist und leite daraus eine obere Schranke für den Grenzwert der Reihe ab.

b. Zeige, die Folge $(a_n)_{n\geq 1}$ mit $a_n = \left(1 + \frac{1}{n}\right)^n$ ist monton wachsend und konvergent.

Hinweis zu Teil b., man kann die Bernoulli-Ungleichung an geeigneter Stelle verwenden.

Aufgabe 12.61.
Bestimme für die folgenden Reihen, ob sie konvergent, absolut konvergent oder divergent sind:

a. $\sum_{n=1}^{\infty} (-1)^n \cdot \left(1 - \frac{1}{n}\right)^n$.

b. $\sum_{n=0}^{\infty} \frac{3^n}{3 \cdot n!}$.

c. $\sum_{n=1}^{\infty}(-1)^n \cdot \frac{1}{\sqrt{n}}$.

d. $\sum_{n=1}^{\infty}\left(\frac{n+2}{n}\right)^n$.

e. $\sum_{n=1}^{\infty}\frac{1}{n+\sqrt{n}}$.

f. $\sum_{n=1}^{\infty}\frac{n!}{n^n}$.

Aufgabe 12.62.
Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ zwei Folgen in $\mathbb{C}$, so daß die Reihen $\sum_{n=0}^{\infty} a_n^2$ und $\sum_{n=0}^{\infty} b_n^2$ absolut konvergent sind. Zeige, dann ist auch die Reihe $\sum_{n=0}^{\infty} a_n \cdot b_n$ absolut konvergent.

Aufgabe 12.63.
Gib ein Beispiel für eine absolut konvergente Reihe $\sum_{n=0}^{\infty} a_n$ mit $\left|\frac{a_{n+1}}{a_n}\right| > 1$ für unendlich viele $n \in \mathbb{N}$. (Wie immer mit Begründung!)

Aufgabe 12.64.
Bestimme den Konvergenzradius der Reihe $\sum_{n=1}^{\infty} t^n$ über $\mathbb{R}$ und untersuche die Reihe auf Konvergenz in den Randpunkten des Konvergenzintervalls.

Aufgabe 12.65.
Bestimme für die folgenden Potenzreihen über $\mathbb{R}$ den Konvergenzradius r und untersuche das Konvergenzverhalten in den Randpunkten des Konvergenzintervalls $(-r, r)$:

a. $\sum_{n=1}^{\infty}\frac{1}{n\cdot 4^n} \cdot t^n$.

b. $\sum_{n=1}^{\infty}\frac{n}{2^n} \cdot t^{n+1}$.

c. $\sum_{n=0}^{\infty}\sqrt{n} \cdot t^n$.

d. $\sum_{n=0}^{\infty} n! \cdot t^n$.

e. $\sum_{n=1}^{\infty}\left(\frac{n+1}{n}\right)^n \cdot t^n$.

f. $\sum_{n=0}^{\infty}(-1)^n \cdot t^n$.

Aufgabe 12.66.
Zeige, wenn $\sum_{n=0}^{\infty} a_n \cdot t^n$ den Konvergenzradius r hat und $c \in \mathbb{R} \setminus \{0\}$ ist, dann hat $\sum_{n=0}^{\infty}\frac{a_n}{c^n} \cdot t^n$ den Konvergenzradius $|c| \cdot r$.

Aufgabe 12.67.
Bestimme die 5-adische Darstellung der rationalen Zahl $\frac{111336}{1000}$ mittels des Algorithmus' aus der Vorlesung.

Aufgabe 12.68.

a. Berechne die 7-adische Darstellung der Zahl 4973.

b. Berechne das Produkt der Zahlen $(2,33\ E2)_5$ und $(1,01\ E1)_5$ im Dezimalsystem.

c. Berechne die 3-adische Zahldarstellung der rationalen Zahl $\frac{123}{75}$.

d. Berechne die Dezimalzahldarstellung des 3-adischen Bruches $(2,12121\ E2)_3$.

e. Berechne die b-adische Zahldarstellung von $\frac{1}{7}$ für $b = 2, 7, 10, 16$.

§ 13 Grenzwerte von Funktionen

Wir werden uns in den folgenden Paragraphen im wesentlichen dem Studium von Abbildungen

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

widmen und nennen diese eher *Funktionen* als Abbildungen. Wir könnten dabei viele Begriffe auch gleich wieder für den Körper der komplexen Zahlen $\mathbb{C}$ statt $\mathbb{R}$ einführen und untersuchen, wollen dies aber zurück stellen, um näher an dem Vorwissen aus der Schulzeit zu bleiben.

A) Häufungspunkte von Teilmengen von $\mathbb{R}$

Definition 13.1 (Häufungspunkte).
Es sei $U \subseteq \mathbb{R}$ eine Teilmenge von $\mathbb{R}$ und $a \in \mathbb{R}$. Wir nennen a einen *Häufungspunkt* von U, wenn

$$\forall\, \varepsilon > 0 \;\; \exists\, x \in U \setminus \{a\} \;:\; 0 < |x - a| < \varepsilon.$$

Man beachte, daß a kein Element von U sein muß.

Bemerkung 13.2 (ε-Umgebung).
Für $\varepsilon > 0$ und $a \in \mathbb{R}$ nennen wir das Intervall

$$U_\varepsilon(a) := (a - \varepsilon, a + \varepsilon) = \{x \in \mathbb{R} \mid |x - a| < \varepsilon\}$$

die ε-*Umgebung* von a.

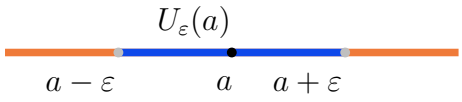


Abbildung 4: Die ε-Umgebung $U_\varepsilon(a)$ von a.

Mit dieser Sprechweise gilt also:

Genau dann ist a ein Häufungspunkt von U, wenn jede ε-Umgebung von a einen von a verschiedenen Punkt aus U enthält.

Beispiel 13.3.
Jede reelle Zahl ist Häufungspunkt von $\mathbb{Q}$.

Dazu seien $a \in \mathbb{R}$ und $\varepsilon > 0$ gegeben. Wir wenden Satz 9.6 an und finden eine rationale Zahl x im Intervall $(a, a + \varepsilon)$, d.h. $|x - a| < \varepsilon$ und $x \neq a$.

Proposition 13.4 (Folgenkriterium für Häufungspunkte).
Ein $a \in \mathbb{R}$ ist genau dann Häufungspunkt von $U \subseteq \mathbb{R}$, wenn es eine Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n \in U \setminus \{a\}$ und $\lim\limits_{n\to\infty} a_n = a$ gibt, d.h. a ist Grenzwert einer Folge in $U \setminus \{a\}$.

Beweis: Ist a ein Häufungspunkt von U und $n \in \mathbb{N}$, so gibt es zu $\varepsilon = \frac{1}{n+1} > 0$ ein $x \in U$ mit $0 < |x - a| < \varepsilon = \frac{1}{n+1}$. Wir wählen a_n als dieses x. Dann konvergiert die Folge $(a_n)_{n\in\mathbb{N}}$ offenbar gegen a nach dem Einschließungssatz, da

$$0 < |a_n - a| < \frac{1}{n+1} \longrightarrow 0.$$

Gibt es umgekehrt eine Folge $(a_n)_{n\in\mathbb{N}}$ in $U \setminus \{a\}$, die gegen a konvergiert, so gibt es für jedes $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ mit $0 < |a_n - a| < \varepsilon$ für alle $n \geq n_\varepsilon$. Setzen wir nun $x = a_{n_\varepsilon} \in U$, so folgt $0 < |x - a| < \varepsilon$ und somit ist a ein Häufungspunkt von U. □

Das folgende Beispiel zeigt, welche Art von Häufungspunkten einer Menge U, die nicht bereits in U liegen, wir typischerweise erwarten.

Beispiel 13.5.
Sind $a, b \in \mathbb{R}$ mit $a < b$, so enthält $[a, b]$ genau die Häufungspunkte von (a, b), d.h. zu den Punkten im Intervall kommen noch die Randpunkte hinzu.

Die analogen Aussagen für halboffene, abgeschlossene und uneigentliche Intervalle gelten ebenfalls und mit analogem Beweis.

Beweis: Ist $c \in \mathbb{R}$ ein Häufungspunkt von (a, b), so gibt es nach dem Folgenkriterium 13.4 eine Folge $(a_n)_{n\in\mathbb{N}}$ in $(a, b) \subset [a, b]$, die gegen c konvergiert, und nach Satz 11.28 ist dann $c \in [a, b]$.
Ist $a \leq c < b$, so gilt für $n \geq 1$

$$(a, b) \ni c + \frac{b - c}{2 \cdot n} \longrightarrow c,$$

also ist c ein Häufungspunkt von (a, b). Analog gilt $(a, b) \ni b - \frac{b-a}{2\cdot n} \longrightarrow b$, so daß auch b ein Häufungspunkt von (a, b) ist. □

B) Grenzwerte von Funktionen

Definition 13.6 (ε-δ-Kriterium für Grenzwerte von Funktionen).
Sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ eine Funktion und a ein Häufungspunkt von U.
Wir nennen $y \in \mathbb{R}$ den *Grenzwert* von f in a, falls

$$\forall\, \varepsilon > 0 \;\; \exists\, \delta_\varepsilon > 0 \;:\; \forall\, x \in U \text{ mit } 0 < |x - a| < \delta_\varepsilon \;\text{ gilt }\; |f(x) - y| < \varepsilon.$$

Wir schreiben dann

$$\lim_{x\to a} f(x) = y$$

oder "$f(x) \longrightarrow y$ für $x \to a$" und sagen, $f(x)$ *konvergiert* gegen y für x gegen a.

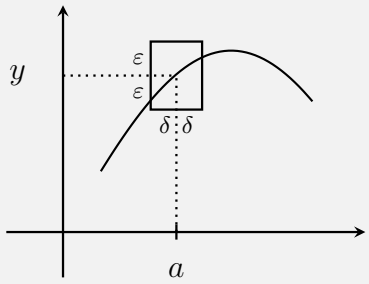


Abbildung 5: ε-δ-Kriterium für Grenzwerte

Proposition 13.7 (Folgenkriterium für Grenzwerte von Funktionen).
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ eine Funktion und a ein Häufungspunkt von U.
Dann sind die beiden folgenden Aussagen gleichwertig:

a. $\lim\limits_{x\to a} f(x) = y$.

b. $\forall\, (a_n)_{n\in\mathbb{N}}$ mit $a_n \in U \setminus \{a\}$ und $\lim\limits_{n\to\infty} a_n = a$ gilt $\lim\limits_{n\to\infty} f(a_n) = y$.

Beweis: **a. $\Longrightarrow$ b.:** Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $U\setminus\{a\}$ mit $\lim\limits_{n\to\infty} a_n = a$. Wir müssen $\lim\limits_{n\to\infty} f(a_n) = y$ zeigen. Dazu sei $\varepsilon > 0$ gegeben.
Wegen $\lim\limits_{x\to a} f(x) = y$ gibt es ein $\delta_\varepsilon > 0$, so daß aus $x \in U$ mit $0 < |x-a| < \delta_\varepsilon$ auch $|f(x) - y| < \varepsilon$ folgt.
Wegen $\lim\limits_{n\to\infty} a_n = a$ gibt es zu δ_ε nun ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $n \geq n_\varepsilon$ auch $|a_n - a| < \delta_\varepsilon$ gilt.
Sei nun $n \geq n_\varepsilon$ dann erfüllt $a_n \in U$ die Bedingung $0 < |a_n - a| < \delta_\varepsilon$ und somit ist auch $|f(a_n) - y| < \varepsilon$. Damit ist $f(a_n) \longrightarrow y$ gezeigt.

b. $\Longrightarrow$ a.: Wir nehmen an, y wäre nicht der Grenzwert von f in a. Dann gilt:

$$\exists\, \varepsilon > 0 \;:\; \forall\, \delta_\varepsilon > 0\; \exists\, x_{\delta_\varepsilon} \in U \;\text{ mit }\; 0 < |x_{\delta_\varepsilon} - a| < \delta_\varepsilon, \;\text{ aber } |f(x_{\delta_\varepsilon}) - y| \geq \varepsilon.$$

Für $n \geq 1$ und $\delta_\varepsilon = \frac{1}{n}$ setzen wir $a_n := x_{\delta_\varepsilon} = x_{\frac{1}{n}} \in U \setminus \{a\}$. Dann gilt

$$0 < |a_n - a| < \frac{1}{n} \longrightarrow 0,$$

so daß $a_n \longrightarrow a$, und zugleich gilt

$$|f(a_n) - y| \geq \varepsilon$$

für alle $n \in \mathbb{N}$. Dies ist ein Widerspruch dazu, daß $f(a_n)$ gegen y konvergieren muß.

□

Beispiel 13.8.

a. Betrachte $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$ und $a = 3$. Für eine Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n \longrightarrow 3$ gilt dann wegen der Grenzwertsätze für Folgen 11.15

$$f(a_n) = a_n^2 = a_n \cdot a_n \longrightarrow 3 \cdot 3 = 9.$$

Mithin ist 9 der Grenzwert von f in 3, d.h.

$$\lim_{x\to 3} x^2 = 9 = f(3).$$

b. Betrachte die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 1, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0 \end{cases}$$

und $a = 0$. Ist nun $(a_n)_{n\in\mathbb{N}}$ eine Folge mit $a_n \longrightarrow 0$ und $a_n \neq 0$, dann gilt

$$f(a_n) = 1 \longrightarrow 1.$$

Mithin ist 1 der Grenzwert von f in 0, d.h.

$$\lim_{x\to 0} f(x) = 1 \neq 0 = f(0).$$

c. Betrachte die Funktion

$$f : \mathbb{R} \setminus \{1\} \longrightarrow \mathbb{R} : x \mapsto \frac{x^2 - 1}{x - 1}$$

und $a = 1$. Da es in $\mathbb{R} \setminus \{1\}$ offenbar eine Folge gibt, die gegen 1 konvergiert, ist a ein Häufungspunkt von $\mathbb{R} \setminus \{1\}$. Sei nun $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R} \setminus \{1\}$ mit $a_n \longrightarrow 1$, so gilt

$$f(a_n) = \frac{a_n^2 - 1}{a_n - 1} = a_n + 1 \longrightarrow 2.$$

Mithin ist 2 der Grenzwert von f in 1, d.h.

$$\lim_{x\to 1} f(x) = 2.$$

Man beachte, daß in diesem Fall $a = 1$ gar nicht im Definitionsbereich von f liegt.

d. Betrachte die Funktion

$$f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x < 0, \\ 1, & \text{falls } x > 0 \end{cases}$$

und $a = 0$. Dann gilt für $a_n := -\frac{1}{n} \longrightarrow 0$ und $f(a_n) = 0 \longrightarrow 0$ sowie $b_n := \frac{1}{n} \longrightarrow 0$ und $f(b_n) = 1 \longrightarrow 1$. Mithin existiert der Grenzwert von f in $a = 0$ nicht.

!!! Warnung !!!

In manchen Büchern und von manchen Dozenten wird der Begriff des Grenzwertes leicht anders definiert (siehe z.B. [Dec10] oder [Gat08])! Wenn f im Punkt a definiert ist, muß bei uns **nicht** notwendig $\lim_{x\to a} f(x) = f(a)$ gelten (siehe Beispiel 13.8 b.), was nach deren Definition gelten muß!

C) Die Grenzwertsätze für Funktionen

Definition 13.9.
Für zwei Funktionen $f : U \longrightarrow \mathbb{R}$ und $g : V \longrightarrow \mathbb{R}$ sowie $c \in \mathbb{R}$ definieren wir

$$c \cdot f : U \longrightarrow \mathbb{R} : x \mapsto c \cdot f(x),$$
$$f + g : U \cap V \longrightarrow \mathbb{R} : x \mapsto f(x) + g(x),$$
$$f - g : U \cap V \longrightarrow \mathbb{R} : x \mapsto f(x) - g(x)$$

und

$$f \cdot g : U \cap V \longrightarrow \mathbb{R} : x \mapsto f(x) \cdot g(x).$$

Falls zudem $g(x) \neq 0$ für $x \in U \cap V$, so definieren wir

$$\frac{f}{g} : U \cap V \longrightarrow \mathbb{R} : x \mapsto \frac{f(x)}{g(x)}.$$

Proposition 13.10 (Grenzwertsätze für Funktionen).
Es seien $f : U \longrightarrow \mathbb{R}$ und $g : U \longrightarrow \mathbb{R}$ zwei Funktionen, a ein Häufungspunkt von U und $c \in \mathbb{R}$.

a. Der Grenzwert von f in a ist eindeutig bestimmt, d.h. falls $\lim_{x\to a} f(x) = y$ und $\lim_{x\to a} f(x) = z$, so ist $y = z$.

b. Wenn $\lim_{x\to a} f(x)$ und $\lim_{x\to a} g(x)$ existieren, so gelten:
   (i) $\lim_{x\to a}(c \cdot f)(x) = c \cdot \lim_{x\to a} f(x)$.
   (ii) $\lim_{x\to a}(f + g)(x) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$.
   (iii) $\lim_{x\to a}(f - g)(x) = \lim_{x\to a} f(x) - \lim_{x\to a} g(x)$.
   (iv) $\lim_{x\to a}(f \cdot g)(x) = \lim_{x\to a} f(x) \cdot \lim_{x\to a} g(x)$.

c. Falls zudem $\lim\limits_{x\to a} f(x) \neq 0$, so ist a ein Häufungspunkt der Menge $V = \{x \in U \mid f(x) \neq 0\}$ und es gilt

$$\lim_{x\to a} \frac{1}{f}(x) = \frac{1}{\lim\limits_{x\to a} f(x)}.$$

Beweis: a. Dies folgt aus dem Folgenkriterium für Grenzwerte von Funktionen 13.7 und der Eindeutigkeit des Grenzwertes bei Folgen 11.8. Genauer, da a ein Häufungspunkt von U ist, gibt es nach Proposition 13.4 eine Folge $(a_n)_{n\in\mathbb{N}}$ in $U \setminus \{a\}$ mit $a_n \longrightarrow a$, und mit den eben erwähnten Sätzen folgt dann

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

b. Analog folgen die Aussagen aus dem Folgenkriterium für Grenzwerte von Funktionen 13.7 und den Grenzwertsätzen für Folgen 11.15 unter Berücksichtigung von Proposition 13.4.

c. Nach Proposition 13.4 gibt es eine Folge $(a_n)_{n\in\mathbb{N}}$ in $U\setminus\{a\}$, die gegen a konvergiert, und nach dem Folgenkriterium 13.7 gilt dann

$$f(a_n) \longrightarrow \lim_{x\to a} f(x) =: y.$$

Wegen $y \neq 0$ gibt es wegen der Grenzwertsätze für Folgen 11.15 ein n_0, so daß $f(a_n) \neq 0$ für alle $n \geq n_0$, so daß $(a_n)_{n\geq n_0}$ eine Folge in V ist mit $a_n \longrightarrow a$. Nach Proposition 13.4 ist dann a ein Häufungspunkt von V. Die Aussage zum Grenzwert folgt dann wieder aus den Grenzwertsätzen für Folgen 11.15 und dem Folgenkriterium 13.7.

□

Definition 13.11.
Ist t eine Veränderliche und sind $a_0, \ldots, a_n \in \mathbb{R}$, so nennen wir einen Ausdruck der Form

$$\sum_{k=0}^{n} a_k \cdot t^k = a_n \cdot t^n + a_{n-1} \cdot t^{n-1} + \ldots + a_1 \cdot t + a_0$$

ein *Polynom* in der Veränderlichen t mit Koeffizienten in $\mathbb{R}$. Ist $a_n \neq 0$, so heißt

$$\deg\left(\sum_{k=0}^{n} a_k \cdot t^k\right) := n$$

der *Grad* des Polynoms, und wir setzen zudem $\deg(0) := -\infty$. Mit

$$\mathbb{R}[t] := \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; n \in \mathbb{N}, a_0, \ldots, a_n \in \mathbb{R} \right\}$$

bezeichnen wir die Menge aller Polynome in der Veränderlichen t mit Koeffizienten in $\mathbb{R}$, so daß der Grad eine Abbildung $\deg : \mathbb{R}[t] \longrightarrow \mathbb{N} \cup \{-\infty\}$ ist.
Für ein Polynom $f = \sum_{k=0}^{n} a_k \cdot t^k \in \mathbb{R}[t]$ und ein $x \in \mathbb{R}$ setzen wir

$$f(x) := \sum_{k=0}^{n} a_k \cdot x^k.$$

Sind $f, g \in \mathbb{R}[t]$ zwei Polynome, $g \neq 0$ nicht das Nullpolynom, so nennen wir die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto f(x)$$

eine *Polynomfunktion* und die Funktion

$$\frac{f}{g} : \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) = 0\} \longrightarrow \mathbb{R} : x \mapsto \frac{f(x)}{g(x)}$$

nennen wir eine *rationale Funktion*.
Ist $h : \mathbb{R} \longrightarrow \mathbb{R}$ irgendeine Funktion, so nennen wir eine reelle Zahl $x \in \mathbb{R}$ mit $h(x) = 0$ eine *Nullstelle* von h.

Bemerkung 13.12.
Man zeigt in der Vorlesung Algebraische Strukturen, daß die Menge der *Nullstellen* von $0 \neq g \in \mathbb{R}[t]$ eine endliche Menge ist. Genauer zeigt man:

$$\big|\{x \in \mathbb{R} \mid g(x) = 0\}\big| \leq \deg(g) < \infty.$$

Beispiel 13.13.

a. Ist $f = \sum_{k=0}^{n} a_k \cdot t^k$ ein Polynom und $a \in \mathbb{R}$, so gilt

$$\lim_{x \to a} f(x) = f(a).$$

Dies folgt aus den Grenzwertsätzen für Funktionen 13.10, da offenbar $\lim_{x \to a} \mathrm{id}(x) = \lim_{x \to a} x = a$ und f sich als endliche Summe von Produkten dieser Funktion mit sich selbst und mit Konstanten schreiben läßt.

b. Für jede rationale Funktion

$$\frac{f}{g} : \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) = 0\} \longrightarrow \mathbb{R} : x \mapsto \frac{f(x)}{g(x)}$$

und jedes $a \in \mathbb{R}$ mit $g(a) \neq 0$ folgt dann aus Teil a. und Satz 13.10 c., daß a ein Häufungspunkt von $\mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) \neq 0\}$ ist und daß

$$\lim_{x \to a} \frac{f}{g}(x) = \frac{f(a)}{g(a)} = \frac{f}{g}(a).$$

D) Uneigentliche Grenzwerte

Definition 13.14 (Grenzwerte für $x \longrightarrow \pm\infty$).
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ und $y \in \mathbb{R}$.

a. Wir nennen U *nach oben unbeschränkt* bzw. *nach unten unbeschränkt*, wenn die Menge $U \cap [0, \infty)$ bzw. $U \cap (-\infty, 0]$ nicht beschränkt ist.

b. Ist U nach oben unbeschränkt, so nennen wir y den *Grenzwert* von f in ∞, wenn

$$\forall\, \varepsilon > 0 \;\; \exists\, s_\varepsilon > 0 \;:\; \forall\, x \in U \text{ mit } x > s_\varepsilon \text{ gilt } |f(x) - y| < \varepsilon.$$

Wir schreiben dann $\lim_{x \to \infty} f(x) = y$.

c. Ist U nach unten unbeschränkt, so nennen wir y den *Grenzwert* von f in $-\infty$, wenn

$$\forall\, \varepsilon > 0 \;\; \exists\, s_\varepsilon < 0 \;:\; \forall\, x \in U \text{ mit } x < s_\varepsilon \text{ gilt } |f(x) - y| < \varepsilon.$$

Wir schreiben dann $\lim_{x \to -\infty} f(x) = y$.

Bemerkung 13.15 (Folgenkriterium und Grenzwertsätze für Grenzwerte in $\pm\infty$).
Das Folgenkriterium für Grenzwerte von Funktionen gilt analog auch für die Grenzwerte in $\pm\infty$. D.h.

$$\lim_{x \to \infty} f(x) = y \iff \forall\, (a_n)_{n \in \mathbb{N}} \text{ mit } a_n \in U \text{ und } a_n \to \infty \text{ gilt } f(a_n) \longrightarrow y$$

und

$$\lim_{x\to-\infty} f(x) = y \iff \forall\, (a_n)_{n\in\mathbb{N}} \text{ mit } a_n \in U \text{ und } a_n \to -\infty \text{ gilt } f(a_n) \longrightarrow y.$$

Zudem gelten auch die Grenzwertsätze für Funktionen 13.10 für Grenzwerte in $\pm\infty$.

Definition 13.16 (Uneigentliche Grenzwerte).
Sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ und a ein Häufungspunkt von U.

a. Wir nennen ∞ den *uneigentlichen Grenzwert* von f in a, wenn

$$\forall\, s > 0 \;\; \exists\, \delta_s > 0 \;:\; \forall\, x \in U \text{ mit } 0 < |x-a| < \delta_s \text{ gilt } f(x) > s.$$

Wir schreiben dann $\lim\limits_{x\to a} f(x) = \infty$.

b. Wir nennen $-\infty$ den *uneigentlichen Grenzwert* von f in a, wenn

$$\forall\, s < 0 \;\; \exists\, \delta_s > 0 \;:\; \forall\, x \in U \text{ mit } 0 < |x-a| < \delta_s \text{ gilt } f(x) < s.$$

Wir schreiben dann $\lim\limits_{x\to a} f(x) = -\infty$.

c. Ist U nach oben unbeschränkt, so nennen wir ∞ den *uneigentlichen Grenzwert* von f in ∞, wenn

$$\forall\, s > 0 \;\; \exists t > 0 \;:\; \forall\, x \in U \text{ mit } x > t \text{ gilt } f(x) > s.$$

Wir schreiben dann $\lim\limits_{x\to\infty} f(x) = \infty$.

d. Ist U nach oben unbeschränkt, so nennen wir $-\infty$ den *uneigentlichen Grenzwert* von f in ∞, wenn

$$\forall\, s < 0 \;\; \exists t > 0 \;:\; \forall\, x \in U \text{ mit } x > t \text{ gilt } f(x) < s.$$

Wir schreiben dann $\lim\limits_{x\to\infty} f(x) = -\infty$.

e. Ist U nach unten unbeschränkt, so nennen wir ∞ den *uneigentlichen Grenzwert* von f in $-\infty$, wenn

$$\forall\, s > 0 \;\; \exists t < 0 \;:\; \forall\, x \in U \text{ mit } x < t \text{ gilt } f(x) > s.$$

Wir schreiben dann $\lim\limits_{x\to-\infty} f(x) = \infty$.

f. Ist U nach unten unbeschränkt, so nennen wir $-\infty$ den *uneigentlichen Grenzwert* von f in $-\infty$, wenn

$$\forall\, s < 0 \;\; \exists t < 0 \;:\; \forall\, x \in U \text{ mit } x < t \text{ gilt } f(x) < s.$$

Wir schreiben dann $\lim\limits_{x\to-\infty} f(x) = -\infty$.

Bemerkung 13.17 (Folgenkriterium und Grenzwertsätze für uneigentliche GWe).
Auch für uneigentliche Grenzwerte gelten naheliegende Folgenkriterien:

a. $\lim\limits_{x\to a} f(x) = \infty \iff \forall\, (a_n)_{n\in\mathbb{N}}$ mit $a_n \in U \setminus \{a\}$ und $a_n \to a$ gilt $f(a_n) \to \infty$.

b. $\lim\limits_{x\to\infty} f(x) = \infty \iff \forall\, (a_n)_{n\in\mathbb{N}}$ mit $a_n \in U$ und $a_n \to \infty$ gilt $f(a_n) \to \infty$.

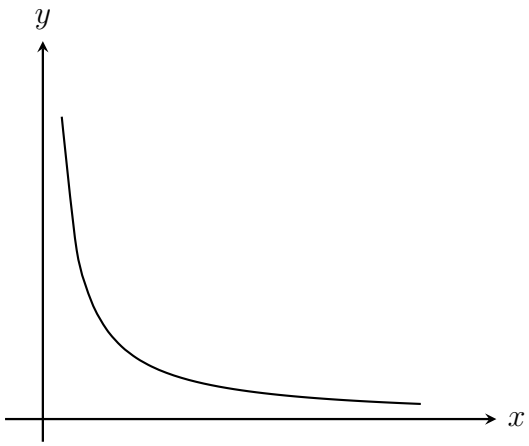
Die übrigen Fälle ergeben sich analog. Außerdem verallgemeinern sich auch die Grenzwertsätze für Funktionen 13.10 auf uneigentliche Grenzwerte in der naheliegenden Weise, wenn wir die Konventionen aus Bemerkung 11.35 berücksichtigen.

Beispiel 13.18.
Für die Funktion $f : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x}$ ist 0 ein Häufungspunkt des Definitionsbereiches, und es gilt

$$\lim_{x\to 0} f(x) = \infty.$$

Zudem ist $(0,\infty)$ nach oben unbeschränkt und

$$\lim_{x\to\infty} f(x) = 0.$$



Beweis: Wenn $(a_n)_{n\in\mathbb{N}}$ eine Folge in $(0,\infty)$ mit $a_n \longrightarrow 0$ ist und $s > 0$, so gibt es ein $n_s \in \mathbb{N}$ mit $a_n < \frac{1}{s}$ für $n \geq n_s$. Damit gilt dann für $n \geq n_s$ aber auch

$$f(a_n) = \frac{1}{a_n} > s,$$

d.h. $f(a_n) \longrightarrow \infty$.

Wenn $(a_n)_{n\in\mathbb{N}}$ eine Folge in $(0,\infty)$ mit $a_n \longrightarrow \infty$ ist und $\varepsilon > 0$, so gibt es ein $n_\varepsilon \in \mathbb{N}$ mit $a_n > \frac{1}{\varepsilon}$ für alle $n \geq n_\varepsilon$. Damit gilt dann für $n \geq n_\varepsilon$ aber auch

$$|f(a_n) - 0| = \frac{1}{a_n} < \varepsilon,$$

d.h. $f(a_n) \longrightarrow 0$. □

Beispiel 13.19.
Es sei $f = \sum_{k=0}^{n} a_k \cdot t^k \in \mathbb{R}[t]$ ein Polynom vom Grad $n \geq 1$. Dann gilt

$$\lim_{x\to\infty} f(x) = \begin{cases} \infty, & \text{falls } a_n > 0, \\ -\infty, & \text{falls } a_n < 0, \end{cases}$$

und

$$\lim_{x\to-\infty} f(x) = \begin{cases} \infty, & \text{falls } (a_n > 0 \text{ und } n \text{ gerade }) \text{ oder } (a_n < 0 \text{ und } n \text{ ungerade}), \\ -\infty, & \text{falls } (a_n < 0 \text{ und } n \text{ gerade }) \text{ oder } (a_n > 0 \text{ und } n \text{ ungerade}). \end{cases}$$

Wir beweisen die Aussage nur für $\lim_{x\to\infty} f(x)$ und $a_n > 0$, da der Rest sich analog zeigen läßt. Hierzu betrachten wir ein beliebiges $x \in \mathbb{R}$ mit

$$x \geq \max\left\{\frac{-2\cdot n\cdot a_0}{a_n}, \frac{-2\cdot n\cdot a_1}{a_n}, \ldots, \frac{-2\cdot n\cdot a_{n-1}}{a_n}, 1\right\}.$$

Dann gilt

$$\frac{a_n\cdot x^n}{2\cdot n} \geq -a_k\cdot x^k$$

für alle $0 \leq k \leq n-1$, und mithin

$$\frac{a_n\cdot x^n}{2} = n\cdot\frac{a_n\cdot x^n}{2\cdot n} \geq -\sum_{k=0}^{n-1} a_k\cdot x^k$$

oder alternativ

$$f(x) = \frac{a_n\cdot x^n}{2} + \left(\frac{a_n\cdot x^n}{2} + \sum_{k=0}^{n-1} a_k\cdot x^k\right) \geq \frac{a_n\cdot x^n}{2}.$$

Da zudem offenbar $\lim_{x\to\infty} \frac{a_n\cdot x^n}{2} = \infty$, muß auch $\lim_{x\to\infty} f(x) = \infty$ gelten. □

E) Rechts- und linksseitige Grenzwerte

Bemerkung 13.20 (Rechts- und linksseitge Grenzwerte).
Sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ und $a \in \mathbb{R}$. Wir definieren

$$U_{<a} := \{x \in U \mid x < a\}$$

und

$$U_{>a} := \{x \in U \mid x > a\}.$$

Ist a ein Häufungspunkt von $U_{<a}$ und existiert der Grenzwert der Einschränkung $f_{|U_{<a}}$ in a, so bezeichnen wir diesen mit

$$\lim_{x\to a-} f(x) := \lim_{x\to a} f_{|U_{<a}}(x)$$

und nennen ihn den *linksseitigen Grenzwert* von f in a.

Analog, ist a ein Häufungspunkt von $U_{>a}$ und existiert der Grenzwert der Einschränkung $f_{|U_{>a}}$ in a, so bezeichnen wir diesen mit

$$\lim_{x\to a+} f(x) := \lim_{x\to a} f_{|U_{>a}}(x)$$

und nennen ihn den *rechtsseitigen Grenzwert* von f in a.

Sei nun a ein Häufungspunkt $U_{<a}$ und von $U_{>a}$. Genau dann existieren links- und rechtsseitiger Grenzwert von f in a und stimmen überein, wenn auch der Grenzwert von f in a existiert und

$$\lim_{x\to a} f(x) = \lim_{x\to a-} f(x) = \lim_{x\to a+} f(x)$$

gilt. In manchen Situationen, insbesondere wenn Funktionen abschnittsweise definiert sind, ist es leichter, die links- und rechtsseitigen Grenzwerte getrennt zu berechnen.

Beweis: Wenn a ein Häufungspunkt einer Teilmenge von U ist, ist es automatisch auch ein Häufungspunkt von U. Es bleibt also zu zeigen, dass der Grenzwert von f in a existiert, wenn links- und rechtsseitiger Grenzwert übereinstimmen.

Sei $y = \lim_{x\to a-} f(x) = \lim_{x\to a+} f(x)$ und sei $\varepsilon > 0$ gegeben. Dann existiert ein $\delta'_\varepsilon > 0$, so daß für alle $x \in U_{<a}$ mit $|x-a| < \delta'_\varepsilon$ auch

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

gilt. Zudem existiert ein $\delta''_\varepsilon > 0$, so daß für alle $x \in U_{>a}$ mit $|x-a| < \delta''_\varepsilon$ auch

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

gilt. Wir setzen nun $\delta_\varepsilon := \min\{\delta'_\varepsilon, \delta''_\varepsilon\} > 0$. Dann gilt für alle $x \in U$ mit $|x-a| < \delta_\varepsilon$ auch

$$|f(x) - y| < \varepsilon.$$

Also ist $y = \lim_{x\to a} f(x)$. □

Beispiel 13.21 (Links- und rechtsseitiger Grenzwert).
Wir betrachten die Funktion

$$f : \mathbb{R}\setminus\{0\} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} x+1, & \text{wenn } x < 0, \\ x^3 + 2x^2 + 1, & \text{wenn } x > 0. \end{cases}$$

Mit $U = \mathbb{R}\setminus\{0\}$ ist dann $U_{<0} = (-\infty, 0)$ und $U_{>0} = (0, \infty)$, so daß 0 ein Häufungspunkt von beiden Mengen ist. Ferner gilt

$$\lim_{x\to 0-} f(x) = \lim_{x\to 0-} x + 1 = 1$$

und

$$\lim_{x\to 0+} f(x) = \lim_{x\to 0+} 3x^3 + 2x^2 + 1 = 1.$$

Links- und rechtsseitger Grenzwert von f in 0 existieren also und stimmen überein. Damit existiert dann auch der Grenzwert von f in 0 und es gilt

$$\lim_{x\to 0} f(x) = 1.$$

Aufgaben

Aufgabe 13.22.
Bestimme die Häufungspunkte der folgenden Mengen:

a. $A = \left\{(-1)^n + \frac{n+1}{n} \;\middle|\; n \in \mathbb{N}\right\}$.

b. $B = \left\{x \in \mathbb{R} \;\middle|\; 3 < |x+1|\right\}$.

Aufgabe 13.23.
Bestimme für die nachfolgenden Mengen jeweils die Menge aller ihrer Häufungspunkte:

a. $A = \left\{(-1)^n + \left(\frac{-1}{n}\right)^{n+1} \;\middle|\; n \in \mathbb{Z}_{\geq 1}\right\}$.

b. $B = \mathbb{N}$.

c. $C = \{x \in \mathbb{Q} \mid x^2 < 2\}$.

Aufgabe 13.24.
Gib ein Beispiel für eine Menge, die genau drei Häufungspunkte hat.

Aufgabe 13.25.
Bestimme die folgenden Grenzwerte oder zeige, daß sie nicht existieren:

a. $\lim\limits_{x\to 3} \frac{x^3-9x}{x-3}$.

b. $\lim\limits_{x\to 5} \frac{2x^4+x^3-3x^2+4}{x^2-5}$.

c. $\lim\limits_{x\to\infty} \sqrt{\frac{2x^3+5x}{(x+1)^3}}$.

d. $\lim\limits_{x\to\infty} \sqrt{x+\sqrt{x}} - \sqrt{x}$.

e. $\lim\limits_{x\to 2} \frac{x^3-2x^2-x+2}{x^2-x-2}$.

f. $\lim\limits_{x\to x_0} \frac{x^n-x_0^n}{x-x_0}$, wobei $n \in \mathbb{N}$ und $x_0 \in \mathbb{R}$ beliebig, aber fest vorgegeben sind.

g. $\lim\limits_{x\to\infty} \left(\sqrt{x+5} - \sqrt{x}\right)$.

Aufgabe 13.26 (Cauchy-Kriterium für Grenzwerte).
Es sei $U \subseteq \mathbb{R}$, $a \in \mathbb{R}$ ein Häugfungspunkt von U und $f : U \longrightarrow \mathbb{R}$ eine Funktion. Zeige, der Grenzwert $\lim\limits_{x \to a} f(x)$ existiert genau dann, wenn

$$\forall\, \varepsilon > 0\ \exists\, \delta_\varepsilon > 0\ :\ \forall\, x, y \in \big(U \cap U_{\delta_\varepsilon}(a)\big) \setminus \{a\} \ \text{ gilt } |f(x) - f(y)| < \varepsilon.$$

§ 14 Stetigkeit

Ausblick 14.0.
In technischen Anwendungen erwartet man gemeinhin, daß kleine Änderungen an den Eingabedaten nur zu kleinen Änderungen beim Ergebnis führen. In der Mathematik entspricht das dem Begriff der Stetigkeit des zugrundeliegenden Prozesses.

A) Stetige Funktionen

Definition 14.1 (ε-δ-Kriterium für Stetigkeit).
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ und $a \in U$.
Wir nennen f *stetig in* a, wenn

$$\forall\, \varepsilon > 0 \;\; \exists\, \delta_\varepsilon > 0 \;:\; \forall\, x \in U \text{ mit } |x-a| < \delta_\varepsilon \text{ gilt } |f(x) - f(a)| < \varepsilon.$$

Die Funktion f heißt *stetig* (auf U), wenn sie stetig in jedem Punkt in U ist.
$\mathcal{C}(U, \mathbb{R}) := \{f : U \longrightarrow \mathbb{R} \mid f \text{ stetig}\}$ ist die Menge der auf U stetigen Funktionen.

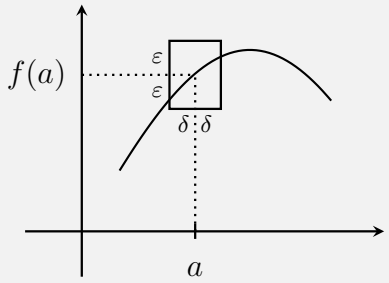


Abbildung 6: ε-δ-Kriterium für Stetigkeit

Bemerkung 14.2.
Für die Stetigkeit einer Funktion in einem Punkt a ist nur das Verhalten von f in einer kleinen ϵ-Umgebung $U_\epsilon(a) = (a-\epsilon, a+\epsilon)$ von a maßgeblich. Wir sagen deshalb auch, daß die Stetigkeit eine *lokale* Eigenschaft ist!

Lemma 14.3 (Stetigkeit in Häufungspunkten).
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ und $a \in U$ ein Häufungspunkt.
Genau dann ist f stetig in a, wenn $\lim\limits_{x \to a} f(x) = f(a)$.

Beweis: Dies folgt unmittelbar aus den Definitionen 13.6 und 14.1. □

Beispiel 14.4 (Polynomfunktionen sind stetig.).

a. Jede Polynomfunktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ ist stetig.
Denn nach Beispiel 13.13 gilt für $a \in \mathbb{R}$ auch $\lim\limits_{x\to a} f(x) = f(a)$.

b. Jede rationale Funktion $\frac{f}{g} : \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) = 0\} \longrightarrow \mathbb{R}$ ist stetig.
Denn nach Beispiel 13.13 ist $a \in \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) \neq 0\}$ ein Häufungspunkt des Definitionsbereiches und $\lim\limits_{x\to a} \frac{f}{g}(x) = \frac{f}{g}(a)$.

c. Die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 1, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0 \end{cases}$$

aus Beispiel 13.8 b. ist nicht stetig in 0, da $\lim\limits_{x\to 0} f(x) = 1 \neq 0 = f(0)$. Aber, f ist stetig in jedem $a \neq 0$, wie man leicht sieht.

d. Ist $f : U \longrightarrow \mathbb{R}$ stetig und $V \subseteq U$, so ist die Einschränkung $f_| : V \longrightarrow \mathbb{R}$ von f auf V offenbar ebenfalls stetig.

e. Ist $f : \mathbb{Z} \longrightarrow \mathbb{R}$ irgendeine Funktion, so ist f stetig! (Kein nützliches Konzept!)
Denn, ist $a \in \mathbb{Z}$ und $\varepsilon > 0$ wählen wir $\delta_\varepsilon := \frac{1}{2}$. Für $x \in \mathbb{Z}$ mit $|x - a| < \delta_\varepsilon = \frac{1}{2}$ muß dann $x = a$ gelten und somit auch $|f(x) - f(a)| = 0 < \varepsilon$.

Satz 14.5 (Folgenkriterium für Stetigkeit).
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ und $a \in U$.
Genau dann ist f stetig in a, wenn

(23) $\quad \forall\, (a_n)_{n\in\mathbb{N}}$ mit $a_n \in U$ und $\lim\limits_{n\to\infty} a_n = a$ gilt $\lim\limits_{n\to\infty} f(a_n) = f(a)$.

Beweis: Der Beweis geht genau wie der Beweis des Folgenkriteriums für Grenzwerte von Funktionen 13.7.

"$\Longrightarrow$": Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in U mit $\lim\limits_{n\to\infty} a_n = a$. Wir müssen $\lim\limits_{n\to\infty} f(a_n) = f(a)$ zeigen. Dazu sei $\varepsilon > 0$ gegeben.
Da f stetig in a ist, gibt es ein $\delta_\varepsilon > 0$, so daß aus $x \in U$ mit $|x - a| < \delta_\varepsilon$ auch $|f(x) - f(a)| < \varepsilon$ folgt.
Wegen $\lim\limits_{n\to\infty} a_n = a$ gibt es zu δ_ε nun ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $n \geq n_\varepsilon$ auch $|a_n - a| < \delta_\varepsilon$ gilt.
Sei nun $n \geq n_\varepsilon$ dann erfüllt $a_n \in U$ die Bedingung $|a_n - a| < \delta_\varepsilon$ und somit ist auch $|f(a_n) - f(a)| < \varepsilon$. Damit ist $f(a_n) \longrightarrow f(a)$ gezeigt.
"$\Longleftarrow$": Wir nehmen an, f wäre nicht stetig in a. Dann gilt:

$$\exists\, \varepsilon > 0 \;:\; \forall\, \delta_\varepsilon > 0 \;\; \exists\, x_{\delta_\varepsilon} \in U \;\text{ mit }\; |x_{\delta_\varepsilon} - a| < \delta_\varepsilon, \;\text{ aber } |f(x_{\delta_\varepsilon}) - f(a)| \geq \varepsilon.$$

Für $n \geq 1$ und $\delta_\varepsilon = \frac{1}{n}$ setzen wir $a_n := x_{\delta_\varepsilon} = x_{\frac{1}{n}} \in U \setminus \{a\}$. Dann gilt

$$0 \leq |a_n - a| < \frac{1}{n} \longrightarrow 0,$$

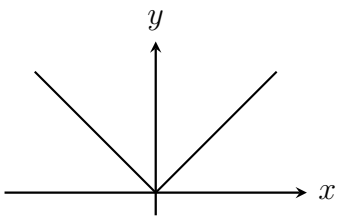
so daß $a_n \longrightarrow a$, und zugleich gilt

$$|f(a_n) - f(a)| \geq \varepsilon$$

für alle $n \in \mathbb{N}$. Dies ist ein Widerspruch dazu, daß $f(a_n)$ gegen $f(a)$ konvergieren muß.

□

Beispiel 14.6 (Die Betragsfunktion ist stetig).
Die Betragsfunktion $|\cdot| : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto |x|$ ist stetig.



Denn für $a \in \mathbb{R}$ und $(a_n)_{n\in\mathbb{N}}$ mit $a_n \longrightarrow a$ gilt aufgrund der Grenzwertsätze für Folgen 11.15 auch $|a_n| \longrightarrow |a|$.

B) Rechnen mit stetigen Funktionen

Proposition 14.7 (Rechenregeln für stetige Funktionen).
Seien $f : U \longrightarrow \mathbb{R}$ und $g : U \longrightarrow \mathbb{R}$ Funktionen, die in $a \in U$ stetig sind, und $c \in \mathbb{R}$.

a. $c \cdot f$, $f + g$, $f - g$ und $f \cdot g$ sind stetig in a.

b. Ist $g(a) \neq 0$, so ist auch $\frac{f}{g} : U \setminus \{x \in U \mid g(x) = 0\} \longrightarrow \mathbb{R}$ stetig in a.

Beweis: Der Beweis folgt aus dem Folgenkriterium für Stetigkeit 14.5 und den Grenzwertsätzen für Folgen 11.15.

Z.B. sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in U mit $a_n \longrightarrow a$, dann gilt

$$(f + g)(a_n) = f(a_n) + g(a_n) \longrightarrow f(a) + g(a) = (f + g)(a),$$

da f und g in a stetig sind. Also ist auch $f + g$ stetig in a. □

Proposition 14.8 (Komposition stetiger Funktionen).
Es seien $f : U \longrightarrow \mathbb{R}$ und $g : V \longrightarrow \mathbb{R}$ Funktionen mit $\mathrm{Im}(f) \subseteq V$ und es sei $a \in U$. Ist f stetig in a und g stetig in $f(a)$, so ist $g \circ f$ stetig in a.

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in U mit $a_n \longrightarrow a$, dann ist $(f(a_n))_{n\in\mathbb{N}}$ eine Folge in V und, da f stetig in a ist, gilt zudem $f(a_n) \longrightarrow f(a)$. Nun ist auch g stetig in $f(a)$, so daß daraus

$$(g \circ f)(a_n) = g(f(a_n)) \longrightarrow g(f(a)) = (g \circ f)(a)$$

folgt. Aufgrund des Folgenkriteriums für Stetigkeit 14.5 ist dann $g \circ f$ stetig in a. □

Beispiel 14.9.
Ist $f : U \longrightarrow \mathbb{R}$ stetig in $a \in U$, so ist auch $|f| : U \longrightarrow \mathbb{R} : x \mapsto |f(x)|$ als Komposition stetiger Funktionen stetig in a.

Definition 14.10 (Stetig fortsetzbar).
Es sei $f : U \longrightarrow \mathbb{R}$ eine stetige Funktion und $a \in \mathbb{R} \setminus U$ ein Häufungspunkt von U. Wir nennen f in a *stetig fortsetzbar*, wenn $\lim\limits_{x\to a} f(x)$ existiert.
In dieser Situation nennen wir

$$g : U \cup \{a\} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } x \neq a, \\ \lim\limits_{z\to a} f(z), & \text{falls } x = a, \end{cases}$$

die *stetige Fortsetzung* von f, und g ist nach Lemma 14.3 stetig in a und damit stetig auf $U \cup \{a\}$.

Beispiel 14.11.

a. Die Funktion

$$f : \mathbb{R} \setminus \{1\} \longrightarrow \mathbb{R} : x \mapsto \frac{x^2 - 1}{x - 1}$$

aus Beispiel 13.8 c. ist in $a = 1$ stetig fortsetzbar, und die stetige Fortsetzung ist

$$g : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x + 1.$$

b. Die Funktion

$$f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x < 0, \\ 1, & \text{falls } x > 0 \end{cases}$$

aus Beispiel 13.8 d. ist in $a = 0$ nicht stetig fortsetzbar, da der Grenzwert von f in 0 nicht existiert.

c. Die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 1, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0 \end{cases}$$

aus Beispiel 13.8 b. ist nach unserer Definition in $a = 0$ nicht stetig fortsetzbar, obwohl der Grenzwert von f in 0 existiert, da 0 bereits zum Definitionsbereich der Funktion gehört!

C) Der Zwischenwertsatz

Satz 14.12 (Zwischenwertsatz).
Eine stetige Funktion $f : [a, b] \longrightarrow \mathbb{R}$ nimmt jeden Wert zwischen $f(a)$ und $f(b)$ an.

Beweis: Für den Beweis können wir $f(a) \leq f(b)$ annehmen. Für $c \in [f(a), f(b)]$ definieren wir eine Funktion

$$g : [a, b] \longrightarrow \mathbb{R} : x \mapsto f(x) - c,$$

und diese ist aufgrund der Proposition 14.7 stetig auf $[a, b]$.

Wir müssen zeigen, daß g eine Nullstelle in $[a, b]$ besitzt.

Dazu wenden wir wie im Beweis des Satzes von Bolzano-Weierstraß 11.26 ein Intervallschachtelungsverfahren an. Wir setzen

$$[a_0, b_0] := [a, b]$$

und betrachten den Punkt

$$x_0 = \frac{a_0 + b_0}{2} \in [a, b].$$

Ist $g(x_0) = 0$, so sind wir fertig. Andernfalls gilt entweder $g(x_0) > 0$ und wir setzen $[a_1, b_1] := [a_0, x_0]$, oder es gilt $g(x_0) < 0$ und wir setzen $[a_1, b_1] := [x_0, b_0]$.

Mit dem neuen Intervall verfahren wir wie mit dem vorherigen. Auf dem Weg finden wir entweder nach endlich vielen Schritten einen Punkt $x_n \in [a, b]$ mit $g(x_n) = 0$, oder wir konstruieren rekursiv eine monoton steigende, beschränkte Folge $(a_n)_{n\in\mathbb{N}}$ in $[a, b]$ und eine monoton fallende, beschränkte $(b_n)_{n\in\mathbb{N}}$ in $[a, b]$ mit

$$b_n - a_n = \frac{b - a}{2^n} \longrightarrow 0. \tag{24}$$

Aufgrund des Monotoniekriteriums für Folgen konvergiert $(a_n)_{n\in\mathbb{N}}$ gegen einen Wert x und $(b_n)_{n\in\mathbb{N}}$ gegen einen Wert y, und wegen (24) gilt dann

$$x = y.$$

Da das Intervall $[a, b]$ abgeschlossen ist, gilt zudem nach Satz 11.28

$$x \in [a, b].$$

Man beachte auch, daß aufgrund der Konstruktion von $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ stets

$$g(a_n) < 0 \quad \text{und} \quad g(b_n) > 0.$$

Für die stetige Funktion g folgt dann aus dem Folgenkriterium 14.5 und Satz 11.17

$$g(x) = \lim_{n\to\infty} g(a_n) \leq 0 \leq \lim_{n\to\infty} g(b_n) = g(x),$$

also $g(x) = 0$. □

Beispiel 14.13 (Nullstellen von Polynomfunktionen).
Ist $f \in \mathbb{R}[t]$ ein Polynom von ungeradem Grad, so besitzt f eine Nullstelle.

Denn nach Beispiel 13.19 gilt, daß $\lim_{x\to\infty} f(x)$ und $\lim_{x\to-\infty} f(x)$ verschiedene Vorzeichen haben, so daß es $a, b \in \mathbb{R}$ mit $f(a) > 0$ und $f(b) < 0$ geben muß. Wenden wir dann den Zwischenwertsatz auf $f_{|[a,b]}$ bzw. $f_{|[b,a]}$ an, so folgt die Behauptung.

D) Beschränktheit stetiger Funktionen

Definition 14.14 (Beschränkte Funktionen).
Eine Funktion $f : U \longrightarrow \mathbb{R}$ heißt *beschränkt*, wenn $\mathrm{Im}(f)$ beschränkt ist.

Proposition 14.15 (Beschränktheit stetiger Funktionen).
Eine stetige Funktion $f : [a, b] \longrightarrow \mathbb{R}$ ist beschränkt.

Beweis: Nehmen wir an, f wäre nicht beschränkt. Dann gibt es für jedes $n \in \mathbb{N}$ ein $a_n \in [a, b]$ mit

$$|f(a_n)| > n.$$

Die Folge $(a_n)_{n\in\mathbb{N}}$ ist beschränkt, da sie im abgeschlossenen Intervall $[a, b]$ liegt, und nach dem Satz von Bolzano-Weierstraß 11.26 gibt es also eine konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ mit Grenzwert c, d.h.

$$a_{n_k} \longrightarrow c.$$

Da das Intervall $[a, b]$ abgeschlossen ist, gilt nach Satz 11.28

$$c \in [a, b].$$

Da f und somit nach Beispiel 14.9 auch $|f|$ stetig auf $[a, b]$ ist, folgt

$$|f(c)| \longleftarrow |f(a_{n_k})| \geq n_k \longrightarrow \infty,$$

was ein offensichtlicher Widerspruch ist. □

Satz 14.16 (Maximum / Minimum stetiger Funktionen).
Eine stetige Funktion $f : [a, b] \longrightarrow \mathbb{R}$ nimmt ihr Maximum und ihr Minimum an, d.h. es gibt $c, d \in [a, b]$, so daß für alle $x \in [a, b]$ gilt

$$f(c) \leq f(x) \leq f(d).$$

Beweis: Nach Proposition 14.15 ist die Menge

$$A := \mathrm{Im}(f) = \{f(x) \mid x \in [a,b]\}$$

beschränkt und somit existiert

$$y := \sup(A) \in \mathbb{R}.$$

Da y die kleinste obere Schranke von A ist, gibt es für jedes $n \geq 1$ ein $a_n \in [a,b]$ mit

$$y - \frac{1}{n} < f(a_n) \leq y.$$

Die Folge $(a_n)_{n\geq 1}$ ist beschränkt, da sie im abgeschlossenen Intervall $[a,b]$ liegt, also besitzt sie nach dem Satz von Bolzano-Weierstraß 11.26 eine konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ mit Grenzwert d. Dann gilt aber

$$y \longleftarrow y - \frac{1}{n_k} < f(a_{n_k}) \leq y \longrightarrow y,$$

so daß aufgrund des Einschachtelungssatzes 11.17 auch

$$f(a_{n_k}) \longrightarrow y$$

gilt. Da f aber stetig ist, folgt dann

$$f(d) = \lim_{k\to\infty} f(a_{n_k}) = y.$$

Die Existenz von c zeigt man analog mit Hilfe von $\inf(A)$. □

Beispiel 14.17.

a. Die Funktion $f : [-1,1] \longrightarrow \mathbb{R} : x \mapsto x^2$ ist beschränkt, und es gilt $f(0) = 0$ ist das Minimum und $f(1) = f(-1) = 1$ ist das Maximum von $\mathrm{Im}(f)$.

b. Die Funktion $f : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x}$ ist nicht beschränkt und nimmt weder ihr Minimum noch ihr Maximum an.

E) Umkehrsatz für streng monotone stetige Funktionen

Definition 14.18 (Monotone Funktionen).
Es sei $f : U \longrightarrow \mathbb{R}$ eine Funktion.

a. f heißt *monoton wachsend*, wenn für $x, y \in U$ aus $x \leq y$ stets $f(x) \leq f(y)$ folgt.

b. f heißt *streng monoton wachsend*, wenn für $x, y \in U$ aus $x < y$ stets $f(x) < f(y)$ folgt.

c. f heißt *monoton fallend*, wenn für $x, y \in U$ aus $x \leq y$ stets $f(x) \geq f(y)$ folgt.

d. f heißt *streng monoton fallend*, wenn für $x, y \in U$ aus $x < y$ stets $f(x) > f(y)$ folgt.

Beispiel 14.19.

a. Die Funktion $f : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto x^n$ ist für jedes $n \geq 1$ streng monoton wachsend, da nach Lemma 8.14 aus $0 \leq x < y$ stets $x^n < y^n$ folgt.

b. Die Funktion

$$f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x < 0, \\ 1, & \text{falls } x > 0 \end{cases}$$

aus Beispiel 13.8 d. ist monoton wachsend, aber nicht streng monoton wachsend.

Bemerkung 14.20.

Ist $f : U \longrightarrow \mathbb{R}$ streng monoton wachsend oder fallend, so ist f injektiv.

Denn, für $x, y \in U$ mit $x \neq y$ gilt $x < y$ oder $x > y$ und somit $f(x) < f(y)$ oder $f(x) > f(y)$, aber in jedem Fall $f(x) \neq f(y)$.

Satz 14.21 (Umkehrsatz für streng monotone stetige Funktionen).
Es seien $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$ mit $a < b$, $f : (a, b) \longrightarrow \mathbb{R}$ sei eine Funktion und es seien $c := \inf(\mathrm{Im}(f)) \in \mathbb{R} \cup \{-\infty\}$ und $d := \sup(\mathrm{Im}(f)) \in \mathbb{R} \cup \{\infty\}$.

a. Ist f streng monoton wachsend und stetig, so gelten:
   (i) $f : (a, b) \longrightarrow (c, d)$ ist bijektiv.
   (ii) $f^{-1} : (c, d) \longrightarrow (a, b)$ ist streng monoton wachsend und stetig.

b. Ist f streng monoton fallend und stetig, so gelten:
   (i) $f : (a, b) \longrightarrow (c, d)$ ist bijektiv.
   (ii) $f^{-1} : (c, d) \longrightarrow (a, b)$ ist streng monoton fallend und stetig.

Beweis: Wir beweisen nur den Fall, daß f streng monoton wachsend ist, da der Beweis für streng monoton fallende Funktionen analog geht.

Zeige: $c, d \notin \mathrm{Im}(f)$**:** Wäre $d \in \mathrm{Im}(f)$, so würde es ein $x \in (a, b)$ geben mit $f(x) = d$. Wegen $x < b$ gibt es ein $x' \in (a, b)$ mit $x < x'$ und somit

$$d = f(x) < f(x') \in \mathrm{Im}(f),$$

im Widerspruch dazu, daß d das Supremum von $\mathrm{Im}(f)$ ist. Analog sieht man, daß $c \notin \mathrm{Im}(f)$.

Zeige: $\mathrm{Im}(f) = (c,d)$**:** Nach Definition von $c = \inf(\mathrm{Im}(f))$ und $d = \sup(\mathrm{Im}(f))$ sowie nach der obigen Vorüberlegung folgt für $y \in \mathrm{Im}(f)$ sofort $c < y < d$, d.h.

$$\mathrm{Im}(f) \subseteq (c,d).$$

Sei nun $y \in (c,d)$. Wegen $y > c = \inf(\mathrm{Im}(f))$ gibt es ein $x_1 \in (a,b)$ mit $y > f(x_1)$, und wegen $y < d = \sup(\mathrm{Im}(f))$ gibt es ein $x_2 \in (a,b)$ mit $y < f(x_2)$. Nach Voraussetzung ist die Einschränkung von f

$$f_| : [x_1, x_2] \longrightarrow \mathbb{R}$$

auf das Intervall $[x_1, x_2]$ stetig als Einschränkung einer stetigen Funktion, und nach dem Zwischenwertsatz 14.12 gibt es wegen $f(x_1) < y < f(x_2)$ dann ein $x \in [x_1, x_2] \subset (c,d)$ mit $y = f(x)$, d.h.

$$(c,d) \subseteq \mathrm{Im}(f).$$

Zeige: $f : (a,b) \longrightarrow (c,d)$ **ist bijektiv:** Nach Bemerkung 14.20 ist die streng monotone Funktion f injektiv, und wie eben gezeigt, ist f surjektiv auf (c,d).

Zeige: $f^{-1} : (c,d) \longrightarrow (a,b)$ **ist streng monoton wachsend:** Seien $y_1, y_2 \in (c,d)$ mit $y_1 < y_2$. Dann gibt es $x_1, x_2 \in (a,b)$ mit $f(x_1) = y_1 < y_2 = f(x_2)$, und da f streng monoton wachsend ist, muß notwendigerweise auch $x_1 < x_2$ gelten. Dann ist aber

$$f^{-1}(y_1) = x_1 < x_2 = f^{-1}(y_2),$$

und f^{-1} ist streng monoton wachsend.

Zeige: $f^{-1} : (c,d) \longrightarrow (a,b)$ **ist stetig:** Seien $y_0 \in (c,d)$ und $\varepsilon > 0$ gegeben. Wir setzen $x_0 := f^{-1}(y_0) \in (a,b)$ und

$$r_\varepsilon := \min\left\{\frac{\varepsilon}{2}, \frac{b - x_0}{2}, \frac{x_0 - a}{2}\right\} > 0.$$

Damit gilt

$$a < x_0 - r_\varepsilon < x_0 < x_0 + r_\varepsilon < b$$

und somit

$$f(x_0 - r_\varepsilon) < y_0 < f(x_0 + r_\varepsilon),$$

da f streng monoton wachsend ist. Für

$$\delta_\varepsilon := \min\{y_0 - f(x_0 - r_\varepsilon), f(x_0 + r_\varepsilon) - y_0\} > 0$$

gilt dann offenbar

$$f(x_0 - r_\varepsilon) \leq y_0 - \delta_\varepsilon < y_0 < y_0 + \delta_\varepsilon \leq f(x_0 + r_\varepsilon),$$

und da f^{-1} streng monoton wachsend ist, folgt für $y \in (y_0 - \delta_\varepsilon, y_0 + \delta_\varepsilon) \subset (c,d)$ deshalb

$$x_0 - r_\varepsilon = f^{-1}(f(x_0 - r_\varepsilon)) < f^{-1}(y) < f^{-1}(f(x_0 + r_\varepsilon)) = x_0 + r_\varepsilon$$

d.h.

$$|f^{-1}(y) - f^{-1}(y_0)| = |x_0 - f^{-1}(y)| < 2 \cdot r_\varepsilon \leq \varepsilon.$$

Also ist f^{-1} stetig in y_0, und damit stetig auf (c, d).

□

Bemerkung 14.22 (Umkehrsatz für streng monotone stetige Funktionen). Ist die Abbildung $f : (a, b) \longrightarrow \mathbb{R}$ im Umkehrsatz 14.21 streng monoton wachsend, so ist

$$c = \inf(\mathrm{Im}(f)) = \lim_{x \to a} f(x) \quad \text{und} \quad d = \sup(\mathrm{Im}(f)) = \lim_{x \to b} f(x),$$

und ist f streng monoton fallend, so ist

$$c = \inf(\mathrm{Im}(f)) = \lim_{x \to b} f(x) \quad \text{und} \quad d = \sup(\mathrm{Im}(f)) = \lim_{x \to a} f(x).$$

Außerdem, falls f stetig in $a \in \mathbb{R}$ bzw. in $b \in \mathbb{R}$ fortgesetzt werden kann, so ist $\lim\limits_{x \to a} f(x) \in \mathbb{R}$ bzw. $\lim\limits_{x \to b} f(x) \in \mathbb{R}$ und f^{-1} wird durch

$$f^{-1}\left(\lim_{x \to a} f(x)\right) = a \quad \text{bzw.} \quad f^{-1}\left(\lim_{x \to b} f(x)\right) = b$$

stetig fortgesetzt. D.h. die Aussagen im Umkehrsatz 14.21 gelten für *halboffene* und *abgeschlossene* Intervalle entsprechend.

Beweis: Wir betrachten nur den Fall f streng monoton wachsend und

$$d := \sup(\mathrm{Im}(f)) \in \mathbb{R} \cup \{\infty\}.$$

1. **Fall:** $d \in \mathbb{R}$**:** Zu $\varepsilon > 0$ gibt es ein $y \in \mathrm{Im}(f)$ mit $y > d - \varepsilon$ und es gibt ein $x_0 \in (a, b)$ mit $f(x_0) = y$.
   Fall 1.1: $b \in \mathbb{R}$**:** Wir setzen nun $\delta_\varepsilon := b - x_0$ und erhalten für $x \in (a, b)$ mit $b - x = |x - b| < \delta_\varepsilon = b - x_0$ notwendigerweise $x_0 < x$ und somit auch $y = f(x_0) < f(x)$, d.h.

   $$|f(x) - d| = d - f(x) < d - y < \varepsilon.$$

   Fall 1.2: $b = \infty$**:** Wir setzen dann $t = \max\{x_0, 1\}$ und erhalten für $x > t$ dann auch $y = f(x_0) < f(x)$, d.h.

   $$|f(x) - d| = d - f(x) < d - y < \varepsilon.$$

   In beiden Fällen ist damit $\lim\limits_{x \to b} f(x) = d$ gezeigt.
2. **Fall:** $d = \infty$**:** Zu $s > 0$ gibt es dann ein $y \in \mathrm{Im}(f)$ mit $y > s$ und wieder gibt es ein $x_0 \in (a, b)$ mit $f(x_0) = y$.
   Fall 1.1: $b \in \mathbb{R}$**:** Wir setzen nun $\delta_\varepsilon := b - x_0$ und erhalten für $x \in (a, b)$ mit $b - x = |x - b| < \delta_\varepsilon = b - x_0$ notwendigerweise $x_0 < x$ und somit auch

   $$f(x) > f(x_0) = y > s.$$

Fall 1.2: $b = \infty$**:** Wir setzen dann $t = \max\{x_0, 1\}$ und erhalten für $x > t$ dann auch

$$f(x) > f(x_0) = y > s.$$

In beiden Fällen ist damit wieder $\lim\limits_{x\to b} f(x) = d$ gezeigt.

Läßt sich nun zudem f in $b \in \mathbb{R}$ stetig fortsetzen, so heißt dies, daß der Grenzwert

$$d := \lim_{x\to b} f(x) \in \mathbb{R}$$

in $\mathbb{R}$ liegt. Da f^{-1} stetig und streng monton wachsend auf (c, d) ist, gilt zudem

$$b = \lim_{x\to d} f^{-1}(x),$$

und somit läßt sich f^{-1} in d durch $f^{-1}(d) = b$ stetig fortsetzen. □

Beispiel 14.23 (Wurzelfunktion).
Für $n \geq 2$ ist die Funktion

$$f : (0, \infty) \longrightarrow \mathbb{R} : x \mapsto x^n$$

nach Beispiel 14.19 streng monoton wachsend und nach Beispiel 14.4 stetig. Zudem gilt

$$\inf(\mathrm{Im}(f)) = 0 \quad \text{und} \quad \sup(\mathrm{Im}(f)) = \infty.$$

Nach dem Umkehrsatz 14.21 gibt es also eine Umkehrfunktion

$$\sqrt[n]{\cdot} : (0, \infty) \longrightarrow (0, \infty) : x \mapsto \sqrt[n]{x}$$

und diese ist streng monoton wachsend und stetig.

Dies ist unter anderem ein alternativer Beweis zu Satz 9.8 für die Existenz von n-ten Wurzeln!

Man beachte zudem, daß wegen Bemerkung 14.22

$$\lim_{x\to 0} \sqrt[n]{x} = 0$$

gilt, so daß die Wurzelfunktion stetig nach 0 fortgesetzt werden kann:

$$\sqrt[n]{\cdot} : [0, \infty) \longrightarrow [0, \infty) : x \mapsto \sqrt[n]{x}.$$

Insbesondere ist auch die Funktion $\sqrt{\cdot} : [0, \infty) \longrightarrow [0, \infty) : x \mapsto \sqrt{x}$ stetig.

Korollar 14.24.
$\lim\limits_{n\to\infty} \sqrt[n]{n} = 1.$

Beweis: Wir müssen zeigen, daß die Folge $(a_n)_{n\geq 2}$ mit $a_n := \sqrt[n]{n} - 1$ eine Nullfolge ist. Da die Funktion $\sqrt[n]{\cdot}$ streng monoton wachsend ist, folgt aus $n > 1$ auch $\sqrt[n]{n} > \sqrt[n]{1} = 1$, und somit $a_n > 0$. Aus dem Binomischen Lehrsatz 7.15 folgt damit

$$n = (a_n + 1)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot a_n^k \cdot 1^{n-k} \geq 1 + \frac{n \cdot (n-1)}{2} \cdot a_n^2 > 1,$$

oder alternativ

$$0 \longleftarrow \frac{2}{n} \geq a_n^2 > 0.$$

Der Einschachtelungssatz 11.17 bedingt dann, daß

$$a_n^2 \longrightarrow 0,$$

und da die Wurzelfunktion stetig in 0 ist, folgt damit

$$a_n = \sqrt{a_n^2} \longrightarrow \sqrt{0} = 0.$$

□

F) Gleichmäßige Stetigkeit

Bemerkung 14.25 (Stetigkeit auf U).
Wir erinnern uns, eine Funktion $f : U \longrightarrow \mathbb{R}$ heißt *stetig auf* U, wenn sie in jedem Punkt $a \in U$ stetig ist, d.h.

$$\forall\, a \in U\ \forall\, \varepsilon > 0\ \exists\, \delta_{\varepsilon,a} > 0\ :\ \forall\, x \in U \text{ mit } |x - a| < \delta_{\varepsilon,a} \text{ gilt } |f(x) - f(a)| < \varepsilon.$$

Wir schreiben diesmal $\delta_{\varepsilon,a}$ statt δ_ε, um zu verdeutlichen, daß wir bei gegebenem $\varepsilon > 0$ zwar in jedem Punkt a ein geeignetes δ finden müssen, daß dieses δ sich mit dem Punkt a aber ändern kann! Es hängt also vom Punkt a ab. In der nächsten Definition wollen wir einen stärkeren Begriff der Stetigkeit einführen, bei dem genau das nicht mehr der Fall ist.

Definition 14.26 (Gleichmäßige Stetigkeit).
Eine Funktion $f : U \longrightarrow \mathbb{R}$ heißt *gleichmäßig stetig* auf U, wenn

$$\forall\, \varepsilon > 0\ \exists\, \delta_\varepsilon > 0\ :\ \forall\, x, y \in U \text{ mit } |x - y| < \delta_\varepsilon \text{ gilt } |f(x) - f(y)| < \varepsilon.$$

Bemerkung 14.27.
Offenbar ist jede auf U gleichmäßig stetige Funktion $f : U \longrightarrow \mathbb{R}$ auch stetig auf U.

Satz 14.28.
Eine stetige Funktion $f : [a, b] \longrightarrow \mathbb{R}$ ist gleichmäßig stetig auf $[a, b]$.

Beweis: Angenommen, f wäre nicht gleichmäßig stetig auf $[a,b]$. Dann gilt:

$$\exists\, \varepsilon > 0 \;:\; \forall\, \delta_\varepsilon > 0 \;:\; \exists\, x_{\delta_\varepsilon}, y_{\delta_\varepsilon} \in [a,b] \text{ mit } |x_{\delta_\varepsilon} - y_{\delta_\varepsilon}| < \delta_\varepsilon, \text{ aber } |f(x_{\delta_\varepsilon}) - f(y_{\delta_\varepsilon})| \geq \varepsilon.$$

Für $n \geq 1$ und $\delta_\varepsilon := \frac{1}{n}$ setzen wir $a_n := x_{\delta_\varepsilon} = x_{\frac{1}{n}}$ und $b_n := y_{\delta_\varepsilon} = y_{\frac{1}{n}}$. Damit erhalten wir zwei beschränkte Folgen $(a_n)_{n\geq 1}$ und $(b_n)_{n\geq 1}$ in $[a,b]$. Nach dem Satz von Bolzano-Weierstraß 11.26 besitzt $(a_n)_{n\geq 1}$ eine konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$, und ebenso besitzt dann $(b_{n_k})_{k\in\mathbb{N}}$ eine konvergente Teilfolge $(b_{n_{k_l}})_{l\in\mathbb{N}}$. Nach Konstruktion gilt

$$0 \leq |a_{n_{k_l}} - b_{n_{k_l}}| \leq \frac{1}{n_{k_l}} \longrightarrow 0,$$

so daß die Grenzwerte von $(a_{n_{k_l}})_{l\in\mathbb{N}}$ und $(b_{n_{k_l}})_{l\in\mathbb{N}}$ wegen des Einschachtelungssatzes 11.17 übereinstimmen müssen, d.h.

$$a_{n_{k_l}} \longrightarrow y \quad \text{und} \quad b_{n_{k_l}} \longrightarrow y.$$

Da das Intervall $[a,b]$ abgeschlossen ist, gilt nach Satz 11.28 zudem

$$y \in [a,b].$$

Da f und die Betragsfunktion stetig sind, folgt damit

$$0 = |f(y) - f(y)| \longleftarrow |f(a_{n_{k_l}}) - f(b_{n_{k_l}})| \geq \varepsilon,$$

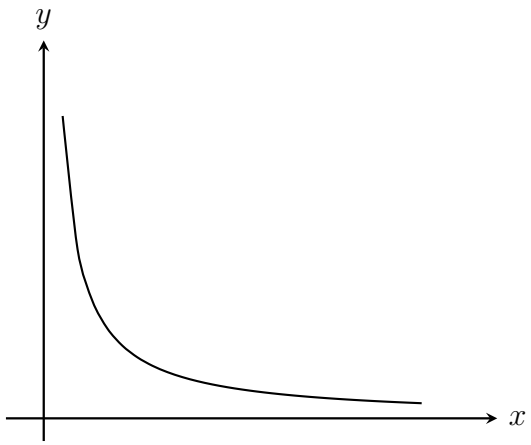
was ein offensichtlicher Widerspruch ist. □

Beispiel 14.29.

a. Die Funktion $f : [0,1] \longrightarrow \mathbb{R} : x \mapsto x^2$ ist gleichmäßig stetig auf $[0,1]$.
Dies folgt aus Satz 14.28. Will man es aus der Definition selbst herleiten, so kann man $\delta_\varepsilon := \frac{\varepsilon}{2}$ zu gegebenem $\varepsilon > 0$ wählen, denn aus $|x-y| < \delta_\varepsilon$ folgt dann

$$|f(x) - f(y)| = |x^2 - y^2| = |x-y| \cdot |x+y| \leq |x-y| \cdot 2 < 2 \cdot \delta_\varepsilon = \varepsilon.$$

b. Die Funktion $f : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x}$ ist *nicht* gleichmäßig stetig auf $(0,\infty)$.



Dazu setzen wir $\varepsilon := 1$ und wählen $\delta > 0$ beliebig. Dann setzen wir $x := \delta$ und $y := \frac{\delta}{1+\delta}$, also $x, y \in (0, \infty)$ mit

$$|x - y| = \delta - \frac{\delta}{1+\delta} < \delta,$$

aber

$$|f(x) - f(y)| = \left| \frac{1}{\delta} - \frac{1+\delta}{\delta} \right| = 1 \geq \varepsilon.$$

Also ist f nicht gleichmäßig stetig auf $(0, \infty)$.
Das Problem liegt darin, daß bei fest vorgegebenem $\varepsilon > 0$ das $\delta_{\varepsilon,a}$, das man für die Stetigkeit in a wählen muß, immer kleiner werden muß, je näher a an 0 liegt, da die Steigung des Graphen von f nahe bei Null immer steiler wird.

Aufgaben

Aufgabe 14.30.
Entscheide, ob die folgende Funktion in $a = -2$ stetig ist

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{4x^2+x+1}{1-x}, & \text{falls } x < -2, \\ |x-1|, & \text{falls } x \geq -2. \end{cases}$$

Aufgabe 14.31.
Bestimme eine reelle Zahl a, so daß die folgende Funktion stetig ist:

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 5x - 1, & \text{falls } x < a, \\ x + 7, & \text{falls } x \geq a. \end{cases}$$

Aufgabe 14.32.
Bestimme eine reelle Zahl b, so daß die folgende Funktion stetig ist:

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} x^2 - b, & \text{falls } x < 0, \\ \sqrt{x} + b, & \text{falls } x \geq 0. \end{cases}$$

Aufgabe 14.33.
Sei $f : U \longrightarrow \mathbb{R}$ stetig, $a \in U$ und $b \in \mathbb{R}$.

a. Zeige, ist $f(a) > b$, so gibt, es ein $\delta > 0$, so dass $f(x) > b$ für alle $x \in U \cap (a - \delta, a + \delta)$.

b. Zeige, ist $f(a) \neq b$, so gibt es ein $\delta > 0$, so dass $f(x) \neq b$ für alle $x \in U \cap (a-\delta, a+\delta)$.

Aufgabe 14.34.
Die Wurzelfunktion $\sqrt{\cdot} : [0,\infty) \longrightarrow \mathbb{R} : x \mapsto \sqrt{x}$ ist gleichmäßig stetig auf $[0,\infty)$.

Aufgabe 14.35.
Verwende die ϵ-δ-Definition der Stetigkeit, um zu zeigen, dass die Funktion $f : [0,1] \longrightarrow \mathbb{R},\ x \longmapsto \sqrt{1-x^3}$ stetig in $[0,1]$ ist.

Aufgabe 14.36 (Lipschitz-Stetigkeit).
Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ eine Funktion und $L \in \mathbb{R}_{>0}$. Zeige, wenn $|f(x)-f(y)| \leq L \cdot |x-y|$ für alle $x, y \in \mathbb{R}$ gilt, so ist f stetig in $\mathbb{R}$.

Aufgabe 14.37.
Sei $f : [0,1] \longrightarrow \mathbb{R}$ eine Funktion definiert durch

$$f(x) := \begin{cases} \frac{2^n}{n!} & \text{für } x = \frac{1}{n} \text{ mit } n \geq 1 \\ 0 & \text{für } x \in [0,1] \backslash \{\frac{1}{n} \mid n \geq 1\} \end{cases}.$$

Bestimme (mit Beweis) sämtliche Punkte auf $[0,1]$, in denen f stetig ist.

Aufgabe 14.38.
Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ stetig und $a, b \in \mathbb{R}$ mit $f(a) \neq b$. Zeige, es gibt ein $\delta > 0$, so dass $f(x) \neq b$ für alle $x \in (a-\delta, a+\delta)$.

Aufgabe 14.39 (Fixpunktsatz von Banach).
Sei $f : [a,b] \longrightarrow \mathbb{R}$ eine stetige Abbildung mit $\mathrm{Im}(f) \subseteq [a,b]$. Zeige, daß f einen Fixpunkt hat, d.h. es gibt ein $c \in [a,b]$ mit $f(c) = c$.

Aufgabe 14.40.
Zeige, ist $f : [0,1] \longrightarrow \mathbb{R}$ eine stetige Funktion mit $f(0) = f(1)$. Zeige, dann gibt es ein $a \in [0, \frac{1}{2}]$ mit $f(a) = f(a + \frac{1}{2})$.

Aufgabe 14.41.
Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ eine stetige Abbildung und $a \in \mathbb{R}_{>0}$ mit $f(x) = f(x+a)$ für alle $x \in \mathbb{R}$. Zeige, daß es ein $b \in (0,a)$ gibt mit $f(b + \frac{a}{2}) = f(b)$.

Aufgabe 14.42 (Stetige Fortsetzbarkeit).

a. Sei $f : (a, b] \longrightarrow \mathbb{R}$ eine stetige Funktion. Zeige, dass f genau dann stetig in a fortsetzbar ist, wenn f gleichmäßig stetig ist.

b. Gibt es eine beschränkte Funktion $f : (0, 1] \longrightarrow \mathbb{R}$, die sich nicht stetig in 0 fortsetzen lässt?

Aufgabe 14.43.

Es seien $f : [a, b] \longrightarrow \mathbb{R}$ stetig auf $[a, b]$ und $g : [b, c] \longrightarrow \mathbb{R}$ stetig auf $[b, c]$ mit $f(b) = g(b)$, so ist auch die Funktion

$$h : [a, c] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} g(x), & \text{falls } x \leq b, \\ h(x), & \text{falls } x > b \end{cases}$$

stetig auf $[a, c]$.

Aufgabe 14.44.

Es sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ eine Funktion und $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$, so daß die Folge $(f(x_n))_{n\in\mathbb{N}}$ konvergiert.

a. Zeige durch ein Beispiel, daß die Folge $(x_n)_{n\in\mathbb{N}}$ nicht konvergieren muß.

b. Zeige, wenn f stetig und streng monoton ist, dann ist $(x_n)_{n\in\mathbb{N}}$ konvergent.

Aufgabe 14.45.

Betrachte die Funktion

$$f : (0, \infty) \longrightarrow \mathbb{R} : x \mapsto \sin\left(\frac{1}{x}\right).$$

a. Ist f stetig auf $(0, \infty)$?

b. Ist f beschränkt auf $(0, \infty)$?

c. Ist f gleichmäßig stetig auf $(0, \infty)$?

Aufgabe 14.46.

Zeige, die Funktion $f : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto \sqrt{x}$ ist gleichmäßig stetig.

Aufgabe 14.47.

Zeige, ist $f : [0, \infty) \longrightarrow [0, \infty)$ gleichmäßig stetig mit $f(0) = 0$, so gibt es eine Konstante $K > 0$ mit

$$f(x) \leq 1 + K \cdot x$$

für alle $x \in [0, \infty)$.

Aufgabe 14.48 (Thomaesche Funktion, die nur an $R \setminus \mathbb{Q}$ stetig ist.).
Für eine rationale Zahl $x = \frac{a}{b}$ mit $a, b > 0$ setzen wir $N(x) = \frac{b}{\operatorname{ggt}(a,b)}$, d.h. $N(x)$ ist der Nenner von x in gekürzter Form und wir setzen $N(0) = 1$. Zeige, die Funktion

$$f : [0,1] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{1}{N(x)}, & \text{falls } x \in \mathbb{Q} \text{ und } x \neq 0, \\ 0, & \text{sonst,} \end{cases}$$

ist genau dann stetig in $x \in [0,1]$, wenn x irrational ist.

§ 15 Konvergenz von Funktionenfolgen

A) Punktweise und gleichmäßige Konvergenz von Funktionenfolgen

Definition 15.1 (Konvergenz von Funktionenfolgen).

a. Für jedes $n \in \mathbb{N}$ sei $f_n : U \longrightarrow \mathbb{R}$ eine Funktion, so nennen wir $(f_n)_{n\in\mathbb{N}}$ eine *Folge von Funktionen* auf U.

b. Wir nennen die Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen *punktweise konvergent auf* U, wenn für jedes $x \in U$ der Grenzwert $\lim\limits_{n\to\infty} f_n(x)$ existiert, d.h. In diesem Fall nennen wir die Funktion

$$f : U \longrightarrow \mathbb{R} : x \mapsto \lim_{n\to\infty} f_n(x)$$

den *Grenzwert* oder die *Grenzfunktion* der Funktionenfolge $(f_n)_{n\in\mathbb{N}}$, und wir sagen auch, daß $(f_n)_{n\in\mathbb{N}}$ *punktweise gegen* f *konvergiert*. Wir schreiben dann

$$f = \lim_{n\to\infty} f_n.$$

Man beachte, $(f_n)_{n\in\mathbb{N}}$ konvergiert auf U genau dann punktweise gegen f, wenn

$$\forall\, x \in U\ \forall\, \varepsilon > 0\ \ \exists\, n_{\varepsilon,x}\ :\ \forall\, n \geq n_{\varepsilon,x} \text{ gilt } |f_n(x) - f(x)| < \varepsilon.$$

c. Wir sagen $(f_n)_{n\in\mathbb{N}}$ *konvergiert gleichmäßig auf* U gegen f, wenn

$$\forall\, \varepsilon > 0\ \ \exists\, n_\varepsilon\ :\ \forall\, n \geq n_\varepsilon \text{ und } \forall\, x \in U \text{ gilt } |f_n(x) - f(x)| < \varepsilon.$$

Bemerkung 15.2.
Konvergiert $(f_n)_{n\in\mathbb{N}}$ auf U gleichmäßig gegen f, so konvergiert die Folge auch punktweise gegen f.

Beispiel 15.3.
Die Folge $f_n : [0,1] \longrightarrow \mathbb{R} : x \mapsto x^n$ konvergiert auf $[0,1]$ punktweise gegen die Funktion

$$f : [0,1] \longrightarrow \mathbb{R} : x \mapsto \lim_{n\to\infty} x^n = \begin{cases} 0, & \text{falls } x < 1, \\ 1, & \text{falls } x = 1. \end{cases}$$

Aber, $(f_n)_{n\in\mathbb{N}}$ konvergiert auf $[0,1]$ *nicht* gleichmäßig gegen f.

Beachte auch, daß die Grenzfunktion nicht stetig in 1 ist, obwohl alle f_n stetig waren!

Um zu sehen, daß die Konvergenz nicht gleichmäßig ist, betrachten wir $\varepsilon := \frac{1}{4} > 0$ und ein beliebiges $n_\varepsilon \in \mathbb{N}$. Setze $n := \max\{n_\varepsilon, 2\} \geq n_\varepsilon$ und $x = \frac{1}{\sqrt[n]{2}} \in [0,1)$, dann gilt

$$|f_n(x) - f(x)| = \frac{1}{2} > \frac{1}{4} = \varepsilon.$$

Satz 15.4 (Gleichmäßige Konvergenz von Potenzreihen).
Es sei $\sum_{k=0}^{\infty} a_k \cdot t^k$ eine Potenzreihe über $\mathbb{R}$ mit Konvergenzradius r und für $n \in \mathbb{N}$ sei

$$f_n : (-r,r) \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{n} a_k \cdot x^k.$$

Dann konvergiert die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ auf $(-r,r)$ punktweise gegen

$$f : (-r,r) \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{\infty} a_k \cdot x^k,$$

und für jedes $0 \leq R < r$ konvergiert $(f_n)_{n\in\mathbb{N}}$ auf $[-R,R]$ gleichmäßig gegen f.

Beweis: Daß die f_n auf $(-r,r)$ punktweise gegen f konvergieren, folgt unmittelbar aus der Definition von f_n und f. Es bleibt also nur, für $0 \leq R < r$ zu zeigen, daß die f_n auf $[-R,R]$ gleichmäßig konvergieren.

Sei dazu $\varepsilon > 0$ vorgegeben. Aus Satz 12.32 wissen wir, daß die Reihe $\sum_{k=0}^{\infty} |a_k| \cdot R^k$ konvergiert, und wegen Lemma 12.6 ist die Folge der Restglieder dann eine Nullfolge, so daß es ein $n_\varepsilon \in \mathbb{N}$ gibt mit

$$\sum_{k=n}^{\infty} |a_k| \cdot R^k = \left| \sum_{k=n}^{\infty} |a_k| \cdot R^k \right| < \varepsilon$$

für alle $n \geq n_\varepsilon$. Sei nun $n \geq n_\varepsilon$ und $x \in [-R,R]$, so gilt

$$|f_n(x) - f(x)| = \left| -\sum_{k=n+1}^{\infty} a_k \cdot x^k \right| \leq \sum_{k=n+1}^{\infty} |a_k| \cdot |x|^k \leq \sum_{k=n+1}^{\infty} |a_k| \cdot R^k < \varepsilon.$$

Man beachte hierbei, daß wir hier mehrfach die Proposition 11.17 a. für die betrachteten Folgen der Partialsummen verwenden. □

Beispiel 15.5.
Die Folge $f_n : (-1,1) \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{n} x^k$ konvergiert auf $(-1,1)$ *nicht* gleichmäßig gegen die geometrische Reihe $f : (-1,1) \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{\infty} x^k$.

Um dies zu sehen, seien $\varepsilon := 1$ gegeben und $n_\varepsilon \in \mathbb{N}$ beliebig. Wir betrachten zunächst die stetige Funktion

$$g : (-1,1) \longrightarrow \mathbb{R} : x \mapsto \frac{x^{n_\varepsilon+1}}{1-x}.$$

Man sieht leicht, daß $\lim_{x\to 1} g(x) = \infty$, so daß es sicher ein $x \in (0,1)$ mit

$$\frac{x^{n_\varepsilon+1}}{1-x} = g(x) \geq 1 = \varepsilon$$

geben muß. Für dieses x gilt nun

$$|f_{n_\varepsilon}(x) - f(x)| = \sum_{k=n_\varepsilon+1}^{\infty} x^k = x^{n_\varepsilon+1} \cdot \sum_{k=0}^{\infty} x^k = \frac{x^{n_\varepsilon+1}}{1-x} \geq \varepsilon.$$

Mithin konvergiert f_n auf $(-1,1)$ nicht gleichmäßig gegen f.

B) Gleichmäßiger Grenzwert stetiger Funktionen

Satz 15.6 (Der gleichmäßige Grenzwert stetiger Funktionen ist stetig.).
Ist $f_n : U \longrightarrow \mathbb{R}$ stetig auf U für $n \in \mathbb{N}$ und konvergiert $(f_n)_{n\in\mathbb{N}}$ auf U gleichmäßig gegen f, so ist f stetig auf U.

Beweis: Seien $a \in U$ und $\varepsilon > 0$ gegeben. Da die f_n gleichmäßig gegen f konvergieren, gilt:

$$\exists\, n_\varepsilon \in \mathbb{N} \;:\; \forall\, n \geq n_\varepsilon \text{ und } \forall\, x \in U \;:\; |f_n(x) - f(x)| < \frac{\varepsilon}{3}.$$

Da zudem f_{n_ε} stetig in a ist, gilt:

$$\exists\, \delta_\varepsilon > 0 \quad \forall\, x \in U \text{ mit } |x-a| < \delta_\varepsilon \text{ gilt } |f_{n_\varepsilon}(x) - f_{n_\varepsilon}(a)| < \frac{\varepsilon}{3}.$$

Sei nun $x \in U$ mit $|x-a| < \delta_\varepsilon$ gegeben, so gilt

$$|f(x) - f(a)| \leq |f(x) - f_{n_\varepsilon}(x)| + |f_{n_\varepsilon}(x) - f_{n_\varepsilon}(a)| + |f_{n_\varepsilon}(a) - f(a)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Mithin ist f stetig in a. □

Korollar 15.7 (Potenzreihen sind stetig.).
Ist $\sum_{n=0}^{\infty} a_n \cdot t^n$ eine Potenzreihe mit Konvergenzradius r, dann ist

$$f : (-r,r) \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} a_n \cdot x^n$$

stetig auf $(-r,r)$.

Beweis: Sei $x \in (-r,r)$ beliebig, so ist $0 \leq R < r$ für $R := \frac{|x|+r}{2}$. Nach Satz 15.4 konvergiert die Folge stetiger Funktionen

$$f_n : [-R,R] \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{n} a_k \cdot x^k$$

auf $[-R, R]$ gleichmäßig gegen f, und nach Satz 15.6 ist f mithin stetig auf $[-R, R]$ und damit insbesondere in $x \in (-R, R)$. □

Beispiel 15.8.
Die Exponentialfunktion, der Sinus und der Cosinus sind stetig auf $\mathbb{R}$.

Dies folgt aus Korollar 15.7 zusammen mit den Sätzen 12.36 und 12.38.

Aufgaben

Aufgabe 15.9.
Für $n \geq 2$ sei $f_n = \sqrt[n]{\cdot} : [0, \infty) \longrightarrow \mathbb{R}$. Beweise oder widerlege die folgenden Aussagen.

a. $(f_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig auf $[0, \infty)$.

b. $(f_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig auf $[1, 100]$.

Aufgabe 15.10.
Finde eine Folge $(f_n)_{n\in\mathbb{N}}$ von stetigen Funktionen $f_n : [0, 1] \longrightarrow \mathbb{R}$, die punktweise gegen die Nullfunktion konvergiert, aber unbeschränkt ist, d.h., so daß zu jedem $c \in \mathbb{R}$ ein $n \in \mathbb{N}$ und ein $x \in [0, 1]$ existiert mit $|f_n(x)| > c$.

Aufgabe 15.11.
Wir betrachten die Funktionenfolge $(f_n)_{n\geq 1}$ mit

$$f_n : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \frac{1}{1 + |x|^n}.$$

a. Zeige, die Funktionenfolge konvergiert punktweise gegen eine Grenzfunktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ und gib die Werte $f(1)$, $f(\frac{1}{2})$ und $f(2)$ an.

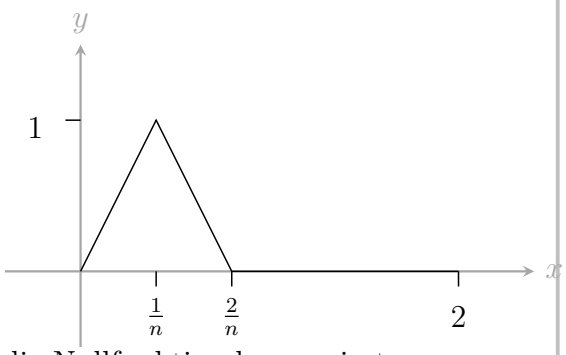
b. Skizziere den Graphen der Grenzfunktion f.

c. Welche der folgenden Grenzwerte existieren?
   (i) $\lim_{x\to 1} f(x)$ für $x \in (1, \infty)$.
   (ii) $\lim_{x\to 1} f(x)$ für $x \in (-1, 1)$.
   (iii) $\lim_{x\to 1} f(x)$ für $x \in \mathbb{R}$.

d. An welchen Stellen $a \in \mathbb{R}$ ist die Funktion f stetig.

Aufgabe 15.12.
Zeige, daß die Folge $(f_n)_{n\geq 1}$ von Funktionen $f_n : [0, 2] \longrightarrow \mathbb{R}$, gegeben durch

$$f_n(x) = \begin{cases} nx, & x \in [0, \frac{1}{n}) \\ 2 - nx, & x \in [\frac{1}{n}, \frac{2}{n}) \\ 0, & x \in [\frac{2}{n}, 2] \end{cases},$$



punktweise, aber nicht gleichmäßig gegen die Nullfunktion konvergiert.

Aufgabe 15.13.
Es sei $g : \mathbb{R} \longrightarrow \mathbb{R}_{\geq 0}$ eine beschränkte, nicht-negative Funktion. Zeige, dass die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{\infty} \frac{g(2^k \cdot x)}{2^k}$$

wohldefiniert und stetig ist.

Aufgabe 15.14. a. Zeige, eine Folge von beschränkten Funktionen $f_n : [a, b] \longrightarrow \mathbb{R}$ konvergiert genau dann gleichmäßig gegen eine Grenzfunktion $f : [a, b] \longrightarrow \mathbb{R}$, wenn

$$\lim_{n\to\infty} \sup_{x\in[a,b]} |f_n(x) - f(x)| = 0.$$

b. Untersuche die Funktionenfolge $(f_n)_{n\geq 1}$ mit $f_n : [0, 1] \longrightarrow \mathbb{R} : x \mapsto e^{\frac{x}{n}}$ auf punktweise und gleichmäßige Konvergenz

Aufgabe 15.15.
Es sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von monoton wachsenden Funktionen $f_n : [a, b] \longrightarrow \mathbb{R}$, die punktweise gegen eine stetige Funktion $f : [a, b] \longrightarrow \mathbb{R}$ konvergieren. Zeige, dann konvergiert die Folge schon gleichmäßig gegen f.

§ 16 Exponentialfunktion, Logarithmus, trigonometrische Funktionen

A) Exponentialfunktionen und Logarithmusfunktionen

Satz 16.1 (Die Exponentialfunktion).
Die Exponentialfunktion

$$\exp : (-\infty, \infty) \longrightarrow (0, \infty)$$

ist stetig, streng monoton wachsend und bijektiv. Insbesondere gelten

$$\lim_{x \to -\infty} \exp(x) = 0 \quad \text{und} \quad \lim_{x \to \infty} \exp(x) = \infty.$$

Beweis: Für $z \in \mathbb{R}$ mit $z > 0$ gilt offenbar

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \geq \frac{z^1}{1!} + \frac{z^0}{0!} = z + 1 > 1, \tag{25}$$

und mit Hilfe der Funktionalgleichung in Satz 12.36 folgt dann

$$\exp(-z) \cdot \exp(z) = \exp(-z + z) = \exp(0) = 1$$

sowie

$$\exp(-z) = \frac{1}{\exp(z)} > 0, \tag{26}$$

die Exponentialfunktion nimmt also nur positive Werte an. Wenden wir die Funktionalgleichung noch mal für $x, y \in \mathbb{R}$ mit $x < y$ wie folgt an

$$\exp(y) = \exp(y - x + x) = \exp(y - x) \cdot \exp(x) \overset{(25),(26)}{>} 1 \cdot \exp(x) = \exp(x),$$

so erhalten wir, daß exp streng monoton wachsend ist. Da exp nach Beispiel 15.8 zudem stetig ist, können wir den Umkehrsatz für streng monotone Funktionen 14.21 anwenden und erhalten, daß

$$\exp : (-\infty, \infty) \longrightarrow (c, d)$$

auch bijektiv ist, wobei

$$c = \inf(\mathrm{Im}(f)) \overset{14.22}{=} \lim_{x \to -\infty} \exp(x)$$

und

$$d = \sup(\mathrm{Im}(f)) \overset{14.22}{=} \lim_{x \to \infty} \exp(x).$$

Nun gilt für $x > 0$ aber

$$\exp(x) \overset{(25)}{\geq} x + 1 \overset{x \to \infty}{\longrightarrow} \infty,$$

so daß $d = \lim_{x \to \infty} \exp(x) = \infty$ folgt.

Mit (26) und aus den Grenzwertsätzen für uneigentliche Grenzwerte von Funktionen folgt zudem

$$\exp(-x) = \frac{1}{\exp(x)} \overset{x\to\infty}{\longrightarrow} \frac{1}{\infty} = 0,$$

d.h. $c = \lim\limits_{x\to-\infty} \exp(x) = 0$. □

Bemerkung 16.2.
Die Zahl $e = \exp(1)$ ist irrational und es gilt $2 < e < 3$ (siehe auch Beispiel 18.36).

Beweis: Aus (25) wissen wir

$$\exp(1) > 1 + 1 = 2.$$

Wegen

$$\frac{1}{k!} = \frac{1}{1\cdot 2\cdot 3\cdot \ldots \cdot k} \le \frac{1}{1\cdot 2\cdot 2\cdot \ldots \cdot 2} = \frac{1}{2^{k-1}}$$

für $k \ge 1$ folgt unter Berücksichtigung der endlichen geometrischen Reihe 7.12

$$\begin{aligned} s_n &:= \sum_{k=0}^{n} \frac{1}{k!} \le 1 + 1 + \frac{1}{2} + \frac{1}{6} + \sum_{k=4}^{n} \frac{1}{2^{k-1}} \\ &= \frac{8}{3} + \frac{1}{8}\cdot \sum_{k=0}^{n-4} \frac{1}{2^k} = \frac{8}{3} + \frac{1}{8}\cdot \frac{1-\left(\frac{1}{2}\right)^{n-3}}{1-\frac{1}{2}} \\ &\le \frac{8}{3} + \frac{1}{8}\cdot \frac{1}{1-\frac{1}{2}} = \frac{8}{3} + \frac{1}{4} = \frac{35}{12} \end{aligned}$$

für alle $n \in \mathbb{N}$. Grenzwertbildung liefert deshalb $\exp(1) = \lim\limits_{n\to\infty} s_n \le \frac{35}{12} < 3$.

Wir müssen nun noch zeigen, daß e keine rationale Zahl sein kann. Nehmen wir dazu an, es gelte $e = \frac{p}{q} \in \mathbb{Q}$ mit $p, q \in \mathbb{N}$. Wegen $2 < e < 3$ muß $q \ge 2$ sein. Wir betrachten nun die Zahlen

$$a := \frac{q!}{0!} + \frac{q!}{1!} + \frac{q!}{2!} + \ldots + \frac{q!}{q!} \in \mathbb{Z}$$

und

$$b := \sum_{n=q+1}^{\infty} \frac{q!}{n!} = q!\cdot e - a = (q-1)!\cdot p - a \in \mathbb{Z}. \tag{27}$$

Da $q \ge 2$ ist, folgt für $n > q$

$$\frac{q!}{n!} = \frac{1}{(q+1)\cdot(q+2)\cdots n} \le \frac{1}{3\cdot 3\cdots 3} = \frac{1}{3^{n-q}}$$

und deshalb

$$b = \sum_{n=q+1}^{\infty} \frac{q!}{n!} \le \sum_{n=q+1}^{\infty} \frac{1}{3^{n-q}} = \sum_{n=1}^{\infty} \frac{1}{3^n} = \frac{1}{1-\frac{1}{3}} - 1 = \frac{1}{2}$$

Da aber aufgrund der Definition von b auch

$$b = \sum_{n=q+1}^{\infty} \frac{q!}{n!} > 0$$

gilt, kann b keine ganze Zahl sein, im Widerspruch zu (27). Der Widerspruch kommt von unserer Annahme, daß e eine rationale Zahl wäre. □

Definition 16.3 (Natürlicher Logarithmus).
Die Umkehrabbildung der Exponentialfunktion wird mit

$$\ln : (0, \infty) \longrightarrow (-\infty, \infty)$$

bezeichnet und *(natürlicher) Logarithmus* genannt.

Satz 16.4 (Natürlicher Logarithmus).
Der natürliche Logarithmus

$$\ln : (0, \infty) \longrightarrow (-\infty, \infty)$$

ist stetig, streng monoton wachsend und bijektiv. Insbesondere gelten

$$\lim_{x \to 0} \ln(x) = -\infty \quad \text{und} \quad \lim_{x \to \infty} \ln(x) = \infty.$$

Beweis: Die Aussagen folgen aus dem Umkehrsatz für streng monotone Funktionen 14.21 und Satz 16.1 unter Berücksichtigung von Bemerkung 14.22. □

Bemerkung 16.5.
Man beachte, daß aus

$$\exp(0) = 1 \quad \text{und} \quad \exp(1) = e$$

unmittelbar

$$\ln(1) = 0 \quad \text{und} \quad \ln(e) = 1$$

folgt. Die Graphen der Exponentialfunktion und des natürlichen Logarithmus sind in der folgenden Abbildung dargestellt.

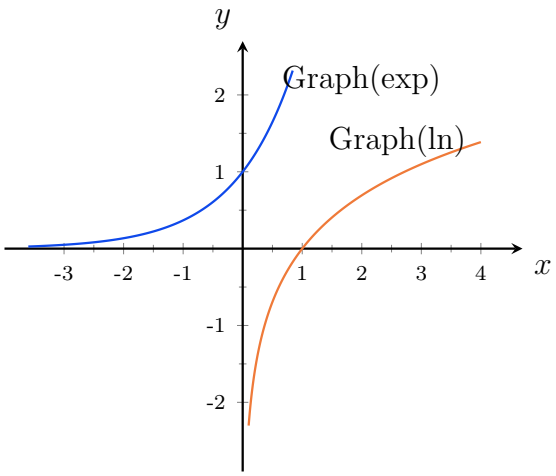


Abbildung 7: Die Graphen von exp und ln

Definition 16.6 (Exponentialfunktion zur Basis a).
Für $a \in \mathbb{R}$ mit $a > 0$ und $x \in \mathbb{R}$ definieren wir

$$a^x := \exp\big(x \cdot \ln(a)\big).$$

Man beachte, daß damit $e^x = \exp(x \cdot \ln(e)) = \exp(x \cdot 1) = \exp(x)$ gilt, so daß die neue Definition im Fall $a = e$ mit der Definition aus Bemerkung 12.37 übereinstimmt.

Satz 16.7 (Exponential- und Logarithmusfunktion zur Basis a).
Es sei $a \in \mathbb{R}$ mit $a > 0$ und $a \neq 1$.

a. Die Abbildung

$$\exp_a : \mathbb{R} \longrightarrow (0, \infty) : x \mapsto a^x$$

heißt *Exponentialfunktion zur Basis* a, ist stetig, bijektiv und
- streng monoton wachsend, falls $a > 1$, und
- streng monoton fallend, falls $a < 1$.

b. Die Umkehrabbildung

$$\log_a : (0, \infty) \longrightarrow \mathbb{R}$$

von $\exp_a$ heißt *Logarithmus zur Basis* a, ist stetig, bijektiv und
- streng monoton wachsend, falls $a > 1$, und
- streng monoton fallend, falls $a < 1$.

Beweis: Für $a > 1$ ist $\ln(a) > 0$, da $\ln(1) = 0$ und ln streng monoton wachsend, so daß aus $x < y$ auch

$$\ln(a) \cdot x < \ln(a) \cdot y$$

und damit

$$\exp_a(x) = \exp\big(x \cdot \ln(a)\big) < \exp\big(y \cdot \ln(a)\big) = \exp_a(y)$$

folgt. $\exp_a$ ist dann also streng monoton wachsend. Außerdem gilt

$$\lim_{x\to\infty} x \cdot \ln(a) = \infty \quad \text{sowie} \quad \lim_{x\to-\infty} x \cdot \ln(a) = -\infty,$$

und da exp stetig ist folgt dann auch

$$\lim_{x\to\infty} \exp_a(x) = \lim_{x\to\infty} \exp\big(x \cdot \ln(a)\big) = \infty$$

sowie

$$\lim_{x\to-\infty} \exp_a(x) = \lim_{x\to-\infty} \exp\big(x \cdot \ln(a)\big) = 0.$$

Aus dem Umkehrsatz für streng monotone Funktionen 14.21 folgt dann, weil $\exp_a$ stetig, und streng monoton wachsend ist, daß $\exp_a$ auch bijektiv ist. Zudem folgen die entsprechenden Aussagen über die Umkehrfunktion $\log_a$ für $a > 1$.

Den Fall $a < 1$ beweist man analog, da dann $\ln(a) < 0$ gilt. □

B) Potenz- und Logarithmusgesetze

Korollar 16.8 (Potenzgesetze).
Seien $a, b, x, y \in \mathbb{R}$ mit $a, b > 0$.

a. $a^{x+y} = a^x \cdot a^y$.

b. $a^{x\cdot y} = (a^x)^y$.

c. $(a \cdot b)^x = a^x \cdot b^x$.

d. $a^{-x} = \frac{1}{a^x}$.

e. Für $n \in \mathbb{Z}$ stimmen die Definitionen von a^n in 7.9 und 16.6 überein.

f. Für $p, q \in \mathbb{Z}$ mit $q \geq 2$ gilt $a^{\frac{p}{q}} = \sqrt[q]{a^p}$.
Insbesondere stimmen die Definitionen von $a^{\frac{1}{q}}$ in 9.8 und 16.6 überein.

Beweis:

a. Mit Hilfe der Funktionalgleichung für die Exponentialfunktion sieht man:

$$\begin{aligned} a^{x+y} &= \exp\big((x+y) \cdot \ln(a)\big) = \exp\big(x \cdot \ln(a) + y \cdot \ln(a)\big) \\ &= \exp\big(x \cdot \ln(a)\big) \cdot \exp\big(y \cdot \ln(a)\big) = a^x \cdot a^y. \end{aligned}$$

b. Wegen $a^x = \exp(x \cdot \ln(a))$ gilt auch

$$\ln(a^x) = \ln\big(\exp(x \cdot \ln(a))\big) = x \cdot \ln(a)$$

und damit

$$(a^x)^y = \exp\big(y \cdot \ln(a^x)\big) = \exp\big(x \cdot y \cdot \ln(a)\big) = a^{x \cdot y}.$$

c. Wir verwenden in der folgenden Gleichung bereits ein Logarithmusgesetz 16.9, dessen Beweis unabhängig von diesem Potenzgesetz ist:

$$\begin{aligned}(a \cdot b)^x &= \exp\big(x \cdot \ln(a \cdot b)\big) \overset{16.9b.}{=} \exp\big(x \cdot (\ln(a) + \ln(b))\big) \\ &= \exp\big(x \cdot \ln(a)\big) \cdot \exp\big(x \cdot \ln(b)\big) = a^x \cdot b^x.\end{aligned}$$

d. Die Gleichung $a^{-x} = \frac{1}{a^x}$ folgt unmittelbar aus

$$a^x \cdot a^{-x} = a^{x-x} = a^0 = \exp(0 \cdot \ln(a)) = \exp(0) = 1.$$

e. Mit Induktion nach $n \geq 1$ und a. sieht man, daß $a^n = \prod_{k=1}^{n} a$. Wir haben bereits gesehen, daß zudem $a^0 = 1$ gilt, und aus d. folgt dann

$$a^{-n} = \frac{1}{a^n} = \frac{1}{\prod_{k=1}^{n} a} = \prod_{k=1}^{n} \frac{1}{a}.$$

Die Definitionen von a^n in 7.9 und 16.6 stimmen also überein.

f. Mit Satz 9.8 folgt $a^{\frac{p}{q}} = \sqrt[q]{a^p}$ aus

$$\big(a^{\frac{p}{q}}\big)^q \overset{b.}{=} a^{\frac{p}{q} \cdot q} = a^p.$$

□

Korollar 16.9 (Logarithmusgesetze).
Seien $a, x, y \in \mathbb{R}_{>0}$ mit $a \neq 1$ und $z \in \mathbb{R}$.

a. $\log_a(x) = \frac{\ln(x)}{\ln(a)}$.

b. $\log_a(x \cdot y) = \log_a(x) + \log_a(y)$.

c. $\log_a(x^z) = z \cdot \log_a(x)$.

d. $\log_a(\frac{x}{y}) = \log_a(x) - \log_a(y)$.

Beweis:

a. Falls $a \neq 1$, so gilt

$$\exp_a\left(\frac{\ln(x)}{\ln(a)}\right) = \exp\left(\frac{\ln(x)}{\ln(a)} \cdot \ln(a)\right) = \exp\big(\ln(x)\big) = x = \exp_a\big(\log_a(x)\big),$$

und da $\exp_a$ injektiv ist, gilt dann auch

$$\frac{\ln(x)}{\ln(a)} = \log_a(x).$$

b. Es gilt

$$\begin{aligned}\exp_a\big(\log_a(x \cdot y)\big) &= x \cdot y = \exp_a\big(\log_a(x)\big) \cdot \exp_a\big(\log_a(y)\big) \\ &\overset{16.8a.}{=} \exp_a\big(\log_a(x) + \log_a(y)\big),\end{aligned}$$

und da $\exp_a$ injektiv ist, folgt somit

$$\log_a(x \cdot y) = \log_a(x) + \log_a(y).$$

c. Falls $a \neq 1$ und $x > 0$, so ist

$$x^z = \exp\big(z \cdot \ln(x)\big) \overset{a.}{=} \exp\big(z \cdot \log_a(x) \cdot \ln(a)\big) = \exp_a\big(z \cdot \log_a(x)\big)$$

definiert. Wenden wir auf beiden Seiten die Funktion $\log_a$ an, so erhalten wir

$$\log_a\big(x^z\big) = \log_a\big(\exp_a\big(z \cdot \log_a(x)\big)\big) = z \cdot \log_a(x).$$

d. Es gilt

$$\log_a\left(\frac{x}{y}\right) \overset{b.}{=} \log_a(x) + \log_a(y^{-1}) \overset{c.}{=} \log_a(x) - \log_a(y).$$

□

C) Die Zahl π

Wir wollen uns nun den trigonometrischen Funktionen zuwenden. Dazu führen wir zunächst die Zahl π als kleinste positive Nullstelle des Sinus ein.

Satz 16.10 (Definition der Zahl π.).
Der Sinus besitzt eine kleinste positive Nullstelle, die wir π nennen, und für alle $x \in (0, \pi)$ gilt $\sin(x) > 0$.

Beweis: Wir wählen ein $x \in (0, 4]$ und setzen $a_n := \frac{x^{2n+1}}{(2n+1)!}$ für $n \geq 1$. Die Folge $(a_n)_{n\geq 1}$ ist monoton fallend, denn

$$a_{n+1} = \frac{x^{2n+3}}{(2n+3)!} = \frac{x^{2n+1}}{(2n+1)!} \cdot \frac{x \cdot x}{(2n+3) \cdot (2n+2)} \leq \frac{x^{2n+1}}{(2n+1)!} \cdot \frac{16}{20} < a_n,$$

und da die Reihe $\sum_{n=1}^{\infty}(-1)^n \cdot a_n = \sin(x) - x$ absolut konvergiert, muß die Folge $(a_n)_{n\geq 1}$ auch eine Nullfolge sein. Aus dem Beweis des Leibniz-Kriteriums erfüllen die Partialsummen der Reihe $\sum_{n=1}^{\infty}(-1)^n \cdot a_n$ dann insbesondere

$$s_1 \leq \sum_{n=1}^{\infty}(-1)^n \cdot a_n \leq s_4,$$

und damit

$$x - \frac{x^3}{6} = x + s_1 \leq \sin(x) \leq x + s_4. \tag{28}$$

Wenden wir dies für $x = 1$ an, so erhalten wir

$$\sin(1) \geq 1 - \frac{1}{6} > 0,$$

und wenden wir die Aussage für $x = 4$ an, so erhalten wir

$$\sin(4) \leq 4 + s_4(4) = -\frac{268}{405} < 0.$$

Da der Sinus auf dem abgeschlossenen Intervall $[1, 4]$ stetig ist mit $\sin(1) > 0$ und $\sin(4) < 0$, muß er nach dem Zwischenwertsatz 14.12 eine Nullstelle besitzen, das heißt, die Menge

$$A := \{x \in [1, 4] \mid \sin(x) = 0\}$$

ist nicht leer und nach unten beschränkt. Dann existiert aber ihr Infimum

$$\pi := \inf(A).$$

Nach Bemerkung 11.22 gibt es eine monoton fallende Folge $(b_n)_{n\in\mathbb{N}}$ in A, die gegen das Infimum $\pi = \inf(A)$ konvergiert. Da der Sinus stetig ist, gilt dann auch

$$0 = \sin(b_n) \longrightarrow \sin(\pi),$$

also $\sin(\pi) = 0$, d.h. π ist eine Nullstelle des Sinus.

Wir müssen nun noch zeigen, daß $\sin(x) > 0$ für alle $x \in (0, \pi)$. Für $x \in [1, \pi)$ ist dies der Fall, da entweder aus $\sin(x) = 0$ oder aus $\sin(x) < 0$ und $\sin(1) > 0$ mit Hilfe des Zwischenwertsatzes 14.12 die Existenz einer kleineren Nullstelle des Sinus als π im Intervall $[1, 4]$ folgen würde. Für $x \in (0, 1)$ folgt aber aus (28)

$$\sin(x) \geq x - \frac{x^3}{6} > x - \frac{x}{6} \geq \frac{5x}{6} > 0,$$

da $x^3 < x$. Also haben wir $\sin(x) > 0$ für alle $x \in (0, \pi)$ gezeigt, so daß π die kleinste positive Nullstelle des Sinus ist. □

Bemerkung 16.11 (Approximation von π).
Aus dem Beweis von Satz 16.10 wissen wir bislang nur, daß $1 < \pi < 4$ gilt. Wir werden später sehen (siehe Aufgabe 18.44), daß man die Zahl π approximieren kann durch

$$3,14159\ldots.$$

D) Der Cosinus

Satz 16.12 (Monotonie des Cosinus).
Der Cosinus $\cos : [0,\pi] \longrightarrow [-1,1]$ ist streng monoton fallend und bijektiv.

Beweis: Es seien $x, y \in [0,\pi]$ mit $x < y$. Aus dem Additionstheorem für den Cosinus sowie der Tatsache, daß der Cosinus eine gerade und der Sinus eine ungerade Funktion ist (siehe Satz 12.38), folgen

$$\begin{aligned}\cos(y) &= \cos\left(\frac{y+x}{2} + \frac{y-x}{2}\right)\\ &= \cos\left(\frac{y+x}{2}\right)\cdot\cos\left(\frac{y-x}{2}\right) - \sin\left(\frac{y+x}{2}\right)\cdot\sin\left(\frac{y-x}{2}\right)\end{aligned}$$

und

$$\begin{aligned}\cos(x) &= \cos\left(\frac{y+x}{2} - \frac{y-x}{2}\right)\\ &= \cos\left(\frac{y+x}{2}\right)\cdot\cos\left(-\frac{y-x}{2}\right) - \sin\left(\frac{y+x}{2}\right)\cdot\sin\left(-\frac{y-x}{2}\right)\\ &= \cos\left(\frac{y+x}{2}\right)\cdot\cos\left(\frac{y-x}{2}\right) + \sin\left(\frac{y+x}{2}\right)\cdot\sin\left(\frac{y-x}{2}\right).\end{aligned}$$

Subtrahieren wir die beiden Gleichungen voneinander, so erhalten wir

$$\cos(x) - \cos(y) = 2\cdot\sin\left(\frac{y+x}{2}\right)\cdot\sin\left(\frac{y-x}{2}\right) \overset{16.10}{>} 0,$$

da mit $x, y \in [0,\pi]$ und $x < y$ auch

$$0 < \frac{y+x}{2} < \pi$$

und

$$0 < \frac{y-x}{2} < \pi$$

gelten muß.

Somit haben wir gezeigt, daß der Cosinus auf dem Intervall $[0,\pi]$ streng monoton fallend ist. Aus dem Umkehrsatz 14.21 folgt damit, daß

$$\cos : [0,\pi] \longrightarrow [\lim_{x\to\pi}\cos(x), \lim_{x\to 0}\cos(x)]$$

bijektiv ist. Da der Cosinus stetig ist, gilt nun

$$\lim_{x\to 0}\cos(x) = \cos(0) = \sum_{n=0}^{\infty}(-1)^n\cdot\frac{0^{2n}}{(2n)!} = 1$$

und

$$\lim_{x\to\pi}\cos(x) = \cos(\pi).$$

Aus Satz 12.38 wissen wir zudem, daß

$$\cos(\pi)^2 = 1 - \sin(\pi)^2 = 1 - 0 = 1$$

gilt, so daß $\cos(\pi) \in \{1, -1\}$. Da der Cosinus auf dem Intervall $[0, \pi]$ streng monoton fallend mit $\cos(0) = 1$ ist, muß somit $\cos(\pi) = -1$ gelten. □

E) Eigenschaften des Sinus' und des Cosinus'

Satz 16.13 (Eigenschaften des Sinus' und des Cosinus').

a. Für $x \in \mathbb{R}$ gelten

$$\sin(x + \pi) = -\sin(x) \quad \text{und} \quad \cos(x + \pi) = -\cos(x).$$

b. Für $x \in \mathbb{R}$ gelten zudem

$$\sin(x) \in [-1, 1] \quad \text{und} \quad \cos(x) \in [-1, 1].$$

c. Wir können folgende Werte des Sinus und Cosinus explizit angeben:

x	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	π	$\frac{3\cdot\pi}{2}$	2π
$\sin(x)$	0	$\frac{1}{\sqrt{2}}$	1	0	-1	0
$\cos(x)$	1	$\frac{1}{\sqrt{2}}$	0	-1	0	1

d. Sinus und Cosinus sind 2π-periodisch, d.h. für $x \in \mathbb{R}$ gelten

$$\sin(x + 2\pi) = \sin(x) \quad \text{und} \quad \cos(x + 2\pi) = \cos(x).$$

e. Die Perioden von Sinus und Cosinus sind um $\frac{\pi}{2}$ verschoben, d.h. für $x \in \mathbb{R}$ gelten

$$\sin\left(x + \frac{\pi}{2}\right) = \cos(x)$$

und

$$\cos\left(x - \frac{\pi}{2}\right) = \sin(x).$$

f. Der Sinus

$$\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \longrightarrow [-1, 1]$$

ist streng monoton wachsend und bijektiv.

g. Die Nullstellen des Sinus sind genau die ganzzahligen Vielfachen von π, d.h.

$$\sin(x) = 0 \iff x \in \{k \cdot \pi \mid k \in \mathbb{Z}\},$$

und für den Cosinus gilt mithin

$$\cos(x) = 0 \iff x \in \left\{\frac{\pi}{2} + k \cdot \pi \;\middle|\; k \in \mathbb{Z}\right\}.$$

Beweis:

a. Aus den Additionstheoremen 12.38 erhalten wir

$$\begin{aligned}\sin(x+\pi) &= \sin(x)\cdot\cos(\pi)+\sin(\pi)\cdot\cos(x)\\ &= \sin(x)\cdot(-1)+0\cdot\cos(x) = -\sin(x)\end{aligned}$$

und

$$\begin{aligned}\cos(x+\pi) &= \cos(x)\cdot\cos(\pi)-\sin(x)\cdot\sin(\pi)\\ &= \cos(x)\cdot(-1)-\sin(x)\cdot 0 = -\cos(x).\end{aligned}$$

b. Nach Satz 12.38 gilt

$$|\sin(x)| \le \sqrt{\cos(x)^2+\sin(x)^2} = 1$$

und

$$|\cos(x)| \le \sqrt{\cos(x)^2+\sin(x)^2} = 1.$$

c. Die Werte für $x = 0$ folgen unmittelbar aus der Definition von Sinus und Cosinus als Potenzreihen

$$\sin(0) = \sum_{n=0}^{\infty}(-1)^n\cdot\frac{0^{2n+1}}{(2n+1)!} = 0$$

und

$$\cos(0) = \sum_{n=0}^{\infty}(-1)^n\cdot\frac{0^{2n}}{(2n)!} = 1.$$

Die Werte für $x = \pi$ folgen aus Satz 16.10 und Satz 16.12 oder alternativ aus Teil a.. Mit Hilfe der Additionstheoreme 12.38 folgt dann

$$-1 = \cos(\pi) = \cos\left(\frac{\pi}{2}+\frac{\pi}{2}\right) = \cos\left(\frac{\pi}{2}\right)^2-\sin\left(\frac{\pi}{2}\right)^2$$

und somit

$$0 \le \cos\left(\frac{\pi}{2}\right)^2 = \sin\left(\frac{\pi}{2}\right)^2-1 \overset{b.}{\le} 0.$$

Damit müssen notwendigerweise

$$\cos\left(\frac{\pi}{2}\right) = 0$$

und

$$\sin\left(\frac{\pi}{2}\right) \in \{-1,1\}$$

gelten. Da wir aber aus Satz 16.10 wissen, daß der Sinus auf dem Intervall $(0,\pi)$ strikt positiv ist, folgt

$$\sin\left(\frac{\pi}{2}\right) = 1.$$

Aus dem Additionstheorem für den Cosinus erhalten wir dann

$$0 = \cos\left(\frac{\pi}{2}\right) = \cos\left(\frac{\pi}{4}+\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right)^2-\sin\left(\frac{\pi}{4}\right)^2$$

und damit

$$\cos\left(\frac{\pi}{4}\right)^2 = \sin\left(\frac{\pi}{4}\right)^2.$$

Aus Satz 12.38 wissen wir zudem

$$1 = \cos\left(\frac{\pi}{4}\right)^2 + \sin\left(\frac{\pi}{4}\right)^2 = 2 \cdot \cos\left(\frac{\pi}{4}\right)^2,$$

und damit

$$\cos\left(\frac{\pi}{4}\right), \sin\left(\frac{\pi}{4}\right) \in \left\{-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right\}.$$

Nach Satz 16.12 ist der Cosinus auf $[0, \pi]$ streng monoton fallend mit Nullstelle bei $\frac{\pi}{2}$, also muß $\cos\left(\frac{\pi}{4}\right)$ positiv sein, und aus Satz 16.10 wissen wir, daß auch $\sin\left(\frac{\pi}{4}\right)$ positiv ist. Mithin gilt

$$\sin\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}.$$

Die übrigen Werte folgen, indem wir Teil a. auf die bisherigen Ergebnisse anwenden.

d. Durch Anwenden der Additionstheoreme 12.38 erhalten wir wie in Teil a.

$$\begin{aligned}\sin(x + 2\pi) &= \sin(x) \cdot \cos(2\pi) + \sin(2\pi) \cdot \cos(x)\\ &= \sin(x) \cdot 1 + 0 \cdot \cos(x) = \sin(x)\end{aligned}$$

und

$$\begin{aligned}\cos(x + 2\pi) &= \cos(x) \cdot \cos(2\pi) - \sin(x) \cdot \sin(2\pi)\\ &= \cos(x) \cdot 1 - \sin(x) \cdot 0 = \cos(x).\end{aligned}$$

e. Den Beweis überlassen wir dem Leser als Übungsaufgabe.

f. Aus Teil a. und e. folgt

$$\sin(x) = \cos\left(x - \frac{\pi}{2}\right) = -\cos\left(x + \frac{\pi}{2}\right),$$

so daß die Aussage aus Satz 16.12 folgt.

g. Aus Teil a. und $\sin(\pi) = 0$ folgt mit Induktion, daß $\sin(k \cdot \pi) = 0$ für alle $k \in \mathbb{Z}$. Ist umgekehrt $x \in \mathbb{R}$ eine Nullstelle des Sinus, so gibt es eine ganze Zahl $k \in \mathbb{Z}$, so daß

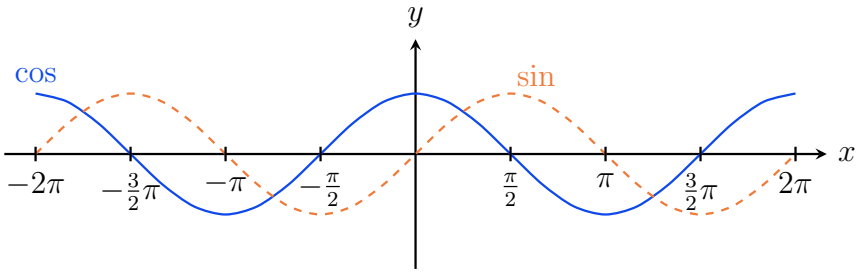
$$0 \leq x - k \cdot \pi < \pi,$$

und da der Sinus im Intervall $(0, \pi)$ keine Nullstelle besitzt, muß mithin $x = k \cdot \pi$ gelten. Aus Teil e. folgt dann die Aussage für die Nullstellen des Cosinus.

□

Bemerkung 16.14.

Aus Satz 16.13 können wir den Verlauf der Graphen des Sinus und des Cosinus im wesentlichen herleiten:



F) Polarkoordinaten

Bemerkung 16.15 (Polarkoordinaten).
Ist $x \in \mathbb{R}$ und betrachten wir ein rechtwinkliges Dreieck mit den Kathetenlängen $\sin(x)$ und $\cos(x)$, so folgt wegen

$$\cos(x)^2 + \sin(x)^2 = 1 = 1^2$$

aus dem Satz von Pythagoras, daß die Hypothenuse die Seitenlänge 1 besitzt. D.h. der Punkt

$$(\cos(x), \sin(x)) = \cos(x) + i \cdot \sin(x) = \exp(i \cdot x) \in \mathbb{C}$$

liegt auf dem Einheitskreis und wir nennen x den Winkel im Bogenmaß, den der Strahl vom Ursprung durch diesen Punkt mit der x-Achse einschließt (siehe Abbildung 8).

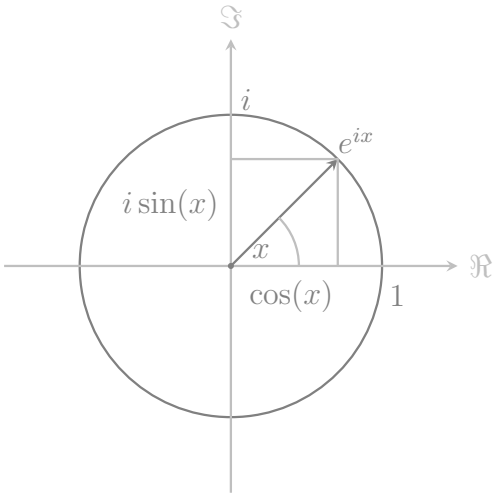


Abbildung 8: Polarkoordinaten von $e^{i \cdot x} = \cos(x) + i \cdot \sin(x)$

Ist umgekehrt $z = a + ib = (a, b)$ ein Punkt auf dem Einheitskreis, so folgt aus $1 = a^2 + b^2$ sofort, daß $a \in [-1, 1]$ liegt. Da der Cosinus bijektiv auf dem Intervall $[0, \pi]$ mit Bild $[-1, 1]$ ist, gibt es genau ein $x \in [0, \pi]$ mit $a = \cos(x)$, und es gilt

$$\sin(x)^2 = 1 - \cos(x)^2 = 1 - a^2 = b^2,$$

d.h. $b \in \{-\sin(x), \sin(x)\} = \{\sin(-x), \sin(x)\}$. Wegen $a = \cos(x) = \cos(-x)$ finden wir also ein $y \in [-\pi, \pi]$ mit

$$z = a + ib = \cos(y) + i \cdot \sin(y) = e^{i \cdot y},$$

d.h. jeder Punkt auf dem Einheitskreis hat die Gestalt $z = e^{i \cdot y}$ mit $y \in \mathbb{R}$. Genauer kann man sogar sagen, daß es genau ein solches $y \in [-\pi, \pi)$ gibt.

Wir haben damit die Behauptung aus Bemerkung 10.9 gezeigt, daß jede komplexe Zahl z sich schreiben läßt als

$$z = |z| \cdot e^{i \cdot \arg z},$$

und wir können $\arg(z)$ im Intervall $[-\pi, \pi)$ eindeutig wählen. Wir nennen diese Darstellung die *Polarkoordinatendarstellung* von z.

Außerdem haben wir damit auch Bemerkung 10.11 gezeigt, daß nämlich jede komplexe Zahl eine n-te Wurzel besitzt, da wir dazu nur die Polarkoordinatendarstellung von z benötigt haben.

Man beachte, daß für $n \geq 2$ die Zahlen

$$e^{\frac{2 \cdot k \cdot \pi \cdot i}{n}} \quad \text{mit} \quad k = 0, \ldots, n-1$$

genau die n-ten Wurzeln aus 1 sind. Man nennt sie auch die *n-ten Einheitswurzeln*.

Daß sie in der Tat n-te Wurzeln von 1 sind, folgt unmittelbar aus

$$\left(e^{\frac{2 \cdot k \cdot \pi \cdot i}{n}}\right)^n = e^{2 \cdot k \cdot \pi \cdot i} = \cos(2 \cdot k \cdot \pi) + i \cdot \sin(2 \cdot k \cdot \pi) = 1.$$

Und daß es keine weiteren n-ten Wurzeln geben kann, folgt aus der Tatsache, daß jede n-te Wurzel eine Nullstelle des Polynoms $t^n - 1$ ist und dieses nach Bemerkung 13.12 höchstens n verschiedene Nullstellen besitzen kann.

G) Weitere trigonometrische Funktionen

Definition 16.16 (Tangens und Cotangens).
Die Funktion

$$\tan : \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k \cdot \pi \;\middle|\; k \in \mathbb{Z} \right\} \longrightarrow \mathbb{R} : x \mapsto \frac{\sin(x)}{\cos(x)}$$

heißt *Tangens* und die Funktion

$$\cot : \mathbb{R} \setminus \{k \cdot \pi \mid k \in \mathbb{Z}\} \longrightarrow \mathbb{R} : x \mapsto \frac{\cos(x)}{\sin(x)}$$

heißt *Cotangens*.

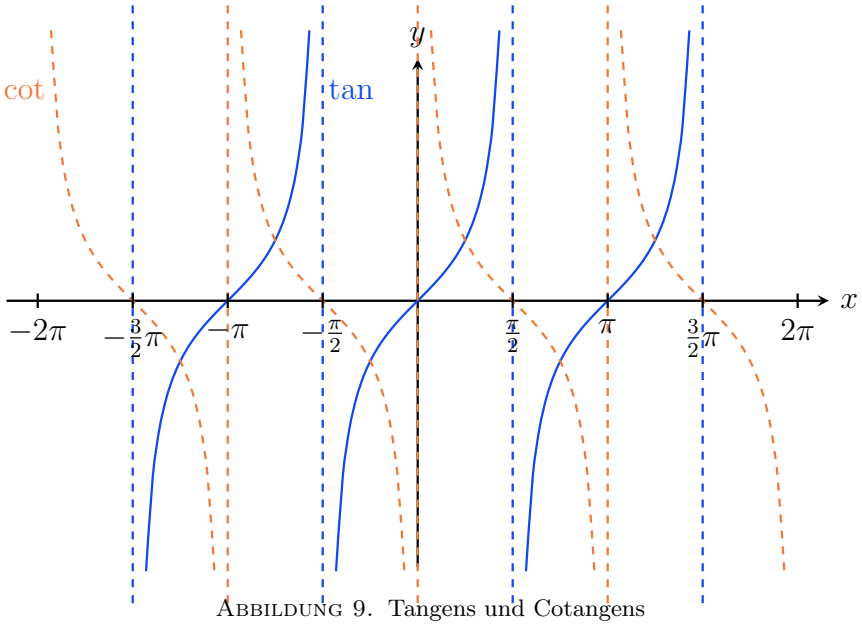


Abbildung 9. Tangens und Cotangens

Satz 16.17 (Tangens und Cotangens).

a. Für $x \in \mathbb{R}$ gelten

$$\tan(-x) = -\tan(x) \quad \text{und} \quad \cot(-x) = -\cot(x)$$

und

$$\tan(x+\pi) = \tan(x) \quad \text{und} \quad \cot(x+\pi) = \cot(x).$$

b. Der Tangens ist auf jedem der Intervalle $\left(-\frac{\pi}{2} + k \cdot \pi, \frac{\pi}{2} + k \cdot \pi\right)$, $k \in \mathbb{Z}$, streng monoton wachsend, stetig, bijektiv mit Bild $\mathbb{R}$ und punktsymmetrisch zu seiner Nullstelle $k \cdot \pi$.

c. Der Cotangens ist auf jedem der Intervalle $(k \cdot \pi, (k+1) \cdot \pi)$, $k \in \mathbb{Z}$, streng monoton fallend, stetig, bijektiv mit Bild $\mathbb{R}$ und punktsymmetrisch zu seiner Nullstelle $\frac{2k+1}{2} \cdot \pi$.

Beweis:

a. Die Aussagen folgen unmittelbar aus den entsprechenden Aussagen für den Sinus und Cosinus in Satz 12.38 und Satz 16.13.

b. Für $0 \leq x < y < \frac{\pi}{2}$ folgt aus der Monotonie des Cosinus (Satz 16.12) und des Sinus (Satz 16.13)

$$\cos(x) > \cos(y) > 0$$

und

$$0 < \sin(x) < \sin(y),$$

so daß mithin

$$\tan(x) = \frac{\sin(x)}{\cos(x)} < \frac{\sin(y)}{\cos(y)} = \tan(y).$$

Der Tangens ist auf dem Intervall $\left[0, \frac{\pi}{2}\right)$ also streng monoton wachsend. Wegen $\tan(-x) = -\tan(x)$ ist der Tangens aber punktsymmetrisch zum Ursprung und somit auch streng monoton wachsend auf $\left(-\frac{\pi}{2}, 0\right]$, also streng monoton wachsend auf $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.

Wegen der Stetigkeit von Sinus und Cosinus erhalten wir zudem

$$\lim_{x \to \frac{\pi}{2}} \tan(x) = \frac{\lim\limits_{x \to \frac{\pi}{2}} \sin(x)}{\lim\limits_{x \to \frac{\pi}{2}} \cos(x)} = \frac{1}{0} = \infty$$

und

$$\lim_{x \to -\frac{\pi}{2}} \tan(x) = \frac{\lim\limits_{x \to -\frac{\pi}{2}} \sin(x)}{\lim\limits_{x \to -\frac{\pi}{2}} \cos(x)} = \frac{-1}{0} = -\infty.$$

Aus dem Umkehrsatz 14.21 folgt dann, daß die stetige Funktion

$$\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \longrightarrow (-\infty, \infty) = \mathbb{R}$$

bijektiv ist. Aus Teil a. folgt die Aussage zur Punktsymmetrie, und zudem die Aussage für die verschobenen Intervalle.

c. Die Aussage wird analog zur Aussage in Teil b. bewiesen.

□

Satz 16.18 (Arcusfunktionen).

a. Die Funktion $\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \longrightarrow [-1, 1]$ besitzt eine stetige, streng monoton wachsende Umkehrabbildung, die wir *Arcussinus* nennen,

$$\arcsin : [-1, 1] \longrightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

b. Die Funktion $\cos : [0, \pi] \longrightarrow [-1, 1]$ besitzt eine stetige, streng monoton fallende Umkehrabbildung, die wir *Arcuscosinus* nennen,

$$\arccos : [-1, 1] \longrightarrow [0, \pi].$$

c. Die Funktion $\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \longrightarrow \mathbb{R}$ besitzt eine stetige, streng monoton wachsende Umkehrabbildung, die wir *Arcustangens* nennen,

$$\arctan : \mathbb{R} \longrightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

d. Die Funktion $\cot : (0, \pi) \longrightarrow \mathbb{R}$ besitzt eine stetige, streng monoton fallende Umkehrabbildung, die wir *Arcuscotangens* nennen,

$$\operatorname{arccot} : \mathbb{R} \longrightarrow (0, \pi).$$

Beweis: Die Aussagen folgen aus dem Umkehrsatz 14.21 zusammen mit den Monotonieaussagen in den Sätzen 16.12, 16.13 und 16.17. □

Aufgaben

Aufgabe 16.19.
Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ stetig mit $f(x+y) = f(x)\cdot f(y)$ für alle $x, y \in \mathbb{R}$ und $f(1) = a > 0$. Zeige, dann ist $f = \exp_a$.

Aufgabe 16.20.
Löse die folgenden Gleichungen für $x \in \mathbb{R}$:

a. $\log_2(x) - \log_2(x-6) = 3$.

b. $4^x + 4 = 2^{x+2} + 2^x$.

c. $\log_4(x+2) - \log_4(x-2) = \frac{1}{2}$.

d. $2^{3-x} \cdot 3^{x-1} = 6^{2x-3}$.

Aufgabe 16.21.
Computer speichern Daten in Form von Bits, d.h. in Variablen, die nur die Wert 0 oder 1 annehmen können.

a. Wie viele Bits benötigt ein Computer, um die Zahl 65^{365} zu speichern?

b. Ein Algorithmus benötige bei einem Eingabedatum mit n Bits $n\log_2(n)$ Rechenschritte. Berechne die Laufzeit in Rechenschritten für die Wert $n = 10$, $n = 256$ und $n = 1000$.

c. Was ist die kleinste Eingabe in Bits, für die der Algorithmus mindestens 100 Rechenschritte benötigt?

Aufgabe 16.22.

a. Zeige, ist $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge mit $0 \neq a_n \in \mathbb{R}_{>-1}$ für alle $n \in \mathbb{N}$, so gilt
$$\lim_{n\to\infty} (1+a_n)^{\frac{1}{a_n}} = e.$$

b. Für $x \in \mathbb{R}$ zeige $\lim_{n\to\infty} \left(1+\frac{x}{n}\right)^n = e^x$.

c. Für $a \in \mathbb{R}_{>0}$ zeige $\lim_{n\to\infty} n\cdot(\sqrt[n]{a}-1) = \ln(a)$.

Aufgabe 16.23.
Bestimme das Bild der folgenden Funktionen f und entscheide, ob diese injektiv sind. Falls die Umkehrfunktion $f^{-1} : \mathrm{Im}(f) \longrightarrow \mathbb{R}$ existiert, ist sie dann stetig?

a. $f : [-1, 1] \longrightarrow \mathbb{R} : x \mapsto \log_2(x^2 + 1)$.

b. $f : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto \sum_{n=1}^{20} \frac{x}{n} \cdot e^{nx^2+n}$.

Aufgabe 16.24 (Additionstheoreme für Tangens und Arcustangens).

a. Zeige das folgende Additionstheorem für den Tangens:

$$\tan(x + y) = \frac{\tan(x) + \tan(y)}{1 - \tan(x) \cdot \tan(y)},$$

wobei $x, y, x + y \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ gelten soll.

b. Folgere daraus das folgende Additionstheorem für den Arcustangens:

$$\arctan(x) + \arctan(y) = \arctan\left(\frac{x + y}{1 - xy}\right).$$

c. Zeige die Gleichung

$$4 \cdot \arctan\left(\frac{1}{5}\right) - \arctan\left(\frac{1}{239}\right) = \frac{\pi}{4}.$$

Aufgabe 16.25.

a. Zeige, für alle $x \in \mathbb{R}$ gilt die Gleichung

$$\sin(3x) = 3 \cdot \sin(x) - 4 \cdot \sin^3(x).$$

b. Zeige mit Hilfe der Formel in Teil a. die Gleichung

$$\sin\left(\frac{\pi}{6}\right) = \frac{1}{2}.$$

Aufgabe 16.26 (Rechenregeln für die Arcusfunktionen).
Beweise die folgenden Gleichungen:

a. $\sin\big(\arccos(x)\big) = \sqrt{1 - x^2}$ für $x \in [-1, 1]$.

b. $\cos\big(\arcsin(x)\big) = \sqrt{1 - x^2}$ für $x \in [-1, 1]$.

c. $\sin\big(\arctan(x)\big) = \frac{x}{\sqrt{1+x^2}}$ für $x \in \mathbb{R}$.

d. $\cos\big(\arctan(x)\big) = \frac{1}{\sqrt{1+x^2}}$ für $x \in \mathbb{R}$.

e. $\arcsin(x) = \arctan\left(\frac{x}{\sqrt{1-x^2}}\right)$ für $x \in (-1, 1)$.

f. $\arccos(x) = \frac{\pi}{2} - \arctan\left(\frac{x}{\sqrt{1-x^2}}\right)$ für $x \in (-1, 1)$.

g. $\arcsin(x) = 2 \cdot \arctan\left(\frac{x}{1+\sqrt{1-x^2}}\right)$ für $x \in (-1, 1)$.

Aufgabe 16.27 (Eingangsfrage aus dem JEE advanced 2022).
Berechne die folgenden Ausdruck

$$\frac{3}{2} \cdot \arccos\left(\sqrt{\frac{2}{2+\pi^2}}\right) + \frac{1}{4} \cdot \arcsin\left(\frac{2 \cdot \sqrt{2} \cdot \pi}{2+\pi^2}\right) + \arctan\left(\frac{\sqrt{2}}{\pi}\right).$$

§ 17 Differenzierbarkeit

A) Der Differenzenquotient

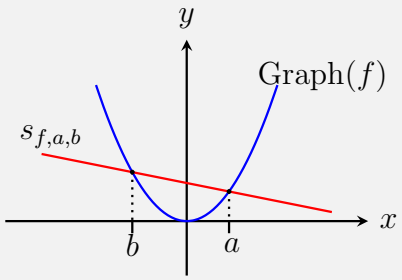
Definition 17.1 (Differenzenquotient).
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ und $a \in U$. Die Funktion

$$\mathrm{Diff}_{f,a} : U \setminus \{a\} \longrightarrow \mathbb{R} : x \mapsto \frac{f(x) - f(a)}{x - a}$$

heißt der *Differenzenquotient* von f an der Stelle a.
Für ein festes b ist der Wert des Differenzenquotienten $\mathrm{Diff}_{f,a}(b)$ die *Steigung* der *Sekante* $s_{f,a,b}$ an den Graphen von f durch die Punkte $(b, f(b))$ und $(a, f(a))$, deren Geradengleichung durch

$$\begin{aligned} y =& \frac{f(b) - f(a)}{b - a} \cdot x + \frac{f(a) \cdot b - f(b) \cdot a}{b - a} \\ =& \frac{f(b) - f(a)}{b - a} \cdot x + f(a) - \frac{f(b) - f(a)}{b - a} \cdot a \end{aligned} \tag{29}$$

gegeben ist.



Beispiel 17.2.
Ist $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^n$ mit $n \geq 1$, so ist

$$\mathrm{Diff}_{f,a}(x) = \frac{x^n - a^n}{x - a} = x^{n-1} + a \cdot x^{n-2} + a^2 \cdot x^{n-3} + \ldots + a^{n-2} \cdot x + a^{n-1}$$

für $x \in \mathbb{R} \setminus \{a\}$.

B) Differenzierbarkeit und die Ableitung

Definition 17.3.
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ eine Funktion und $a \in U$. Wir nennen f *differenzierbar in* a, wenn a ein Häufungspunkt von U ist und der Grenzwert

$$\lim_{x \to a} \mathrm{Diff}_{f,a}(x) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \in \mathbb{R}$$

des Differenzenquotienten in a existiert. In diesem Fall schreiben wir

$$f'(a) := \frac{\partial f}{\partial x}(a) := \lim_{x \to a} \frac{f(x) - f(a)}{x - a},$$

und nennen diesen Grenzwert die *Ableitung* von f an der Stelle a.
Wir nennen die Funktion f *differenzierbar (auf U)*, wenn f in jedem Punkt von U differenzierbar ist. In diesem Fall nennen wir die Funktion

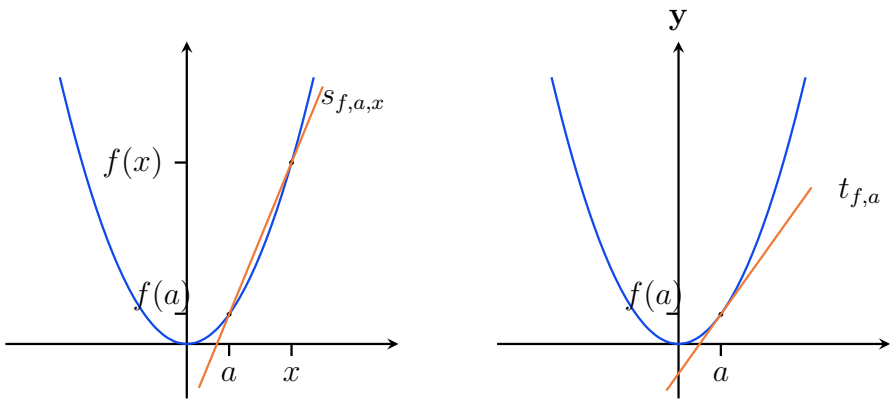
$$f' : U \longrightarrow \mathbb{R} : x \mapsto f'(x)$$

die *Ableitung* von f. Beachte auch, daß dann insbesondere jeder Punkt von U ein Häufungspunkt von U sein muß!

Bemerkung 17.4.
Der Definition liegt die Idee zugrunde, daß sich die Sekante $s_{f,a,x}$ für $x \to a$ einer Geraden annähert, die im Punkt $(a, f(a))$ den Graphen von f berührt und ihn optimal *linear approximiert*. Diese Gerade wollen wir die *Tangente* $t_{f,a}$ von f in a nennen, und der Grenzwert des Differenzenquotienten, d.h. die Steigung von $s_{f,a,x}$ konvergiert dann für $x \to a$ gegen die Steigung der Tangenten. D.h. die Tangente an den Graphen von f im Punkt $(a, f(a))$ hat die Geradengleichung

$$y = f'(a) \cdot x + \big(f(a) - a \cdot f'(a)\big) = f(a) + f'(a) \cdot (x - a),$$

die sich aus (29) ergibt, indem man den Grenzwert für $b = x$ gegen a betrachtet.



Beispiel 17.5.
Die folgenden Funktionen sind alle auf ganz $\mathbb{R}$ definiert, und dort ist jeder Punkt ein Häufungspunkt!

a. Eine konstante Funktion $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto c$ ist differenzierbar auf $\mathbb{R}$ und die Ableitung ist die Nullfunktion

$$f' : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto 0,$$

da für jedes $a \in \mathbb{R}$ der Differenzenquotient $\mathrm{Diff}_{f,a}$ die Nullfunktion ist und somit der Grenzwert $f'(a) = 0$ existiert.

b. Für $n \in \mathbb{N}$ ist die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^n$$

differenzierbar auf $\mathbb{R}$ mit Ableitung

$$f' : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto n \cdot x^{n-1},$$

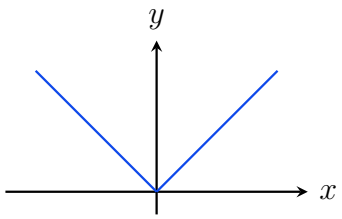
da sich für $a \in \mathbb{R}$ aus Beispiel 17.2 folgendes ergibt:

$$f'(a) = \lim_{x \to a} \left(x^{n-1} + a \cdot x^{n-2} + a^2 \cdot x^{n-3} + \ldots + a^{n-2} \cdot x + a^{n-1} \right) = n \cdot a^{n-1}.$$

c. Die Betragsfunktion

$$|\cdot| : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto |x|$$

ist in $a = 0$ *nicht* differenzierbar. In jedem anderen Punkt a ist sie jedoch differenzierbar mit $f'(a) = -1$ falls $a < 0$ und $f'(a) = 1$ falls $a > 0$.



Anschaulich bedeutet die Nicht-Differenzierbarkeit im Punkt $a = 0$, daß man am Graphen im Ursprung keine klare Tangente findet.
Um die Nicht-Differenzierbarkeit in $a = 0$ zu sehen, betrachten wir die Nullfolge $\left(\frac{(-1)^n}{n}\right)_{n \geq 1}$. Die zugehörige Folge der Werte des Differenzenquotienten

$$\left(\mathrm{Diff}_{f,0} \left(\frac{(-1)^n}{n} \right) \right)_{n \geq 1} = \left(\frac{\frac{1}{n}}{\frac{(-1)^n}{n}} \right)_{n \geq 1} = \left((-1)^n \right)_{n \geq 1}$$

ist nicht konvergent. Mithin existiert der Grenzwert des Differenzenquotienten in $a = 0$ nicht, und somit ist die Funktion in $a = 0$ nicht differenzierbar.
Außerdem, ist $a < 0$ und x nahe bei a, so ist auch $x < 0$ und mithin

$$\mathrm{Diff}_{f,a}(x) = \frac{|x| - |a|}{x - a} = \frac{-x + a}{x - a} = -1 \xrightarrow{x \to a} -1,$$

und analog ist für $a > 0$ und x nahe bei a auch $x > 0$, so daß

$$\mathrm{Diff}_{f,a}(x) = \frac{|x| - |a|}{x - a} = \frac{x - a}{x - a} = 1 \xrightarrow{x \to a} 1.$$

Damit ist auch gezeigt, daß die Ableitung in allen Punkten $a \neq 0$ existiert.

Bemerkung 17.6.

a. Wie bei der Stetigkeit wollen wir auch bei der Differenzierbarkeit anmerken, daß es sich um eine *lokale* Eigenschaft der Funktion handelt. D.h. sie ist punktweise definiert und hängt nur vom Verhalten der Funktion in einer sehr kleinen ε-Umgebung des betrachteten Punktes a ab!

b. Ist a ein Häufungspunkt von U und $f : U \longrightarrow \mathbb{R}$, so ist f genau dann in a differenzierbar, wenn der Grenzwert

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \in \mathbb{R}$$

existiert.
Um dies zu sehen ersetzt man im Differenzenquotienten einfach $x-a$ durch h.

c. Ist a ein Häufungspunkt von U und $f : U \longrightarrow \mathbb{R}$, so ist f genau dann in a differenzierbar, wenn es eine Zahl $c \in \mathbb{R}$ und eine Funktion $\rho : U \longrightarrow \mathbb{R}$ gibt, so daß

$$f(x) = f(a) + c\cdot(x-a) + \rho(x) \quad \text{und} \quad \lim_{x\to a} \frac{\rho(x)}{|x-a|} = 0$$

gilt.
Man beachte, ist f differenzierbar in a, so wählt man $c = f'(a)$ und $\rho(x) = \big(\operatorname{Diff}_{f,a}(x) - f'(a)\big)\cdot(x-a)$. Umgekehrt, wenn c und ρ existieren, so ist $\operatorname{Diff}_{f,a}(x) = c + \frac{\rho(x)}{x-a}$ und der Grenzwert des Differenzenquotienten existiert nach Voraussetzung. Wir erwähnen diese äquivalente Formulierung der Differenzierbarkeit an dieser Stelle, da sie für die Verallgemeinerung des Begriffes in der mehrdimensionalen Analysis von Vorteil ist. Die Bedeutung der Bedingung $\lim_{x\to a} \frac{\rho(x)}{|x-a|} = 0$ ist, daß die Funktion ρ, die den Unterschied des Differenzenquotienten und der Ableitung beschreibt, sehr schnell gegen Null konvergiert für $x \to a$, jedenfalls schneller als die lineare Funktion $x-a$.

C) Differenzierbarkeit und Stetigkeit

Satz 17.7 (Differenzierbar impliziert stetig.).
Ist $f : U \longrightarrow \mathbb{R}$ differenzierbar in a, so ist f stetig in a.

Beweis: Da nach Voraussetzung a ein Häufungspunkt von U ist, müssen wir nach Lemma 14.3 nur zeigen, daß $\lim_{x\to a} f(x) = f(a)$ oder alternativ $\lim_{x\to a}(f(x)-f(a)) = 0$ gilt. Nach Voraussetzung existiert der Grenzwert

$$f'(a) = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} \in \mathbb{R},$$

und da die Identität stetig ist, gilt zudem $\lim_{x\to a}(x-a)=0$. Mithin erhalten wir aus den Grenzwertsätzen für Funktionen 13.10

$$\begin{aligned}\lim_{x\to a}\big(f(x)-f(a)\big) &= \lim_{x\to a}\frac{f(x)-f(a)}{x-a}\cdot(x-a)\\ &= \lim_{x\to a}\frac{f(x)-f(a)}{x-a}\cdot\lim_{x\to a}(x-a) = f'(a)\cdot 0 = 0.\end{aligned}$$

Also ist f stetig in a. □

Beispiel 17.8.

a. Die Umkehrung von Satz 17.7 gilt nicht, wie das Beispiel der Betragsfunktion zeigt, die stetig in $a=0$ ist (siehe Beispiel 14.6), ohne dort differenzierbar zu sein (siehe Beispiel 17.5).

b. Die Funktion

$$f:\mathbb{R}\longrightarrow\mathbb{R}:x\mapsto\begin{cases}x^2, & \text{wenn } x\notin\mathbb{Q},\\ 0, & \text{wenn } x\in\mathbb{Q}\end{cases}$$

ist im Punkt $a=0$ differenzierbar, weil

$$\mathrm{Diff}_{f,0}(x)=\begin{cases}\frac{x^2-0}{x-0}=x, & \text{wenn } x\notin\mathbb{Q},\\ \frac{0-0}{x-0}=0, & \text{wenn } x\in\mathbb{Q}\end{cases}$$

und deshalb

$$\lim_{x\to 0}\mathrm{Diff}_{f,0}(x)=0$$

existiert. Die Funktion ist aber in keinem anderen Punkt $0\neq a\in\mathbb{R}$ differenzierbar, weil sie in den anderen Punkten nicht mal stetig ist; denn für $a\neq 0$ enthält jede noch so kleine Umgebung von a Punkte, an denen der Funktionswert 0 ist und Punkte, an denen der Funktionswert größer als a^2 ist. f ist somit ein Beispiel für einen Funktion, die in genau einem Punkt differenzierbar ist.

D) Ableitungsregeln – Linearität, Produkt- und Quotientenregel

Proposition 17.9 (Linearität der Ableitung).
Seien $f:U\longrightarrow\mathbb{R}$ und $g:U\longrightarrow\mathbb{R}$ in $a\in U$ differenzierbar und sei $c,d\in\mathbb{R}$. Dann ist $c\cdot f+d\cdot g$ differenzierbar in a mit $(c\cdot f+d\cdot g)'(a)=c\cdot f'(a)+d\cdot g'(a)$.

Beweis: Wir beachten zunächst, daß nach Voraussetzung a ein Häufungspunkt von U ist und daß U jeweils der Definitionsbereich der Funktionen ist. Dann folgt aus den

Grenzwertsätzen für Funktionen 13.10, daß der Grenzwert

$$\begin{aligned}(c\cdot f+d\cdot g)'(a) =& \lim_{x\to a}\frac{c\cdot f(x)+d\cdot g(x)-c\cdot f(a)-d\cdot g(a)}{x-a}\\ =& c\cdot\lim_{x\to a}\frac{f(x)-f(a)}{x-a}+d\cdot\lim_{x\to a}\frac{g(x)-g(a)}{x-a}\\ =& c\cdot f'(a)+d\cdot g'(a)\end{aligned}$$

existiert. □

Beispiel 17.10 (Polynomfunktionen sind differenzierbar.).
Ist $f:\mathbb{R}\longrightarrow\mathbb{R}:x\mapsto\sum_{k=0}^{n}a_k\cdot x^k$ eine Polynomfunktion, so ist f differenzierbar auf $\mathbb{R}$ mit

$$f':\mathbb{R}\longrightarrow\mathbb{R}:x\mapsto\sum_{k=1}^{n}k\cdot a_k\cdot x^{k-1}.$$

Dies folgt unmittelbar aus Proposition 17.9 und Beispiel 17.5.

Proposition 17.11 (Produktregel).
Seien $f:U\longrightarrow\mathbb{R}$ und $g:U\longrightarrow\mathbb{R}$ in $a\in U$ differenzierbar, so ist $f\cdot g$ differenzierbar in a mit

$$(f\cdot g)'(a)=f'(a)\cdot g(a)+f(a)\cdot g'(a).$$

Beweis: Wir beachten, daß nach Voraussetzung a ein Häufungspunkt von U ist und daß U der Definitionsbereich von $f\cdot g$ ist. Der Differenzenquotient von $f\cdot g$ an der Stelle a genügt der Gleichung

$$\begin{aligned}\mathrm{Diff}_{f\cdot g,a}(x) =& \frac{f(x)\cdot g(x)-f(a)\cdot g(a)}{x-a}\\ =& \frac{f(x)\cdot g(x)-f(a)\cdot g(x)+f(a)\cdot g(x)-f(a)\cdot g(a)}{x-a}\\ =& \frac{f(x)\cdot g(x)-f(a)\cdot g(x)}{x-a}+\frac{f(a)\cdot g(x)-f(a)\cdot g(a)}{x-a}\\ =& \frac{f(x)-f(a)}{x-a}\cdot g(x)+f(a)\cdot\frac{g(x)-g(a)}{x-a}.\end{aligned}$$

Da f und g in a differenzierbar sind und da g nach Satz 17.7 zudem stetig in a ist, existiert damit der Grenzwert

$$(f\cdot g)'(a)=\lim_{x\to a}\mathrm{Diff}_{f\cdot g,a}(x)=f'(a)\cdot g(a)+f(a)\cdot g'(a)$$

aufgrund der Grenzwertsätze für Funktionen 13.10. □

Proposition 17.12 (Quotientenregel).
Seien $f:U\longrightarrow\mathbb{R}$ und $g:U\longrightarrow\mathbb{R}$ in $a\in U$ differenzierbar mit $g(a)\neq 0$, so ist

auch $\frac{f}{g} : \{x \in U \mid g(x) \neq 0\} \longrightarrow \mathbb{R}$ differenzierbar in a mit

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a) \cdot g(a) - f(a) \cdot g'(a)}{g(a)^2}.$$

Beweis: Wir müssen zunächst einmal zeigen, daß a ein Häufungspunkt der Menge

$$V := \{x \in U \mid g(x) \neq 0\}$$

ist. Wegen Satz 17.7 ist g stetig in a, und da a ein Häufungspunkt von U ist gilt somit

$$\lim_{x \to a} g(x) = g(a) \neq 0.$$

Dann folgt aus Proposition 13.10 c. aber bereits, daß a auch ein Häufungspunkt von V ist.

Ferner gilt für den Differenzenquotienten

$$\mathrm{Diff}_{\frac{1}{g},a}(x) = \frac{\frac{1}{g(x)} - \frac{1}{g(a)}}{x - a} = -\frac{g(x) - g(a)}{x - a} \cdot \frac{1}{g(x) \cdot g(a)}.$$

Da g in a differenzierbar und stetig ist, existiert damit der Grenzwert

$$\left(\frac{1}{g}\right)'(a) = \lim_{x \to a} \mathrm{Diff}_{\frac{1}{g},a}(x) = -\lim_{x \to a} \frac{g(x) - g(a)}{x - a} \cdot \lim_{x \to a} \frac{1}{g(x) \cdot g(a)} = -\frac{g'(a)}{g(a)^2}.$$

Wenden wir nun die Produktformel 17.11 auf $f \cdot \frac{1}{g}$ an, so folgt das Ergebnis. □

Aus der Quotientenregel und Beispiel 17.10 folgt die folgende Aussage.

Beispiel 17.13.
Jede rationale Funktion $\frac{f}{g} : \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) = 0\} \longrightarrow \mathbb{R}$ ist differenzierbar.
Z.B. gilt für $h : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x^n}$ mit $n \geq 1$ für die Ableitung

$$h' : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} : x \mapsto -\frac{n \cdot x^{n-1}}{x^{2n}} = -\frac{n}{x^{n+1}}.$$

□

E) Ableitungsregeln – Ableitung der Umkehrfunktion und Kettenregel

Aufgrund des Umkehrsatzes für streng monotone Funktionen 14.21 sowie Bemerkung 14.22 wissen wir, daß eine stetige und streng monotone Funktion auf einem Intervall eine stetige Umkehrfunktion besitzt. Dabei kann das Intervall offen, halboffen oder abgeschlossen sein und es kann auch ein uneigentliches Intervall sein. Wir wenden uns nun der Frage zu, ob die Umkehrfunktion differenzierbar ist, wenn f differenzierbar ist.

Satz 17.14 (Ableitung der Umkehrfunktion).
Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \longrightarrow \mathbb{R}$ sei stetig und streng monoton (wachsend oder fallend). Ist f differenzierbar in a und ist $f'(a) \neq 0$, so ist die Umkehrfunktion

$$f^{-1} : f(I) \longrightarrow I$$

differenzierbar in $b := f(a)$ und es gilt

$$\left(f^{-1}\right)'(b) = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))}.$$

Beweis: Aus dem Umkehrsatz 14.21 sowie Bemerkung 14.22 wissen wir, daß die Umkehrfunktion

$$f^{-1} : f(I) \longrightarrow I$$

existiert und daß sie stetig und bijektiv ist.

Wir betrachten nun eine Folge $(y_n)_{n\in\mathbb{N}}$ in $f(I) \setminus \{b\}$ mit $y_n \longrightarrow b$. Da f^{-1} stetig ist, gilt dann

$$x_n := f^{-1}(y_n) \longrightarrow f^{-1}(b) = a.$$

Da f^{-1} bijektiv ist, ist $(x_n)_{n\in\mathbb{N}}$ somit eine Folge in $I \setminus \{a\}$, die gegen a konvergiert. Aufgrund der Grenzwertsätze für Folgen 11.15 erhalten wir dann

$$\frac{f^{-1}(y_n) - f^{-1}(b)}{y_n - b} = \frac{x_n - a}{f(x_n) - f(a)} = \frac{1}{\mathrm{Diff}_{f,a}(x_n)} \longrightarrow \frac{1}{f'(a)},$$

und wegen des Folgenkriteriums für Grenzwerte 13.7 existiert somit der Grenzwert

$$\left(f^{-1}\right)'(b) = \lim_{y\to b} \mathrm{Diff}_{f^{-1},b}(y) = \lim_{y\to b} \frac{f^{-1}(y) - f^{-1}(b)}{y - b} = \frac{1}{f'(a)}.$$

□

Beispiel 17.15.
Für $n \geq 2$ ist die Funktion

$$f : [0,\infty) \longrightarrow [0,\infty) : x \mapsto x^n$$

streng monoton wachsend und stetig nach Beispiel 14.23 mit der Wurzelfunktion als Umkehrfunktion und mit der Ableitung

$$f' : [0,\infty) \longrightarrow [0,\infty) : x \mapsto n \cdot x^{n-1}.$$

Da $f'(x) \neq 0$ für $x \neq 0$, folgt aus Satz 17.14, daß die Wurzelfunktion

$$\sqrt[n]{\cdot} : [0,\infty) \longrightarrow [0,\infty) : y \mapsto y^{\frac{1}{n}}$$

auf dem Intervall $(0,\infty)$ differenzierbar ist mit Ableitung

$$\left(\sqrt[n]{\cdot}\right)' : (0,\infty) \to \mathbb{R} : y \mapsto \frac{1}{n \cdot \left(\sqrt[n]{y}\right)^{n-1}} = \frac{1}{n} \cdot y^{\frac{1}{n}-1}.$$

Im Falle von $n = 2$ erhalten wir insbesondere

$$\left(\sqrt{y}\right)' = \frac{1}{2 \cdot \sqrt{y}}.$$

Wir wollen nun noch zeigen, daß die Wurzelfunktion $\sqrt[n]{\cdot}$ in $a = 0$ in der Tat *nicht* differenzierbar ist!

Dazu betrachten wir die Nullfolge $\left(\frac{1}{k^n}\right)_{k\in\mathbb{N}}$ und die zugehörigen Werte des Differenzenquotienten

$$\operatorname{Diff}_{\sqrt[n]{\cdot},0}\left(\frac{1}{k^n}\right) = \frac{\frac{1}{k} - 0}{\frac{1}{k^n} - 0} = k^{n-1} \longrightarrow \infty$$

für $k \longrightarrow \infty$, da $n \geq 2$. Also existiert der Grenzwert des Differenzenquotienten in $a = 0$ nicht, und somit ist die Wurzelfunktion dort auch nicht differenzierbar.

Proposition 17.16 (Kettenregel – äußere Ableitung × innere Ableitung). Es seien $f : U \longrightarrow \mathbb{R}$ und $g : V \longrightarrow \mathbb{R}$ Funktionen mit $\operatorname{Im}(f) \subseteq V$ und es sei $a \in U$. Ist f differenzierbar in a und g differenzierbar in $f(a)$, so ist $g \circ f$ differenzierbar in a mit

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a).$$

Beweis: Wir definieren auf V eine Funktion durch

$$h : V \longrightarrow \mathbb{R} : y \mapsto \begin{cases} \operatorname{Diff}_{g,f(a)}(y) = \frac{g(y)-g(f(a))}{y-f(a)}, & \text{falls } y \neq f(a), \\ g'(f(a)), & \text{falls } y = f(a). \end{cases}$$

Da g in $f(a)$ differenzierbar ist, gilt dann $\lim\limits_{y\to f(a)} h(y) = h(f(a))$, d.h. h ist stetig in $f(a)$. Außerdem gilt für alle $y \in V$

$$h(y) \cdot \big(y - f(a)\big) = g(y) - g\big(f(a)\big). \tag{30}$$

Wir beachten nun noch, daß nach Proposition 14.8 $\lim\limits_{x\to a} h(f(x)) = h(f(a)) = g'(f(a))$ gilt, da die Funktion h stetig in $f(a)$ und die Funktion f nach Satz 17.7 stetig in a ist. Damit erhalten wir dann, daß der Grenzwert

$$\begin{aligned}(g \circ f)'(a) &= \lim_{x\to a} \frac{g(f(x)) - g(f(a))}{x - a} \overset{(30)}{=} \lim_{x\to a} \frac{h(f(x)) \cdot (f(x) - f(a))}{x - a} \\ &= \lim_{x\to a} h(f(x)) \cdot \lim_{x\to a} \frac{f(x) - f(a)}{x - a} = g'(f(a)) \cdot f'(a)\end{aligned}$$

existiert. □

Beispiel 17.17.
Die Funktion

$$h : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sqrt{x^2 + 1}$$

läßt sich schreiben als $g \circ f$ mit

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2 + 1$$

und $g = \sqrt{\cdot}$. Also ist h differenzierbar auf $\mathbb{R}$ mit Ableitung

$$h'(x) = g'(f(x)) \cdot f'(x) = \frac{1}{2 \cdot \sqrt{x^2+1}} \cdot 2x = \frac{x}{\sqrt{x^2+1}}$$

in x.

F) Stetige Differenzierbarkeit

Abschließend wollen wir in diesem Abschnitt noch einige Begriffe einführen, die im folgenden nützlich sein werden.

Definition 17.18.
Es sei $U \subseteq \mathbb{R}$ und $f : U \longrightarrow \mathbb{R}$ eine Funktion.

a. Wir nennen f *stetig differenzierbar*, wenn f differenzierbar auf U und f' stetig auf U ist.

b. Wir definieren die *k-fache Differenzierbarkeit* und die *k-te Ableitung* von f rekursiv. f heißt *1-fach differenzierbar* auf U, wenn f auf U differenzierbar ist, und $f^{(1)} := f'$ heißt die *erste Ableitung* von f. Für $k > 1$ heißt f *k-fach differenzierbar*, wenn $f^{(k-1)}$ differenzierbar ist, und $f^{(k)} := \left(f^{(k-1)}\right)'$ heißt dann die *k-te Ableitung* von f. Wir schreiben auch $f^{(0)} := f$, $f'' := f^{(2)}$ und $f''' := f^{(3)}$.

c. f heißt *k-fach stetig differenzierbar*, wenn f k-fach differenzierbar auf U und zudem $f^{(k)}$ stetig auf U ist. Mit

$$\mathcal{C}^k(U, \mathbb{R}) := \{f : U \longrightarrow \mathbb{R} \mid f \text{ ist } k\text{-fach stetig differenzierbar}\}$$

bezeichnen wir die Menge der k-fach stetig differenzierbaren Funktionen auf U.

d. f heißt *unendlich oft differenzierbar* auf U, wenn $f \in \mathcal{C}^k(U, \mathbb{R})$ für alle $k \geq 1$. Wir bezeichnen mit

$$\mathcal{C}^\infty(U, \mathbb{R}) := \{f : U \longrightarrow \mathbb{R} \mid f \text{ ist unendlich oft differenzierbar}\}$$

die Menge der unendlich oft differenzierbaren Funktionen auf U.

Beispiel 17.19 (Nicht stetig differenzierbare Funktion).

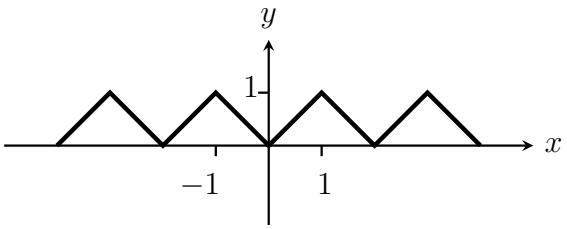
a. Die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} x^2 \cdot \sin\left(\frac{1}{x}\right), & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0, \end{cases}$$

ist differenzierbar auf $\mathbb{R}$, die Ableitung ist aber nicht stetig in $a = 0$. Der Beweis ist ein Übungsaufgabe, für die man unter anderem Korollar 18.22 benötigt.

b. Leitet man eine Polynomfunktion oder eine rationale Funktion ab, so erhält man wieder eine Polynomfunktion oder eine rationale Funktion mit dem jeweils gleichen Definitionsbreich. Da diese wieder differenzierbar sind, sehen wir, daß Polynomfunktionen und rationale Funktionen unendlich oft differenzierbar sind.

Bemerkung 17.20 (Überall stetig, nirgendwo differenzierbar).
Betrachte die periodische Funktion $g : \mathbb{R} \longrightarrow \mathbb{R}$, deren Graph in folgendem Bild dargestellt ist.



Die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} \frac{g(2^n \cdot x)}{2^n}$$

ist ein Beispiel für eine Funktion, die in jedem Punkt stetig (siehe Aufgabe 15.13) und in keinem Punkt differenzierbar ist! Funktionen, die wie f gebildet werden, nennt man auch *Weiterstraß-Funktionen.*

Aufgaben

Aufgabe 17.21.
Bestimme mittels der Ableitungsregeln die Ableitung der folgenden Funktionen:

a. $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto 10 + x \cdot \cos(x)$.

b. $g : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto e^{-3x} \cdot \left(x^3 - 2x\right)^2$.

Aufgabe 17.22.
Für $n \in \{0, 1, 2\}$ sei

$$f_n : [0, \infty) \longrightarrow \mathbb{R},\ x \longmapsto \begin{cases} x^n \cdot \sin\left(\frac{1}{x}\right), & \text{für } x > 0, \\ 0, & \text{für } x = 0. \end{cases}$$

Welche der Funktionen sind stetig in $a = 0$, differenzierbar in $a = 0$, stetig differenzierbar auf $[0, \infty)$?

Aufgabe 17.23 (Leibnitz-Regel).
Sei $n \geq 1$, $U \subseteq \mathbb{R}$ und $f, g \in \mathcal{C}^n(U, \mathbb{R})$ zwei n-fach differenzierbare Funktionen mit gleichem Definitionsbereich. Zeige:

$$(f \cdot g)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} \cdot f^{(k)} \cdot g^{(n-k)}.$$

Aufgabe 17.24.
Mit $\lfloor x \rfloor = \max\{z \in \mathbb{Z} \mid z \leq x\}$ sei die Abrundung der reellen Zahl $x \in \mathbb{R}$ bezeichnet. Skizziere den Graphen der Funktion

$$f : (-1, 1) \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{1}{\lfloor \frac{1}{x} \rfloor}, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0, \end{cases}$$

schematisch, überprüfe, an welchen Stellen die Funktion differenzierbar ist, und bestimme dort die Ableitung.

Aufgabe 17.25.
Bestimme für die folgenden Funktionen $f : I \longrightarrow \mathbb{R}$ ein maximales Intervall I, auf dem sie differenzierbar sind, und berechne ihre Ableitungen:

a. $f(x) = x^{\sqrt{x}}$.

b. $f(x) = \log_3(e^{2x+1})$.

c. $f(x) = x \cdot \ln(x) - x$.

d. $f(x) = (\cos(x))^{\sin(x)}$.

e. $f(x) = e^{3x} \cdot \big(\tan(x) - \sin(3x^2) + 3\big) \cdot x$.

f. $f(x) = e^{e^x}$.

Aufgabe 17.26.
Für eine auf U differenzierbare Funktion $f : U \longrightarrow \mathbb{R}$ ohne Nullstellen nennen wir $L(f) := \frac{f'}{f}$ die logarithmische Ableitung von f.
Zeige die folgenden Rechenregeln für die logarithmische Ableitung:

a. $L(f_1 \cdot \ldots \cdot f_n) = L(f_1) + \ldots + L(f_n)$.

b. $L\left(\frac{f}{g}\right) = L(f) - L(g)$.

c. $L\big(\exp(f(x))\big) = f'(x)$.

Aufgabe 17.27.
Es sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ eine Funktion, die in a differenzierbar ist und es seien $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ zwei Folgen mit $x_n < a < y_n$ und $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = a$. Dann gilt

$$\lim_{n\to\infty} \frac{f(y_n) - f(x_n)}{y_n - x_n} = f'(a).$$

Man zeige zudem, daß die Aussage i.a. nicht mehr richtig ist, wenn man auf die Voraussetzung $x_n < a < y_n$ verzichtet.

Aufgabe 17.28.
Zeige, daß die Funktion f in Bemerkung 17.20 in keinem Punkt ihres Definitionsbereiches differenzierbar ist.

Aufgabe 17.29.
Es sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ die Funktion aus Bemerkung 17.20 und

$$h : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x \cdot f(x).$$

Zeige, h ist stetig auf ganz $\mathbb{R}$, aber nur im Punkt $a = 0$ differenzierbar.

§ 18 Der Mittelwertsatz und seine Anwendungen

A) Notwendige Bedingung für Extremstellen

Definition 18.1 (Extremstellen).
Es sei $f : U \longrightarrow \mathbb{R}$ eine Funktion und $a \in U$.

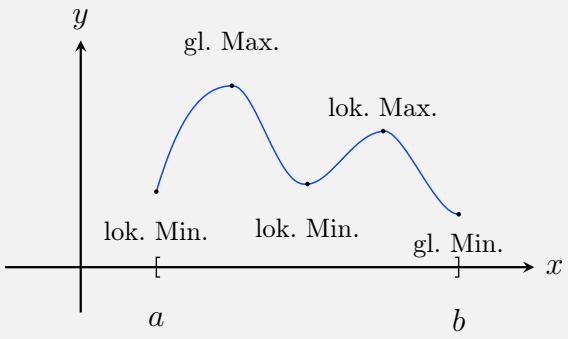
a. f hat in a ein *globales Maximum*, wenn $f(a) \geq f(x)$ für alle $x \in U$.

b. f hat in a ein *lokales Maximum*, wenn es ein $\delta > 0$ gibt, so daß $f(a) \geq f(x)$ für alle $x \in U \cap U_\delta(a)$.

c. f hat in a ein *globales Minimum*, wenn $f(a) \leq f(x)$ für alle $x \in U$.

d. f hat in a ein *lokales Minimum*, wenn es ein $\delta > 0$ gibt, so daß $f(a) \leq f(x)$ für alle $x \in U \cap U_\delta(a)$.

e. a heißt *Extremstelle* und $f(a)$ *Extremum* von f, wenn f in a ein lokales Minimum oder ein lokales Maximum hat.



Proposition 18.2 (Notwendige Bedingung für eine Extremstelle: $f'(c) = 0$).
Ist $f : (a, b) \longrightarrow \mathbb{R}$ in einer Extremstelle c differenzierbar, so ist $f'(c) = 0$.

Beweis: Wir können ohne Einschränkung annehmen, daß c ein lokales Maximum ist, da der Beweis für ein lokales Minimum dann durch Übergang von f zu $-f$ folgt.

Nach Definition gibt es ein $\delta > 0$, so daß $f(c) \geq f(x)$ für alle $x \in (a, b) \cap (c - \delta, c + \delta)$. Ersetzen wir δ durch $\min\{\delta, b-c, c-a\}$ so können wir annehmen, daß $(c-\delta, c+\delta) \subseteq (a, b)$. Wir betrachten nun die Folgen $(a_n)_{n\geq 2}$ und $(b_n)_{n\geq 2}$ mit

$$a_n := c - \frac{\delta}{n} < c$$

und

$$b_n := c + \frac{\delta}{n} > c.$$

Dann konvergiert die Folge $(a_n)_{n\geq 2}$ von links gegen c, und die Folge $(b_n)_{n\geq 2}$ konvergiert von rechts gegen c. Nun betrachten wir den Grenzwert des Differenzenquotienten von f in c für die beiden Folgen und berücksichtigen, daß stets $f(a_n) - f(c) \leq 0$ und $f(b_n) - f(c) \leq 0$ gilt und daß außerdem $a_n - c < 0$ und $b_n - c > 0$ gilt:

$$0 \leq \frac{f(a_n) - f(c)}{a_n - c} \longrightarrow f'(c)$$

und

$$0 \geq \frac{f(b_n) - f(c)}{b_n - c} \longrightarrow f'(c).$$

Für den Grenzwert $f'(c)$ gilt also

$$0 \leq f'(c) \leq 0,$$

und mithin

$$f'(c) = 0.$$

□

Beispiel 18.3. a. Ist $n \geq 2$ gerade, so nimmt die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^n$$

in $a = 0$ ein globales Minimum an, und es gilt auch

$$f'(0) = n \cdot 0^{n-1} = 0.$$

b. Die Funktion $f : (-1, 1) \longrightarrow \mathbb{R} : x \mapsto x^3$ hat in Null *keine* Extremstelle, da $f(x) < 0$ für $x < 0$ und $f(x) > 0$ für $x > 0$, dennoch gilt $f'(0) = 3 \cdot 0^2 = 0$. Die Bedingung $f'(c) = 0$ für eine Extremstelle c ist also *notwendig*, aber sie ist *nicht hinreichend*.

Bemerkung 18.4.
Selbst wenn f auf dem abgeschlossenen Intervall $[a, b]$ definiert und dort überall differenzierbar ist, macht Proposition 18.2 *keine* Aussagen über die Ableitung in den *Randpunkten* a und b, falls diese Extremstellen sind!

Die Funktion $f : [-1, 1] \longrightarrow \mathbb{R} : x \mapsto x^3$ nimmt in $a = -1$ ihr globales Minimum und in $a = 1$ ihr globales Maximum an, aber die Ableitungen $f'(-1) = 3 = f'(1)$ sind beide nicht Null.

B) Der Satz von Rolle und der Mittelwertsatz

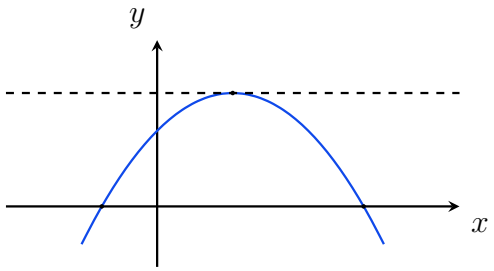
Satz 18.5 (Satz von Rolle).
Ist $a < b$ und ist $f : [a, b] \longrightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b) mit $f(a) = f(b)$, so gibt es ein $c \in (a, b)$ mit $f'(c) = 0$.

Beweis: Ist f konstant auf dem Intervall $[a, b]$ so ist $f'(c) = 0$ für jedes $c \in (a, b)$. Wir können also annehmen, daß es ein $y \in (a, b)$ mit $f(y) \neq f(a) = f(b)$ gibt.

Wir betrachten zunächst den Fall, daß $f(y) > f(a) = f(b)$ gilt. Da f stetig auf dem abgeschlossenen Intervall $[a, b]$ ist, nimmt f dort nach Satz 14.16 sein Maximum an, d.h. es gibt ein $c \in [a, b]$ mit $f(c) \geq f(x)$ für alle $x \in [a, b]$. c ist also eine Extremstelle, und wegen $f(y) > f(a) = f(b)$, muß $c \in (a, b)$ gelten, so daß wir Proposition 18.2 anwenden können und $f'(c) = 0$ erhalten.

Der Fall $f(y) < f(a) = f(b)$ geht analog, da dann ein globales Minimum von f in (a, b) existiert. □

Bemerkung 18.6.
Der Satz von Rolle besagt insbesondere, daß die Ableitung zwischen zwei Nullstellen einer differenzierbaren Funktion mindestens einmal Null werden muß, und im Beweis haben wir gesehen, daß das daran liegt, daß die Funktion dort eine Extremstelle besitzt.



Satz 18.7 (Mittelwertsatz).
Ist $a < b$ und ist $f : [a, b] \longrightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b), so gibt es ein $c \in (a, b)$ mit $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Beweis: Die Funktion

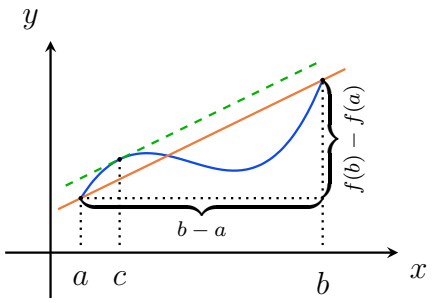
$$g : [a, b] \longrightarrow \mathbb{R} : x \mapsto f(x) - \frac{f(b) - f(a)}{b - a} \cdot (x - a)$$

ist stetig auf $[a, b]$ und differenzierbar auf (a, b). Außerdem gilt $g(a) = f(a) = g(b)$. Aus dem Satz von Rolle 18.5 folgt die Existenz eines $c \in (a, b)$ mit

$$0 = g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}.$$

□

Bemerkung 18.8.
Der Mittelwertsatz besagt, daß zwischen a und b ein c liegt, in dem die Steigung der Tangente $t_{f,c}$ an den Graphen von f mit der Steigung der Sekante $s_{f,a,b}$ durch a und b übereinstimmt.



Beispiel 18.9.
Betrachten wir die Funktion

$$f : [-1, 1] \longrightarrow \mathbb{R} : x \mapsto x^3.$$

Aus dem Mittelwertsatz folgt, daß es ein $c \in (-1, 1)$ geben muß, so daß die Tangente an den Graphen von f im Punkt (c, c^3) die Steigung

$$\frac{f(1) - f(-1)}{1 - (-1)} = \frac{2}{2} = 1$$

hat. Da wir die Ableitungsfunktion kennen, können wir versuchen, c zu bestimmen. Es muß gelten

$$1 = f'(c) = 3 \cdot c^2.$$

Wir finden also zwei solcher Stellen:

$$c = \frac{1}{\sqrt{3}} \quad \text{und} \quad c = -\frac{1}{\sqrt{3}}.$$

Korollar 18.10 (Allgemeiner Mittelwertsatz).
Ist $a < b$ und sind $f, g : [a, b] \longrightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b), so gibt es ein $c \in (a, b)$ mit $f'(c) \cdot \big(g(b) - g(a)\big) = g'(c) \cdot \big(f(b) - f(a)\big)$.

Beweis: Wir betrachten die Funktion

$$h : [a, b] \longrightarrow \mathbb{R} : x \mapsto f(x) \cdot \big(g(b) - g(a)\big) - g(x) \cdot \big(f(b) - f(a)\big).$$

h ist auf $[a, b]$ stetig und auf (a, b) differenzierbar mit

$$h(a) = f(a) \cdot g(b) - g(a) \cdot f(b) = h(b).$$

Aus dem Satz von Rolle folgt, daß es ein $c \in (a,b)$ gibt mit

$$0 = h'(c) = f'(c) \cdot \big(g(b) - g(a)\big) - g'(c) \cdot \big(f(b) - f(a)\big).$$

□

Bemerkung 18.11 (Geometrische Interpretation).
Wollen wir den allgemeinen Mittelwertsatz geometrisch interpretieren, brauchen wir einen Vorgriff auf die mehrdimensionale Analysis. Sind f und g zwei Funktionen wie in Korollar 18.10, dann definiert

$$\varphi : [a,b] \longrightarrow \mathbb{R}^2 : t \mapsto \big(f(t), g(t)\big)^t$$

eine differenzierbare Abbildung vom Intervall $[a,b]$ in die Ebene $\mathbb{R}^2$. Das Bild der Abbildung φ ist eine ebene Kurve mit den Endpunkten $(f(a), g(a))^t$ und $(f(b), g(b))^t$ (siehe Abbildung 10).

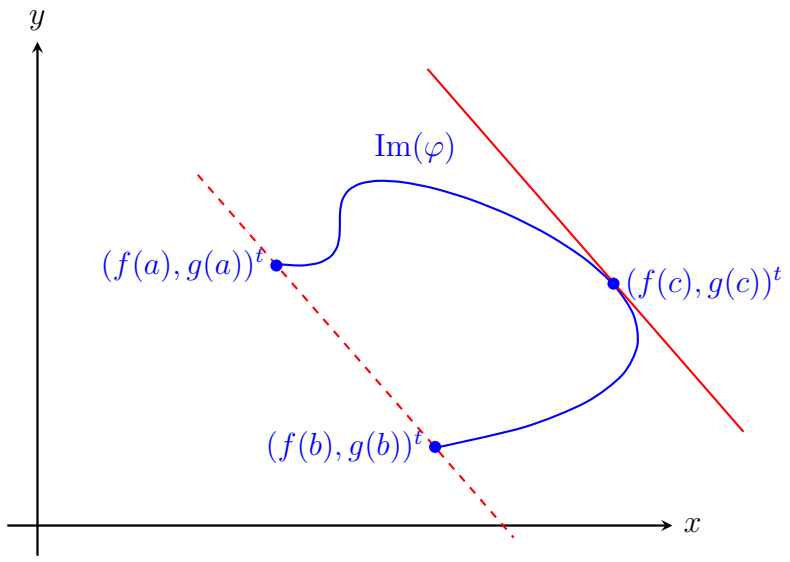


Abbildung 10: Geometrische Interpretation des Allgemeinen Mittelwertsatzes

Die Richtung der Tangente an die Kurve im Punkt $(f(c), g(c))^t$ kann man mit Hilfe des Gradienten bestimmen, denn wenn der Vektor

$$v(c) = \nabla\varphi(c) = \big(f'(c), g'(c)\big)^t$$

nicht der Nullvektor ist, ist er ein Richtungsvektor der Tagente. Schauen wir uns nun die Gleichung

$$\frac{g'(c)}{f'(c)} = \frac{g(b) - g(a)}{f(b) - f(a)}$$

an, die wir aus dem Allgemeinen Mittelwertsatz erhalten, so ist die linke Seite die Steigung der Tangente an die Kurve im Punkt $(f(c), g(c))^t$ und die rechte Seite ist die Steigung der Geraden durch die beiden Endpunkte der Geraden. Der Allgemeine

Mittelwertsatz sagt mithin, dass es einen Punkt auf der Kurve gibt, in dem die Tangete parallel zur Geraden durch die beiden Endpunkte ist.

C) Anwendungen des Mittelwertsatzes

Wir wollen uns nun den Anwendungen des Mittelwertsatzes zuwenden.

C.1) Konstante Funktionen

Proposition 18.12.
Es sei $f : [a, b] \longrightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b) mit $f'(x) = 0$ für alle $x \in (a, b)$, dann ist f eine konstante Funktion.

Beweis: Sei $x \in (a, b]$, so ist f stetig auf $[a, x]$ und differenzierbar auf (a, x). Aus dem Mittelwertsatz folgt dann, daß es ein $c \in (a, x)$ gibt mit

$$\frac{f(x) - f(a)}{x - a} = f'(c) = 0.$$

Also gilt $f(x) = f(a)$, und dies gilt für alle $x \in (a, b]$. □

C.2) Monotonie und Ableitung

Mit Hilfe der Ableitung läßt sich bei differenzierbaren Funktionen ein hinreichendes Kriterium für Monotonie angeben.

Proposition 18.13 (Hinreichendes Kriterium für Monotonie).
Es sei $a < b$ und $f : [a, b] \longrightarrow \mathbb{R}$ sei stetig auf $[a, b]$ und differenzierbar auf (a, b).

a. Ist $f'(x) > 0$ für alle $x \in (a, b)$, so ist f streng monoton wachsend auf $[a, b]$.

b. Ist $f'(x) < 0$ für alle $x \in (a, b)$, so ist f streng monoton fallend auf $[a, b]$.

Beweis: a. Es seien $x, y \in [a, b]$ mit $x < y$ gegeben. Dann ist f stetig auf $[x, y]$ und differenzierbar auf (x, y). Aus dem Mittelwertsatz folgt deshalb, daß es ein $c \in (x, y)$ gibt mit

$$f(y) - f(x) = f'(c) \cdot (y - x) > 0.$$

Also ist f streng monoton wachsend.

b. Der Beweis geht analog zum ersten Teil.

□

Beispiel 18.14.
Betrachte für $n \geq 1$ die Funktion

$$f : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto x^n.$$

Für die Ableitung gilt

$$f'(x) = n \cdot x^{n-1} > 0$$

für alle $x \in (0, \infty)$. Mithin ist die Funktion streng monoton wachsend auf jedem Intervall $[0, b] \subseteq [0, \infty)$ und mithin auf $[0, \infty)$. Dies ist ein alternativer Beweis der Aussage in Beispiel 14.19.

C.3) Hinreichende Bedingung für Extremstellen

Proposition 18.15 (Hinreichende Bedingung für eine Extremstelle).
Es sei $f : (a, b) \longrightarrow \mathbb{R}$ eine zweifach differenzierbare Funktion und $c \in (a, b)$.

a. Falls $f'(c) = 0$ und $f''(c) < 0$, so ist c ein lokales Maximum.

b. Falls $f'(c) = 0$ und $f''(c) > 0$, so ist c ein lokales Minimum.

Beweis:

b. Nach Voraussetzung ist

$$\lim_{x \to c} \frac{f'(x)}{x - c} = \lim_{x \to c} \frac{f'(x) - f'(c)}{x - c} = f''(c) > 0.$$

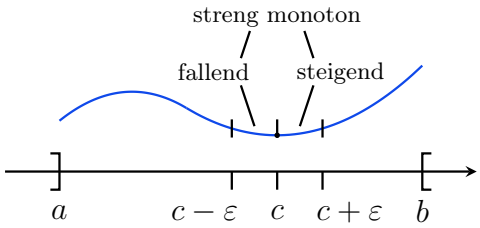
Zu $\varepsilon := \frac{f''(c)}{2} > 0$ gibt es dann ein $\delta_\varepsilon > 0$, so daß

$$-\frac{f''(c)}{2} = -\varepsilon < \frac{f'(x)}{x - c} - f''(c) < \varepsilon = \frac{f''(c)}{2}$$

für alle $x \in (a, b)$ mit $|x - c| < \delta_\varepsilon$. Insbesondere folgt für diese x dann

$$\frac{f'(x)}{x - c} > -\frac{f''(c)}{2} + f''(c) = \frac{f''(c)}{2} > 0. \tag{31}$$

Wir können ohne Einschränkung annehmen, daß δ_ε so klein ist, daß $a < c - \delta_\varepsilon < c + \delta_\varepsilon < b$ gilt.
Für $x \in (c - \delta_\varepsilon, c)$ folgt aus (31) dann $f'(x) < 0$, und nach Proposition 18.13 ist f dann streng monoton fallend auf dem Intervall $[c - \delta_\varepsilon, c]$.
Analog folgt für $x \in (c, c + \delta_\varepsilon)$ aus (31) $f'(x) > 0$ und aus Proposition 18.13 folgt, daß f streng monoton wachsend auf dem Intervall $[c, c + \delta_\varepsilon]$ ist.
Insbesondere heißt das, daß $f(c) \leq f(x)$ für alle $x \in [c - \delta_\varepsilon, c + \delta_\varepsilon]$, so daß f in c ein Minimum besitzt.



a. Die Aussage beweist man analog.

□

Beispiel 18.16.
Wir betrachten die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^3 - 3x^2 + 2.$$

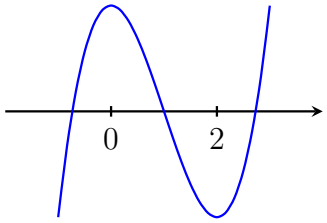
Um mögliche Extremstellen zu finden, müssen wir die Nullstellen der ersten Ableitung

$$f'(x) = 3x^2 - 6x$$

finden. Das ist für $x = 0$ und $x = 2$ der Fall. In diesen Punkten schauen wir uns die zweite Ableitung

$$f''(x) = 6x - 6$$

an. Aus $f''(0) = -6 < 0$ folgt, daß in $x = 0$ ein Maximum vorliegt, und aus $f''(2) = 6 > 0$ folgt, daß in $x = 2$ ein Minimum vorliegt.



Bemerkung 18.17 (Hinreichende Bedingung für Extremstellen).
Anstatt zweifacher Differenzierbarkeit und der Bedingung an die zweite Ableitung kann man auch einfach fordern, daß die erste Ableitung in c einen Vorzeichenwechsel hat, wie wir ihn im Beweis von Proposition 18.15 aus den Bedingungen an $f''(c)$ ableiten.

C.4) Vertauschbarkeit von Grenzwert und Ableitung

Satz 18.18 (Vertauschbarkeit von Grenzwert und Ableitung).
Ist $(f_n)_{n\in\mathbb{N}}$ eine Funktionenfolge stetig differenzierbarer Funktionen auf dem Intervall $[a, b]$, so daß $(f_n)_{n\in\mathbb{N}}$ punktweise gegen f und $(f_n')_{n\in\mathbb{N}}$ gleichmäßig gegen g konvergiert, dann ist f stetig differenzierbar auf $[a, b]$ mit Ableitung $f' = g$.

Beweis: Nach Voraussetzung sind die f_n' stetig auf $[a,b]$, so daß die Grenzfunktion g als gleichmäßiger Grenzwert stetiger Funktionen nach Satz 15.6 ebenfalls stetig ist.

Sei nun $\varepsilon > 0$ und $c \in [a,b]$ gegeben, so müssen wir ein $\delta_\varepsilon > 0$ finden, so daß für alle $c \neq x \in [a,b]$ mit $|x-c| < \delta_\varepsilon$ auch

$$|\operatorname{Diff}_{f,c}(x) - g(c)| = \left|\frac{f(x)-f(c)}{x-c} - g(c)\right| < \varepsilon$$

gilt, d.h. $g(c)$ ist der Grenzwert des Differenzenquotienten von f in c.

Da $(f_n')_{n\in\mathbb{N}}$ gleichmäßig gegen g konvergiert, gibt es ein $n_\varepsilon \in \mathbb{N}$, so daß

$$|f_n'(x) - g(x)| < \frac{\varepsilon}{3} \tag{32}$$

für alle $n \geq n_\varepsilon$ und alle $x \in [a,b]$ gilt.

Da g stetig in c ist, gibt es zudem ein $\delta_\varepsilon > 0$, so daß für alle $x \in [a,b]$ mit $|x-c| < \delta_\varepsilon$ auch

$$|g(x) - g(c)| < \frac{\varepsilon}{3} \tag{33}$$

gilt.

Sei nun $c \neq x \in [a,b]$ mit $|x-c| < \delta_\varepsilon$ gegeben. Für $n \geq n_\varepsilon$ können wir den Mittelwertsatz 18.7 auf die differenzierbare Funktion f_n anwenden und finden somit ein y zwischen x und c mit

$$\frac{f_n(x)-f_n(c)}{x-c} = f_n'(y) \tag{34}$$

und da y zwischen x und c liegt, gilt auch $|y-c| \leq |x-c| < \delta_\varepsilon$.

Setzen wir die obigen Ergebnisse nun zusammen, so erhalten wir

$$\begin{aligned}\left|\frac{f_n(x)-f_n(c)}{x-c} - g(c)\right| &\overset{(34)}{=} |f_n'(y) - g(c)| \\ &\leq |f_n'(y) - g(y)| + |g(y) - g(c)| \\ &\overset{(32)(33)}{<} \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\cdot\varepsilon}{3}\end{aligned}$$

für alle $n \geq n_\varepsilon$. Lassen wir nun n gegen unendlich laufen, so erhalten wir für den Grenzwert

$$\left|\frac{f(x)-f(c)}{x-c} - g(c)\right| = \lim_{n\to\infty}\left|\frac{f_n(x)-f_n(c)}{x-c} - g(c)\right| \leq \frac{2\cdot\varepsilon}{3} < \varepsilon.$$

Damit haben wir gezeigt, daß f in einem beliebigen Punkt c des Intervalls $[a,b]$ differenzierbar ist und daß $f' = g$ gilt. Da wir bereits wissen, daß g stetig ist, ist f mithin stetig differenzierbar auf $[a,b]$. □

Bemerkung 18.19.

a. Die Aussage in Satz 18.18 besagt, daß

$$\lim_{n\to\infty}(f_n') = \big(\lim_{n\to\infty} f_n\big)',$$

d.h. die Grenzwertbildung für die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ vertauscht mit der Ableitung!
Auf die Differenzenquotienten zurückgeführt, bedeutet dies

$$\lim_{n\to\infty}\lim_{x\to a}\frac{f_n(x)-f_n(a)}{x-a} = \lim_{x\to a}\lim_{n\to\infty}\frac{f_n(x)-f_n(a)}{x-a}.$$

Hier vertauschen zwei Grenzwertprozesse! Das ist eine Besonderheit!

b. Man kann in Satz 18.18 auf die Bedingung, daß die Ableitungen f_n' stetig sind, verzichten. Der Beweis wird dann aber etwas technischer.

c. Auch wenn wir in Satz 18.18 nur die punktweise Konvergenz für die Folge $(f_n)_{n\in\mathbb{N}}$ gefordert haben, erzwingt die gleichmäßige Konvergenz der Folge $(f_n')_{n\in\mathbb{N}}$ letztlich die gleichmäßige Konvergenz der Folge $(f_n)_{n\in\mathbb{N}}$.

C.5) Ableitung von Potenzreihen

Korollar 18.20 (Ableitung von Potenzreihen).
Ist $\sum_{n=0}^{\infty} a_n\cdot t^n$ eine Potenzreihe über $\mathbb{R}$ mit Konvergenzradius $r>0$, dann ist die Funktion

$$f:(-r,r)\longrightarrow\mathbb{R}:x\mapsto\sum_{n=0}^{\infty}a_n\cdot x^n$$

differenzierbar auf $(-r,r)$ und die Ableitung in $x\in(-r,r)$ ist gegeben durch

$$f'(x)=\sum_{n=1}^{\infty}n\cdot a_n\cdot x^{n-1},$$

d.h. durch die formale Ableitung $\sum_{n=1}^{\infty} n\cdot a_n\cdot t^{n-1}$ der Potenzreihe.

Beweis: Aus Aufgabe 12.53 wissen wir, daß die beiden Reihen $\sum_{n=0}^{\infty} a_n\cdot t^n$ und ihre formale Ableitung $\sum_{n=1}^{\infty} n\cdot a_n\cdot t^{n-1}$ den gleichen Konvergenzradius r besitzen. Insbesondere definiert letztere Reihe eine Funktion

$$g:(-r,r)\longrightarrow\mathbb{R}:x\mapsto\sum_{n=1}^{\infty}n\cdot a_n\cdot x^{n-1},$$

die nach Korollar 15.7 stetig ist.

Sei nun $a\in(-r,r)$ gegeben. Wir wollen zeigen, daß f in a differenzierbar ist mit

$$f'(a)=g(a).$$

Dazu setzen wir $R := \frac{r+|a|}{2} < r$, so daß $a \in (-R, R)$ liegt. Die Folge $(f_n)_{n\in\mathbb{N}}$ mit

$$f_n : [-R, R] \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{n} a_k \cdot x^k$$

konvergiert nach Satz 15.4 auf dem abgeschlossenen Intervall $[-R, R]$ gleichmäßig gegen die Funktion f. Nach Beispiel 17.10 sind die f_n differenzierbar mit stetiger Ableitung

$$f_n' : [-R, R] \longrightarrow \mathbb{R} : x \mapsto \sum_{k=1}^{n} k \cdot a_k \cdot x^{k-1}.$$

Die Folge $(f_n')_{n\in\mathbb{N}}$ der Ableitungen konvergiert dann wieder nach Satz 15.4 auf $[-R, R]$ gleichmäßig gegen g. Da die Voraussetzungen von Satz 18.18 erfüllt sind, ist f auf $[-R, R]$ differenzierbar mit $f' = g$. Insbesondere ist f also in a differenzierbar mit $f'(a) = g(a)$. □

Da wir die Aussage des Korollars auch auf die formale Ableitung der Potenzreihe anwenden können, erhalten wir durch Induktion die folgende Aussage.

Korollar 18.21 (Differenzierbarkeit von Potenzreihen).
Eine durch eine Potenzreihe definierte Funktion ist auf ihrem Konvergenzbereich unendlich oft differenzierbar.

Die Exponentialfunktion, der Sinus und der Cosinus sind also differenzierbar.

C.6) Ableitungen einiger spezieller Funktionen

Korollar 18.22 (Ableitungen wichtiger Funktionen).

a. Die Exponentialfunktion

$$\exp : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

ist unendlich oft differenzierbar auf $\mathbb{R}$ mit Ableitung

$$\exp'(x) = \exp(x).$$

b. Der Sinus

$$\sin : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!}$$

ist unendlich oft differenzierbar auf $\mathbb{R}$ mit Ableitung

$$\sin'(x) = \cos(x).$$

c. Der Cosinus

$$\cos : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n}}{(2n)!}$$

ist unendlich oft differenzierbar auf $\mathbb{R}$ mit Ableitung

$$\cos'(x) = -\sin(x).$$

d. Für $a \in \mathbb{R}_{>0}$ ist die Exponentialfunktion zur Basis a

$$\exp_a : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \exp\big(x \cdot \ln(a)\big)$$

stetig differenzierbar auf $\mathbb{R}$ mit Ableitung

$$\exp_a'(x) = \ln(a) \cdot \exp_a(x).$$

e. Für $a \in \mathbb{R}_{>0}$ mit $a \neq 1$ ist die Logarithmusfunktion zur Basis a

$$\log_a : (0, \infty) \longrightarrow \mathbb{R}$$

stetig differenzierbar auf $(0, \infty)$ mit Ableitung

$$\log_a'(x) = \frac{1}{x \cdot \ln(a)}.$$

Insbesondere gilt für die Ableitung des natürlichen Logarithmus

$$\ln'(x) = \frac{1}{x}.$$

f. Der Tangens

$$\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \longrightarrow \mathbb{R} : x \mapsto \frac{\sin(x)}{\cos(x)}$$

ist auf $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ stetig differenzierbar mit Ableitung

$$\tan'(x) = \frac{1}{\cos^2(x)}.$$

g. Der Cotangens

$$\cot : (0, \pi) \longrightarrow \mathbb{R} : x \mapsto \frac{\cos(x)}{\sin(x)}$$

ist stetig differenzierbar auf $(0, \pi)$ mit Ableitung

$$\cot'(x) = -\frac{1}{\sin^2(x)}.$$

h. Der Arcustangens ist auf $\mathbb{R}$ stetig differenzierbar mit

$$\arctan'(x) = \frac{1}{1 + x^2}.$$

i. Der Arcuscotangens ist auf $\mathbb{R}$ stetig differenzierbar mit

$$\operatorname{arccot}'(x) = -\frac{1}{1 + x^2}.$$

j. Der Arcussinus ist auf $(-1,1)$ differenzierbar mit

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}.$$

k. Der Arcuscosinus ist auf $(-1,1)$ differenzierbar mit

$$\arccos'(x) = -\frac{1}{\sqrt{1-x^2}}.$$

Beweis:

a. exp ist nach Korollar 18.20 und 18.21 unendlich oft differenzierbar auf $\mathbb{R}$ mit Ableitung

$$\exp'(x) = \sum_{n=1}^{\infty} n \cdot \frac{x^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \exp(x).$$

b. sin ist nach Korollar 18.20 und 18.21 unendlich oft differenzierbar auf $\mathbb{R}$ mit Ableitung

$$\sin'(x) = \sum_{n=0}^{\infty} (-1)^n (2n+1) \cdot \frac{x^{2n}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = \cos(x).$$

c. cos ist nach Korollar 18.20 und 18.21 unendlich oft differenzierbar auf $\mathbb{R}$ mit Ableitung

$$\begin{aligned}\cos'(x) &= \sum_{n=1}^{\infty} (-1)^n (2n) \cdot \frac{x^{2n-1}}{(2n)!} = -\sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^{2n-1}}{(2n-1)!} \\ &= -\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = -\sin(x).\end{aligned}$$

d. Aus der Kettenregel erhalten wir, daß die Exponentialfunktion zur Basis a differenzierbar ist mit Ableitung

$$\exp_a'(x) = \ln(a) \cdot \exp_a(x),$$

und diese Funktion ist offenbar wieder stetig.

e. Aus dem Satz zur Ableitung von Umkehrfunktionen 17.14 folgt, daß $\log_a$ auf $(0,\infty)$ differenzierbar ist, da die Ableitung der Exponentialfunktion $\exp_a$ nie Null wird. Für die Ableitung erhalten wir zudem

$$\log_a'(x) = \frac{1}{\exp_a'(\log_a(x))} = \frac{1}{\ln(a) \cdot \exp_a(\log_a(x))} = \frac{1}{\ln(a) \cdot x}.$$

Zudem ist die Ableitung offenbar stetig.

f. Aus der Quotientenregel erhalten wir, daß der Tangens in $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ differenzierbar ist mit

$$\tan'(x) = \frac{\sin'(x)\cdot\cos(x) - \sin(x)\cdot\cos'(x)}{\cos(x)^2} = \frac{\sin(x)^2 + \cos(x)^2}{\cos(x)^2} = \frac{1}{\cos(x)^2}.$$

Als Quotient stetiger Funktionen ist die Ableitung insbesondere stetig.

g. Aus der Quotientenregel erhalten wir, daß der Cotangens in $x \in (0,\pi)$ differenzierbar ist mit

$$\cot'(x) = \frac{\cos'(x)\cdot\sin(x) - \cos(x)\cdot\sin'(x)}{\sin(x)^2} = \frac{-cos(x)^2 - \sin(x)^2}{\sin(x)^2} = -\frac{1}{\sin(x)^2}.$$

Als Quotient stetiger Funktionen ist die Ableitung insbesondere stetig.

h. Aus dem Satz zur Ableitung von Umkehrfunktionen 17.14 zusammen mit Satz 16.18 folgt, daß arctan auf $\mathbb{R}$ differenzierbar ist, da die Ableitung des Tangens nie Null wird auf $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Für die Ableitung erhalten wir zudem

$$\begin{aligned}\arctan'(x) =& \frac{1}{\tan'(\arctan(x))} = \frac{1}{\frac{1}{\cos^2(\arctan(x))}}\\ =& \frac{1}{\frac{\sin^2(\arctan(x))+\cos^2(\arctan(x))}{\cos^2(\arctan(x))}} = \frac{1}{\frac{\sin^2(\arctan(x))}{\cos^2(\arctan(x))}+1}\\ =& \frac{1}{\tan(\arctan(x))^2+1} = \frac{1}{x^2+1}.\end{aligned}$$

Die Ableitung ist zudem offenbar stetig.

i. Aus dem Satz zur Ableitung von Umkehrfunktionen 17.14 zusammen mit Satz 16.18 folgt, daß arccot auf $\mathbb{R}$ differenzierbar ist, da die Ableitung des Cotangens nie Null wird auf $(0,\pi)$. Für die Ableitung erhalten wir zudem

$$\begin{aligned}\operatorname{arccot}'(x) =& \frac{1}{\cot'(\operatorname{arccot}(x))} = \frac{1}{-\frac{1}{\sin^2(\operatorname{arccot}(x))}}\\ =& -\frac{1}{\frac{\sin^2(\operatorname{arccot}(x))+\cos^2(\operatorname{arccot}(x))}{\sin^2(\operatorname{arccot}(x))}} = -\frac{1}{1+\frac{\cos^2(\operatorname{arccot}(x))}{\sin^2(\operatorname{arccot}(x))}}\\ =& -\frac{1}{1+\cot(\operatorname{arccot}(x))^2} = -\frac{1}{1+x^2}.\end{aligned}$$

Die Ableitung ist zudem offenbar stetig.

j. Aus dem Satz zur Ableitung von Umkehrfunktionen 17.14 zusammen mit Satz 16.18 folgt, daß arcsin auf $(-1,1)$ differenzierbar ist, da die Ableitung des Sinus nie Null wird auf $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Für die Ableitung erhalten wir zudem unter Berücksichtigung

der Tatsache, daß der Cosinus auf $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ positiv ist:

$$\begin{aligned}\arcsin'(x) =& \frac{1}{\sin'(\arcsin(x))} = \frac{1}{\cos(\arcsin(x))} \\ =& \frac{1}{\sqrt{\cos^2(\arcsin(x))}} = \frac{1}{\sqrt{1-\sin^2(\arcsin(x))}} \\ =& \frac{1}{\sqrt{1-x^2}}.\end{aligned}$$

Die Ableitung ist zudem offenbar stetig.

k. Aus dem Satz zur Ableitung von Umkehrfunktionen 17.14 zusammen mit Satz 16.18 folgt, daß arccos auf $(-1, 1)$ differenzierbar ist, da die Ableitung des Cosinus nie Null wird auf $(0, \pi)$. Für die Ableitung erhalten wir zudem unter Berücksichtigung der Tatsache, daß der Sinus auf $(0, \pi)$ positiv ist:

$$\begin{aligned}\arccos'(x) =& \frac{1}{\cos'(\arccos(x))} = \frac{1}{-\sin(\arccos(x))} \\ =& -\frac{1}{\sqrt{\sin^2(\arccos(x))}} = -\frac{1}{\sqrt{1-\cos^2(\arccos(x))}} \\ =& -\frac{1}{\sqrt{1-x^2}}.\end{aligned}$$

Die Ableitung ist zudem offenbar stetig.

□

Bemerkung 18.23.
Schaut man sich die Ableitungen der Funktionen in Korollar 18.22 d.-k. an, so kann man leicht durch Induktion zeigen, daß jede der Funktionen auf ihrem Definitionsbereich unendlich oft differenzierbar ist.

Beispiel 18.24.
Für $a \in \mathbb{R}$ ist die Funktion

$$f : (0, \infty) \longrightarrow \mathbb{R} : x \mapsto x^a$$

unendlich oft differenzierbar auf $(0, \infty)$ mit

$$f'(x) = a \cdot x^{a-1}.$$

Um dies zu sehen, beachten wir, daß $f(x) = \exp(a \cdot \ln(x))$ gilt, so daß f nach Korollar 18.22 die Verkettung zweier differenzierbarer Funktionen ist. Aus der Kettenregel 17.16 folgt dann

$$f'(x) = \exp'\left(a \cdot \ln(x)\right) \cdot \frac{a}{x} = \exp\left(a \cdot \ln(x)\right) \cdot \frac{a}{x} = x^a \cdot \frac{a}{x} = a \cdot x^{a-1}.$$

Daß f sogar unendlich oft differenzierbar ist, folgt dann mit Induktion aus der Tatsache, daß f' eine Funktion der gleichen Gestalt ist.

C.7) Die Regeln von de l'Hôpital

Bemerkung 18.25 (Anwendung der Grenzwertsätze für Funktionen).
Es gibt Situationen, in denen man mit Hilfe der Grenzwersätze Grenzwerte der Form

$$\lim_{x\to c}\frac{f(x)}{g(x)}=\frac{\lim\limits_{x\to c}f(x)}{\lim\limits_{x\to c}g(x)}$$

leicht bestimmen kann, egal ob es sich um eigentliche oder uneigentliche Grenzwerte handelt:

- $\lim\limits_{x\to 0}\frac{\cos(x)}{x^2+1}=\frac{\lim\limits_{x\to 0}\cos(x)}{\lim\limits_{x\to 0}x^2+1}=\frac{1}{1}=1.$
- $\lim\limits_{x\to 0}\frac{\exp(x)}{\sin^2(x)}=\frac{\lim\limits_{x\to 0}\exp(x)}{\lim\limits_{x\to 0}\sin^2(x)}=\frac{1}{0}=\infty.$
- $\lim\limits_{x\to 0}\frac{x^2+1}{\ln(x)}=\frac{\lim\limits_{x\to 0}x^2+1}{\lim\limits_{x\to 0}\ln(x)}=\frac{1}{-\infty}=0.$
- $\lim\limits_{x\to \infty}\frac{\arctan(x)}{1+\frac{1}{x}}=\frac{\lim\limits_{x\to \infty}\arctan(x)}{\lim\limits_{x\to \infty}1+\frac{1}{x}}=\frac{\frac{\pi}{2}}{1}=\frac{\pi}{2}.$
- $\lim\limits_{x\to \infty}\frac{1+\frac{1}{x}}{\exp(x)}=\frac{\lim\limits_{x\to \infty}1+\frac{1}{x}}{\lim\limits_{x\to \infty}\exp(x)}=\frac{1}{\infty}=0.$

Aber es gibt auch Situationen, in denen die Grenzwertsätze nicht weiterhelfen:

$$\lim_{x\to 0}\frac{\sin(x)}{\sqrt{x}}\overset{?}{=}\frac{\lim\limits_{x\to 0}\sin(x)}{\lim\limits_{x\to 0}\sqrt{x}}\overset{?}{=}\frac{0}{0}\qquad\text{oder}\qquad\lim_{x\to\infty}\frac{\ln(x)}{x}\overset{?}{=}\frac{\lim\limits_{x\to\infty}\ln(x)}{\lim\limits_{x\to\infty}x}\overset{?}{=}\frac{\infty}{\infty}.$$

In solchen Situationen könnnen u.U. die Regeln von de l'Hôpital helfen.

Im folgenden Satz soll $[-\infty,\infty]:=\mathbb{R}\cup\{-\infty,\infty\}$ bezeichnen.

Satz 18.26 (Regeln von de l'Hôpital).
Seien $a,b\in[-\infty,\infty]$ mit $a<b$, $f,g:(a,b)\longrightarrow\mathbb{R}$ differenzierbar und $c\in[a,b]$. Ferner gelte $g'(x)\neq 0$ für alle $x\in(a,b)$ und $\lim\limits_{x\to c}\frac{f'(x)}{g'(x)}$ existiere eigentlich oder uneigentlich.

a. Falls $\lim\limits_{x\to c}f(x)=\lim\limits_{x\to c}g(x)=0$, so gilt $\lim\limits_{x\to c}\frac{f(x)}{g(x)}=\lim\limits_{x\to c}\frac{f'(x)}{g'(x)}$.

b. Falls $\lim\limits_{x\to c}g(x)\in\{\infty,-\infty\}$, so gilt $\lim\limits_{x\to c}\frac{f(x)}{g(x)}=\lim\limits_{x\to c}\frac{f'(x)}{g'(x)}$.

Beweis: Wir beschränken uns im Beweis auf den Fall $c\in\mathbb{R}$ und $k:=\lim\limits_{x\to c}\frac{f'(x)}{g'(x)}\in\mathbb{R}$. Die Fälle $c\in\{-\infty,\infty\}$ oder $k\in\{-\infty,\infty\}$ beweist man analog.

Sei $\varepsilon>0$ gegeben, so müssen wir ein $\delta_\varepsilon>0$ finden, so daß

$$\left|\frac{f(x)}{g(x)}-k\right|<\varepsilon \tag{35}$$

für alle $c\neq x\in(a,b)$ mit $|x-c|<\delta_\varepsilon$.

Da $\frac{f'(x)}{g'(x)}$ gegen k konvergiert für x gegen c, gibt es ein $\delta'_\varepsilon > 0$, so daß für alle $c \neq z \in (a,b)$ mit $|z-c| < \delta'_\varepsilon$ auch

$$\left|\frac{f'(z)}{g'(z)} - k\right| < \frac{\varepsilon}{2} \tag{36}$$

gilt.

Wir betrachten nun $c \neq x, y \in (a,b) \cap (c-\delta'_\varepsilon, c+\delta'_\varepsilon)$ mit $x \neq y$ und wenden den allgemeinen Mittelwertsatz 18.10 an. Dann gibt es ein z zwischen x und y mit

$$f'(z) \cdot \big(g(x) - g(y)\big) = g'(z) \cdot \big(f(x) - f(y)\big). \tag{37}$$

Da z zwischen x und y liegt, gilt auch

$$z \in (a,b) \cap (c-\delta'_\varepsilon, c+\delta'_\varepsilon). \tag{38}$$

Nach Voraussetzung ist $g'(z) \neq 0$, und wir behaupten, daß auch $g(x) - g(y) \neq 0$ gilt, da es nach dem Satz von Rolle 18.5 sonst ein w zwischen x und y geben würde mit $g'(w) = 0$, was aber nach Voraussetzung nicht möglich ist. Damit können wir Gleichung (37) auch in der folgenden Form schreiben:

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(z)}{g'(z)}. \tag{39}$$

a. Wir betrachten nun den Fall, daß $\lim\limits_{x\to c} f(x) = \lim\limits_{x\to c} g(x) = 0$. Ist $c \in (a,b)$, so folgt aus der Stetigkeit von f und g automatisch $f(c) = 0 = g(c)$. Ist $c \notin (a,b)$, so können wir f und g in c stetig fortsetzen durch $f(c) = 0 = g(c)$.
Aus (36), (38) und (39) folgt

$$\left|\frac{f(x) - f(y)}{g(x) - g(y)} - k\right| = \left|\frac{f'(z)}{g'(z)} - k\right| < \frac{\varepsilon}{2},$$

und dies gilt für alle $c \neq x, y \in (a,b) \cap (c-\delta'_\varepsilon, c+\delta'_\varepsilon)$ mit $x \neq y$. Da die Funktionen f und g nun stetig in c mit Funktionswert 0 sind, können wir y gegen c gehen lassen und erhalten im Grenzwert

$$\left|\frac{f(x)}{g(x)} - k\right| = \left|\frac{f(x) - f(c)}{g(x) - g(c)} - k\right| = \lim_{y\to c}\left|\frac{f(x) - f(y)}{g(x) - g(y)} - k\right| \leq \frac{\varepsilon}{2} < \varepsilon.$$

Dies gilt für alle $c \neq x \in (a,b) \cap (c-\delta'_\varepsilon, c+\delta'_\varepsilon)$, so daß wir mit $\delta_\varepsilon := \delta'_\varepsilon$ unsere Aussage in diesem Fall bewiesen haben.

b. Wir können annehmen, daß f nicht konstant 0 in einer kleinen Umgebung von c ist, da sonst auch $k = 0$ gilt und (35) sicher erfüllt ist. Deshalb können wir ein $c \neq y \in (a,b) \cap (c-\delta'_\varepsilon, c+\delta'_\varepsilon)$ festhalten mit $f(y) \neq 0$, und wegen $\lim\limits_{x\to c} g(x) = \pm\infty$ können wir auch $g(y) \neq 0$ annehmen.
Aus $\lim\limits_{x\to c} g(x) = \pm\infty$ folgt $\lim\limits_{x\to c} \frac{1}{g(x)} = 0$, und deshalb gibt es ein $\delta''_\varepsilon > 0$, so daß

$$\left|\frac{1}{g(x)}\right| < \frac{\varepsilon}{4 \cdot |f(y)|} \tag{40}$$

für alle $x \in (a,b) \cap (c-\delta''_\varepsilon, c+\delta''_\varepsilon)$ mit $x \neq c, y$.
Aus (36), (38) und (39) folgt, daß

$$\left|\frac{f(x)-f(y)}{g(x)-g(y)}\right| \leq \left|\frac{f(x)-f(y)}{g(x)-g(y)} - k\right| + |k| < \frac{\varepsilon}{2} + |k| =: s \tag{41}$$

für alle $x \in (a,b) \cap (c-\delta'_\varepsilon, c+\delta'_\varepsilon)$ mit $x \neq c, y$, d.h. der Ausdruck ist auf dem angegebenen Intervall nach oben beschränkt.
Wegen $\lim\limits_{x\to c} \frac{1}{g(x)} = 0$ gibt es ein $\delta'''_\varepsilon > 0$ mit

$$\left|\frac{1}{g(x)}\right| < \frac{\varepsilon}{4 \cdot |g(y)| \cdot s} \tag{42}$$

für alle $x \in (a,b) \cap (c-\delta'''_\varepsilon, c+\delta'''_\varepsilon)$ mit $x \neq c, y$.
Nun setzen wir $\delta_\varepsilon := \min\{\delta'_\varepsilon, \delta''_\varepsilon, \delta'''_\varepsilon, |y-c|\}$ und betrachten ein beliebiges $c \neq x \in (a,b) \cap (c-\delta_\varepsilon, c+\delta_\varepsilon)$. (35) gilt dann auch in diesem Fall wegen

$$\begin{aligned}
\left|\frac{f(x)}{g(x)} - k\right| &= \left|\frac{f(y)}{g(x)} + \frac{f(x)-f(y)}{g(x)} - k\right| \\
&= \left|\frac{f(y)}{g(x)} + \frac{f(x)-f(y)}{g(x)-g(y)} \cdot \frac{g(x)-g(y)}{g(x)} - k\right| \\
&= \left|\frac{f(y)}{g(x)} - \frac{f(x)-f(y)}{g(x)-g(y)} \cdot \frac{g(y)}{g(x)} + \frac{f(x)-f(y)}{g(x)-g(y)} - k\right| \\
&\leq \left|\frac{f(y)}{g(x)}\right| + \left|\frac{f(x)-f(y)}{g(x)-g(y)} \cdot \frac{g(y)}{g(x)}\right| + \left|\frac{f(x)-f(y)}{g(x)-g(y)} - k\right| \\
&= \underbrace{\left|\frac{f(y)}{g(x)}\right|}_{\substack{(40)\\ <\frac{\varepsilon}{4}}} + \underbrace{\left|\frac{f(x)-f(y)}{g(x)-g(y)}\right|}_{\substack{(41)\\ \leq s}} \cdot \underbrace{\left|\frac{g(y)}{g(x)}\right|}_{\substack{(42)\\ <\frac{\varepsilon}{4\cdot s}}} + \underbrace{\left|\frac{f(x)-f(y)}{g(x)-g(y)} - k\right|}_{\substack{(36)(38)(39)\\ <\frac{\varepsilon}{2}}} < \varepsilon.
\end{aligned}$$

□

Bemerkung 18.27 (Die Regeln von de l'Hôpital).

a. Wenn c in Satz 18.26 eine Nullstelle der Funktionen f und g ist und diese lokal in c beide stetig differenzierbar sind, dann kann man einen sehr einfachen Beweis für die Aussage des Satzes geben:

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{\frac{f(x)-f(c)}{x-c}}{\frac{g(x)-g(c)}{x-c}} = \frac{\lim\limits_{x\to c} \frac{f(x)-f(c)}{x-c}}{\lim\limits_{x\to c} \frac{g(x)-g(c)}{x-c}} = \frac{f'(c)}{g'(c)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}.$$

Man beachte auch, daß es hier nicht reicht, daß f und g in c differenzierbar sind, da daraus noch nicht einmal die Existenz des Grenzwertes auf der rechten Seite folgt. Dazu kann man sich z.B.

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} x^2 \cdot \sin\left(\frac{1}{x}\right), & \text{für } x \neq 0, \\ 0, & \text{für } x = 0 \end{cases}$$

aus Beispiel 17.19 anschauen. Wir haben dort gezeigt, daß diese Funktion in $c = 0$ differenzierbar ist, daß die Ableitung aber nicht stetig ist. Als zweite Funktion wählen wir den Sinus $g = \sin$. Dann haben f und g in c eine Nullstelle und sind auf ganz $\mathbb{R}$ differenzierbar, aber der Grenzwert

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = \lim_{x\to 0} \frac{2 \cdot x \cdot \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)}{\cos(x)}$$

existiert nicht, da $\lim_{x\to 0} \cos\left(\frac{1}{x}\right)$ nicht existiert, wie man leicht sieht.

b. Im Beweis von Satz 18.26 haben wir gesehen, daß die Funktion g auf dem Intervall $[a, b]$ keinen Wert *zweimal* annehmen kann, da wegen des Satzes von Rolle 18.5 die Ableitung ansonsten auch einmal Null würde. Insbesondere zeigt das, daß g höchstens eine Nullstelle haben kann! Die Bedingung $g'(x) \neq 0$ für alle $x \in (a, b)$ erzwingt also, daß auch $g(x)$ im wesentlichen ungleich Null ist.

c. Ist g' stetig, so muß g' auf (a, b) entweder stets positiv oder stets negativ sein. Aus Proposition 18.13 folgt dann, daß g *streng monoton* auf dem Intervall (a, b) sein muß. Das zeigt, für welchen Typ von Funktionen g man die Regeln von de l'Hôpital überhaupt nur anwenden kann!

d. Man beachte, daß die zweite Regel von de l'Hôpital 18.26 nur in der Situation $\lim_{x\to c} f(x) = \pm\infty$, d.h.

$$\frac{\lim_{x\to c} f(x)}{\lim_{x\to c} g(x)} = \frac{\pm\infty}{\pm\infty},$$

interessant ist, um den Grenzwert $\lim_{x\to c} \frac{f(x)}{g(x)}$ zu bestimmen, da für $\lim_{x\to c} f(x) = k \in \mathbb{R}$ schon aus den normalen Grenzwertsätzen

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \frac{\lim_{x\to c} f(x)}{\lim_{x\to c} g(x)} = \frac{k}{\pm\infty} = 0$$

folgen würde!

Beispiel 18.28.

a. Wir betrachten die Funktionen $f = \sin$ und $g = \sqrt{\cdot}$ auf dem Intervall $(0, \infty)$. Dort sind beide differenzierbar mit

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} g(x) = 0$$

und

$$g'(x) = \frac{1}{2 \cdot \sqrt{x}} \neq 0$$

für alle $x \in (0, \infty)$. Aus der ersten Regel von de l'Hôpital 18.26 folgt dann

$$\lim_{x\to 0} \frac{\sin(x)}{\sqrt{x}} = \lim_{x\to 0} \frac{f'(x)}{g'(x)} = \lim_{x\to 0} \cos(x) \cdot 2 \cdot \sqrt{x} = \cos(0) \cdot 2 \cdot \sqrt{0} = 0.$$

b. Wir betrachten die Funktionen

$$\ln : (0,\infty) \longrightarrow \mathbb{R}$$

und

$$g : (0,\infty) \longrightarrow (0,\infty) : x \mapsto x^a = \exp\big(a \cdot \ln(x)\big)$$

für ein festes $a \in \mathbb{R}_{>0}$. Nach Korollar 18.22 und Beispiel 18.24 sind beide Funktionen differenzierbar auf $(0,\infty)$.
Da sowohl $\exp(x)$, als auch $\ln(x)$ für $x \to \infty$ gegen ∞ divergieren und da a positiv ist, folgt aus den Grenzwertsätzen für uneigentliche Grenzwerte 13.17

$$\lim_{x\to\infty} g(x) = \lim_{x\to\infty} \exp\big(a \cdot \ln(x)\big) = \infty.$$

Außerdem gilt nach Beispiel 18.24

$$g'(x) = a \cdot x^{a-1} \neq 0$$

für alle $x \in (0,\infty)$. Aus der zweiten Regel von de l'Hôpital 18.26 folgt dann

$$\lim_{x\to\infty} \frac{\ln(x)}{x^a} = \lim_{x\to\infty} \frac{\ln'(x)}{g'(x)} = \lim_{x\to\infty} \frac{\frac{1}{x}}{a \cdot x^{a-1}} = \lim_{x\to\infty} \frac{1}{a \cdot g(x)} = \frac{1}{\infty} = 0.$$

C.8) Wachstum der Exponentialfunktion

Korollar 18.29 (Wachstum der Exponentialfunktion).
Die Exponentialfunktion wächst schneller als jede Polynomfunktion, d.h. ist $f = \sum_{k=0}^{n} a_k \cdot t^k$ ein Polynom über $\mathbb{R}$, so gilt

$$\lim_{x\to\infty} \frac{f(x)}{\exp(x)} = 0.$$

Beweis: Wir bezeichnen die zu f gehörige Polynomfunktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{n} a_k \cdot x^k$$

wieder mit f. Dann ist f differenzierbar und die Ableitung von f ist die Polynomfunktion zum Polynom

$$f' = \sum_{k=1}^{n} k \cdot a_k \cdot t^{k-1}.$$

Ist $f = a_0$ konstant, so folgt die Aussage aus den Grenzwertsätzen,

$$\lim_{x\to\infty} \frac{f(x)}{\exp(x)} = \frac{a_0}{\lim\limits_{x\to\infty} \exp(x)} = 0,$$

da $\lim\limits_{x\to\infty} \exp(x) = \infty$.

Für ein allgemeines Polynom $f \neq 0$ führt man den Beweis am besten durch Induktion nach dem Grad n des Polynoms. Den Fall $n = 0$ haben wir bereits betrachtet. Ist $n \neq 0$, so können wir die zweite Regel von de l'Hôpital 18.26 anwenden, da $\exp'(x) = \exp(x) \neq 0$ für alle $x \in \mathbb{R}$ und da $\lim\limits_{x\to\infty} \exp(x) = \infty$ gilt. Damit erhalten wir

$$\lim_{x\to\infty} \frac{f(x)}{\exp(x)} = \lim_{x\to\infty} \frac{f'(x)}{\exp'(x)} = \lim_{x\to\infty} \frac{f'(x)}{\exp(x)},$$

aber die rechte Seite ist dann Null nach Induktion. □

Bemerkung 18.30.
Korollar 18.29 besagt, daß die Exponentialfunktion asymptotisch für x gegen ∞ erheblich viel schneller wächst als jede Polynomfunktion. Das ist der Grund, weshalb Algorithmen mit exponentiellem Laufzeitverhalten wesentlich schlechter sind als Algorithmen mit polynomialem Laufzeitverhalten, zumindest wenn man Eingabedaten beliebiger Größe zulassen möchte. Für konkrete, hinreichend kleine Eingabedaten muß das nicht der Fall sein. Zudem gibt es Probleme, wie etwa die Berechnung von Gröbnerbasen, für die nur Algorithmen mit exponentiellem Laufzeitverhalten bekannt sind. Damit muß man dann ggf. leben.

C.9) Der Satz von Taylor

Definition 18.31 (Taylorpolynome).
Es sei $f : U \longrightarrow \mathbb{R}$ und $a \in U$.
Ist f n-fach differenzierbar, so nennen wir das Polynom

$$T^n_{f,a} := \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} \cdot (t-a)^k$$

das n-te *Taylorpolynom* von f mit Entwicklungspunkt a.
Ist f unendlich oft differenzierbar, so nennen wir die Potenzreihe

$$T_{f,a} := \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} \cdot (t-a)^k$$

die *Taylorreihe* von f mit Entwicklungspunkt a oder die *Taylor-Entwicklung* von f im Punkt a. Beachte, stets gilt $T^n_{f,a}(a) = T_{f,a}(a) = f(a)$.

Bemerkung 18.32 (Tangenten und das 1. Taylorpolynome).
Ist $f : \mathbb{R} \longrightarrow \mathbb{R}$ differenzierbar, so ist die Gleichung der Tangente an den Graphen von f im Punkt $\big(a, f(a)\big)$ gegeben durch

$$y = f'(a) \cdot (x-a) + f(a) = T^1_{f,a}(x).$$

D.h. das erste Taylorpolynom von f mit Entwicklungspunkt a ist die optimale lineare Approximation der Funktion f lokal in a.

Die Idee ist nun, daß mit steigendem n die Taylorpolynome $T^n_{f,a}$ immer bessere Approximationen von f lokal in a sein werden, und daß im Grenzwert dann die Taylorreihe vielleicht sogar mit f übereinstimmt. Das wird nicht immer aber doch oft der Fall sein – siehe Beispiel 18.34! Funktionen, für die das gilt, nennt man *analytisch* in a.

Man kann die Theorie der Differenzierbarkeit statt für Funktionen auf $\mathbb{R}$ auch für Funktionen auf $\mathbb{C}$ einführen. In der Vorlesung Einführung in die Funktionentheorie wird das getan, und dort zeigt man, daß über $\mathbb{C}$ jede einmal auf $\mathbb{C}$ differenzierbare Funktion schon analytisch ist, d.h. durch eine Potenzreihe gegeben und damit unendlich oft differenzierbar ist! Die komplexen Zahlen verhalten sich also weit unkomplizierter als die reellen Zahlen.

Beispiel 18.33 (Potenzreihen als Taylorreihen).
Ist $\sum_{n=0}^{\infty} a_n \cdot t^n$ eine Potenzreihe auf $\mathbb{R}$ mit Konvergenzradius $r > 0$, so ist die Funktion

$$f : (-r, r) \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} a_n \cdot x^n$$

nach Korollar 18.21 unendlich oft differenzierbar, und mittels Induktion nach n zeigt man, daß

$$f^{(n)}(0) = n! \cdot a_n.$$

Damit stimmt f also mit seiner Taylorreihe

$$T_{f,0} = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} \cdot t^n = \sum_{n=0}^{\infty} a_n \cdot t^n$$

auf dem Konvergenzbereich $(-r, r)$ überein, und die Taylorpolynome

$$T^n_{f,0} = \sum_{k=0}^{n} a_k \cdot t^k$$

definieren eine Folge von Funktionen, die auf jedem abgeschlossenen Intervall $[-R, R] \subseteq (-r, r)$ gleichmäßig gegen f konvergieren.

Beispiel 18.34.
Die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \exp\left(-\frac{1}{x^2}\right), & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0, \end{cases}$$

ist unendlich oft differenzierbar mit $f^{(n)}(0) = 0$ für alle $n \in \mathbb{N}$. Insbesondere ist die Taylorreihe von f also Null,

$$T_{f,0} = 0.$$

In diesem Fall stimmt die Taylorreihe also nur im Punkt $x = 0$ mit der Funktion überein, da $f(x) \neq 0$ für alle $x \neq 0$. Der Beweis ist Aufgabe 18.43.

Der Satz von Taylor sagt etwas darüber aus, wie gut das n-te Taylorpolynom f approximiert.

Satz 18.35 (Satz von Taylor – Restglied nach Lagrange).
Sei I ein Intervall, $f : I \longrightarrow \mathbb{R}$ eine $n+1$-fach differenzierbare Funktion und $x, a \in I$. Dann gibt es ein c zwischen x und a, so daß

$$f(x) - T^n_{f,a}(x) = \frac{f^{(n+1)}(c)}{(n+1)!} \cdot (x-a)^{n+1}.$$

Wir nennen die rechte Seite auch das *Restglied* des n-ten Taylorpolynoms.

Beweis: Wir können ohne Einschränkung $x > a$ annehmen.

Dann definieren wir eine reelle Zahl

$$z := \frac{\big(f(x) - T^n_{f,a}(x)\big) \cdot (n+1)!}{(x-a)^{n+1}} \in \mathbb{R}$$

und eine Funktion $g : [a, x] \longrightarrow \mathbb{R}$ durch

$$\begin{aligned} g(y) :=& f(x) - T^n_{f,y}(x) - \frac{z}{(n+1)!} \cdot (x-y)^{n+1} \\ =& f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(y)}{k!} \cdot (x-y)^k - \frac{z}{(n+1)!} \cdot (x-y)^{n+1} \\ =& f(x) - f(y) - \sum_{k=1}^{n} \frac{f^{(k)}(y)}{k!} \cdot (x-y)^k - \frac{z}{(n+1)!} \cdot (x-y)^{n+1} \end{aligned}$$

für $y \in [a, x]$ — man beachte hier, daß g eine Funktion in der Veränderlichen y ist, während x konstant ist!

Nach Voraussetzung ist f $n+1$-fach differenzierbar auf I, so daß die Funktion g differenzierbar auf $[a, x]$ ist, und mit Hilfe der Produktregel erhalten wir für die Ableitung

$$\begin{aligned} g'(y) =& - f'(y) - \sum_{k=1}^{n} \left(\frac{f^{(k+1)}(y)}{k!} \cdot (x-y)^k - \frac{f^{(k)}(y)}{k!} \cdot k \cdot (x-y)^{k-1} \right) + \frac{z \cdot (n+1)}{(n+1)!} \cdot (x-y)^n \\ =& - f'(y) - \sum_{k=1}^{n} \left(\frac{f^{(k+1)}(y)}{k!} \cdot (x-y)^k - \frac{f^{(k)}(y)}{(k-1)!} \cdot (x-y)^{k-1} \right) + \frac{z}{n!} \cdot (x-y)^n \\ =& - f'(y) - \left(\frac{f^{(n+1)}(y)}{n!} \cdot (x-y)^n - f'(y) \right) + \frac{z}{n!} \cdot (x-y)^n \\ =& \frac{z}{n!} \cdot (x-y)^n - \frac{f^{(n+1)}(y)}{n!} \cdot (x-y)^n = \frac{z - f^{(n+1)}(y)}{n!} \cdot (x-y)^n, \end{aligned}$$

wobei wir beachten, daß die Summe in der zweiten Zeile eine Teleskopsumme ist, so daß nur die Randsummanden übrig bleiben.

Zudem folgt aus der Definition von z

$$g(a) = f(x) - T^n_{f,a}(x) - \frac{z}{(n+1)!} \cdot (x-a)^{n+1} = 0,$$

und aus der Definition des Taylorpolynoms folgt

$$g(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!} \cdot (x-x)^k - \frac{z}{(n+1)!} \cdot (x-x)^{n+1} = f(x) - f(x) = 0.$$

Wir können also den Satz von Rolle 18.5 anwenden und finden ein $c \in (a,x)$ mit

$$0 = g'(c) = \frac{z - f^{(n+1)}(c)}{n!} \cdot (x-c)^n.$$

Da $x - c \neq 0$, muß

$$f^{(n+1)}(c) = z = \frac{\big(f(x) - T^n_{f,a}(x)\big) \cdot (n+1)!}{(x-a)^{n+1}}$$

gelten, und damit

$$f(x) - T^n_{f,a}(x) = \frac{f^{(n+1)}(c)}{(n+1)!} \cdot (x-a)^{n+1}.$$

□

Beispiel 18.36 (Näherungswert für die Eulersche Zahl e).
Wir betrachten die Funktion $f = \exp$, $x = 1$ und $a = 0$. Dann ist $f^{(n+1)} = \exp$ und das n-te Taylorpolynom erfüllt

$$T^n_{\exp,0}(1) = T^n_{f,a}(1) = \sum_{k=0}^{n} \frac{1}{k!}.$$

Mit Hilfe des Satzes von Taylor finden wir ein $c \in (0,1)$ mit

$$\left| e - \sum_{k=0}^{n} \frac{1}{k!} \right| = \left| \exp(1) - T^n_{\exp,0}(1) \right| = \frac{|\exp(c)|}{(n+1)!} \cdot |1-0|^{n+1} < \frac{e}{(n+1)!} < \frac{3}{(n+1)!},$$

wenn wir ausnutzen, daß die Exponentialfunktion streng monoton wachsend und positiv auf $[0,1]$ ist. Wenden wir diese Abschätzung mit $n = 6$ an, so erhalten wir

$$\left| e - \frac{1957}{720} \right| < \frac{1}{1680} < \frac{1}{1000}.$$

Die Dezimalzahldarstellung von $\frac{1957}{720}$ stimmt also bis zur dritten Nachkommastelle mit der Zahl e überein, und daraus ersehen wir:

$$e = 2,718\ldots.$$

Beispiel 18.37 (Taylor-Entwicklung des natürlichen Logarithmus).
Wir wissen, daß der natürliche Logarithmus

$$\ln : (0,\infty) \longrightarrow \mathbb{R}$$

unendlich oft differenzierbar mit Ableitung

$$\ln' : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x}$$

ist. Eine einfache Induktion zeigt, daß für $n \geq 1$ dann

$$\ln^{(n)} : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto (-1)^{n-1} \cdot \frac{(n-1)!}{x^n}$$

gilt. Das n-te Taylorpolynom mit Entwicklungspunkt 1 ist mithin

$$T^n_{\ln,1}(x) = \sum_{k=0}^{n} \frac{\ln^{(k)}(1)}{k!} \cdot (x-1)^k = \sum_{k=1}^{n} (-1)^{k-1} \cdot \frac{(x-1)^k}{k}.$$

Der Betrag aller Ableitungen von ln ist auf dem Intervall $[1,2]$ streng monoton fallend, so daß insbesondere

$$|\ln^{(n+1)}(c)| \leq |\ln^{(n+1)}(1)| = n!$$

für jedes $c \in [1,2]$ gilt. Mit dem Satz von Taylor finden wir zu $x \in [1,2]$ ein c zwischen 1 und $x \leq 2$, so daß

$$|\ln(x) - T^n_{\ln,1}(x)| = \frac{|\ln^{(n+1)}(c)|}{(n+1)!} \cdot |(x-1)|^{n+1} \leq \frac{1}{n+1}.$$

Auf dem Intervall $[1,2]$ konvergiert die durch die Taylorpolynome definierte Funktionenfolge mithin gleichmäßig gegen die Funktion ln, und zugleich konvergiert sie dort gleichmäßig gegen die durch die Taylorreihe definierte Funktion, d.h. für $x \in [1,2]$ gilt

$$\ln(x) = \sum_{n=1}^{\infty} (-1)^{n-1} \cdot \frac{(x-1)^n}{n}.$$

Werten wir diese Gleichheit in $x = 2$ aus, so erhalten wir den Wert für die alternierende harmonische Reihe als

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -\ln(2).$$

C.10) Allgemeinere Bedingung für Extremstellen

Satz 18.38 (Allgemeinere Bedingung für Extremstellen).
Es sei $f : (a,b) \longrightarrow \mathbb{R}$ eine $n+1$-fach stetig differenzierbare Funktion mit $n \in \mathbb{N}$ und es sei $c \in (a,b)$ mit $f'(c) = f''(c) = \ldots = f^{(n)}(c) = 0$ und $f^{(n+1)}(c) \neq 0$.

a. Falls n gerade ist, so ist c keine Extremstelle von f.

b. Falls n ungerade und $f^{(n+1)}(c) < 0$, so ist c ein lokales Maximum.

c. Falls n ungerade und $f^{(n+1)}(c) > 0$, so ist c ein lokales Minimum.

Beweis: Wir können uns ohne Einschränkung auf den Fall $f^{(n+1)}(c) > 0$ beschränken.

Weil $f^{(n+1)}$ stetig ist und in c nicht den Wert Null annimt, muß es nach Aufgabe 14.33 ein $\varepsilon > 0$ geben, so daß

$$f^{(n+1)}(x) \neq 0$$

für alle $x \in (c-\varepsilon, c+\varepsilon)$. Da $f^{(n+1)}$ stetig ist, garantiert der Zwischenwertsatz 14.12 dann, daß $f^{(n+1)}$ in der Tat strikt positiv auf diesem Intervall ist.

Wir betrachten das n-te Taylorpolynom

$$T^n_{f,c} = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!} \cdot (t-c)^k = f(c) + \sum_{k=1}^{n} \frac{0}{k!} \cdot (t-c)^k = f(c)$$

von f mit Entwicklungspunkt c sowie ein $c \neq x \in (c-\varepsilon, c+\varepsilon)$. Aus dem Satz von Taylor erhalten wir ein d_x zwischen c und x, so daß

$$f(x) - f(c) = f(x) - T^n_{f,c}(x) = \frac{f^{(n+1)}(d_x)}{(n+1)!} \cdot (x-c)^{n+1}.$$

Man beachte, daß $d_x \in (c-\varepsilon, c+\varepsilon)$ liegen muß, so daß $f^{(n+1)}(d_x) > 0$ ist.

Ist n ungerade, so ist auch $(x-c)^{n+1}$ stets positiv, und wir erhalten

$$f(x) - f(c) \geq 0$$

für alle $x \in (c-\varepsilon, c+\varepsilon)$, d.h. c ist ein Minimum.

Ist n gerade, so wechselt $(x-c)^{n+1}$ in c das Vorzeichen, d.h.

$$f(x) - f(c) \begin{cases} < 0 & \text{falls } x < c, \\ > 0 & \text{falls } x > c. \end{cases}$$

In c liegt also keine Extremstelle vor. □

Im Fall $n = 1$ lautet die Bedingung für eine Extremstelle einfach $f'(c) = 0$ und $f''(c) \neq 0$ und stimmt mit der Bedingung in Proposition 18.15 überein. Aber wir haben diesmal vorausgesetzt, daß die zweite Ableitung *stetig* ist. Darauf konnten wir in Proposition 18.15 verzichten.

Beispiel 18.39.
Betrachten wir die Funktion $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^4$, dann ist $a = 0$ offenbar ein globales Minimum, da $f(x) > 0$ für alle $x \neq 0$ gilt. Proposition 18.15 hätte uns dieses Minimum nicht offenbart, da

$$f'(0) = 4 \cdot 0^3 = 0, \quad f''(0) = 12 \cdot 0^2 = 0 \quad \text{und} \quad f'''(0) = 24 \cdot 0 = 0$$

gilt. Aber wegen $f^{(4)}(0) = 24 > 0$ weist Satz 18.38 $a = 0$ als ein wenn auch nur lokales Minimum aus.

Bemerkung 18.40.
Man beachte, daß die Funktion in Beispiel 18.34 im Punkt $a = 0$ ein globales Minimum besitzt und unendlich oft differenzierbar ist, daß aber alle Ableitungen im Punkt $a = 0$ verschwinden! Satz 18.38 ist also nicht immer anwendbar, um Extremstellen zu finden.

Aufgaben

Aufgabe 18.41.
Sei $f : (0, 1] \longrightarrow \mathbb{R}$ eine differenzierbare Funktion mit beschränkter Ableitung. Zeige, dass f dann stetig in 0 fortsetzbar ist.

Aufgabe 18.42.
Berechne die Ableitungen der folgenden Funktionen $f : (a, \infty) \longrightarrow \mathbb{R}$.

a. $f(x) = \ln\left(\frac{\ln(x)}{x}\right)$ mit $a = 1$.

b. $f(x) = \frac{x^2+4}{x-4}$ mit $a = 4$.

c. $f(x) = \sqrt{e^{\cos(\sqrt{x})}}$ mit $a = 0$.

Aufgabe 18.43.
Zeige, daß für die Funktion

$$f : \mathbb{R} \longmapsto \mathbb{R},\ x \longmapsto \begin{cases} \exp\left(-\frac{1}{x^2}\right) & \text{für } x \neq 0, \\ 0 & \text{für } x = 0. \end{cases}$$

die folgenden Aussagen gelten:

a. Für alle $n \geq 1$ gibt es ein Polynom $p_n \in \mathbb{R}[t]$, so daß für $x \neq 0$ gilt:

$$f^{(n)}(x) = \frac{p_n(x)}{x^{3 \cdot 2^{n-1}}} \cdot \exp\left(-\frac{1}{x^2}\right).$$

b. Für alle $k \in \mathbb{N}$ gilt $\lim\limits_{x \to 0} \frac{\exp(-\frac{1}{x^2})}{x^k} = 0$.

c. Für alle $n \in \mathbb{N}$ gilt $f^{(n)}(0) = 0$.

d. $f \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ und $T_{f,0} = 0$.

Aufgabe 18.44 (Näherungsweise Berechnung von π).
Betrachte die Funktion $f = \arctan$ auf $\mathbb{R}$.

a. Berechne das dritte Taylorpolynom $T^3_{f,0}$ von arctan mit Entwicklungspunkt 0.

b. Benutze $T^3_{f,0}$ und Aufgabe 16.24 c., um $\frac{\pi}{4}$ und damit π näherungsweise zu bestimmen. Zeige dabei, daß die in der Näherung bis auf zwei Nachkommastellen exakt ist mit

$$\pi = 3,14\ldots.$$

Aufgabe 18.45.

Berechne die folgenden Grenzwerte:

a. $\lim_{x\to\pi} \frac{\cos(x)+1}{x^2-\pi^2}$ mit $x < \pi$.

b. $\lim_{x\to 1} \left(\frac{1}{x-1} - \frac{1}{\ln(x)}\right)$ mit $x > 1$.

c. $\lim_{x\to 0} x^x$ mit $x > 0$.

Aufgabe 18.46.

Bestimme alle Extrema der Funktion $f : [0,1] \longrightarrow \mathbb{R}$, $x \longmapsto (1-x)\cdot\sqrt{1+9x^2}$.

Aufgabe 18.47.

Berechne für die Funktion

$$f : \mathbb{R} \longrightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), x \longmapsto \arctan(x)$$

das zweite Taylor-Polynom $T^2_{f,0}$ um 0 und gib eine Abschätzung für das Restglied $|f(x) - T^2_{f,0}(x)|$ auf dem Intervall $[-1,1]$ an.

Aufgabe 18.48.

Berechne für die Funktion

$$f : \left(-\frac{1}{2}, \frac{1}{2}\right) \longrightarrow \mathbb{R}, x \longmapsto \frac{\cos(x)}{1-(2x)^4}$$

das vierte Taylor-Polynom $T^4_{f,0}$ um 0.

Hinweis: mit etwas Überlegung kann man die Berechnung aller vier Ableitungen von f vermeiden.

Aufgabe 18.49.

Beweise oder widerlege durch eine Gegenbeispiel die folgenden Aussagen:

a. Sei $f : [a,b] \longrightarrow \mathbb{R}$ eine stetige und auf (a,b) differenzierbare Funktion, so daß $\lim\limits_{x\to a} f'(x)$ existiert. Dann ist f differenzierbar in a.

b. Sei $f \in \mathcal{C}^1([a,b],\mathbb{R})$ eine stetig differenzierbare streng monoton wachsende Funktion. Dann gilt $f'(x) > 0$ für alle $x \in [a,b]$.

Aufgabe 18.50 (Binomialreihe).
Zeige mit Hilfe der Taylorentwicklung des natürlichen Logarithmus in Beispiel 18.37 für alle $a \in \mathbb{R} \setminus \mathbb{N}$ und $x \in (-1,1)$ die Gleichheit

$$(1+x)^a = \sum_{n=0}^{\infty} \binom{a}{n} \cdot x^n.$$

Siehe auch Aufgabe 12.59.

Aufgabe 18.51.
Zeige, daß die Funktion $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto 2e^x + e^{2x} - 1$ streng monoton wachsend ist, bestimme das Bild von f sowie die Umkehrfunktion und deren Ableitung.

Aufgabe 18.52 (Vorsicht bei den Regeln von de l'Hôpital).
Wir betrachten die beiden Funktionen

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x + \sin(x) \cdot \cos(x)$$

und

$$g : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto f(x) \cdot \exp(\sin(x)).$$

Man zeige, daß $\lim_{x\to\infty} g(x) = \infty$ gilt und daß auch der Grenzwert $\lim_{x\to\infty} \frac{f'(x)}{g'(x)}$ existiert, daß der Grenzwert $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ aber trotzdem nicht mit der Regel von de l'Hôpital bestimmt werden kann. Welche Voraussetzung von Satz 18.26 ist verletzt?

Aufgabe 18.53.
Bestimme alle Extremstellen von $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^3 - x^2 - 8x + 1$.

Aufgabe 18.54.
Zeige, daß die Funktion

$$f_a : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^3 - 3a^2x + 7$$

für keinen Parameterwert $a \in \mathbb{R}_{>0}$ drei nicht-negative Nullstellen hat.

Aufgabe 18.55.
Bestimme eine monoton steigende, differenzierbare Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} 0, & \text{für } x < -1, \\ 4, & \text{für } x > 1. \end{cases}$$

Aufgabe 18.56.
Bestimme für die Funktion $f : [a, b] \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x}$ mit $0 < a < b$ einen Wert $c \in (a, b)$ mit

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Aufgabe 18.57.
Berechne die folgenden Grenzwerte mittels der Regeln von de l'Hôpital:

a. $\lim_{x\to 0} \frac{1-\cos(x)}{x}$.

b. $\lim_{x\to -1} \frac{x^2-1}{x^3+2x+3}$.

c. $\lim_{x\to \infty} \frac{x^2+\ln(x)}{x^2-\ln(x)}$.

Aufgabe 18.58.
Berechne die folgenden Grenzwerte mittels der Regeln von de l'Hôpital:

a. $\lim_{x\to 0} \frac{20\cdot x^{1300}}{\tan(x^{1300})}$.

b. $\lim_{x\to 0} \frac{e^x-x-1}{1-\cos(x)}$.

c. $\lim_{x\to 0} \frac{a^x-1}{x}$ mit $1 \neq a \in \mathbb{R}_{>0}$.

Aufgabe 18.59.
Es sei $a \in \mathbb{R}$ mit $a > 1$ und

$$f : (0, \infty) \longrightarrow \mathbb{R} : x \mapsto x \cdot \log_a(x).$$

Untersuche, auf welchen Teilintervallen des Definitionsbereichs die Funktion streng monton wächst bzw. fällt, und untersuche das Grenzverhalten für $x \to 0$.

Aufgabe 18.60.
Gegeben seien reelle Zahlen $0 \leq a_1 < a_2 < \ldots < a_n$ sowie die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{i=1}^{n} (x - a_i)^2.$$

Bestimme den minimalen Wert von f und die Stelle $x \in \mathbb{R}$, an der dieser Wert angenommen wird.

Aufgabe 18.61.
Berechne das dritte Taylorpolynom der folgenden Funktionen $f : \mathbb{R} \longrightarrow \mathbb{R}$ im angegebenen Entwicklungspunkt a:

a. $f(x) = \sin(x)$, $a = \frac{\pi}{2}$.

b. $f(x) = e^{4x^2+x}$, $a = 0$.

Hinweis: Mit etwas Überlegung braucht man in Teil b. keine Ableitung zu berechnen!

§ 19 Das Riemann-Integral

Wir werden in diesem Abschnitt im wesentlichen nur Funktionen betrachten, die auf einem *abgeschlossenen Intervall* $[a, b]$ definiert und die dort *beschränkt* sind.

A) Obersummen und Untersummen

Definition 19.1 (Zerlegungen eines Intervalls).
Es seien $a, b \in \mathbb{R}$ mit $a < b$. Ein Tupel $Z = (x_0, \ldots, x_n)$ mit $n \geq 1$ heißt eine *Zerlegung* des Intervalls $[a, b]$, falls

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b.$$

Die Zahl $l(Z) := \max\{x_i - x_{i-1} \mid i = 1, \ldots, n\}$ heißt die *Länge* oder *Feinheit* der Zerlegung, die Menge $\operatorname{supp}(Z) := \{x_0, \ldots, x_n\}$ ihr *Träger*, die Zahl $|Z| := n$ ihre *Mächtigkeit* und die x_i ihre *Stützpunkte*.
Eine zweite Zerlegung $Z' = (y_0, \ldots, y_m)$ von $[a, b]$ heißt *Verfeinerung* von Z, falls

$$\{x_0, \ldots, x_n\} \subseteq \{y_0, \ldots, y_m\}.$$

Zu zwei Zerlegungen $Z = (x_0, \ldots, x_n)$ und $Z' = (y_0, \ldots, y_m)$ definieren wir

$$Z * Z' := (z_0, \ldots, z_k),$$

indem wir die Elemente der Vereinigung $\operatorname{supp}(Z) \cup \operatorname{supp}(Z') = \{z_0, \ldots, z_k\}$ der Größe nach ordnen. Sind Z und Z' Zerlegungen des gleichen Intervalls, so nennen wir $Z * Z'$ ihre *gemeinsame Verfeinerung*.

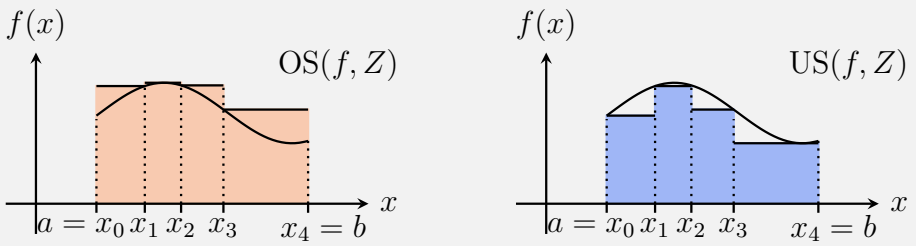
Beispiel 19.2.
Die Tupel $Z = (0, 1, 3, 5)$ und $Z' = (0, 2, 5)$ sind Zerlegungen von $[0, 5]$ der Länge 2 bzw. 3, und die gemeinsame Verfeinerung von Z und Z' ist $Z * Z' = (0, 1, 2, 3, 5)$.

Definition 19.3 (Obersummen und Untersummen).
Sei $f : [a, b] \to \mathbb{R}$ beschränkt, $a < b$, und $Z = (x_0, \ldots, x_n)$ eine Zerlegung von $[a, b]$.
Wir definieren die *Obersumme* von f bezüglich Z als

$$\operatorname{OS}(f, Z) := \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \sup\{f(x) \mid x \in [x_{i-1}, x_i]\},$$

und die *Untersumme* von f bezüglich Z als

$$\operatorname{US}(f, Z) := \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}.$$



Beispiel 19.4.

Wir betrachten die Identität $\mathrm{id} : [0,1] \longrightarrow \mathbb{R} : x \mapsto x$ auf dem Intervall $[0,1]$ sowie die folgende äquidistante Zerlegung der Länge $\frac{1}{n}$

$$Z^n = (x_0, \ldots, x_n) = \left(0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right).$$

Auf einem Teilintervall $[x_{i-1}, x_i] = \left[\frac{i-1}{n}, \frac{i}{n}\right]$ gilt dann

$$m_i := \inf\left\{ \mathrm{id}(x) \;\middle|\; x \in [x_{i-1}, x_i] \right\} = x_{i-1} = \frac{i-1}{n}$$

und

$$M_i := \sup\left\{ \mathrm{id}(x) \;\middle|\; x \in [x_{i-1}, x_i] \right\} = x_i = \frac{i}{n}.$$

Für die Unter- und Obersumme von id bezüglich Z ergibt sich unter Berücksichtigung von Beispiel 7.11 damit

$$\mathrm{US}(\mathrm{id}, Z^n) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot m_i = \sum_{i=1}^{n} \frac{1}{n} \cdot \frac{i-1}{n} = \frac{n \cdot (n-1)}{2 \cdot n^2} = \frac{1}{2} - \frac{1}{2n}$$

und

$$\mathrm{OS}(\mathrm{id}, Z^n) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot M_i = \sum_{i=1}^{n} \frac{1}{n} \cdot \frac{i}{n} = \frac{n \cdot (n+1)}{2 \cdot n^2} = \frac{1}{2} + \frac{1}{2n}.$$

Lemma 19.5.

Es sei $f : [a,b] \longrightarrow \mathbb{R}$ beschränkt mit $|f(x)| \leq M$ für alle $x \in [a,b]$, $a < b$.

a. Ist Z' eine Verfeinerung der Zerlegung Z von $[a,b]$, so gelten

$$0 \leq \mathrm{US}(f, Z') - \mathrm{US}(f, Z) \leq 2 \cdot M \cdot l(Z) \cdot \big(|Z'| - |Z|\big)$$

und

$$0 \leq \mathrm{OS}(f, Z) - \mathrm{OS}(f, Z') \leq 2 \cdot M \cdot l(Z) \cdot \big(|Z'| - |Z|\big).$$

Insbesondere gilt also

$$\mathrm{US}(f, Z) \leq \mathrm{US}(f, Z') \leq \mathrm{OS}(f, Z') \leq \mathrm{OS}(f, Z).$$

b. Für je zwei Zerlegungen Z und Z' von $[a,b]$ gilt

$$\mathrm{US}(f,Z) \le \mathrm{OS}(f,Z').$$

c. Es gelten

$$-M\cdot(b-a) \le \mathrm{US}(f,Z) \le \mathrm{OS}(f,Z) \le (b-a)\cdot M.$$

Beweis:

a. Es sei $Z=(x_0,\dots,x_n)$, und wir setzen für $i=1,\dots,n$ wieder

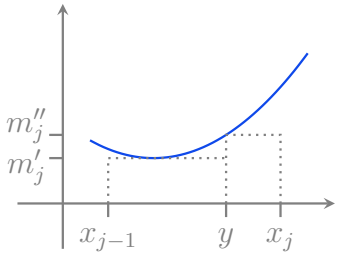
$$m_i := \inf\left\{f(x) \;\middle|\; x\in[x_{i-1},x_i]\right\}.$$

Wir zeigen die Aussage zu den Untersummen für den Fall, daß Z' einen Punkt mehr enthält als Z. Sei also $Z'=(x_0,\dots,x_{j-1},y,x_j,\dots,x_n)$. Dann gilt

$$m'_j := \inf\left\{f(x) \;\middle|\; x\in[x_{j-1},y]\right\} \ge m_j$$

und

$$m''_j := \inf\left\{f(x) \;\middle|\; x\in[y,x_j]\right\} \ge m_j.$$



Daraus ergibt sich

$$\begin{aligned}
\mathrm{US}(f,Z) &= \sum_{i\neq j}(x_i-x_{i-1})\cdot m_i + (x_j-x_{j-1})\cdot m_j\\
&= \sum_{i\neq j}(x_i-x_{i-1})\cdot m_i + (y-x_{j-1})\cdot m_j + (x_i-y)\cdot m_j\\
&\le \sum_{i\neq j}(x_i-x_{i-1})\cdot m_i + (y-x_{j-1})\cdot m'_j + (x_i-y)\cdot m''_j\\
&= \mathrm{US}(f,Z').
\end{aligned}$$

Für die Differenz der beiden Terme erhalten wir

$$\begin{aligned}
0 \le{}& \mathrm{US}(f,Z') - \mathrm{US}(f,Z)\\
={}& (y-x_{j-1})\cdot(m'_j-m_j) + (x_j-y)\cdot(m''_j-m_j)\\
\le{}& (y-x_{j-1})\cdot(M+M) + (x_j-y)\cdot(M+M)\\
={}& (x_j-x_{j-1})\cdot 2\cdot M \le 2\cdot M\cdot l(Z).
\end{aligned}$$

Für eine beliebige Verfeinerung Z' von Z wenden wir dann Induktion an und erhalten die Formel

$$0 \leq \mathrm{US}(f, Z') - \mathrm{US}(f, Z) \leq 2 \cdot M \cdot l(Z) \cdot \big(|Z'| - |Z|\big)$$

Die Aussage für Obersummen zeigt man analog.

b. Wir betrachten die gemeinsame Verfeinerung $Z * Z' = (y_0, \ldots, y_k)$. Wegen

$$m_i := \inf\big\{f(x) \;\big|\; x \in [y_{i-1}, y_i]\big\} \leq \sup\big\{f(x) \;\big|\; x \in [y_{i-1}, y_i]\big\} =: M_i$$

folgt dann

$$\begin{aligned}\mathrm{US}(f, Z) &\overset{a.}{\leq} \mathrm{US}(f, Z * Z') = \sum_{i=1}^{k} (y_i - y_{i-1}) \cdot m_i \\ &\leq \sum_{i=1}^{k} (y_i - y_{i-1}) \cdot M_i = \mathrm{OS}(f, Z * Z') \overset{a.}{\leq} \mathrm{OS}(f, Z').\end{aligned}$$

c. Die Aussage folgt aus a., da Z eine Verfeinerung der Zerlegung (a, b) ist und da $M \geq \sup\big\{f(x) \;\big|\; x \in [a, b]\big\} \geq \inf\big\{f(x) \;\big|\; x \in [a, b]\big\} \geq -M$.

□

Beispiel 19.6.
In Beispiel 19.4 gilt $\mathrm{US}(\mathrm{id}, Z^n) = \frac{1}{2} - \frac{1}{2n} \leq \frac{1}{2} + \frac{1}{2n} = \mathrm{OS}(\mathrm{id}, Z^n)$.

B) Riemann-integrierbare Funktionen

Da die Menge der Obersummen und die Menge der Untersummen nach Lemma 19.5 c. beschränkt sind, können wir ihr Infimum und ihr Supremum betrachten.

Definition 19.7 (Riemann-integrierbar).
Es sei $f : [a, b] \longrightarrow \mathbb{R}$ beschränkt, $a < b$. Wir definieren das *Oberintegral*

$$\mathrm{OI}(f) := \inf\big\{\,\mathrm{OS}(f, Z) \;\big|\; Z \text{ Zerlegung von } [a, b]\big\}$$

von f und das *Unterintegral*

$$\mathrm{UI}(f) := \sup\big\{\,\mathrm{US}(f, Z) \;\big|\; Z \text{ Zerlegung von } [a, b]\big\}$$

von f. Wegen Lemma 19.5 b. und Lemma 8.16 gilt

$$\mathrm{UI}(f) \leq \mathrm{OI}(f).$$

Wir nennen f (Riemann-)*integrierbar* auf $[a, b]$, falls $\mathrm{UI}(f) = \mathrm{OI}(f)$. Dann heißt

$$\int_a^b f(x)\, dx := \mathrm{OI}(f) \in \mathbb{R}$$

das *Integral* von f auf $[a, b]$.

Beispiel 19.8.
Aus Beispiel 19.4 wissen wir für $\mathrm{id} : [0,1] \longrightarrow \mathbb{R}$

$$\frac{1}{2} - \frac{1}{2n} = \mathrm{US}(\mathrm{id}, Z^n) \leq \mathrm{UI}(\mathrm{id}) \leq \mathrm{OI}(\mathrm{id}) \leq \mathrm{OS}(\mathrm{id}, Z^n) = \frac{1}{2} + \frac{1}{2n}$$

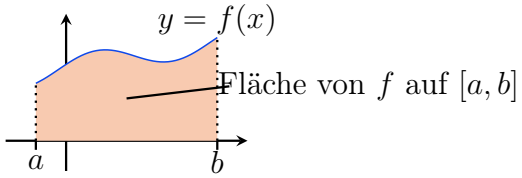
für alle $n \in \mathbb{N}$. Bilden wir nun den Grenzwert für n gegen unendlich, so erhalten wir

$$\frac{1}{2} = \lim_{n\to\infty} \mathrm{US}(\mathrm{id}, Z^n) \leq \mathrm{UI}(\mathrm{id}) \leq \mathrm{OI}(\mathrm{id}) \leq \lim_{n\to\infty} \mathrm{OS}(\mathrm{id}, Z^n) = \frac{1}{2},$$

d.h. id ist integrierbar auf $[0,1]$ mit

$$\int_0^1 x\,dx = \mathrm{OI}(\mathrm{id}) = \mathrm{UI}(\mathrm{id}) = \frac{1}{2}.$$

Bemerkung 19.9 (Das Riemann-Integral als Flächeninhalt).
Wenn die Funktion nur nicht-negative Werte annimmt, dann sind die Untersummen von f nach oben beschränkt durch den Flächeninhalt I der Fläche, die der Graph von f mit der x-Achse einschließt, und die Obersummen von f sind durch diesen nach unten beschränkt. Aufgrund der Definition von $\mathrm{OI}(f)$ als Infimum und $\mathrm{UI}(f)$ als Supremum gilt also stets $\mathrm{UI}(f) \leq I \leq \mathrm{OI}(f)$. Daß f integrierbar ist, bedeutet mithin nichts anderes, als daß das Integral $\int_a^b f(x)\,dx$ den Flächeninhalt der Fläche beschreibt, die der Graph von f auf dem Intervall $[a,b]$ mit der x-Achse einschließt.



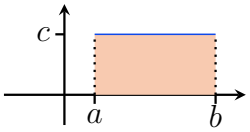
Beispiel 19.10.

a. Jede konstante Funktion $f : [a,b] \longrightarrow \mathbb{R} : x \mapsto c$ ist integrierbar mit

$$\int_a^b f(x)\,dx = (b-a)\cdot c.$$

Denn dann gilt für jede Zerlegung $Z = (x_0, \ldots, x_n)$ von $[a,b]$ bereits

$$\mathrm{OS}(f,Z) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot c = (b-a)\cdot c = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot c = \mathrm{US}(f,Z).$$



b. Die *Dirichletsche Sprungfunktion*

$$f : [0,1] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x \in \mathbb{Q}, \\ 1, & \text{falls } x \notin \mathbb{Q} \end{cases}$$

ist *nicht* integrierbar auf $[0,1]$. Denn ist $Z = (x_0, \ldots, x_n)$ eine beliebige Zerlegung von $[0,1]$, so gibt es im Intervall $[x_{i-1}, x_i]$ sowohl eine rationale Zahl, als auch eine irrationale. Mithin gilt

$$\mathrm{US}(f,Z) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot 0 = 0$$

und

$$\mathrm{OS}(f,Z) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot 1 = 1$$

für jede Zerlegung Z, so daß

$$\mathrm{UI}(f) = 0 < 1 = \mathrm{OI}(f).$$

C) Das Riemannsche Integrabilitätskriterium

Satz 19.11 (Riemannsches Integrabilitätskriterium).
Sei $f : [a,b] \to \mathbb{R}$ beschränkt, $a < b$. Genau dann ist f integrierbar auf $[a,b]$, wenn

$$\forall\, \varepsilon > 0 \;\; \exists\, Z \text{ Zerlegung von } [a,b] \;:\; \mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon.$$

Beweis: $\Longrightarrow$: Sei zunächst f integrierbar auf $[a,b]$ und sei $\varepsilon > 0$ gegeben. Aufgrund der Definition von $\mathrm{UI}(f)$ als Supremum und $\mathrm{OI}(f)$ als Infimum und wegen Proposition 8.15 gibt es Zerlegungen Z' und Z'' von $[a,b]$ mit

$$\mathrm{OI}(f) + \frac{\varepsilon}{2} > \mathrm{OS}(f,Z') \overset{19.5}{\geq} \mathrm{OS}(f, Z' * Z'')$$

und

$$\mathrm{UI}(f) - \frac{\varepsilon}{2} < \mathrm{US}(f,Z'') \overset{19.5}{\leq} \mathrm{US}(f, Z' * Z'').$$

Damit erhalten wir mit $Z = Z' * Z''$ und wegen $\mathrm{UI}(f) = \mathrm{OI}(f)$

$$\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \left(\mathrm{OI}(f) + \frac{\varepsilon}{2}\right) - \left(\mathrm{UI}(f) - \frac{\varepsilon}{2}\right) = \varepsilon.$$

$\Longleftarrow$: Für $\varepsilon := \frac{1}{n}$ mit $n \geq 1$ gibt es eine Zerlegung Z^n von $[a,b]$ mit

$$\frac{1}{n} > \mathrm{OS}(f,Z^n) - \mathrm{US}(f,Z^n) \geq \mathrm{OI}(f) - \mathrm{UI}(f) \geq 0.$$

Da die linke Seite der Ungleichung für $n \to \infty$ gegen Null konvergiert, folgt im Grenzwert

$$0 = \lim_{n\to\infty} \frac{1}{n} \geq \mathrm{OI}(f) - \mathrm{UI}(f) \geq 0,$$

also $\mathrm{OI}(f) = \mathrm{UI}(f)$. Mithin ist f integrierbar auf $[a,b]$. □

Das folgende Verschärfung des Riemannschen Integrabilitätskriteriums sagt, daß für integrierbare Funktionen Untersummen und Obersummen schon beliebig nahe beieinander und damit beim Wert des Integrals liegen, wenn nur die Länge der Zerlegung hinreichend klein gewählt ist.

Lemma 19.12 (Verschärfung des Riemannschen Integrabilitätskriteriums).
Ist $f : [a,b] \longrightarrow \mathbb{R}$ integrierbar, $a < b$, so gilt:

$$\forall\, \varepsilon > 0\ \exists\, \delta_\varepsilon > 0\ :\ \forall\ Z \text{ Zerlegung mit } l(Z) < \delta_\varepsilon \text{ gilt } \mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon.$$

Beweis: Sei $\varepsilon > 0$ gegeben. Aus dem Riemannschen Integrabilitätskriterium erhalten wir eine Zerlegung Z' von $[a,b]$, so daß

$$\mathrm{OS}(f,Z') - \mathrm{US}(f,Z') < \frac{\varepsilon}{2}. \tag{43}$$

Wir setzen nun

$$\delta_\varepsilon := \frac{\varepsilon}{8 \cdot |Z'| \cdot M} > 0,$$

wobei $M := \sup\big\{|f(x)| \;\big|\; x \in [a,b]\big\}$. Ist Z eine Zerlegung von $[a,b]$ mit $l(Z) < \delta_\varepsilon$, so folgt aus aus Lemma 19.5 und $|Z * Z'| - |Z| \le |Z'|$

$$\mathrm{OS}(f,Z) - \mathrm{OS}(f,Z*Z') \le 2 \cdot M \cdot l(Z) \cdot (|Z*Z'| - |Z|) < 2 \cdot M \cdot \delta_\varepsilon \cdot |Z'| = \frac{\varepsilon}{4} \tag{44}$$

und

$$\mathrm{US}(f,Z*Z') - \mathrm{US}(f,Z) \le 2 \cdot M \cdot l(Z) \cdot (|Z*Z'| - |Z|) < 2 \cdot M \cdot \delta_\varepsilon \cdot |Z'| = \frac{\varepsilon}{4}. \tag{45}$$

Da $Z * Z'$ eine Verfeinerung von Z' ist, folgt aus (43) zusammen mit Lemma 19.5

$$\mathrm{OS}(f,Z*Z') - \mathrm{US}(f,Z*Z') \le \mathrm{OS}(f,Z') - \mathrm{US}(f,Z') < \frac{\varepsilon}{2}. \tag{46}$$

Insgesamt erhalten wir damit

$$\begin{aligned}\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) &= \mathrm{OS}(f,Z) - \mathrm{OS}(f,Z*Z') + \mathrm{OS}(f,Z*Z') - \mathrm{US}(f,Z*Z') \\ &\quad + \mathrm{US}(f,Z*Z') - \mathrm{US}(f,Z) \overset{(44)(46)(45)}{<} \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon.\end{aligned}$$

□

D) Zwei Klassen integrierbarer Funktionen

Im folgenden wollen wir das Riemannsche Integrabilitätskriterium anwenden, um von zwei wichtigen Klassen von Funktionen zu zeigen, daß sie integrierbar sind.

Satz 19.13 (Stetige Funktionen sind integrierbar.).
Ist $f : [a,b] \longrightarrow \mathbb{R}$ stetig, $a < b$, so ist f integrierbar auf $[a,b]$.

Beweis: Da f stetig auf dem abgeschlossenen Intervall $[a,b]$ ist, ist f dort beschränkt nach Proposition 14.15 und gleichmäßig stetig nach Satz 14.28.

Sei nun $\varepsilon > 0$ gegeben. Da f gleichmäßig stetig auf $[a,b]$ ist, gibt es $\delta_\varepsilon > 0$, so daß

$$|f(x) - f(y)| < \frac{\varepsilon}{b-a} \tag{47}$$

für alle $x, y \in [a,b]$ mit $|x-y| < \delta_\varepsilon$. Wir wählen nun eine Zerlegung $Z = (x_0, \ldots, x_n)$ mit Länge $l(Z) < \delta_\varepsilon$. Da f stetig auf $[x_{i-1}, x_i]$ ist, existieren $y_i, z_i \in [x_{i-1}, x_i]$ mit

$$f(y_i) = \sup\left\{f(y) \;\middle|\; y \in [x_{i-1}, x_i]\right\}$$

und

$$f(z_i) = \inf\left\{f(y) \;\middle|\; y \in [x_{i-1}, x_i]\right\},$$

und wegen $|y_i - z_i| \leq |x_i - x_{i-1}| < \delta_\varepsilon$ folgt aus (47) zudem

$$0 \leq f(y_i) - f(z_i) < \frac{\varepsilon}{b-a}.$$

Damit erhalten wir insbesondere

$$\begin{aligned} \mathrm{OS}(f,Z) - \mathrm{US}(f,Z) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \big(f(y_i) - f(z_i)\big) \\ &< \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \frac{\varepsilon}{b-a} = (b-a) \cdot \frac{\varepsilon}{b-a} = \varepsilon. \end{aligned}$$

Somit ist f integrierbar nach dem Riemannschen Integrabilitätskriterium 19.11. □

Beispiel 19.14.
Wir betrachten die Funktion

$$f : [0,1] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x = 0, \\ 1, & \text{falls } x \neq 0 \end{cases}$$

und die Zerlegung $Z^n = \left(0, \frac{1}{n}, 1\right)$ für $n \geq 1$. Dann gilt

$$\mathrm{US}(f, Z^n) = \left(\frac{1}{n} - 0\right) \cdot 0 + \left(1 - \frac{1}{n}\right) \cdot 1 = 1 - \frac{1}{n}$$

und

$$\mathrm{OS}(f, Z^n) = \left(\frac{1}{n} - 0\right) \cdot 1 + \left(1 - \frac{1}{n}\right) \cdot 1 = 1.$$

Wir erhalten also

$$1 \longleftarrow 1 - \frac{1}{n} = \mathrm{US}(f, Z^n) \leq \mathrm{UI}(f) \leq \mathrm{OI}(f) \leq \mathrm{OS}(f, Z^n) = 1.$$

Mithin ist f auf $[0,1]$ integrierbar mit

$$\int_0^1 f(x)\,dx = 1.$$

Dies zeigt, daß eine Funktion nicht stetig sein muß, um integrierbar zu sein.

Proposition 19.15 (Monotone Funktionen sind integrierbar.).
Ist $f : [a,b] \longrightarrow \mathbb{R}$ monoton wachsend oder fallend, $a < b$, so ist f integrierbar.

Beweis: Wir können ohne Einschränkung annehmen, daß f monoton wachsend und nicht konstant ist. Insbesondere ist $f(b) > f(a)$. Außerdem ist f beschränkt, da $f(a) \le f(x) \le f(b)$ für alle $x \in [a,b]$.

Sei $\varepsilon > 0$ gegeben. Wir wählen eine natürliche Zahl n so, daß

$$\frac{1}{n} < \frac{\varepsilon}{(b-a)\cdot(f(b)-f(a))}, \tag{48}$$

und betrachten die Zerlegung $Z = (x_0, \ldots, x_n)$ mit

$$x_i := a + i \cdot \frac{(b-a)}{n}.$$

Da f monoton wachsend ist, ist

$$\sup\big\{f(x) \;\big|\; x \in [x_{i-1}, x_i]\big\} = f(x_i)$$

und

$$\inf\big\{f(x) \;\big|\; x \in [x_{i-1}, x_i]\big\} = f(x_{i-1}).$$

Für die Ober- und Untersumme von f bezüglich Z folgt damit

$$\begin{aligned}
\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \big(f(x_i) - f(x_{i-1})\big) \\
&= \sum_{i=1}^{n} \frac{b-a}{n} \cdot \big(f(x_i) - f(x_{i-1})\big) \\
&= \frac{b-a}{n} \cdot \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big) \\
&= \frac{b-a}{n} \cdot \big(f(x_n) - f(x_0)\big) \\
&= \frac{b-a}{n} \cdot \big(f(b) - f(a)\big) \overset{(48)}{<} \varepsilon.
\end{aligned}$$

Somit ist f integrierbar nach dem Riemannschen Integrabilitätskriterium 19.11. □

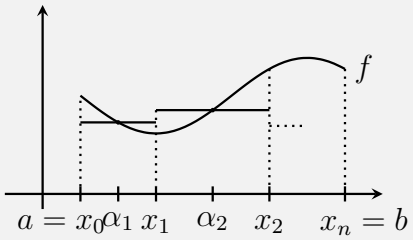
Beispiel 19.16.
Die Funktion in Beispiel 19.14 ist monoton wachsend und deshalb nach Proposition 19.15 auch integrierbar. 19.15 sagt aber nichts über den Wert des Integrals aus!

E) Riemannsche Zwischensummen

Definition 19.17 (Riemannsche Zwischensummen).
Sei $f : [a, b] \to \mathbb{R}$ beschränkt, $a < b$, und $Z = (x_0, \ldots, x_n)$ eine Zerlegung von $[a, b]$.
Erfüllt $\alpha = (\alpha_1, \ldots, \alpha_n)$ die Bedingung $\alpha_i \in [x_{i-1}, x_i]$ für $i = 1, \ldots, n$, so nennen wir

$$\mathrm{ZS}(f, Z, \alpha) := \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot f(\alpha_i)$$

die *Riemannsche Zwischensumme* von f bezüglich der Zerlegung Z und den *Zwischenpunkten* α.



Das nächste Lemma sagt, daß man Obersummen und Untersummen beliebig gut approximieren kann durch Zwischensummen.

Lemma 19.18 (Approximation von Ober- und Unter- durch Zwischensummen).
Sei $f : [a, b] \longrightarrow \mathbb{R}$ beschränkt, $a < b$, Z eine Zerlegung von $[a, b]$ und $\varepsilon > 0$.

a. Dann gibt es Zwischenpunkte α mit $0 \leq \mathrm{OS}(f, Z) - \mathrm{ZS}(f, Z, \alpha) < \varepsilon$.

b. Dann gibt es Zwischenpunkte β mit $0 \leq \mathrm{ZS}(f, Z, \beta) - \mathrm{US}(f, Z) < \varepsilon$.

Beweis: Sei $Z = (x_0, \ldots, x_n)$ und sei

$$M_i := \sup \big\{ f(x) \;\big|\; x \in [x_{i-1}, x_i] \big\}.$$

Aufgrund der Definition von M_i als Supremum der Funktionswerte auf dem Intervall $[x_{i-1}, x_i]$ gibt es ein $\alpha_i \in [x_{i-1}, x_i]$, so daß

$$f(\alpha_i) > M_i - \frac{\varepsilon}{b - a}.$$

Damit erhalten wir für $\alpha = (\alpha_1, \ldots, \alpha_n)$

$$\begin{aligned} \mathrm{OS}(f,Z) - \mathrm{ZS}(f,Z,\alpha) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \big(M_i - f(\alpha_i)\big) \\ &< \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \frac{\varepsilon}{b-a} \\ &= (x_n - x_0) \cdot \frac{\varepsilon}{b-a} = \varepsilon. \end{aligned}$$

Damit ist a. gezeigt, und b. zeigt man analog. □

F) Riemannsches Folgenkriterium für Integrierbarkeit

Satz 19.19 (Riemannsches Folgenkriterium für Integrierbarkeit).
Es sei $f : [a,b] \longrightarrow \mathbb{R}$ eine beschränkte Funktion, $a < b$, und $I \in \mathbb{R}$.
Genau dann ist f auf $[a,b]$ integrierbar mit $I = \int_a^b f(x)\,dx$, wenn für jede Folge $(Z^n, \alpha^n)_{n\in\mathbb{N}}$ von Zerlegungen von $[a,b]$ und Zwischenpunkten mit $l(Z^n) \to 0$ gilt

$$\mathrm{ZS}(f, Z^n, \alpha^n) \longrightarrow I.$$

Beweis:

$\Longrightarrow$**:** Es sei $(Z^n, \alpha^n)_{n\in\mathbb{N}}$ eine Folge von Zerlegungen von $[a,b]$ mit Zwischenpunkten, so daß $\lim\limits_{n\to\infty} l(Z^n) = 0$, und sei $I = \int_a^b f(x)\,dx$.
Sei $\varepsilon > 0$ gegeben. Wir müssen ein $n_\varepsilon \in \mathbb{N}$ finden, so daß

$$|\mathrm{ZS}(f, Z^n, \alpha^n) - I| < \varepsilon \tag{49}$$

für alle $n \geq n_\varepsilon$.
Da f integrierbar ist, gibt es nach Lemma 19.12 ein $\delta_\varepsilon > 0$, so daß für eine Zerlegung Z von $[a,b]$ aus $l(Z) < \delta_\varepsilon$ auch

$$\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon \tag{50}$$

gilt. Wegen $\lim\limits_{n\to\infty} l(Z^n) = 0$ gibt es ein $n_\varepsilon \in \mathbb{N}$, so daß $l(Z^n) < \delta_\varepsilon$ für $n \geq n_\varepsilon$.
Für $n \geq n_\varepsilon$ leiten wir dann aus (50)

$$\mathrm{ZS}(f, Z^n, \alpha^n) - I \leq \mathrm{OS}(f, Z^n) - I \leq \mathrm{OS}(f, Z^n) - \mathrm{US}(f, Z^n) < \varepsilon$$

her, sowie

$$\mathrm{ZS}(f, Z^n, \alpha^n) - I \geq \mathrm{US}(f, Z^n) - I \geq \mathrm{US}(f, Z^n) - \mathrm{OS}(f, Z^n) > -\varepsilon.$$

Damit ist (49) für $n \geq n_\varepsilon$ erfüllt, und das heißt $\mathrm{ZS}(f, Z^n, \alpha^n) \longrightarrow I$.

$\Longleftarrow$: Wir wollen das Riemannsche Integrabilitätskriterium anwenden.
Sei dazu $\varepsilon > 0$ gegeben. Wir betrachten die Zerlegung $Z^n = (x_0^n, \ldots, x_n^n)$, $n \geq 1$, mit

$$x_i^n := a + i \cdot \frac{(b-a)}{n}$$

für $i = 0, \ldots, n$. Dann gilt

$$\lim_{n\to\infty} l(Z^n) = \lim_{n\to\infty} \frac{b-a}{n} = 0.$$

Wir wollen zunächst

$$\lim_{n\to\infty} \mathrm{OS}(f, Z^n) = I \tag{51}$$

zeigen. Sei dazu $\varepsilon > 0$ gegeben, dann müssen wir ein n_ε finden, so daß

$$|\,\mathrm{OS}(f, Z^n) - I| < \varepsilon$$

für alle $n \geq n_\varepsilon$ gilt.
Mit Lemma 19.18 finden wir zu $n \in \mathbb{N}$ Zwischenpunkte α^n, so daß

$$\mathrm{OS}(f, Z^n) - \mathrm{ZS}(f, Z^n, \alpha^n) < \frac{\varepsilon}{2}.$$

Außerdem gilt nach Voraussetzung

$$\lim_{n\to\infty} \mathrm{ZS}(f, Z^n, \alpha^n) = I,$$

so daß es ein n_ε gibt mit

$$|\,\mathrm{ZS}(f, Z^n, \alpha^n) - I| < \frac{\varepsilon}{2}$$

für alle $n \geq n_\varepsilon$. Sei nun $n \geq n_\varepsilon$ gegeben, dann folgt

$$\begin{aligned} |\,\mathrm{OS}(f, Z^n) - I| \leq & |\,\mathrm{OS}(f, Z^n) - \mathrm{ZS}(f, Z^n, \alpha^n)| + |\,\mathrm{ZS}(f, Z^n, \alpha^n) - I| \\ < & \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Damit ist (51) gezeigt, und analog sieht man

$$\lim_{n\to\infty} \mathrm{US}(f, Z^n) = I$$

Damit erhalten wir dann insgesamt

$$I \longleftarrow \mathrm{US}(f, Z^n) \leq \mathrm{UI}(f) \leq \mathrm{OI}(f) \leq \mathrm{OS}(f, Z^n) \longrightarrow I,$$

so daß aus dem Einschachtelungssatz 11.17

$$I = \mathrm{OI}(f) = \mathrm{UI}(f) = \int_a^b f(x)\,dx$$

und insbesondere die Integrierbarkeit von f folgt.

□

Beispiel 19.20.
Die Funktion $f : [0, b] \longrightarrow \mathbb{R} : x \mapsto x^2$, $b > 0$, ist stetig und mithin integrierbar. Setzen wir

$$x_i := \frac{i \cdot b}{n},$$

so ist $Z^n = (x_0, \ldots, x_n)$ eine Zerlegung von $[0, b]$ mit Zwischenpunkten $\alpha^n = (x_1, \ldots, x_n)$, und es gilt $l(Z^n) \longrightarrow 0$. Um die Zwischensumme berechnen zu können, verwenden wir die Formel

$$\sum_{i=1}^{n} i^2 = \frac{n \cdot (n+1) \cdot (2n+1)}{6}, \tag{52}$$

die man mit Hilfe von Induktion leicht zeigen kann. Damit ergibt sich

$$\begin{aligned} \mathrm{ZS}(f, Z^n, \alpha^n) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot x_i^2 = \sum_{i=1}^{n} \frac{b}{n} \cdot \frac{i^2 \cdot b^2}{n^2} \\ &= \frac{b^3}{n^3} \cdot \sum_{i=1}^{n} i^2 \overset{(52)}{=} \frac{b^3}{n^3} \cdot \frac{n \cdot (n+1) \cdot (2n+1)}{6} \\ &= \frac{b^3}{3} \cdot \left(1 + \frac{1}{n}\right) \cdot \left(1 + \frac{1}{2n}\right) \longrightarrow \frac{b^3}{3}. \end{aligned}$$

Aus dem Riemannschen Folgenkriterium für Integrierbarkeit 19.19 folgt dann

$$\int_0^b x^2 \, dx = \frac{b^3}{3}.$$

G) Rechenregeln für Integrale – Linearität und Monotonie

Korollar 19.21 (Linearität und Monotonie des Integrals).
Seien $f, g : [a, b] \longrightarrow \mathbb{R}$ integrierbar, $a < b$, und $c, d \in \mathbb{R}$.

a. Dann ist $c \cdot f + d \cdot g$ integrierbar auf $[a, b]$ mit

$$\int_a^b (c \cdot f + d \cdot g)(x) \, dx = c \cdot \int_a^b f(x) \, dx + d \cdot \int_a^b g(x) \, dx.$$

b. Ist $f(x) \leq g(x)$ für alle $x \in [a, b]$, so ist auch

$$\int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx.$$

Beweis:

a. Wir beachten zunächst, daß für jede Zerlegung $Z = (x_0, \ldots, x_n)$ von $[a, b]$ mit Zwischenpunkten $\alpha = (\alpha_1, \ldots, \alpha_n)$ offenbar gilt:

$$\begin{aligned}\mathrm{ZS}(cf + dg, Z, \alpha) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot (cf + dg)(\alpha_i) \\ &= c \cdot \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot f(\alpha_i) + d \cdot \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot g(\alpha_i) \\ &= c \cdot \mathrm{ZS}(f, Z, \alpha) + d \cdot \mathrm{ZS}(g, Z, \alpha).\end{aligned}$$

Es sei nun $(Z^n, \alpha^n)_{n \in \mathbb{N}}$ eine Folge von Zerlegungen von $[a, b]$ und Zwischenpunkten mit $l(Z^n) \longrightarrow 0$. Aus den Grenzwertsätzen für Folgen 11.15 und Satz 19.19 folgt dann

$$\mathrm{ZS}(cf + dg, Z^n, \alpha^n) = c \cdot \mathrm{ZS}(f, Z^n, \alpha^n) + d \cdot \mathrm{ZS}(g, Z^n, \alpha^n) \longrightarrow c \cdot \int_a^b f(x)\,dx + d \cdot \int_a^b g(x)\,dx.$$

Das Riemannsche Folgenkriterium für Integrierbarkeit 19.19 liefert dann die Behauptung.

b. Es sei $(Z^n, \alpha^n)_{n \in \mathbb{N}}$ eine Folge von Zerlegungen von $[a, b]$ und Zwischenpunkten mit $l(Z^n) \longrightarrow 0$. Wegen $f(x) \leq g(x)$ für alle $x \in [a, b]$ gilt dann offenbar

$$\int_a^b f(x)\,dx \longleftarrow \mathrm{ZS}(f, Z^n, \alpha^n) \leq \mathrm{ZS}(g, Z^n, \alpha^n) \longrightarrow \int_a^b g(x)\,dx,$$

wobei die Grenzwerte aus dem Riemannschen Folgenkriterium für Integrierbarkeit folgen. Damit gilt dann aber auch für die Grenzwerte

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$

□

Beispiel 19.22.

Aus Beispiel 19.10 und 19.20 erhalten wir aus der Linearität des Integrals

$$\int_0^b 3x^2 + 5\,dx = 3 \cdot \int_0^b x^2\,dx + \int_0^b 5\,dx = b^3 + 5b.$$

H) Rechenregeln für Integrale – Additivität

Bemerkung 19.23 (Aneinanderhängen von Zerlegungen).

Ist $Z' = (x_0, \ldots, x_n)$ eine Zerlegung von $[a, c]$ und $Z'' = (y_0, \ldots, y_m)$ eine Zerlegung von $[c, b]$, so ist $Z' * Z'' = (x_0, \ldots, x_n, y_1, \ldots, y_m)$ eine Zerlegung von $[a, b]$ und sie entsteht durch aneinanderhängen der beiden Zerlegungen. Ist $\alpha = (\alpha_1, \ldots, \alpha_n)$ ein Tupel von Zwischenpunkten von Z' und $\beta = (\beta_1, \ldots, \beta_m)$ ein Tupel von Zwischenpunkten von Z'', so definieren wir $\alpha \sqcup \beta = (\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m)$ und erhalten damit ein Tupel von Zwischenpunkten von $Z' * Z''$.

Außerdem gelten offenbar

$$\begin{aligned}\mathrm{OS}(f, Z' * Z'') &= \mathrm{OS}(f, Z') + \mathrm{OS}(f, Z''),\\ \mathrm{US}(f, Z' * Z'') &= \mathrm{US}(f, Z') + \mathrm{US}(f, Z''),\\ \mathrm{ZS}(f, Z' * Z'', \alpha \sqcup \beta) &= \mathrm{ZS}(f, Z', \alpha) + \mathrm{ZS}(f, Z'', \beta).\end{aligned}$$

Proposition 19.24 (Additivität des Integrals).
Es sei $f : [a, b] \longrightarrow \mathbb{R}$ beschränkt, $a < b$ und $c \in (a, b)$.
Genau dann ist f integrierbar auf $[a, b]$, wenn f integrierbar auf $[a, c]$ und auf $[c, b]$ ist. Zudem gilt dann

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

Beweis: Es sei $\varepsilon > 0$ gegeben.

Ist f integrierbar auf $[a, c]$ und auf $[c, b]$, so gibt es wegen des Riemannschen Integrabilitätskriteriums 19.11 Zerlegungen $Z' = (x_0, \ldots, x_n)$ von $[a, c]$ und $Z'' = (y_0, \ldots, y_m)$ von $[c, b]$, so daß

$$\mathrm{OS}(f, Z') - \mathrm{US}(f, Z') < \frac{\varepsilon}{2}$$

und

$$\mathrm{OS}(f, Z'') - \mathrm{US}(f, Z'') < \frac{\varepsilon}{2}.$$

Dann ist aber $Z = Z' * Z'' = (x_0, \ldots, x_n, y_1, \ldots, y_m)$ eine Zerlegung von $[a, b]$ und

$$\begin{aligned}\mathrm{OS}(f, Z) - \mathrm{US}(f, Z) =& \big(\mathrm{OS}(f, Z') + \mathrm{OS}(f, Z'')\big) - \big(\mathrm{US}(f, Z') + \mathrm{US}(f, Z'')\big)\\ =& \big(\mathrm{OS}(f, Z') - \mathrm{US}(f, Z')\big) + \big(\mathrm{OS}(f, Z'') - \mathrm{US}(f, Z'')\big)\\ <& \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

Das Riemannsche Integrabilitätskriterium 19.11 impliziert also, daß f auf $[a, b]$ integrierbar ist.

Ist umgekehrt f auf $[a, b]$ integrierbar, so gibt es wegen des Riemannschen Integrabilitätskriteriums eine Zerlegung $Z = (x_1, \ldots, x_n)$ von $[a, b]$ mit

$$\mathrm{OS}(f, Z) - \mathrm{US}(f, Z) < \varepsilon.$$

Nach eventueller Verfeinerung können wir ohne Einschränkung annehmen, daß $c = x_j \in \mathrm{supp}(Z)$ ein Stützpunkt von Z ist. Dann ist $Z' := (x_0, \ldots, x_j)$ eine Zerlegung von $[a, c]$ und $Z'' := (x_j, \ldots, x_n)$ eine Zerlegung von $[c, b]$. Außerdem gilt $Z = Z' * Z''$ und

$$\begin{aligned}&\big(\mathrm{OS}(f, Z') - \mathrm{US}(f, Z')\big) + \big(\mathrm{OS}(f, Z'') - \mathrm{US}(f, Z'')\big)\\ &= \big(\mathrm{OS}(f, Z') + \mathrm{OS}(f, Z'')\big) - \big(\mathrm{US}(f, Z') + \mathrm{US}(f, Z'')\big)\\ &= \mathrm{OS}(f, Z) - \mathrm{US}(f, Z) < \varepsilon.\end{aligned}$$

Mithin gilt auch

$$\mathrm{OS}(f, Z') - \mathrm{US}(f, Z') < \varepsilon \quad \text{und} \quad \mathrm{OS}(f, Z'') - \mathrm{US}(f, Z'') < \varepsilon,$$

so daß aus dem Riemannschen Integrabilitätskriterium 19.11 wieder folgt, daß f auf $[a, c]$ und auf $[c, b]$ integrierbar ist.

Wir wählen nun zwei Folgen $(Z^{n\prime}, \alpha^n)_{n\in\mathbb{N}}$ und $(Z^{n\prime\prime}, \beta^n)_{n\in\mathbb{N}}$ von Zerlegungen von $[a, c]$ bzw. von $[c, b]$ mit Zwischenpunkten, so daß $\lim\limits_{n\to\infty} l(Z^{n\prime}) = \lim\limits_{n\to\infty} l(Z^{n\prime\prime}) = 0$. Wie oben können wir die Zerlegungen $Z^{n\prime}$ und $Z^{n\prime\prime}$ zu einer Zerlegung $Z^n := Z^{n\prime} * Z^{n\prime\prime}$ von $[a, b]$ zusammenfügen und ebenfalls die Zwischenpunkte α^n und β^n zu Zwischenpunkten $\gamma^n := \alpha^n \sqcup \beta^n$ von Z^n. Dann gilt $l(Z^n) = \max\{l(Z^{n\prime}), l(Z^{n\prime\prime})\} \longrightarrow 0$, und somit folgt aus dem Folgenkriterium für Integrierbarkeit 19.19

$$\int_a^b f(x)\,dx \longleftarrow \mathrm{ZS}(f, Z^n, \gamma^n) = \mathrm{ZS}(f, Z^{n\prime}, \alpha^n) + \mathrm{ZS}(f, Z^{n\prime\prime}, \beta^n) \longrightarrow \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

□

Beispiel 19.25.
Aus Proposition 19.24 und Beispiel 19.20 erhalten wir für $0 < a < b$

$$\int_a^b x^2\,dx = \int_0^b x^2\,dx - \int_0^a x^2\,dx = \frac{b^3}{3} - \frac{a^3}{3}.$$

I) Rechenregeln für Integrale – Dreiecksungleichung

Die Dreiecksungleichung für Summen liefert mit Induktion, daß

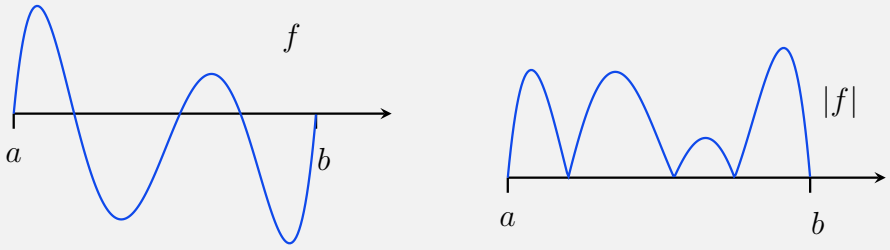
$$|a_1 + \ldots + a_n| \le |a_1| + \ldots + |a_n|$$

gilt. Integrale sind verallgemeinerte Summen, und die Dreiecksungleichung nimmt dann die folgende Gestalt an.

Proposition 19.26 (Dreiecksungleichung für Integrale).
Ist $f : [a, b] \longrightarrow \mathbb{R}$ integrierbar auf $[a, b]$, $a < b$, so ist $|f|$ integrierbar auf $[a, b]$, und es gilt

$$\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f(x)|\,dx.$$

Wir nennen das Integral über $|f|$ auch den *Flächeninhalt*, den der Graph von f auf dem Intervall $[a, b]$ mit der x-Achse einschließt.

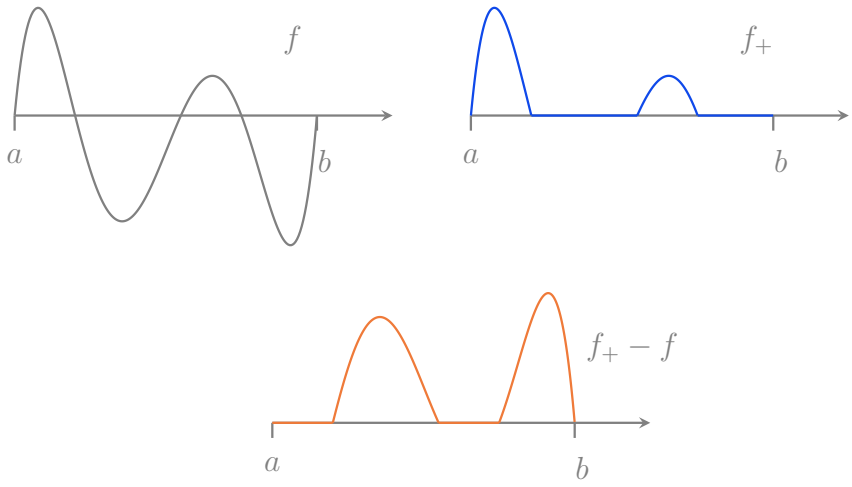


Beweis: Wir betrachten die Funktion

$$f_+ : [a,b] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } f(x) \geq 0, \\ 0, & \text{falls } f(x) < 0. \end{cases}$$

Dann gilt

$$|f| = 2 \cdot f_+ - f.$$



Wir wollen nun mit Hilfe des Riemannschen Integrabilitätskriteriums zeigen, daß f_+ auf $[a,b]$ integrierbar ist. Sei $\varepsilon > 0$ gegeben. Da f auf $[a,b]$ integrierbar ist, gibt es eine Zerlegung $Z = (x_0, \ldots, x_n)$ von $[a,b]$, so daß

$$\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon.$$

Wir behaupten, daß für jede Teilmenge $I \subseteq [a,b]$ die Ungleichung

(53) $\sup\{f_+(x) \mid x \in I\} - \inf\{f_+(x) \mid x \in I\} \leq \sup\{f(x) \mid x \in I\} - \inf\{f(x) \mid x \in I\},$

gilt. Dazu betrachten wir verschiedene Fälle.

1. **Fall:** $f(x) < 0$ **für alle** $x \in I$**:** Dann gilt $f_+ \equiv 0$ auf I, so daß die linke Seite in (53) Null ist. Zugleich gilt

   $$\inf\{f(x) \mid x \in I\} \leq \sup\{f(x) \mid x \in I\} \leq 0,$$

   so daß die rechte Seite von (53) nicht-negativ ist. In diesem Fall gilt (53).
2. **Fall:** $\exists\, y, z \in I$ **mit** $f(y) < 0 \leq f(z)$**:** Also $\sup\{f_+(x) |\ x \in I\} = \sup\{f(x) |\ x \in I\}$ und $\inf\{f_+(x) |\ x \in I\} = 0 > \inf\{f(x) \mid x \in I\}$. Damit gilt die Ungleichung (53).
3. **Fall:** $f(x) \geq 0$ **für alle** $x \in I$**:** Dann ist $f = f_+$ auf I und (53) gilt.

Damit haben wir gezeigt, daß (53) stets erfüllt ist. Für die Differenz der Ober- und Untersumme von f_+ ergibt sich mit $I_i := [x_{i-1}, x_i]$ dann

$$\begin{aligned}\mathrm{OS}(f_+, Z) - \mathrm{US}(f_+, Z) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \big(\sup\{f_+(x) \mid x \in I_i\} - \inf\{f_+(x) \mid x \in I_i\}\big) \\ &\leq \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \big(\sup\{f(x) \mid x \in I_i\} - \inf\{f(x) \mid x \in I_i\}\big) \\ &= \mathrm{OS}(f, Z) - \mathrm{US}(f, Z) < \varepsilon.\end{aligned}$$

Mit Hilfe des Riemannschen Integrabilitätskriteriums 19.11 folgt dann, daß f_+ auf $[a, b]$ integrierbar ist. Aus der Linearität des Integrals 19.21 folgt dann, daß auch

$$|f| = 2 \cdot f_+ - f$$

auf $[a, b]$ integrierbar ist.

Für eine Zerlegung $Z = (x_0, \ldots, x_n)$ mit Zwischenpunkten $\alpha = (\alpha_1, \ldots, \alpha_n)$ gilt

$$|\,\mathrm{ZS}(f, Z, \alpha)| = \left|\sum_{i=1}^{n} (x_i - x_{i-1}) \cdot f(\alpha_i)\right| \leq \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot |f(\alpha_i)| = \mathrm{ZS}(|f|, Z, \alpha).$$

Sei nun $(Z^n, \alpha^n)_{n\in\mathbb{N}}$ eine Folge von Zerlegungen von $[a, b]$ und Zwischenpunkten mit $\lim\limits_{n\to\infty} l(Z^n) = 0$, dann folgt

$$\left|\int_a^b f(x)\,dx\right| \longleftarrow |\,\mathrm{ZS}(f, Z^n, \alpha^n)| \leq \mathrm{ZS}(|f|, Z^n, \alpha^n) \longrightarrow \int_a^b |f(x)|\,dx.$$

Die Ungleichung bleibt für die Grenzwerte erhalten. □

Bemerkung 19.27.
Es sei $f : [a, b] \longrightarrow \mathbb{R}$ integrierbar auf $[a, b]$. Wenden wir Proposition 19.24 zweimal an, so sehen wir, daß f auf jedem Teilintervall $[c, d]$ von $[a, b]$ mit $c < d$ ebenfalls integrierbar ist. Wir definieren nun

$$\int_c^c f(x)\,dx := 0$$

und

$$\int_d^c f(x)\,dx := -\int_c^d f(x)\,dx.$$

Damit müssen die Integrationsgrenzen also nicht mehr verschieden sein, und die untere Integrationsgrenze muß auch nicht mehr die kleinere sein. Die Linearität und Additivität des Integrals verallgemeinern sich dann in naheliegender Weise.

Aufgaben

Aufgabe 19.28.
Wir nennen eine Funktion $f : [a, b] \longrightarrow \mathbb{R}$ *stückweise stetig*, wenn es eine Zerlegung $Z = (x_0, \ldots, x_n)$ von $[a, b]$ gibt, so daß die Funktionen $f_i : [x_{i-1}, x_i] \longrightarrow \mathbb{R} : x \mapsto f(x)$ für $i = 1, \ldots, n$ auf (x_{i-1}, x_i) stetig sind und so daß die Grenzwerte $\lim\limits_{x \to x_{i-1}} f_i(x)$ und $\lim\limits_{x \to x_i} f_i(x)$ in $\mathbb{R}$ existieren.
Zeige, eine stückweise stetige Funktion $f : [a, b] \longrightarrow \mathbb{R}$ ist integrierbar auf $[a, b]$.

Aufgabe 19.29.
Bestimme die folgenden Integrale.

a. $\int_0^{\frac{\pi}{3}} \frac{3}{\cos^2(x)} dx$.

b. $\int_0^1 \frac{1}{1+x} dx$.

c. $\int_{-1}^2 \left(8 \cdot (x-2)^3 + \frac{1}{\sqrt{x+2}}\right) dx$.

d. $\int_0^{\frac{\pi}{4}} \sin(x)\cos(x) dx$.

e. $\int_0^{\frac{1}{4}} x^2 e^{4x} dx$.

Aufgabe 19.30.
Betrachte für $n \in \mathbb{N}$ die Zerlegung $Z^n = (0, \frac{1}{2^n}, \frac{2}{2^n}, \ldots, \frac{2^n-1}{2^n}, 1)$ des Intervalls $[0, 1]$ mit den Zwischenpunkten $\alpha^n = (\frac{1}{2^n}, \frac{2}{2^n}, \ldots, \frac{2^n-1}{2^n}, 1)$. Zeige die folgenden Aussagen.

a. $\mathrm{ZS}(\exp, Z^n, \alpha^n) = (e-1) \cdot e^y \cdot \frac{1}{\frac{e^y-1}{y}}$ für $y = \frac{1}{2^n}$.

b. $\lim\limits_{y \to 0} \frac{e^y-1}{y} = 1$.

c. Berechne $\int_0^1 e^x dx$ mit Hilfe der Zwischensumme aus Aufgabenteil a..

Aufgabe 19.31.
Sei $f : [a, b] \longrightarrow \mathbb{R}$ eine stetige Funktion mit $f(x) \geq 0$ für alle $x \in [a, b]$. Desweiteren existiere ein $c \in [a, b]$ mit $f(c) > 0$. Zeige, daß $\int_a^b f(x)dx > 0$ ist.

Aufgabe 19.32.
Es sei $f : [0, 1] \longrightarrow \mathbb{R}_{\geq 0}$ eine Funktion, so daß es für jedes $\varepsilon > 0$ nur endlich viele Werte $x \in [0, 1]$ gibt mit $f(x) > \varepsilon$. Zeige, f ist integrierbar auf $[0, 1]$ mit $\int_0^1 f(x)\, dx = 0$.

§ 20 Hauptsatz der Differential- und Integralrechnung mit Anwendungen

A) Hauptsatz der Differential- und Integralrechnung

Definition 20.1 (Stammfunktion).
Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \longrightarrow \mathbb{R}$ eine Funktion.
Eine differenzierbare Funktion $F : I \longrightarrow \mathbb{R}$ mit $F' = f$ heißt *Stammfunktion* von f.

Proposition 20.2 (Stammfunktionen sind eindeutig bis auf eine Konstante.).
Sei I ein Intervall, $f : I \longrightarrow \mathbb{R}$ und $F, G : I \longrightarrow \mathbb{R}$ zwei Stammfunktionen von f.
Dann gibt es ein $c \in \mathbb{R}$, so daß $F(x) = G(x) + c$ für alle $x \in I$.

Beweis: Wähle einen Punkt $a \in I$ und setze $c := F(a) - G(a)$. Sei nun $a \neq b \in I$ gegeben, so müssen wir

$$F(b) = G(b) + c$$

zeigen. Wir können ohne Einschränkung annehmen, daß $a < b$ gilt. Nach Voraussetzung ist $F - G$ auf dem Intervall $[a, b] \subseteq I$ differenzierbar, also ist $F - G$ dort auch stetig. Wegen

$$(F - G)'(x) = F'(x) - G'(x) = f(x) - f(x) = 0$$

für alle $x \in [a, b]$ folgt aus Proposition 18.12, daß $F - G$ auf $[a, b]$ konstant ist. Es gilt also insbesondere, daß

$$F(b) - G(b) = F(a) - G(a) = c.$$

□

Beispiel 20.3.
Die Funktion $F : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto \frac{x^3}{3}$ ist eine Stammfunktion von $f : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto x^2$, da $F' = f$. Man beachte, daß wir aus Beispiel 19.20 wissen, daß

$$F(y) = \frac{y^3}{3} = \int_0^y f(x)\, dx.$$

Diese Beobachtung werden wir im folgenden Satz verallgemeinern.

Satz 20.4 (Hauptsatz der Differential- und Integralrechnung).
Es sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \longrightarrow \mathbb{R}$ stetig und $a \in I$.
Dann ist die Funktion $F : I \longrightarrow \mathbb{R} : y \mapsto \int_a^y f(x)\, dx$ eine Stammfunktion von f.

Beweis: Sei $c \in I$ gegeben. Wir müssen zeigen, daß F in c differenzierbar ist mit $F'(c) = f(c)$. Sei dazu wiederum $\varepsilon > 0$ gegeben. Dann müssen wir ein $\delta_\varepsilon > 0$ finden, so daß

$$\left| \frac{F(y) - F(c)}{y - c} - f(c) \right| < \varepsilon$$

für alle $y \in I$ mit $0 < |y - c| < \delta_\varepsilon$ gilt.

Da f stetig in c ist, gibt es ein $\delta_\varepsilon > 0$, so daß

$$|f(x) - f(c)| < \frac{\varepsilon}{2} \tag{54}$$

für alle $x \in I$ mit $|x - c| < \delta_\varepsilon$. Sei nun $c \neq y \in I$ mit $|y - c| < \delta_\varepsilon$, dann gilt

$$\begin{aligned}
\left| \frac{F(y) - F(c)}{y - c} - f(c) \right| &= \left| \frac{1}{y - c} \cdot \left(\int_a^y f(x)\,dx - \int_a^c f(x)\,dx \right) - f(c) \right| \\
&\overset{19.24}{=} \left| \frac{1}{y - c} \cdot \int_c^y f(x)\,dx - f(c) \right| \\
&= \left| \frac{1}{y - c} \cdot \int_c^y f(x)\,dx - \frac{f(c) \cdot (y - c)}{y - c} \right| \\
&= \left| \frac{1}{y - c} \cdot \int_c^y f(x)\,dx - \frac{\int_c^y f(c)\,dx}{y - c} \right| \\
&\overset{19.21}{=} \left| \frac{1}{y - c} \cdot \int_c^y \big(f(x) - f(c) \big)\,dx \right| \\
&\overset{19.26}{\leq} \frac{1}{|y - c|} \cdot \left| \int_c^y \big| f(x) - f(c) \big|\,dx \right| \\
&\overset{(54),19.21}{\leq} \frac{1}{|y - c|} \cdot \left| \int_c^y \frac{\varepsilon}{2}\,dx \right| \\
&= \frac{|y - c|}{|y - c|} \cdot \frac{\varepsilon}{2} < \varepsilon.
\end{aligned}$$

□

Korollar 20.5 (Hauptsatz der Differential- und Integralrechnung).
Es sei $f : [a, b] \longrightarrow \mathbb{R}$ stetig und F sei eine Stammfunktion von f. Dann gilt

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Beweis: Wegen Satz 20.4 und 20.2 gibt eine Konstante $c \in \mathbb{R}$, so daß

$$F(y) = \int_a^y f(x)\,dx + c$$

für alle $y \in I$ gilt. Setzen wir $y = a$ ein, so erhalten wir

$$F(a) = \int_a^a f(x)\,dx + c = 0 + c = c,$$

und mithin gilt insbesondere

$$\int_a^b f(x)\,dx = F(b) - c = F(b) - F(a).$$

□

Bemerkung 20.6.
Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \longrightarrow \mathbb{R}$ stetig.

a. Der Hauptsatz der Differential- und Integralrechnung 20.4 besagt im wesentlichen, daß die Differentiation die Umkehrung der Integration ist.

b. Ist $F : I \longrightarrow \mathbb{R}$ eine Stammfunktion von f und $a, b \in I$, so schreiben wir auch

$$F(x)\big|_a^b := F(b) - F(a) = \int_a^b f(x)\,dx.$$

c. Wir nennen den Ausdruck

$$\int f(x)\,dx$$

ein *unbestimmtes Integral*. Man verwendet ihn gemeinhin, um eine beliebige Stammfunktion F zu bezeichnen, und schreibt dann $F(y) = \int^y f(x)\,dx$.

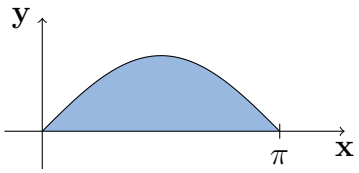
B) Stammfunktionen aus Ableitungen ablesen

Beispiel 20.7 (Einige ausgewählte Stammfunktionen).
In den Abschnitten 17 und 18 haben wir für eine Vielzahl stetig differenzierbarer Abbildungen die Ableitungen kennengelernt. Im Umkehrschluß haben wir damit für die Ableitungsfunktionen auch Stammfunktionen gefunden. Wir wollen für einige wichtige Beispiele von Funktionen f hier die Stammfunktionen F in tabellarischer Form zusammenstellen.

f	$F = \int f(x)\,dx$	f	$F = \int f(x)\,dx$
$\exp$	$\exp$	$\exp_a,\quad a \neq 1$	$\frac{1}{\ln(a)} \cdot \exp_a$
$\cos$	$\sin$	$\sin$	$-\cos$
$x \mapsto \frac{1}{x}$	$\ln$	$x \mapsto x^a,\quad -1 \neq a \in \mathbb{R}$	$x \mapsto \frac{1}{a+1} \cdot x^{a+1}$
$x \mapsto \frac{1}{x^2+1}$	$\arctan$	$\frac{1}{\cos^2}$	$\tan$

Beispiel 20.8 (Flächeninhalt eines Sinusbogens).
Wir können den Flächeninhalt unter einem der Bögen der Sinusfunktion berechnen als

$$\int_0^\pi \sin(x)\,dx = -\cos(x)\big|_0^\pi = -\cos(\pi) + \cos(0) = 1 + 1 = 2.$$



C) Der Mittelwertsatz der Integralrechnung

Korollar 20.9 (Mittelwertsatz der Integralrechnung).
Es sei $f : [a, b] \longrightarrow \mathbb{R}$ stetig, $a < b$. Dann gibt es ein $c \in (a, b)$ mit

$$\int_a^b f(x)\,dx = f(c) \cdot (b - a).$$

Beweis: Die Funktion $F : [a, b] \longrightarrow \mathbb{R} : y \mapsto \int_a^y f(x)\,dx$ ist nach dem Hauptsatz der Differential- und Integralrechnung 20.4 eine Stammfunktion von f und damit differenzierbar. Aus dem Mittelwertsatz der Differentialrechnung erhalten wir deshalb ein $c \in (a, b)$ mit

$$\int_a^b f(x)\,dx = F(b) - F(a) = F'(c) \cdot (b - a) = f(c) \cdot (b - a).$$

□

Wir geben jetzt noch einen alternativen Beweis des Mittelwertsatzes der Integralrechnung, der den Zwischenwertsatz verwendet sowie die Tatsache, daß stetige Funktionen auf abgeschlossenen Intervallen ihr Maximum und Minimum haben. Er hat den Vorteil, daß er sich direkt ins Mehrdimensionale verallgemeinern läßt.

Alternativer Beweis des Mittelwertsatzes 20.9: Da f stetig auf $[a, b]$ ist, nimmt f auf $[a, b]$ sein Minimum und sein Maximum an (siehe Satz 14.16), d.h. es gibt $y, z \in [a, b]$ mit

$$f(y) \leq f(x) \leq f(z)$$

für alle $x \in [a, b]$. Für die Zerlegung $Z = (a, b)$ gilt dann

$$(b - a) \cdot f(y) = \mathrm{US}(f, Z) \leq \int_a^b f(x)\,dx \leq \mathrm{OS}(f, Z) = (b - a) \cdot f(z),$$

und damit

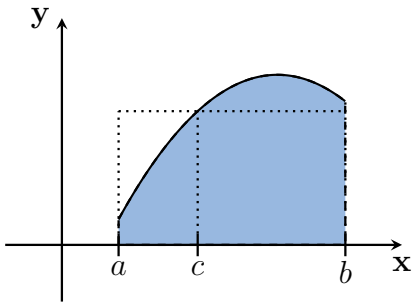
$$f(y) \leq \frac{1}{b - a} \cdot \int_a^b f(x)\,dx \leq f(z).$$

Nach dem Zwischenwertsatz 14.12 nimmt f jeden Wert zwischen $f(y)$ und $f(z)$ für ein geeignetes c zwischen y und z an. Insbesondere gibt es also ein $c \in (a, b)$ mit

$$f(c) = \frac{1}{b - a} \cdot \int_a^b f(x)\,dx.$$

Daraus folgt die Behauptung. □

Bemerkung 20.10 (Geometrische Interpretation des Mittelwertsatzes).
Der Mittelwertsatz besagt, daß das Rechteck mit den Seitenlängen $b-a$ und $f(c)$ den gleichen Flächeninhalt hat, wie die Fläche, die der Graph von f mit der x-Achse einschließt.



D) Partielle Integration

Satz 20.11 (Partielle Integration).
Sind $u, v : [a,b] \longrightarrow \mathbb{R}$ stetig differenzierbar, dann gilt

$$\int_a^b u(x)\cdot v'(x)\,dx = u(x)\cdot v(x)\Big|_a^b - \int_a^b u'(x)\cdot v(x)\,dx.$$

Beweis: Aufgrund der Produktregel gilt $(u\cdot v)'(x) = u(x)\cdot v'(x) + u'(x)\cdot v(x)$ für $x\in[a,b]$, und mithin ist $u\cdot v$ eine Stammfunktion von $x\mapsto u(x)\cdot v'(x)+u'(x)\cdot v(x)$. Da letztere Funktion stetig ist auf $[a,b]$ folgt aus dem Hauptsatz der Differential- und Integralrechnung 20.4 und wegen der Linearität des Integrals 19.21

$$\int_a^b u(x)\cdot v'(x)\,dx + \int_a^b u'(x)\cdot v(x)\,dx = \int_a^b u(x)\cdot v'(x)+u'(x)\cdot v(x)\,dx = u(x)\cdot v(x)\Big|_a^b.$$

Damit ist die Aussage bewiesen. □

Bemerkung 20.12 (Partielle Integration als Umkehrung der Produktregel).
Die partielle Integration ist die Umkehrung der Produktregel. Man wendet sie an, wenn man hofft, das Integral über $u'\cdot v$ leichter berechnen zu können als das über $u\cdot v'$. Auch mit partieller Integration kann man Stammfunktionen berechnen, indem man b durch die Variable y ersetzt und a ignoriert.

Beispiel 20.13 (Stammfunktion von $\cos^2$).
Wir wollen eine Stammfunktion von $\cos^2$ mit Hilfe partieller Integration berechnen. Dazu betrachten wir $u(x)=\cos(x)$ und $v'(x)=\cos(x)$. Dann ist $v(x)=\sin(x)$, und es

gilt

$$\begin{aligned}\int^y \cos^2(x)\,dx &= \int^y u(x)\cdot v'(x)\,dx = u(x)\cdot v(x)\Big|^y - \int^y u'(x)\cdot v(x)\,dx\\ &= \cos(x)\cdot\sin(x)\Big|^y - \int^y -\sin^2(x)\,dx\\ &= \cos(x)\cdot\sin(x)\Big|^y - \int^y \cos^2(x) - 1\,dx\\ &= \cos(x)\cdot\sin(x)\Big|^y - \int^y \cos^2(x)\,dx + \int^y 1\,dx\\ &= \cos(x)\cdot\sin(x)\Big|^y - \int^y \cos^2(x)\,dx + x\Big|^y.\end{aligned}$$

Addieren wir auf beiden Seiten $\int^y \cos^2(x)\,dx$ und teilen durch 2, so erhalten wir

$$\int^y \cos^2(x)\,dx = \frac{1}{2}\cdot\big(y+\cos(y)\cdot\sin(y)\big).$$

E) Der Satz von Taylor

Aus dem Hauptsatz der Differential- und Integralrechnung und unter Anwendung der Methode der partiellen Integration ergibt sich eine Integralform für das Restglied im Satz von Taylor.

Korollar 20.14 (Satz von Taylor – Restglied in Integralform).
Sei $f : I \longrightarrow \mathbb{R}$ $n+1$-fach stetig differenzierbar auf dem Intervall I und $x, a \in I$. Dann gilt

$$f(x) - T^n_{f,a}(x) = \int_a^x \frac{f^{(n+1)}(y)}{n!}\cdot(x-y)^n\,dy.$$

Beweis: Wir führen den Beweis durch Induktion nach n. Für $n = 0$ gilt

$$f(x) - T^0_{f,a}(x) = f(x) - f(a) \overset{20.5}{=} \int_a^x f'(y)\,dy = \int_a^x \frac{f'(y)}{0!}\cdot(x-y)^0\,dy$$

nach dem Hauptsatz der Differential- und Integralrechnung 20.4.

Nun setzen wir voraus, daß $n \geq 1$ ist und daß die Aussage für $n-1$ bereits gezeigt ist und wir wollen sie für n zeigen. Aufgrund der Induktionsvoraussetzung gilt dann

$$f(x) - T^{n-1}_{f,a}(x) = \int_a^x \frac{f^{(n)}(y)}{(n-1)!}\cdot(x-y)^{n-1}\,dy.$$

Wir setzen nun $u(y) := \frac{f^{(n)}(y)}{(n-1)!}$ und $v'(y) := (x-y)^{n-1}$ und wenden partielle Integration an. Die Stammfunktion von v' ist durch $v(y) = \frac{-(x-y)^n}{n}$ gegeben, so daß wir

$$\begin{aligned} f(x) - T_{f,a}^{n-1}(x) &= \int_a^x \frac{f^{(n)}(y)}{(n-1)!} \cdot (x-y)^{n-1}\, dy = \int_a^x u(y) \cdot v'(y)\, dy \\ &= u(y) \cdot v(y)\Big|_a^x - \int_a^x u'(y) \cdot v(y)\, dy \\ &= \frac{f^{(n)}(y)}{(n-1)!} \cdot \frac{-(x-y)^n}{n}\Big|_a^x - \int_a^x \frac{f^{(n+1)}(y)}{(n-1)!} \cdot \frac{-(x-y)^n}{n}\, dy \\ &= \frac{f^{(n)}(a)}{n!} \cdot (x-a)^n + \int_a^x \frac{f^{(n+1)}(y)}{n!} \cdot (x-y)^n\, dy. \end{aligned}$$

Bringen wir den Summanden $\frac{f^{(n)}(a)}{n!} \cdot (x-a)^n$ auf die linke Seite, so erhalten wir

$$f(x) - T_{f,a}^n(x) = f(x) - T_{f,a}^{n-1}(x) - \frac{f^{(n)}(a)}{n!} \cdot (x-a)^n = \int_a^x \frac{f^{(n+1)}(y)}{n!} \cdot (x-y)^n\, dy,$$

und damit folgt die Behauptung mittels des Prinzips der Induktion. □

F) Die Substitutionsregel

Satz 20.15 (Substitutionsregel).
Es sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \longrightarrow \mathbb{R}$ stetig und $\varphi : [a,b] \longrightarrow \mathbb{R}$ stetig differenzierbar mit $\mathrm{Im}(\varphi) \subseteq I$. Dann gilt

$$\int_{\varphi(a)}^{\varphi(b)} f(x)\, dx = \int_a^b f\big(\varphi(x)\big) \cdot \varphi'(x)\, dx.$$

Beweis: Da φ stetig ist, nimmt φ sein Minimum und sein Maximum an, d.h. es gibt $y, z \in [a,b]$ mit $\varphi(y) \leq \varphi(x) \leq \varphi(z)$ für alle $x \in [a,b]$, und mithin ist

$$\mathrm{Im}(\varphi) = \big[\varphi(y), \varphi(z)\big] \subseteq I$$

ein Intervall. Als stetige Funktion besitzt f nach dem Hauptsatz der Differential- und Integralrechnung 20.4 auf diesem Intervall eine Stammfunktion F. Aus der Kettenregel 17.16 folgt dann

$$(F \circ \varphi)'(x) = F'\big(\varphi(x)\big) \cdot \varphi'(x) = f\big(\varphi(x)\big) \cdot \varphi'(x),$$

so daß $F \circ \varphi$ eine Stammfunktion von $x \mapsto f\big(\varphi(x)\big) \cdot \varphi'(x)$ auf $[a,b]$ ist. Dann können wir Korollar 20.5 anwenden und erhalten

$$\int_a^b f\big(\varphi(x)\big) \cdot \varphi'(x)\, dx \overset{20.5}{=} (F \circ \varphi)(b) - (F \circ \varphi)(a) = F\big(\varphi(b)\big) - F\big(\varphi(a)\big) \overset{20.5}{=} \int_{\varphi(a)}^{\varphi(b)} f(x)\, dx.$$

□

Bemerkung 20.16 (Die Substitutionsregel als Umkehrung der Kettenregel).

a. Die Substitutionsregel ist die Umkehrung der Kettenregel.

b. Es ist üblich, bei der Formel für die Substitutionsregel auf der linken Seite statt der Variablen x die Variable z zu verwenden, so daß die Formel folgende Gestalt hat:

$$\int_{\varphi(a)}^{\varphi(b)} f(z)\,dz = \int_a^b f\big(\varphi(x)\big)\cdot\varphi'(x)\,dx.$$

Man sagt dann, daß man $\varphi(x)$ durch z substituiert oder umgekehrt, je nachdem ob man die linke durch die rechte Seite ausrechnen will oder umgekehrt. Man schreibt $z=\varphi(x)$.

Diese Schreibweise kann man nutzen, um sich für die Substitution eine Eselsbrücke zu bauen. In Anlehnung an die Schreibweise $\varphi'=\frac{\partial\varphi}{\partial x}$ kann man mit $z=\varphi(x)$ dann auch

$$\varphi'(x)\,dx=\frac{dz}{dx}\,dx=dz$$

schreiben. Damit wird aus der Substitutionsformel ohne Integralgrenzen dann

$$\int f(\varphi(x))\cdot\varphi'(x)\,dx=\int f(z)\cdot\frac{dz}{dx}\,dx=\int f(z)\,dz.$$

c. Man kann mit Hilfe der Substitutionsregel auch Stammfunktionen ausrechnen, indem man die Integrationsgrenze b durch die Variable y ersetzt und a ignoriert.

Beispiel 20.17 (Stammfunktion von $x\mapsto x\cdot\exp(x^2)$).
Wir wollen das Integral $\int_a^b x\cdot\exp\big(x^2\big)\,dx$ für $a,b\in\mathbb{R}$ berechnen. Dazu substituieren wir $z=x^2$, d.h. wir betrachten $\varphi:\mathbb{R}\longrightarrow\mathbb{R}:x\mapsto x^2$, $\varphi'(x)=2x$ und $f:\mathbb{R}\longrightarrow\mathbb{R}:z\mapsto\frac{\exp(z)}{2}$. Da zudem f eine Stammfunktion von f ist, folgt damit

$$\begin{aligned}\int_a^b x\cdot\exp\big(x^2\big)\,dx&=\int_a^b f\big(\varphi(x)\big)\cdot\varphi'(x)\,dx=\int_{\varphi(a)}^{\varphi(b)} f(z)\,dz\\&=\int_{a^2}^{b^2}\frac{\exp(z)}{2}\,dz=\frac{\exp\big(b^2\big)}{2}-\frac{\exp\big(a^2\big)}{2}.\end{aligned}$$

Beispiel 20.18 (Stammfunktion von $\tan$).
Wir wollen eine Stammfunktion für den Tangens auf dem Intervall $\big(-\frac{\pi}{2},\frac{\pi}{2}\big)$ bestimmen. Dazu substituieren wir $z=\cos(x)$, d.h. $\varphi(x)=\cos(x)$, $\varphi'(x)=-\sin(x)$ und $f(z)=-\frac{1}{z}$. Dann erhalten wir

$$\begin{aligned}\int^y\tan(x)\,dx&=\int^y-\frac{1}{\cos(x)}\cdot\big(-\sin(x)\big)\,dx=\int^y f\big(\varphi(x)\big)\cdot\varphi'(x)\,dx\\&=\int^{\cos(y)} f(z)\,dz=\int^{\cos(y)}-\frac{1}{z}\,dz\\&=-\ln(z)\Big|^{\cos(y)}=-\ln\big(\cos(y)\big).\end{aligned}$$

Also ist $-\ln\circ\cos$ eine Stammfunktion von $\tan$ auf $\big(-\frac{\pi}{2},\frac{\pi}{2}\big)$.

Beispiel 20.19 (Stammfunktion von $x \mapsto \sqrt{1-x^2}$).
Wir wollen mit Hilfe von Substitution eine Stammfunktion für die stetige Funktion

$$f : [-1,1] \longrightarrow \mathbb{R} : z \mapsto \sqrt{1-z^2}$$

bestimmen. Dazu substituieren wir $z = \sin(x)$, d.h. $\varphi(x) = \sin(x)$, $\varphi'(x) = \cos(x)$ und $b = \arcsin(y)$. Dann definiert

$$\begin{aligned}
F(y) = \int^{y} \sqrt{1-z^2}\,dz &= \int^{\varphi(b)} f(z)\,dz = \int^{b} f\big(\varphi(x)\big)\cdot\varphi'(x)\,dx \\
&= \int^{\arcsin(y)} \sqrt{1-\sin^2(x)}\cdot\cos(x)\,dx \\
&= \int^{\arcsin(y)} \sqrt{\cos^2(x)}\cdot\cos(x)\,dx \\
&= \int^{\arcsin(y)} \cos^2(x)\,dx \\
&\overset{20.13}{=} \frac{1}{2}\cdot\big(x+\cos(x)\cdot\sin(x)\big)\Big|^{\arcsin(y)} \\
&= \frac{1}{2}\cdot\big(x+\sqrt{\cos^2(x)}\cdot\sin(x)\big)\Big|^{\arcsin(y)} \\
&= \frac{1}{2}\cdot\big(x+\sqrt{1-\sin^2(x)}\cdot\sin(x)\big)\Big|^{\arcsin(y)} \\
&= \frac{\arcsin(y)+y\cdot\sqrt{1-y^2}}{2}.
\end{aligned}$$

eine Stammfunktion von f auf $[-1,1]$.

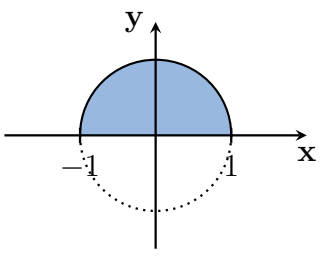
Beispiel 20.20 (Flächeninhalt eines Kreises).
Wir wollen nun den Flächeninhalt eines Kreises berechnen. Die obere Hälfte des Einheitskreises mit dem Ursprung als Mittelpunkt

$$K_1(0) := \big\{(x,y) \in \mathbb{R}^2 \;\big|\; x^2+y^2 = 1\big\}$$

ist die Fläche, die der Graph der Funktion

$$f : [-1,1] \longrightarrow \mathbb{R} : x \mapsto \sqrt{1-x^2}$$

mit der x-Achse einschließt.



Mithin ist das Integral

$$\int_{-1}^{1} \sqrt{1-x^2}\,dx \overset{20.19}{=} \left.\frac{\arcsin(x) + x\cdot\sqrt{1-x^2}}{2}\right|_{-1}^{1}$$
$$= \frac{\arcsin(1)}{2} - \frac{\arcsin(-1)}{2} = \frac{\pi}{4} - \left(-\frac{\pi}{4}\right) = \frac{\pi}{2}$$

die Hälfte des Flächeninhaltes des Einheitskreises.

Der Kreis mit Radius r und dem Ursprung als Mittelpunkt ist

$$K_r(0) := \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; x^2+y^2 = r^2\right\},$$

und sein Flächeninhalt ist entsprechend das Integral

$$2\cdot\int_{-r}^{r} \sqrt{r^2-x^2}\,dx.$$

Um dieses zu berechnen, substituieren wir $z = \frac{x}{r}$, d.h. $\varphi(x) = \frac{x}{r}$, $\varphi'(x) = \frac{1}{r}$, $\varphi(r) = 1$, $\varphi(-r) = -1$ und $f(z) = \sqrt{1-z^2}$. Wir erhalten dann

$$2\cdot\int_{-r}^{r} \sqrt{r^2-x^2}\,dx = 2\cdot r^2\cdot\int_{-r}^{r} \sqrt{1-\frac{x^2}{r^2}}\cdot\frac{1}{r}\,dx$$
$$= 2\cdot r^2\cdot\int_{-1}^{1} \sqrt{1-z^2}\,dz = 2\cdot\frac{\pi}{2}\cdot r^2 = \pi\cdot r^2$$

als Flächeninhalt von $K_r(0)$.

G) Partialbruchzerlegung

Bemerkung 20.21 (Partialbruchzerlegung).
Jede rationale Funktion $r = \frac{f}{g}$ läßt sich schreiben als

$$r = \frac{f}{g} = h + \frac{p}{q}, \tag{55}$$

wobei h, p und q Polynome sind mit $\deg(p) < \deg(q)$. Dies folgt sofort mittels einer einfachen Polynomdivision, wie sie in der Vorlesung Algebraische Strukturen eingeführt wird.

Nicht offensichtlich ist, daß sich der Bruch $\frac{p}{q}$ als Summe von Ausdrücken der Form

$$\frac{A}{(t-a)^k} \quad \text{und} \quad \frac{Bt+C}{(t^2+bt+c)^k} \tag{56}$$

für geeignete $A, B, C, a, b, c \in \mathbb{R}$ mit $4b - a^2 > 0$ schreiben läßt. Genauer kann man zeigen, daß wenn $(t-a)^n$ bzw. $(t^2+bt+c)^n$ die höchste Potenz von $t-a$ bzw. von t^2+bt+c ist, die das Polynom q teilt, so kommen in der Summe Ausdrücke der Form (56) für $k = 1, \ldots, n$ vor. Eine solche Darstellung nennt man dann die *Partialbruchzerlegung*

von $\frac{p}{q}$. Wir werden unten in einem Beispiel sehen, wie man diese unter Umständen finden kann.

Eine rationale Funktion wie in (55) ist stetig auf ihrem Definitionsbereich, der eine Vereinigung von endlich vielen Intervallen ist. Mithin ist sie auf allen abgeschlossenen Teilintervallen ihres Definitionsbereiches integrierbar. Um nun eine Stammfunktion von r zu bestimmen, reicht es im wesentlichen, Funktionen der Form (56) zu integrieren. Dies ist mit Hilfe unserer bisherigen Methoden vergleichsweise einfach, sei in der allgemeinen Form aber dem Leser als Übungsaufgabe überlassen.

Beispiel 20.22 (Integration mit Partialbruchzerlegung).

a. Wir wollen das folgende Integral berechnen:

$$\int_1^2 \frac{3x^5+3x^4+6x^2+x-2}{x^3+x^2}\,dx.$$

Polynomdivision von $3t^5+3t^4+6t^2+t-2$ durch t^3+t^2 liefert

$$r=\frac{3t^5+3t^4+6t^2+t-2}{t^3+t^2}=3t^2+\frac{6t^2+t-2}{t^3+t^2}=f+\frac{p}{q}.$$

Dabei faktorisiert q als

$$q=t^2\cdot(t+1).$$

Das Prinzip der Partialbruchzerlegung läßt uns nun nach Zahlen $A,B,C\in\mathbb{R}$ suchen, so daß

$$\frac{6t^2+t-2}{t^3+t^2}=\frac{A}{t}+\frac{B}{t^2}+\frac{C}{t+1}$$

gilt. Bringen wir die rechte Seite auf den Hauptnenner, so erhalten wir

$$\begin{aligned}\frac{6t^2+t-2}{t^3+t^2}&=\frac{A\cdot t\cdot(t+1)}{t^3+t^2}+\frac{B\cdot(t+1)}{t^3+t^2}+\frac{C\cdot t^2}{t^3+t^2}\\&=\frac{(A+C)\cdot t^2+(A+B)\cdot t+B}{t^3+t^2}.\end{aligned}$$

Ein Koeffizientenvergleich der Polynome im Zähler der beiden Seiten führt zu den Gleichungen:

$$A+C=6,\quad A+B=1\quad\text{und}\quad B=-2.$$

Daraus lesen wir ohne Schwierigkeiten

$$A=3,\quad B=-2\quad\text{und}\quad C=3$$

ab. Für unser Integral ergibt sich daraus

$$\begin{aligned}\int_1^2 \frac{3x^5+3x^4+6x^2+x-2}{x^3+x^2}\,dx &= \int_1^2 3x^2\,dx + \int_1^2 \frac{3}{x}\,dx + \int_1^2 \frac{-2}{x^2}\,dx + \int_1^2 \frac{3}{x+1}\,dx \\ &= x^3 + 3\cdot\ln(x) + \frac{2}{x} + 3\cdot\ln(x+1)\Big|_1^2 \\ &= 8 + 3\cdot\ln(2) + 1 + 3\cdot\ln(3) - 1 - 3\cdot\ln(1) - 2 - 3\cdot\ln(2) \\ &= 3\cdot\ln(3) + 6.\end{aligned}$$

b. Wie sieht eine Stammfunktion zu folgender Funktion aus

$$x \mapsto \frac{2x^3+x^2+1}{x^2\cdot(1+x^2)^2}?$$

Das Prinzip der Partialbruchzerlegung läßt uns nach reellen Zahlen $A, B, C, D, E, F \in \mathbb{R}$ suchen, so daß

$$\frac{2t^3+t^2+1}{t^2\cdot(1+t^2)^2} = \frac{A}{t} + \frac{B}{t^2} + \frac{Ct+D}{1+t^2} + \frac{Et+F}{(1+t^2)^2}$$

gilt. Bringen wir die rechte Seite auf den Hauptnenner, so erhalten wir

$$\begin{aligned}&\frac{2t^3+t^2+1}{t^2\cdot(1+t^2)^2} = \\ &= \frac{(A+C)\cdot t^5 + (B+D)\cdot t^4 + (2A+C+E)\cdot t^3 + (2B+D+F)\cdot t^2 + A\cdot t + B}{t^2\cdot(1+t^2)^2}.\end{aligned}$$

Durch Koeffizientenvergleich erhalten wir dann

$$A = 0,\ B = 1,\ C = 0,\ D = -1,\ E = 2 \quad \text{und} \quad F = 0,$$

d.h.

$$\frac{2t^3+t^2+1}{t^2\cdot(1+t^2)^2} = \frac{1}{t^2} - \frac{1}{1+t^2} + \frac{2t}{(1+t^2)^2}.$$

Damit erhalten wir als Stammfunktion

$$\begin{aligned}F(y) &= \int^y \frac{2x^3+x^2+1}{x^2\cdot(1+x^2)^2}\,dx = \int^y \frac{1}{x^2}\,dx - \int^y \frac{1}{1+x^2}\,dx + \int^y \frac{2x}{(1+x^2)^2}\,dx \\ &= -\frac{1}{y} - \arctan y - \frac{1}{1+y^2},\end{aligned}$$

wobei wir zur Berechnung des letzten Integrals die Substitution $z = 1 + x^2$ vornehmen und so

$$\int^y \frac{2x}{(1+x^2)^2}\,dx = \int^{1+y^2} \frac{1}{z^2}\,dz = -\frac{1}{z}\Big|^{1+y^2} = -\frac{1}{1+y^2}$$

erhalten.

H) Vertauschbarkeit von Grenzwert und Integration

Satz 20.23 (Vertauschbarkeit von Grenzwert und Integration).
Die Folge $(f_n)_{n\in\mathbb{N}}$ stetiger Funktionen $f_n : [a,b] \longrightarrow \mathbb{R}$ konvergiere auf $[a,b]$, $a<b$, gleichmäßig gegen $f : [a,b] \longrightarrow \mathbb{R}$.
Dann ist auch die Grenzfunktion f stetig auf $[a,b]$ und für alle $y \in [a,b]$ gilt

$$\int_a^y f(x)\,dx = \lim_{n\to\infty} \int_a^y f_n(x)\,dx. \tag{57}$$

D.h. der Grenzwert der Stammfunktionen der f_n ist eine Stammfunktion von f.

Beweis: Nach Satz 15.6 ist f als gleichmäßiger Grenzwert stetiger Funktionen stetig. Da stetige Funktionen nach Satz 19.13 integrierbar sind, existieren die Integrale in (57). Es bleibt, für $y \in [a,b]$ zu zeigen, daß

$$\lim_{n\to\infty} \int_a^y f_n(x)\,dx = \int_a^y f(x)\,dx.$$

Sei dazu $\varepsilon > 0$ gegeben. Da $(f_n)_{n\in\mathbb{N}}$ auf $[a,b]$ gleichmäßig gegen f konvergiert, gibt es ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $n \geq n_\varepsilon$ und für alle $x \in [a,b]$

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2\cdot(b-a)}. \tag{58}$$

Dann gilt für $n \geq n_\varepsilon$ auch

$$\left|\int_a^y f_n(x)\,dx - \int_a^y f(x)\,dx\right| \overset{19.26}{\leq} \int_a^y |f_n(x) - f(x)|\,dx \overset{19.21,(58)}{\leq} \int_a^y \frac{\varepsilon}{2\cdot(b-a)}\,dx \leq \frac{\varepsilon}{2} < \varepsilon.$$

Damit ist die Behauptung gezeigt. □

Bemerkung 20.24.

a. Wie in Bemerkung 18.19 wollen wir wieder darauf hinweisen, daß wir in Satz 20.23 gezeigt haben, daß zwei Grenzwertprozesse vertauschen. Es gilt nämlich

$$\lim_{n\to\infty}\lim_{m\to\infty} \mathrm{ZS}(f_n, Z_m, \alpha^m) = \lim_{n\to\infty}\int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx = \lim_{m\to\infty}\lim_{n\to\infty} \mathrm{ZS}(f_n, Z_m, \alpha^m).$$

b. Ersetzen wir in Satz 20.23 die Voraussetzung *stetig* durch *integrierbar*, so wird auch die Grenzfunktion f nur noch *integrierbar* sein, es gilt aber nach wie vor

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx.$$

c. Wir können mit Satz 20.23 und dem Hauptsatz der Differential- und Integralrechnung 20.4 einen wesentlich kürzeren Beweis der Vertauschbarkeit von Grenzwert und Ableitung 18.18 geben:

Die Funktionen f_n in Satz 18.18 sind jeweils Stammfunktion von f_n', und da die f_n' stetig sind, folgt aus dem Hauptsatz der Differential- und Integralrechnung 20.5

$$f_n(y) - f_n(a) = \int_a^y f_n'(x)\,dx.$$

Bilden wir auf beiden Seiten den Grenzwert, so folgt mit Satz 20.23

$$f(y) - f(a) \longleftarrow f_n(y) - f_n(a) = \int_a^y f_n'(x)\,dx \longrightarrow \int_a^y g(x)\,dx.$$

Also ist

$$f(y) = f(a) + \int_a^y g(x)\,dx$$

für $y \in [a,b]$, so daß f nach dem Hauptsatz der Differential- und Integralrechnung 20.4 differenzierbar ist auf $[a,b]$ mit

$$f'(y) = 0 + g(y) = g(y)$$

für alle $y \in [a,b]$.

I) Integration von Potenzreihen

Korollar 20.25 (Integration von Potenzreihen).
Es sei $\sum_{n=0}^{\infty} a_n \cdot t^n$ ein Potenzreihe über $\mathbb{R}$ mit Konvergenzradius $r > 0$.
Dann ist die Funktion $f : (-r,r) \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} a_n \cdot x^n$ auf jedem Intervall $[a,b] \subset (-r,r)$ integrierbar und

$$F : (-r,r) \longrightarrow \mathbb{R} : y \mapsto \sum_{n=0}^{\infty} \frac{a_n}{n+1} \cdot y^{n+1}$$

ist eine Stammfunktion von f. Sie entsteht durch gliedweises Integrieren.

Beweis: Die Folge stetiger Funktionen $(f_n)_{n\in\mathbb{N}}$ mit

$$f_n : (-r,r) \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{n} a_k \cdot x^k$$

konvergiert auf $[a,b]$ gleichmäßig gegen f (siehe Satz 15.4). Also ist f nach Satz 15.6 stetig auf $[a,b]$. Für $y \in (-r,r)$ gilt nach Satz 20.23 zudem

$$\int_0^y f(x)\,dx = \lim_{n\to\infty} \int_0^y f_n(x)\,dx \overset{19.21}{=} \lim_{n\to\infty} \sum_{k=0}^{n} a_k \cdot \int_0^y x^k\,dx \overset{20.7}{=} \lim_{n\to\infty} \sum_{k=0}^{n} \frac{a_k}{k+1} \cdot y^{k+1}.$$

Damit ist die Behauptung gezeigt. □

Man kann durch gliedweises Integrieren Stammfunktionen berechnen oder auch Potenzreihendarstellungen von Funktionen aus der Potenzreihendarstellung ihrer Ableitungen herleiten, wie wir im folgenden Beispiel sehen werden.

Beispiel 20.26 (Reihenentwicklung durch gliedweise Integration).

a. Die Potenzreihe zur Exponentialfunktion ist

$$\sum_{n=0}^{\infty} \frac{t^n}{n!}.$$

Durch gliedweises Integrieren erhalten wir die Potenzreihe

$$\sum_{n=0}^{\infty} \frac{1}{n+1} \cdot \frac{t^{n+1}}{n!} = \sum_{n=1}^{\infty} \frac{t^n}{n!},$$

und diese definiert eine Stammfunktion von exp. Sie unterscheidet sich von der bereits bekannten Stammfunktion exp von exp um die Konstante 1.

b. Die Funktion $f : (-1, \infty) \longrightarrow \mathbb{R} : x \mapsto \ln(1+x)$ ist differenzierbar mit Ableitung

$$f' : (-1, \infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{1+x}.$$

Mit Hilfe der geometrischen Reihe sehen wir, daß

$$f'(x) = \frac{1}{1-(-x)} = \sum_{n=0}^{\infty} (-x)^n = \sum_{n=0}^{\infty} (-1)^n \cdot x^n$$

für alle $x \in (-1, 1)$ gilt. f' ist dort also durch die Potenzreihe $\sum_{n=0}^{\infty} (-1)^n \cdot t^n$ gegeben. Durch gliedweises Integrieren finden wir eine Potenzreihendarstellung einer Stammfunktion von f':

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} \cdot t^{n+1} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \cdot t^n$$

auf dem Intervall $(-1, 1)$. Da auch f auf $(-1, 1)$ eine Stammfunktion von f' ist und zwei Stammfunktionen sich nur um eine Konstante c unterscheiden, werten wir f und diese Potenzreihe in $a = 0$ aus, um c zu bestimmen. Wir erhalten damit

$$c = f(0) - \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \cdot 0^n = \ln(1) - 0 = 0.$$

Wir haben also eine Potenzreihendarstellung für f auf $(-1, 1)$ gefunden; für $x \in (-1, 1)$ gilt

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \cdot x^n.$$

Daraus ergibt sich dann die Potenzreihendarstellung für den natürlichen Logarithmus

$$\ln(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \cdot (x-1)^n$$

für alle $x \in (0, 2)$, d.h. die Taylorreihenentwicklung des natürlichen Logarithmus aus Beispiel 18.37 gilt auf ganz $(0, 2)$, und sie gilt auch im Punkt $a = 2$, wie wir dort bereits gesehen haben.

Aufgaben

Aufgabe 20.27 (Nicht Riemann-integrierbare Funktion mit Stammfunktion).
Zeige, die Funktion

$$F : [0,1] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} x^{\frac{3}{2}} \cdot \sin\left(\frac{1}{x}\right), & \text{wenn } x > 0, \\ 0, & \text{wenn } x = 0 \end{cases}$$

ist Stammfunktion einer nicht Riemann-integrierbaren Funktion.

Aufgabe 20.28.
Sei $f : (-1,1) \longrightarrow \mathbb{R}$, $x \longmapsto \arctan(x)$.

a. Finde die Taylorreihenentwicklung von f mit Hilfe gliedweiser Integration.

b. Bestimme eine Reihendarstellung für $\frac{\pi}{4}$ und π.

Aufgabe 20.29.
Es sei I ein Intervall und $f : I \to \mathbb{R}$ eine stetig differenzierbare Funktion ohne Nullstelle. Zeige, daß $\ln(|f|)$ eine Stammfunktion von $\frac{f'}{f}$ auf I ist.

Aufgabe 20.30.
Berechne mittels Partialbruchzerlegung eine Stammfunktion von

$$x \mapsto \frac{x^3}{x^4 - 4x^3 + 8x^2 - 16x + 16}.$$

Aufgabe 20.31.
Berechne die folgenden bestimmten und unbestimmten Integrale:

a. $\int_{1\frac{1}{4}}^{0} \sqrt{8x+2}\, dx$.

b. $\int_{1}^{2} \frac{x^4+2x}{x^5+5x^2-2}\, dx$.

c. $\int x^2 - \sin(x) + e^{3x}\, dx$.

d. $\int \sin(x) \cdot e^x\, dx$.

e. $\int_{0}^{\frac{\pi}{2}} x^2 \cdot \cos(x)\, dx$.

f. $\int_{2}^{5} \frac{3x}{x^2+1}\, dx$.

g. $\int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \ln\left(\sqrt{\sin(x)}\right) \cdot \cos(x)\, dx$.

h. $\int e^{\sqrt{x}}\, dx$.

Aufgabe 20.32.
Ist $f : [0,1] \longrightarrow \mathbb{R}$ stetig, so gibt es ein $c \in (0,1)$ mit $\int_0^1 f(x) \cdot x^2 \, dx = \frac{f(c)}{3}$.

Aufgabe 20.33.
Berechne die folgenden bestimmten und unbestimmten Integrale:

a. $\int_0^{2\pi} \cos(kx) \, dx$ für $k \in \mathbb{N}$.

b. $\int_0^{\frac{\pi}{4}} \sin(x)\cos(x) dx$.

c. $\int \frac{1}{\sin(x)} \, dx$, substituiere $z = \tan\left(\frac{x}{2}\right)$.

Aufgabe 20.34.
Berechne mittels Partialbruchzerlegung eine Stammfunktion von

$$x \mapsto \frac{2x^5 + 7x^4 + 9x^3 + 9x^2 + 6x - 5}{x^4 + 4x^3 + 5x^2 + 4x + 4}.$$

Aufgabe 20.35.
Es sei $f : [a,b] \longrightarrow \mathbb{R}$ eine stetig differenzierbare Funktion und

$$g : \mathbb{R} \longrightarrow \mathbb{R} : y \mapsto \int_a^b f(x) \cdot \sin(y \cdot x) \, dx.$$

Zeige mit Hilfe partieller Integration $\lim\limits_{y \to \infty} g(y) = 0$ und $\lim\limits_{y \to -\infty} g(y) = 0$.

§ 21 Uneigentliche Integrale

A) Uneigentliche Integrale

Definition 21.1 (Uneigentliche Integrale).

a. Es seien $a \in \mathbb{R}$ und $b \in \mathbb{R} \cup \{\infty\}$ mit $a < b$, und $f : [a, b) \longrightarrow \mathbb{R}$ sei auf $[a, y]$ integrierbar für alle $y \in (a, b)$. Falls der Grenzwert

$$\int_a^b f(x)\,dx := \lim_{y \to b} \int_a^y f(x)\,dx \in \mathbb{R} \cup \{\infty, -\infty\}$$

existiert, so nennen wir ihn ein *uneigentliches Integral*. Ist $\int_a^b f(x)\,dx \in \mathbb{R}$, so nennen wir das uneigentliche Integral *konvergent*.

b. Es seien $b \in \mathbb{R}$ und $a \in \mathbb{R} \cup \{-\infty\}$ mit $a < b$, und $f : (a, b] \longrightarrow \mathbb{R}$ sei auf $[y, b]$ integrierbar für alle $y \in (a, b)$. Falls der Grenzwert

$$\int_a^b f(x)\,dx := \lim_{y \to a} \int_y^b f(x)\,dx \in \mathbb{R} \cup \{\infty, -\infty\}$$

existiert, so nennen wir ihn ein *uneigentliches Integral*. Ist $\int_a^b f(x)\,dx \in \mathbb{R}$, so nennen wir das uneigentliche Intgral *konvergent*.

c. Es seien $a \in \mathbb{R} \cup \{-\infty\}$ und $b \in \mathbb{R} \cup \{\infty\}$ mit $a < b$ und für $f : (a, b) \longrightarrow \mathbb{R}$ gebe es ein $c \in (a, b)$, so daß die uneigentlichen Integrale $\int_c^b f(x)\,dx$ und $\int_a^c f(x)\,dx$ konvergent sind, dann nennen wir auch

$$\int_a^b f(x)\,dx := \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

ein *konvergentes uneigentliches Integral*. Aus der Additivität des Integrals folgt, daß die Definition der linken Seite unabhängig von der Wahl von c ist.

Beispiel 21.2.

a. Das folgende uneigentliche Integral ist konvergent mit Grenzwert 1:

$$\int_0^\infty \exp(-x)\,dx = \lim_{y \to \infty} \int_0^y \exp(-x) = \lim_{y \to \infty} -\exp(-x)\Big|_0^y = \lim_{y \to \infty} 1 - \exp(-y) = 1.$$

b. Es sei $a \in \mathbb{R}$ mit $a \neq 1$. Dann gilt für das uneigentliche Integral

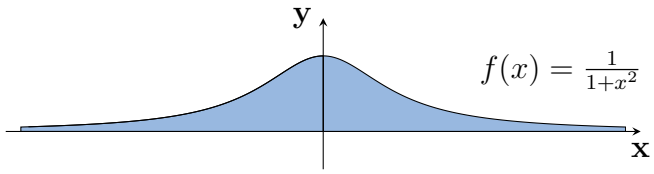
$$\begin{aligned} \int_1^\infty \frac{1}{x^a}\,dx &= \lim_{y \to \infty} \int_1^y x^{-a}\,dx = \lim_{y \to \infty} \frac{x^{-a+1}}{-a+1}\Big|_1^y \\ &= \lim_{y \to \infty} \frac{y^{1-a}}{1-a} - \frac{1}{1-a} = \begin{cases} \infty, & \text{falls } a < 1, \\ \frac{1}{a-1}, & \text{falls } a > 1. \end{cases} \end{aligned}$$

c. Das folgende uneigentliche Integral ist konvergent mit Grenzwert 2:

$$\int_0^1 \frac{1}{\sqrt{1-x}}\,dx = \lim_{y\to 1} -2\cdot\sqrt{1-x}\Big|_0^y = 2 - \lim_{y\to 1} 2\cdot\sqrt{1-y} = 2.$$

d. Der Flächeninhalt unter dem Graphen von $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \frac{1}{1+x^2}$ ist π:

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx &= \int_0^{\infty} \frac{1}{1+x^2}\,dx + \int_{-\infty}^{0} \frac{1}{1+x^2}\,dx \\ &= \lim_{y\to\infty} \arctan(x)\Big|_0^y + \lim_{y\to-\infty} \arctan(x)\Big|_y^0 \\ &= \lim_{y\to\infty} \arctan(y) - \lim_{y\to-\infty} \arctan(y) = \frac{\pi}{2} + \frac{\pi}{2} = \pi.\end{aligned}$$



Bemerkung 21.3.
Die *Linearität* und die *Monotonie* des Integrals (siehe Korollar 19.21) sowie die *Additivität* des Integrals (siehe Proposition 19.24) und die *Dreiecksungleichung* für Integrale (siehe Proposition 19.26) gelten analog auch für uneigentliche Integrale. Der Beweis folgt unmittelbar aus den entsprechenden Aussagen für Integrale zusammen mit den Grenzwertsätzen 13.10.

B) Integralkriterium für Reihen

Lemma 21.4 (Monotoniekriterium für uneigentliche Integrale).
Es sei $a \in \mathbb{R}$ und $b \in \mathbb{R}\cup\{\infty\}$ mit $a < b$, und es sei $f : [a,b) \longrightarrow \mathbb{R}_{\geq 0}$ integrierbar auf $[a,y]$ für alle $y \in (a,b)$. Gibt es ein $s \in \mathbb{R}$ mit $\int_a^y f(x)\,dx < s$ für alle $y \in (a,b)$, so ist $\int_a^b f(x)\,dx$ konvergent.

Beweis: Die Funktion

$$F : [a,b) \longrightarrow \mathbb{R} : y \mapsto \int_a^y f(x)\,dx$$

ist monoton wachsend, da f nur nicht-negative Werte annimmt.

Wir betrachten nun eine monoton wachsende Folge $(a_n)_{n\in\mathbb{N}}$ und die zugehörige Folge $\big(F(a_n)\big)_{n\in\mathbb{N}}$ von Funktionswerten. Diese ist monoton wachsend und beschränkt durch s, mithin ist sie konvergent, d.h. es gibt ein $I \in \mathbb{R}$ mit

$$\int_a^{a_n} f(x)\,dx = F(a_n) \longrightarrow I.$$

Zu gegebenem $\varepsilon > 0$ gibt es also ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $n \geq n_\varepsilon$ gilt

$$I - F(a_n) = |F(a_n) - I| < \varepsilon.$$

Wir wollen nun zeigen, daß $\int_a^b f(x)\,dx$ gegen I konvergiert und unterscheiden dazu die beiden Fälle $b = \infty$ und $b \in \mathbb{R}$.

1. **Fall:** $b = \infty$**:** Für $\varepsilon > 0$ setzen wir nun $s_\varepsilon := a_{n_\varepsilon}$, so daß für $y \in [a, \infty)$ mit $y > a_{n_\varepsilon}$ dann

   $$|F(y) - I| = I - F(y) \leq I - F(a_{n_\varepsilon}) < \varepsilon,$$

   gilt, da F monoton wachsend ist. Mithin gilt

   $$\int_a^\infty f(x)\,dx = \lim_{y\to\infty} F(y) = I.$$

2. **Fall:** $b \in \mathbb{R}$**:** Für $\varepsilon > 0$ setzen wir $\delta_\varepsilon := b - a_{n_\varepsilon} > 0$, so daß für $y \in [a, b)$ mit $b - y = |y - b| < \delta_\varepsilon = b - a_{n_\varepsilon}$ auch $y > a_{n_\varepsilon}$ und damit

   $$|F(y) - I| = I - F(y) \leq I - F(a_{n_\varepsilon}) < \varepsilon,$$

   gilt, da F monoton wachsend ist. Mithin gilt wieder

   $$\int_a^\infty f(x)\,dx = \lim_{y\to\infty} F(y) = I.$$

□

Satz 21.5 (Integralkriterium für Reihen).
Es sei $a \in \mathbb{N}$ und $f : [a, \infty) \longrightarrow \mathbb{R}_{\geq 0}$ sei monoton fallend und auf $[a, c]$ integrierbar für alle $c \in [a, \infty)$. Dann gilt:

$$\int_a^\infty f(x)\,dx \text{ ist konvergent} \quad \Longleftrightarrow \quad \sum_{n=a}^{\infty} f(n) \text{ ist konvergent.}$$

In dieser Situation gilt zudem

$$\sum_{n=a+1}^{\infty} f(n) \;\leq\; \int_a^\infty f(x)\,dx \;\leq\; \sum_{n=a}^{\infty} f(n).$$

Beweis: Für $m \in \mathbb{N}$ mit $m > a$ betrachten wir die Zerlegung

$$Z_m := (a, a+1, a+2, \ldots, m)$$

des Intervalls $[a, m]$. Da f auf $[a, m]$ monoton fallend ist, erhalten wir

$$\sum_{n=a+1}^{m} f(n) = \mathrm{US}(f, Z_m) \leq \int_a^m f(x)\,dx \leq \mathrm{OS}(f, Z_m) = \sum_{n=a}^{m-1} f(n). \tag{59}$$

Ist die Reihe $\sum_{n=a}^{\infty} f(n)$ konvergent und $y \in (a, \infty)$, so wählen wir ein $m \in \mathbb{N}$ mit $y \leq m$ und aus (59) folgt dann

$$\int_a^y f(x)\,dx \leq \int_a^m f(x)\,dx \leq \sum_{n=a}^{m-1} f(n) \leq \sum_{n=a}^{\infty} f(n),$$

so daß das Integral $\int_a^\infty f(x)\,dx$ nach dem Monotoniekriterium 21.4 konvergent ist.

Ist umgekehrt das Integral $\int_a^\infty f(x)\,dx$ konvergent, so ist die Folge der Partialsummen

$$s_m := \sum_{n=a}^{m} f(n) \overset{(59)}{\leq} f(a) + \int_a^m f(x)\,dx \leq f(a) + \int_a^\infty f(x)\,dx$$

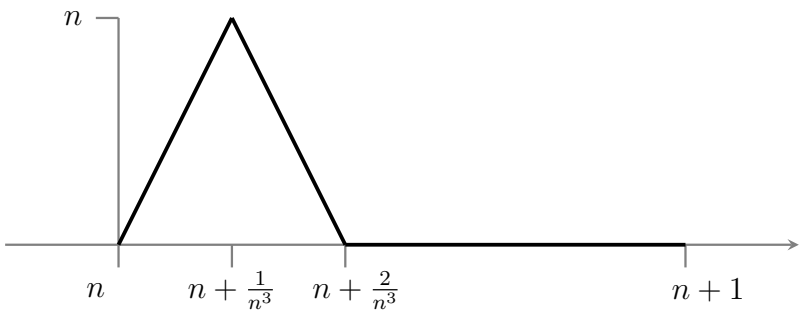
monoton wachsend und beschränkt, mithin konvergent.

Die Abschätzung für die Grenzwerte der Reihen und des Integrals folgt unmittelbar aus (59), indem man m gegen unendlich gehen läßt. □

Beispiel 21.6.
Ist $a \in \mathbb{R}$ mit $a > 1$, so ist die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^a}$ konvergent, nach Satz 21.5 und Teil b. von Beispiel 21.2 .

Bemerkung 21.7.
Wenn $\int_a^\infty f(x)\,dx$ konvergiert, so muß *nicht* $\lim\limits_{x\to\infty} f(x) = 0$ gelten!

Wir betrachten eine Funktion $f : [1, \infty) \longrightarrow \mathbb{R}$, die auf dem Intervall $[n, n+1]$ den folgenden Graphen besitzt:



Dann ist

$$\int_n^{n+1} f(x)\,dx = \frac{1}{2} \cdot \frac{2}{n^3} \cdot n = \frac{1}{n^2}$$

der Flächeninhalt des obigen Dreiecks. Also ist das uneigentliche Integral

$$\int_1^\infty f(x)\,dx = \lim_{n\to\infty} \int_1^{n+1} f(x)\,dx = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

konvergent, aber $\lim\limits_{x\to\infty} f(x) \neq 0$, da

$$f\left(n + \frac{1}{n^3}\right) = n \longrightarrow \infty.$$

Aufgaben

Aufgabe 21.8.
Berechne den Wert des uneigentlichen Integrals $\int_{-\infty}^{0} \frac{e^x}{1+e^{2x}}\,dx$.

Aufgabe 21.9.
Untersuche, für welche $t \in \mathbb{R}$ das uneigentliche Integral $\int_0^\infty x \cdot e^{-tx}\,dx$ konvergiert und bestimme für diese den Wert des Integrals.

Aufgabe 21.10.

a. Zeige, für $y \in (0,\infty)$ ist das uneigentliche Integral
$$\int_0^\infty x^{y-1} \cdot \exp(-x)\,dx$$
konvergent.

b. Die Funktion $\Gamma : (0,\infty) \longrightarrow \mathbb{R} : y \mapsto \int_0^\infty x^{y-1} \cdot \exp(-x)\,dx$ erfüllt die Funktionalgleichung
$$\Gamma(y+1) = y \cdot \Gamma(y)$$
für $y \in (0,\infty)$.

c. Für $n \in \mathbb{N}$ gilt $\Gamma(n+1) = n!$.

Aufgabe 21.11.

a. Zeige, für jedes $y \in (0,\infty)$ ist die Funktion
$$g : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x} - \frac{1}{x+y}$$
streng monoton fallend.

b. Zeige mit Hilfe des Integralkriteriums für Reihen, daß die Reihe
$$\sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+y}\right)$$
für jedes $y \in (0,\infty)$ konvergiert.

c. Zeige, die Folge $(f_n)_{n\in\mathbb{N}}$ mit
$$f_n : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x^2} + \sum_{k=1}^{n} \frac{1}{(x+k)^2}$$

konvergiert auf jedem Intervall $[\delta, \infty)$ mit $\delta > 0$ gleichmäßig gegen

$$f : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x^2} + \sum_{k=1}^{\infty} \frac{1}{(x+k)^2}.$$

d. Zeige, die Funktion

$$F : (0,\infty) \longrightarrow \mathbb{R} : y \mapsto -\frac{1}{y} + \sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+y} \right)$$

ist eine Stammfunktion von f auf $(0,\infty)$ und erfüllt für alle $y \in (0,\infty)$ die Bedingungen $F'(y) > 0$ und $F(y+1) - F(y) = \frac{1}{y}$.

e. Zeige, ist $G : (0,\infty) \longrightarrow \mathbb{R}$ eine differenzierbare Funktion mit $G'(y) \geq 0$ und $G(y+1) - G(y) = \frac{1}{y}$ für alle $y \in (0,\infty)$, so unterscheiden sich F und G nur um eine Konstante.

f. Zeige, die differenzierbare Funktion

$$H : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \big(\ln\big(\Gamma(x)\big)\big)' = \frac{\Gamma'(x)}{\Gamma(x)}$$

erfüllt die Bedingungen $H'(y) \geq 0$ und $H(y+1) - H(y) = \frac{1}{y}$ für alle $y \in (0,\infty)$. Man nennt H auch die *Digammafunktion*.

Hinweis zu Teil e.: setze $F_n(y) = -\frac{1}{y} + \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+y} \right)$ und zeige $G(y) - F_n(y) = G(y+n+1) - \sum_{k=1}^{n} \frac{1}{k}$; nutze dies, um $(G-F)'(y) \geq 0$ für alle y zu zeigen; zeige ferner, daß $G - F$ 1-periodisch ist und folgere dann, daß $G - F$ konstant ist.

Aufgabe 21.12.
Zeige, $\int_0^1 x^t \cdot \ln(x)^n \, dx = \frac{(-1)^n \cdot n!}{(1+t)^{n+1}}$ für alle $t > -1$ und für alle $n \in \mathbb{N}$.

Aufgabe 21.13 (Cauchy-Kriterium für uneigentliche Integrale).
Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a,b) \longrightarrow \mathbb{R}$ sei auf $[a,y]$ integrierbar für alle $y \in (a,b)$. Zeige, genau dann existiert das uneigentliche Integral $\int_a^b f(x)\, dx$, wenn

$$\forall\, \varepsilon > 0 \; \exists\, c_\varepsilon \in [a,b) \;:\; \forall\, c_\varepsilon < s < t < b \;\text{ gilt }\; \left| \int_s^t f(x)\, dx \right| < \varepsilon.$$

Aufgabe 21.14 (Uneigentliche Integrale beschränkter Funktionen).
Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a,b) \longrightarrow \mathbb{R}$ sei eine beschränkte Funktion, die für alle $y \in (a,b)$ auf $[a,y)$ integrierbar ist. Zeige, dann ist das uneigentliche Integral $\int_a^b f(x)\, dx$ konvergent.

Teil B

Mathematik für Informatik 2

Kapitel III

Algebraische Strukturen

Wir werden in diesem Abschnitt einige algebraische Strukturen betrachten, die wir teilweise bereits im Grundlagenkapitel kennengelernt haben. Vor allem die Struktur der Gruppe werden uns nun genauer anschauen.

§ 22 Gruppen und Homomorphismen

Im ersten Teilabschnitt wiederholen wir noch mal Begriffe und Aussagen, die wir bereits aus Abschnitt 7 kennen.

A) Gruppen

Die grundlegendste und wichtigste algebraische Struktur ist die *Gruppe*.

Definition 22.1.

a. Eine *Gruppe* ist ein Paar $(G, \cdot)$ bestehend aus einer *nicht-leeren* Menge G und einer zweistelligen Operation "$\cdot$", d. h. einer Abbildung

$$\cdot : G \times G \to G : (g, h) \mapsto g \cdot h,$$

so daß die folgenden *Gruppenaxiome* gelten:

G1: $(g \cdot h) \cdot k = g \cdot (h \cdot k) \quad \forall\, g, h, k \in G,$ ("Assoziativgesetz")

G2: $\exists\, e \in G \;:\; \forall\, g \in G \;:\; e \cdot g = g,$ ("Existenz eines Neutralen")

G3: $\forall\, g \in G\, \exists\, g' \in G \;:\; g' \cdot g = e.$ ("Existenz von Inversen")

Ein Element mit der Eigenschaft von e nennt man ein *neutrales Element* der Gruppe G. Ein Element mit der Eigenschaft von g' nennt man ein *Inverses zu* g.

b. Eine Gruppe $(G, \cdot)$ heißt *abelsch* oder *kommutativ*, wenn $(G, \cdot)$ zudem noch dem folgenden Axiom genügt:
G4: $g \cdot h = h \cdot g \quad \forall\, g, h \in G$ ("Kommutativgesetz")

c. Eine Gruppe $(G, \cdot)$ heißt *endlich*, falls $|G| < \infty$, und sonst *unendlich*. $|G|$ heißt die *Ordnung* von G.

Bemerkung 22.2.
Für manche Anwendungen ist der Begriff der *Gruppe* stärker als wünschenswert, da mehr Axiome gefordert werden, als die betrachteten Strukturen hergeben. Man kann den Begriff in zweifacher Weise abschwächen. Sei dazu wieder eine Menge G zusammen mit einer zweistelligen Operation "$\cdot$" auf G gegeben.

a. Erfüllt das Paar $(G, \cdot)$ nur das Axiom G1 so nennt man $(G, \cdot)$ eine *Halbgruppe*.

b. Wir nennen $(G, \cdot)$ ein *Monoid*, falls nur die Axiome G1 und G2' gelten:
G2': $\exists\, e \in G \;:\; \forall\, g \in G \;:\; e \cdot g = g \cdot e = g.$ ("Existenz eines Neutralen")

Man beachte, daß bei Monoiden für die Neutralen eine stärkere Bedingung gefordert wird, als bei Gruppen. Inwieweit sie wirklich stärker ist, werden wir in Lemma 22.4 sehen. Die Begriffe *abelsch*, *endlich*, *unendlich* und *Ordnung* führt man für Halbgruppen und Monoide in der gleichen Weise ein wie für Gruppen. In dieser Vorlesung werden wir uns aber nicht weiter mit speziellen Eigenschaften von Halbgruppen oder Monoiden beschäftigen. Wir erwähnen die Begriffe nur der Vollständigkeit halber. □

Bevor es sinnvoll ist, sich mit Eigenschaften einer neuen Struktur wie den eben eingeführten Gruppen zu beschäftigen, sollte man gute Beispiele kennen. Schließlich möchte man keine Aussagen über die leere Menge oder auch nur eine vollkommen uninteressante Struktur treffen.

Beispiel 22.3.

a. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ und $(\mathbb{C}, +)$ mit der üblichen Addition als Gruppenoperation sind abelsche Gruppen. Die Zahl Null erfüllt jeweils die Rolle eines neutralen Elements, und zu einer Zahl g existiert mit $-g$ ein inverses Element.

b. $(\mathbb{Q} \setminus \{0\}, \cdot)$, $(\mathbb{R} \setminus \{0\}, \cdot)$ und $(\mathbb{C} \setminus \{0\}, \cdot)$ mit der üblichen Multiplikation als Gruppenoperation sind ebenfalls abelsche Gruppen. Die Zahl 1 ist jeweils ein neutrales Element, und zur Zahl g existiert als inverses Element die Zahl $\frac{1}{g}$.

c. $(\mathbb{Z} \setminus \{0\}, \cdot)$ ist hingegen nur ein (abelsches) Monoid mit der Zahl Eins als neutralem Element. Das Axiom G3 ist nicht erfüllt, da nur die Zahlen $g = 1$ und $g = -1$ in $\mathbb{Z} \setminus \{0\}$ ein Inverses $\frac{1}{g}$ besitzen.

d. $(\mathbb{N}, +)$ ist ebenfalls nur ein (abelsches) Monoid mit der Zahl Null als neutralem Element, da zu $g > 0$ kein Inverses $-g$ in $\mathbb{N}$ existiert.

e. Die einfachste Gruppe ist die *einelementige Gruppe* $G = \{e\}$, deren Gruppenoperation durch $e \cdot e = e$ definiert ist.

f. Ist M eine nicht-leere Menge, so ist die Menge

$$\mathrm{Sym}(M) := \{f : M \longrightarrow M \mid f \text{ ist bijektiv}\}$$

mit der Komposition von Abbildungen als Gruppenoperation eine Gruppe. Die Assoziativität von "$\circ$" haben wir in Proposition 3.11 gezeigt, die Identität ist das neutrale Element und in Satz 3.12 haben wir gezeigt, daß jede bijektive Abbildung ein Inverses besitzt. Wir nennen $(\mathrm{Sym}(M), \circ)$ die *symmetrische Gruppe* auf M. Enthält M mehr als zwei Elemente, so ist $\mathrm{Sym}(M)$ nicht abelsch.

Man beachte, daß in allen obigen Beispielen ein eindeutig bestimmtes neutrales Element sowie zu jedem g ein eindeutig bestimmtes inverses Element existiert. Dies ist kein Zufall. Die Eindeutigkeit kann aus den Gruppenaxiomen hergeleitet werden.

Lemma 22.4.
Es sei $(G, \cdot)$ eine Gruppe.

a. Das neutrale Element $e \in G$ ist eindeutig bestimmt und hat zusätzlich die Eigenschaft:
$$g \cdot e = g \quad \forall\, g \in G.$$

b. Sei $g \in G$. Das inverse Element g' zu g ist eindeutig bestimmt und hat zusätzlich die Eigenschaft:
$$g \cdot g' = e.$$

Beweis: Da wir für das Paar $(G, \cdot)$ die Axiome G1-G3 aus Definition 22.1 voraussetzen, gibt es ein neutrales Element $e \in G$, und zu beliebigem, aber fest gegebenem $g \in G$ gibt es ein Inverses $g' \in G$.

Wir wollen zunächst zeigen, daß für dieses e und dieses g' die in a. und b. geforderten zusätzlichen Eigenschaften gelten.

Da $(G, \cdot)$ eine Gruppe ist, gibt es ein $g'' \in G$ mit

$$g'' \cdot g' = e. \tag{60}$$

Also folgt:

$$g \cdot g' \overset{G2}{=} e \cdot (g \cdot g') \overset{(60)}{=} (g'' \cdot g') \cdot (g \cdot g') \overset{G1}{=} g'' \cdot (g' \cdot (g \cdot g')) \tag{61}$$
$$\overset{G1}{=} g'' \cdot ((g' \cdot g) \cdot g') \overset{G3}{=} g'' \cdot (e \cdot g') \overset{G2}{=} g'' \cdot g' \overset{(60)}{=} e.$$

Damit ist gezeigt, daß g' die zusätzliche Eigenschaft in b. erfüllt, und wir erhalten:

$$g \cdot e \overset{G3}{=} g \cdot (g' \cdot g) \overset{G1}{=} (g \cdot g') \cdot g \overset{(61)}{=} e \cdot g \overset{G2}{=} g. \tag{62}$$

Nun war aber g ein beliebiges Element in G, so daß damit die zusätzliche Eigenschaft von e in a. gezeigt ist.

Sei nun $\tilde{e} \in G$ irgendein Element mit der Eigenschaft des Neutralen, d.h. so daß

$$\tilde{e} \cdot h = h \tag{63}$$

für alle $h \in G$. Wir müssen zeigen, daß $e = \tilde{e}$ gilt. Da wir bereits wissen, daß e die zusätzliche Eigenschaft in a. erfüllt, können wir diese, d.h. (62), mit $\tilde{e}$ in der Rolle von g anwenden, und anschließend (63) mit e in der Rolle von h:

$$\tilde{e} \overset{(62)}{=} \tilde{e} \cdot e \overset{(63)}{=} e.$$

Schließlich müssen wir noch zeigen, wenn $\tilde{g}' \in G$ ein weiteres inverses Element zu g ist, d.h. wenn

$$\tilde{g}' \cdot g = e \tag{64}$$

gilt, dann ist schon $g' = \tilde{g}'$. Wenden wir das bislang Gezeigte an, so gilt:

$$\tilde{g}' \overset{(62)}{=} \tilde{g}' \cdot e \overset{(61)}{=} \tilde{g}' \cdot (g \cdot g') \overset{G1}{=} (\tilde{g}' \cdot g) \cdot g' \overset{(64)}{=} e \cdot g' \overset{G2}{=} g'.$$

Damit sind alle Aussagen des Lemmas bewiesen. □

Notation 22.5.
Statt $(G, \cdot)$ schreiben wir häufig nur G, sofern keine Unklarheiten über die Operation bestehen. Außerdem schreiben wir, für $g, h \in G$, statt $g \cdot h$ oft verkürzt gh. Das neutrale Element bezeichnen wir auch mit 1 statt mit e, oder mit 1_G bzw. e_G, wenn wir hervorheben wollen, in welcher Gruppe es das Neutrale ist. Und das zu $g \in G$ existierende, eindeutig bestimmte inverse Element wird mit g^{-1} bezeichnet, oder mit g_G^{-1}, wenn wir wieder hervorheben wollen, in welcher Gruppe es das Inverse zu g ist. Zudem definieren wir Potenzen von Gruppenelementen ganz allgemein durch $g^0 := e$ und dann für $n > 0$ rekursiv $g^n := g \cdot g^{n-1}$ und schließlich $g^{-n} := (g^{-1})^n$.

Ist die Gruppe abelsch, so bezeichnet man die Operation oft mit $+$ anstatt mit $\cdot$. In diesem Fall verwenden wir die Bezeichnung 0 (bzw. 0_G) für das neutrale Element und $-g$ (bzw. $-g_G$) für das zu $g \in G$ eindeutig bestimmte Inverse. In diesem Fall schreiben wir $n \cdot g$ statt g^n. □.

Lemma 22.6.
Sei $(G,\cdot)$ eine Gruppe, $g,h,a,b \in G$, $m,n \in \mathbb{Z}$. Dann gelten:

a. $\left(g^{-1}\right)^{-1} = g$ und $(gh)^{-1} = h^{-1}g^{-1}$.

b. $g^n \cdot g^m = g^{n+m}$ und $(g^m)^n = g^{m\cdot n}$.

c. In G gelten die *Kürzungsregeln*:

(i) $ga = gb \;\Rightarrow\; a = b$, und

(ii) $ag = bg \;\Rightarrow\; a = b$.

Beweis: a. Um die erste Gleichheit zu zeigen, reicht es wegen der Eindeutigkeit des Inversen zu g^{-1} zu zeigen, daß g die Eigenschaft *des* Inversen zu g^{-1} besitzt. Beim Beweis können wir die Gruppenaxiome sowie die in Lemma 22.4 bewiesenen zusätzlichen Eigenschaften des Inversen anwenden:

$$g \cdot g^{-1} \overset{\text{Lem. 22.4b.}}{=} e.$$

Also ist g ein Inverses zu g^{-1}, und damit gilt wie angedeutet wegen der Eindeutigkeit des Inversen zu g^{-1}:

$$\left(g^{-1}\right)^{-1} = g.$$

Analog ist nach Voraussetzung $(gh)^{-1}$ ein Inverses zu gh, und es reicht wegen der Eindeutigkeit des Inversen zu gh zu zeigen, daß $h^{-1}g^{-1}$ ebenfalls die Eigenschaft eines Inversen zu gh hat:

$$\begin{aligned}\left(h^{-1}g^{-1}\right)\cdot(gh) &\overset{G1}{=} h^{-1}\cdot\left(g^{-1}\cdot(gh)\right) \overset{G1}{=} h^{-1}\cdot\left((g^{-1}\cdot g)\cdot h\right)\\ &\overset{G3}{=} h^{-1}\cdot(e\cdot h) \overset{G2}{=} h^{-1}\cdot h \overset{G3}{=} e.\end{aligned}$$

Mithin ist $h^{-1}g^{-1}$ ein Inverses zu gh, und somit

$$(gh)^{-1} = h^{-1}g^{-1}.$$

b. Die notwendigen Fallunterscheidungen überlassen wir dem Leser als Übung.

c. Die erste Kürzungsregel folgt durch Multiplikation mit dem Inversen zu g von links:

$$\begin{aligned}a \overset{G2}{=} e\cdot a \overset{G3}{=} \left(g^{-1}\cdot g\right)\cdot a &\overset{G1}{=} g^{-1}\cdot(g\cdot a)\\ &\overset{Vor.}{=} g^{-1}\cdot(g\cdot b) \overset{G1}{=} \left(g^{-1}\cdot g\right)\cdot b \overset{G3}{=} e\cdot b \overset{G2}{=} b.\end{aligned}$$

Entsprechend folgt die zweite Kürzungsregel durch Multiplikation mit g^{-1} von rechts und unter Berücksichtigung der zusätzlichen Eigenschaft des Inversen aus Lemma 22.4. Die Details überlassen wir dem Leser.

□

B) Untergruppen

Ein wichtiges Prinzip in der Mathematik ist es, zu jeder betrachteten Struktur auch ihre *Unter-* oder *Teilstrukturen* zu betrachten. Für eine Menge sind das einfach ihre Teilmengen, für eine Gruppe werden es ihre *Untergruppen* sein – das sind Teilmengen, die die zusätzliche Struktur *respektieren* und mittels dieser selbst wieder eine Gruppe sind.

Definition 22.7.
Sei $(G, \cdot)$ eine Gruppe. Eine nicht-leere Teilmenge $U \subseteq G$ heißt *Untergruppe* von G, wenn für alle $u, v \in U$ stets $uv \in U$ und $u_G^{-1} \in U$ gilt.
Wir schreiben einfach $U \leq G$, um auszudrücken, dass U eine Untergruppe von G. Man bezeichnet die obigen Eigenschaften als die *Abgeschlossenheit* von U bezüglich der Gruppenoperation und der Inversenbildung.

Bevor wir uns Beispiele von Untergruppen anschauen, wollen wir zeigen, daß jede Untergruppe selbst wieder eine Gruppe ist.

Proposition 22.8 (Untergruppen sind Gruppen.).
Ist $(G, \cdot)$ eine Gruppe und $U \leq G$ eine Untergruppe von G, dann ist $(U, \cdot)$ wieder eine Gruppe. Dabei stimmen das Neutrale in U und G ebenso überein wie das Inverse von $g \in U$ in U und in G.

Beweis: Da $uv \in U$ für alle $u, v \in U$, ist das Bild von $U \times U$ unter der Abbildung "$\cdot$" in der Tat in U enthalten. Es bleibt also, die Axiome G1-G3 nachzuprüfen. Dabei überträgt sich G1 von der größeren Menge G auf die Teilmenge U. Da $U \neq \emptyset$, existiert ein $u \in U$. Nach Voraussetzung gilt dann aber $u_G^{-1} \in U$ und damit $e_G = u_G^{-1} u \in U$. Da aber $e_G u = u$ für alle $u \in U$, ist auch G2 erfüllt und es gilt $e_U = e_G$. Ferner haben wir bereits bemerkt, daß für $u \in U$ auch $u_G^{-1} \in U$, und es gilt

$$u_G^{-1} \cdot u = e_G = e_U.$$

Somit ist auch G3 erfüllt und die Inversen von u in U bzw. in G stimmen überein. □

Wir wollen nun Beispiele von Untergruppen betrachten. Man interessiert sich im Übrigen auch deshalb für die Untergruppen einer Gruppe, weil die Kenntnis dieser wichtige Informationen über die Struktur der Gruppe selbst liefert.

Beispiel 22.9.

a. Ist $(G, \cdot)$ eine Gruppe mit neutralem Element e_G, so sind die beiden Teilmengen $\{e_G\}$ und G von G stets Untergruppen. Man nennt sie deshalb auch die *trivialen*

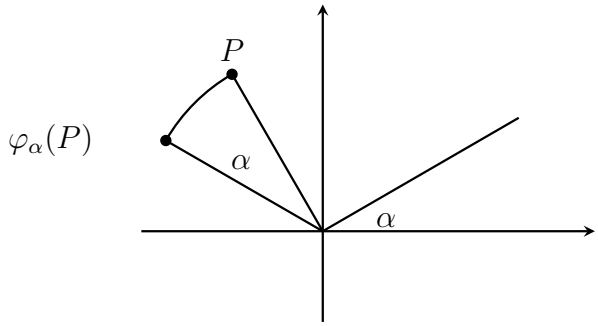
Untergruppen. Sie geben keine zusätzliche Information über die Struktur der Gruppe selbst.

b. $(\{-1,1\},\cdot)$ ist eine Untergruppe von $(\mathbb{Q}\setminus\{0\},\cdot)$, wie unmittelbar aus der Definition folgt.

c. Für $\alpha\in\mathbb{R}$ bezeichnet

$$\varphi_\alpha:\mathbb{R}^2\to\mathbb{R}^2:(x,y)\to\big(\cos(\alpha)\cdot x-\sin(\alpha)\cdot y,\sin(\alpha)\cdot x+\cos(\alpha)\cdot y\big)$$

die Drehung der Ebene $\mathbb{R}^2$ um den Nullpunkt um den Winkel α im Bogenmaß.

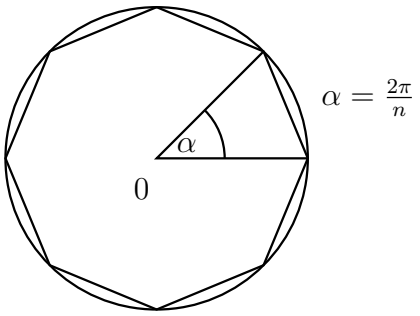


Offensichtlich gilt $\varphi_\alpha\circ\varphi_\beta=\varphi_{\alpha+\beta}$ für $\alpha,\beta\in\mathbb{R}$, und für $\alpha\in\mathbb{R}$ ist somit $\varphi_{-\alpha}=(\varphi_\alpha)^{-1}$, da $\varphi_0=\mathrm{id}_{\mathbb{R}^2}$. Insbesondere ist φ_α also bijektiv für jedes $\alpha\in\mathbb{R}$. Damit folgt, daß

$$\mathrm{SO}(2):=\{\varphi_\alpha:\mathbb{R}^2\to\mathbb{R}^2\mid\alpha\in\mathbb{R}\}$$

eine Untergruppe von $\mathrm{Sym}\big(\mathbb{R}^2\big)$ ist.

d. Sei $E_n\subset\mathbb{R}^2$ das *reguläre n-Eck*.



Wir setzen

$$U:=\big\{\varphi_\alpha\in\mathrm{SO}(2)\;\big|\;\varphi_\alpha(E_n)=E_n\big\}$$

und zeigen, dass $(U,\circ)$ eine Untergruppe von $\big(\mathrm{SO}(2),\circ\big)$ ist:
Für $\varphi_\alpha,\varphi_\beta\in U$ gilt

$$(\varphi_\alpha\circ\varphi_\beta)(E_n)=\varphi_\alpha\big(\varphi_\beta(E_n)\big)=\varphi_\alpha(E_n)=E_n$$

und

$$\varphi_\alpha^{-1}(E_n)=\varphi_\alpha^{-1}\big(\varphi_\alpha(E_n)\big)=\big(\varphi_\alpha^{-1}\circ\varphi_\alpha\big)(E_n)=\mathrm{id}_{\mathbb{R}^2}(E_n)=E_n.$$

Also gilt $\varphi_\alpha\circ\varphi_\beta\in U$ und $\varphi_\alpha^{-1}\in U$, und da $\mathrm{id}_{\mathbb{R}^2}=\varphi_0\in U$, ist $U\neq\emptyset$ und folglich ist U eine Untergruppe von $\mathrm{SO}(2)$.

Man überzeugt sich leicht, daß U aus allen Drehungen φ_α mit $\alpha = k \cdot \frac{2\pi}{n}$, $k = 0, \ldots, n-1$, besteht. Insbesondere gilt also, $|U| = n$.

e. Sei $n \in \mathbb{Z}$. Dann ist $n\mathbb{Z} := \{nz \mid z \in \mathbb{Z}\}$, die Menge aller Vielfachen von n, eine Untergruppe von $(\mathbb{Z}, +)$:
Seien $nz, nz' \in n\mathbb{Z}$, dann gilt $nz+nz' = n(z+z') \in n\mathbb{Z}$ und $-(nz) = n\cdot(-z) \in n\mathbb{Z}$. Da ferner $\emptyset \neq n\mathbb{Z} \subset \mathbb{Z}$, folgt die Behauptung.

f. Für zwei ganze Zahlen $m, n \in \mathbb{Z}$ gilt $m\mathbb{Z} \subseteq n\mathbb{Z}$ genau dann, wenn m ein Vielfaches von n ist.

g. Die Inklusionen $\mathbb{Z} \subset \mathbb{Q}$, $\mathbb{Z} \subset \mathbb{R}$ und $\mathbb{Q} \subset \mathbb{R}$ machen die Teilmenge bezüglich der Addition als Gruppenstruktur jeweils zu einer Untergruppe.

Bemerkung 22.10.
Wie verhalten sich Untergruppen gegenüber Mengenoperationen wie z.B. der Vereinigung?

Betrachten wir die Gruppe $(\mathbb{Z}, +)$ und ihre Untergruppen $2\mathbb{Z} \leq \mathbb{Z}$ sowie $3\mathbb{Z} \leq \mathbb{Z}$. Die Menge $U = 2\mathbb{Z} \cup 3\mathbb{Z}$ ist nicht abgeschlossen bezüglich der Gruppenoperation, denn

$$2 + 3 = 5 \notin 2\mathbb{Z} \cup 3\mathbb{Z},$$

da 5 weder durch 2 noch durch 3 teilbar ist. Mithin ist die Vereinigung von Untergruppen im allgemeinen keine Untergruppe mehr.

Im Gegensatz zur Vereinigung verhalten sich Gruppen bei der Bildung von Schnittmengen gut.

Lemma 22.11 (Durchschnitt von Untergruppen).
Es sei $(G, \cdot)$ eine Gruppe, I eine beliebige Indexmenge und $U_i \leq G$ für $i \in I$. Dann gilt

$$\bigcap_{i \in I} U_i \leq G.$$

Beweis: Wir überlassen den Beweis dem Leser. □

Wir verwenden die gute Schnitteigenschaft der Untergruppen um den Makel der schlechten Vereinigung auszutilgen, und betrachten bei Gruppen statt der Vereinigung von Untergruppen stets das *Erzeugnis* der Vereinigung als kleinste Untergruppe, die die Vereinigung enthält.

Definition 22.12 (Erzeugnis).
Es sei $(G,\cdot)$ eine Gruppe und $M \subseteq G$ eine Teilmenge. Das *Erzeugnis* von M ist die Untergruppe

$$\langle M\rangle = \bigcap_{M\subseteq U\leq G} U,$$

d.h. der Schnitt über alle Untergruppen von G, die M enthalten. Ist $M = \{g_1,\dots,g_n\}$, so schreiben wir wir statt $\langle\{g_1,\dots,g_n\}\rangle$ in aller Regel nur $\langle g_1,\dots,g_n\rangle$.

Eine solche Definition ist nützlich, da sie frei Haus liefert, daß das Erzeugnis eine Untergruppe ist und daß es die *kleinste* Untergruppe ist, die M enthält. Sie ist aber wenig hilfreich, wenn man bei gegebenem M entscheiden soll, welche Elemente letztlich im Erzeugnis liegen. Dies wird durch die folgende Proposition beschrieben, die auch den Begriff *Erzeugnis* rechtfertigt, da die Elemente von $\langle M\rangle$ in der Tat durch die Elemente von M *erzeugt* werden.

Proposition 22.13.
Es sei $(G,\cdot)$ eine Gruppe und $M \subseteq G$ eine Teilmenge. Dann gilt

$$\langle M\rangle = \left\{g_1^{\alpha_1}\cdots g_n^{\alpha_n} \mid n\geq 0, g_1,\dots,g_n\in M, \alpha_1,\dots,\alpha_n\in\mathbb{Z}\right\}.$$

Beweis: Wir geben zunächst der rechten Seite der Gleichung einen Namen,

$$N = \left\{g_1^{\alpha_1}\cdots g_n^{\alpha_n} \mid n\geq 0, g_1,\dots,g_n\in M, \alpha_1,\dots,\alpha_n\in\mathbb{Z}\right\},$$

und zeigen, daß $N \subseteq \langle M\rangle$. Falls $U \leq G$, so daß $M \subseteq U$, dann ist $g_1^{\alpha_1}\cdots g_n^{\alpha_n} \in U$ für alle $g_i \in M$ und $\alpha_1,\dots,\alpha_n \in \mathbb{Z}$. Also ist $N \subseteq U$, und damit $N \subseteq \langle M\rangle$.

Es bleibt zu zeigen, daß $\langle M\rangle \subseteq N$. Dafür reicht es aber zu zeigen, daß $N \leq G$ mit $M \subseteq N$. Da das leere Produkt nach Konvention das neutrale Element e_G ist, ist N nicht leer. Seien nun also $h = g_1^{\alpha_1}\cdots g_n^{\alpha_n}, h' = g_{n+1}^{\alpha_{n+1}}\cdots g_m^{\alpha_m} \in N$ zwei beliebige Elemente in N, dann gilt

$$h\cdot h' = g_1^{\alpha_1}\cdots g_m^{\alpha_m} \in N$$

und

$$h^{-1} = g_n^{-\alpha_n}\cdots g_1^{-\alpha_1} \in N.$$

Also ist $N \leq G$. Da aber $M \subseteq N$ ohnehin gilt, folgt die Behauptung. □

Beispiel 22.14.
Ist $(G,\cdot)$ eine Gruppe und $g \in G$, so folgt aus Proposition 22.13

$$\langle g\rangle = \{g^k \mid k\in\mathbb{Z}\}.$$

Wenden wir dies auf die Gruppe $(\mathbb{Z}, +)$ und eine Zahl $n \in \mathbb{Z}$ an, so ist das Erzeugnis der Menge $M = \{n\}$ die Untergruppe

$$n\mathbb{Z} = \{n \cdot z \mid z \in \mathbb{Z}\} = \langle n \rangle = \langle M \rangle.$$

Man beachte, daß in der additiven Gruppe $(\mathbb{Z}, +)$ aus "g^k" ein "$k \cdot g$" geworden ist.

Definition 22.15.
Eine Gruppe $(G, \cdot)$ heißt *zyklisch*, wenn sie von einem Element erzeugt wird, d.h. wenn es ein $g \in G$ gibt, so daß $G = \langle g \rangle$.

Proposition 22.16.
$U \subseteq \mathbb{Z}$ ist genau dann eine Untergruppe von $(\mathbb{Z}, +)$, wenn es eine ganze Zahl $n \geq 0$ gibt mit $U = n\mathbb{Z} = \langle n \rangle$. Insbesondere ist jede Untergruppe von $(\mathbb{Z}, +)$ zyklisch.

Beweis: Aus Beispiel 22.9 und Beispiel 22.14 wissen wir, daß die Mengen der Form $n\mathbb{Z} = \langle n \rangle$ Untergruppen von $(\mathbb{Z}, +)$ sind.

Sei nun $U \leq \mathbb{Z}$ eine Untergruppe von $(\mathbb{Z}, +)$ so bleibt zu zeigen, daß es ein $n \geq 0$ gibt mit $U = n\mathbb{Z}$. Ist $U = \{0\}$, so können wir $n = 0$ wählen. Ist $U \neq \{0\}$, so gibt es eine Zahl $0 \neq z \in U$ und eine der Zahlen $z \in U$ oder $-z \in U$ ist positiv. Also ist die Teilmenge

$$\{m \in \mathbb{N} \mid 0 \neq m \in U\}$$

der natürlichen Zahlen nicht leer und besitzt wegen des Archimedischen Prinzips somit ein kleinstes Element, d.h. die Zahl

$$n := \min\{z \in U \mid z > 0\} \in U$$

existiert.

Wir wollen zeigen, daß $U = \langle n \rangle$ gilt. Wegen $n \in U$ folgt $\langle n \rangle \subseteq U$ unmittelbar aus der Definition des Erzeugnisses. Sei nun umgekehrt $u \in U$ gegeben. Division von u durch n mit Rest liefert ganze Zahlen $q, r \in \mathbb{Z}$ mit

$$u = q \cdot n + r$$

und

$$0 \leq r < n. \tag{65}$$

Da sowohl u als auch n Elemente von U sind, folgt

$$r = u - q \cdot n \in U,$$

und da n die kleinste *echt positive* Zahl in U ist, bedingt (65), daß $r = 0$ gilt. Mithin ist $u = q \cdot n \in \langle n \rangle$. Damit haben wir $U = \langle n \rangle$ gezeigt. □

C) Gruppenhomomorphismen

Immer wenn man eine Struktur auf einer Menge definiert hat, spielen die *strukturerhaltenden Abbildungen* eine besondere Rolle. Diese werden (Struktur-)*Morphismen* oder (Struktur-)*Homomorphismen* genannt.

Definition 22.17.
Es seien $(G,\cdot)$ und $(H,*)$ zwei Gruppen. Eine Abbildung $\alpha : G \to H$ heißt *Gruppenhomomorphismus*, falls für alle $g,h \in G$ gilt:

$$\alpha(g \cdot h) = \alpha(g) * \alpha(h).$$

Wieder wollen wir uns zunächst Beispiele anschauen.

Beispiel 22.18.

a. Ist $(G,\cdot)$ eine Gruppe und $U \leq G$ eine Untergruppe, dann ist die kanonische Inklusion $i_U : U \to G$ ein Gruppenhomomorphismus, da für $g,h \in U$ gilt $i_U(g\cdot h) = g \cdot h = i_U(g) \cdot i_U(h)$.

b. Sei $a \in \mathbb{R}$ und $m_a : (\mathbb{R},+) \to (\mathbb{R},+) : g \mapsto ag$ die Multiplikation mit a, dann ist m_a ein Gruppenhomomorphismus, da für $g,h \in \mathbb{R}$ gilt

$$m_a(g+h) = a(g+h) = ag + ah = m_a(g) + m_a(h).$$

c. Ist $(G,\cdot)$ eine Gruppe und $g \in G$, so hat man Abbildungen

$$R_g : G \to G : h \mapsto hg \qquad \text{(die "}\textit{Rechtstranslation}\text{")}$$

und

$$L_g : G \to G : h \mapsto gh \qquad \text{(die "}\textit{Linkstranslation}\text{")}$$

Für $g \neq e$ gilt jedoch wegen der Kürzungsregel

$$L_g(g \cdot g) = g^3 \neq g^4 = L_g(g) \cdot L_g(g)$$

und entsprechend für R_g. Also sind L_g und R_g für $g \neq e$ *keine* Gruppenhomomorphismen. Man sieht leicht, daß L_g und R_g bijektive Abbildungen sind, mit Inverser $L_{g^{-1}}$ bzw. $R_{g^{-1}}$.

d. Ist $(G,\cdot)$ eine Gruppe und $g \in G$, so definiert man

$$i_g : G \to G : h \mapsto ghg^{-1} =: h^g.$$

i_g heißt *innerer Automorphismus* oder *Konjugation* mit g und ist ein bijektiver Gruppenhomomorphismus:

Für $h, k \in G$ gilt:

$$\begin{aligned} i_g(hk) =& g(hk)g^{-1} = g\big(hek\big)g^{-1} = g\big(h\big(g^{-1}g\big)k\big)g^{-1} \\ =& \big(ghg^{-1}\big)\big(gkg^{-1}\big) = i_g(h) \cdot i_g(k), \end{aligned}$$

also ist i_g ein Gruppenhomomorphismus. Außerdem gilt für ein beliebiges $h \in G$:

$$(i_g \circ i_{g^{-1}})(h) = g\Big(g^{-1}h\big(g^{-1}\big)^{-1}\Big)g^{-1} = \big(gg^{-1}\big)h\big(gg^{-1}\big) = ehe = h = \mathrm{id}_G(h),$$

also ist $i_g \circ i_{g^{-1}} = \mathrm{id}_G$. Analog sieht man $i_{g^{-1}} \circ i_g = \mathrm{id}_G$, und folglich ist i_g bijektiv mit Inverser $i_{g^{-1}}$ nach Satz 3.12.

Mit der Notation aus obigem Beispiel ist offenbar $i_g = R_g \circ L_{g^{-1}}$. Die Komposition zweier Nicht-Homomorphismen kann also durchaus ein Homomorphismus sein. Das folgende Lemma sagt, daß umgekehrt die Komposition zweier Homomorphismen stets wieder ein Homomorphismus ist.

Lemma 22.19.
Sind $\alpha_1 : (G_1, \cdot) \to (G_2, *)$ und $\alpha_2 : (G_2, *) \to (G_3, \times)$ Gruppenhomomorphismen, so ist auch $\alpha_2 \circ \alpha_1 : (G_1, \cdot) \to (G_3, \times)$ ein Gruppenhomomorphismus.

Beweis: Seien $g, h \in G_1$, dann gilt:

$$\begin{aligned} (\alpha_2 \circ \alpha_1)(g \cdot h) &= \alpha_2\big(\alpha_1(g \cdot h)\big) = \alpha_2\big(\alpha_1(g) * \alpha_1(h)\big) = \alpha_2\big(\alpha_1(g)\big) \times \alpha_2\big(\alpha_1(h)\big) \\ &= (\alpha_2 \circ \alpha_1)(g) \times (\alpha_2 \circ \alpha_1)(h). \end{aligned}$$

□

Definition 22.20.
Sei $\alpha : (G, \cdot) \to (H, *)$ ein Gruppenhomomorphismus.

a. Wir nennen α einen *Isomorphismus*, falls α bijektiv ist.

b. Wir nennen die Gruppen $(G, \cdot)$ und $(H, *)$ *isomorph*, wenn es einen Isomorphismus $\alpha : G \to H$ gibt. Wir schreiben dann kurz $G \cong H$.

Beispiel 22.21.

a. In Beispiel 22.18 ist m_a ein Homomorphismus. Zudem ist m_a ein Isophismus mit Inverser $m_{\frac{1}{a}}$ genau dann wenn $a \neq 0$.

b. Ist $(G, \cdot)$ eine Gruppe und $g \in G$, dann ist die Konjugation i_g mit g nach Beispiel 22.18 ein Isomorphismus mit Inverser $i_{g^{-1}}$.

c. Die Abbildung $\det : \big(\mathrm{Gl}_2(\mathbb{R}), \circ\big) \longrightarrow (\mathbb{R} \setminus \{0\}, \cdot)$ aus Aufgabe 22.32 ist ein surjektiver Homomorphismus.

Der Umstand, daß die Gruppenhomomorphismen die Gruppenstruktur *erhalten*, hat einige einfache, aber ungemein wichtige Auswirkungen.

Proposition 22.22.
Es sei $\alpha : (G, \cdot) \to (H, *)$ ein Gruppenhomomorphismus. Dann gelten:

a. $\alpha(e_G) = e_H$.

b. $\alpha\big(g^{-1}\big) = \big(\alpha(g)\big)^{-1}$ für $g \in G$.

c. $\alpha\big(g^n\big) = \big(\alpha(g)\big)^n$ für $g \in G$ und $n \in \mathbb{Z}$.

d. Ist α bijektiv, so ist $\alpha^{-1} : H \to G$ ein Gruppenhomomorphismus.

e. Ist $U \leq G$, dann ist $\alpha(U) \leq H$. $\alpha(U)$ heißt das *Bild* von U unter α.

f. Ist $V \leq H$, dann ist $\alpha^{-1}(V) \leq G$. $\alpha^{-1}(V)$ heißt das *Urbild* von V unter α.

g. $\mathrm{Im}(\alpha) := \alpha(G)$, das *Bild* von α, ist eine Untergruppe von H.

h. $\mathrm{Ker}(\alpha) := \alpha^{-1}(e_H)$, der *Kern* von α, ist eine Untergruppe von G.

Beweis: a. Es gilt

$$e_H * \alpha(e_G) = \alpha(e_G) = \alpha(e_G \cdot e_G) = \alpha(e_G) * \alpha(e_G).$$

Mit Hilfe der Kürzungsregel 22.6 folgt dann $e_H = \alpha(e_G)$.

b. Für $g \in G$ gilt:

$$\alpha\big(g^{-1}\big) * \alpha(g) = \alpha\big(g^{-1} \cdot g\big) = \alpha(e_G) = e_H.$$

Wegen der Eindeutigkeit der Inversen in H folgt die Behauptung.

c. Es sei $g \in G$ und $n \in \mathbb{Z}$. Den Fall $n \geq 0$ beweisen wir mit Hilfe von Induktion nach n. Ist $n = 0$, so folgt die Behauptung aus a., und ist $n > 0$, so gilt nach Definition und mittels Induktion nach n

$$\alpha(g^n) = \alpha(g^{n-1} \cdot g) = \alpha(g^{n-1}) \cdot \alpha(g) \overset{Ind.}{=} \alpha(g)^{n-1} \cdot \alpha(g) = \alpha(g)^n. \tag{66}$$

Ist nun $n < 0$, so ist $-n > 0$ und es gilt wegen der Potenzgesetze

$$\alpha(g^n) = \alpha\big((g^{-1})^{-n}\big) \overset{(66)}{=} \alpha\big(g^{-1}\big)^{-n} \overset{b.}{=} \big(\alpha(g)^{-1}\big)^{-n} = \alpha(g)^n.$$

d. Ist $\alpha : G \to H$ bijektiv, so existiert die Umkehrabbildung $\alpha^{-1} : H \to G$. Seien $u, v \in H$. Setze $g := \alpha^{-1}(u)$ und $h := \alpha^{-1}(v)$, also $u = \alpha(g)$ und $v = \alpha(h)$. Dann gilt:

$$\alpha^{-1}(u * v) = \alpha^{-1}\big(\alpha(g) * \alpha(h)\big) = \alpha^{-1}\big(\alpha(g \cdot h)\big) = g \cdot h = \alpha^{-1}(u) \cdot \alpha^{-1}(v).$$

Also ist α^{-1} ein Gruppenhomomorphismus.

e. Sind $u, v \in \alpha(U)$, dann existieren $g, h \in U$ mit $\alpha(g) = u$ und $\alpha(h) = v$. Da $g \cdot h \in U$, gilt:

$$u * v = \alpha(g) * \alpha(h) = \alpha(g \cdot h) \in \alpha(U).$$

Außerdem gilt $g^{-1} \in U$ und somit:

$$u^{-1} = \big(\alpha(g)\big)^{-1} = \alpha\big(g^{-1}\big) \in \alpha(U).$$

Da zudem $\alpha(e_G) \in \alpha(U)$, also $\alpha(U) \neq \emptyset$, folgt, daß $\alpha(U)$ eine Untergruppe von H ist.

f. Seien $g, h \in \alpha^{-1}(V)$, so gilt $\alpha(g \cdot h) = \alpha(g) * \alpha(h) \in V$, da V eine Untergruppe ist. Also gilt $g \cdot h \in \alpha^{-1}(V)$. Außerdem gilt $\alpha\big(g^{-1}\big) = \big(\alpha(g)\big)^{-1} \in V$, wieder da V eine Untergruppe ist. Somit liegt auch g^{-1} in $\alpha^{-1}(V)$. Da das Urbild von V unter α ferner nicht leer ist, weil wegen $\alpha(e_G) = e_H \in V$ gilt, daß $e_G \in \alpha^{-1}(V)$, folgt wieder, daß $\alpha^{-1}(V)$ eine Untergruppe von G ist.

g. Dies folgt aus e., da G eine Untergruppe von G ist.

h. Dies folgt aus f., da $\{e_H\}$ eine Untergruppe von H ist.

□

Nach Definition muß man für die Injektivität einer Abbildung nachprüfen, daß jedes Element im Bild nur ein Urbild hat. Bei Gruppenhomomorphismen gibt es ein einfacheres Kriterium.

Lemma 22.23.
Ein Gruppenhomomorphismus $\alpha : (G, \cdot) \to (H, *)$ ist genau dann injektiv, wenn $\mathrm{Ker}(\alpha) = \{e_G\}$.

Beweis: Ist α injektiv, so ist $\alpha^{-1}(e_H)$ höchstens einelementig, und wegen $\alpha(e_G) = e_H$ gilt dann $\mathrm{Ker}(\alpha) = \alpha^{-1}(e_H) = \{e_G\}$.

Gilt umgekehrt $\mathrm{Ker}(\alpha) = \{e_G\}$, und sind $g, h \in G$ mit $\alpha(g) = \alpha(h)$, so folgt wegen:

$$e_H = \alpha(g) * \big(\alpha(h)\big)^{-1} = \alpha(g) * \alpha\big(h^{-1}\big) = \alpha\big(g \cdot h^{-1}\big),$$

daß $g \cdot h^{-1} = e_G$, also $g = h$. Somit ist α injektiv. □

Aufgaben

Aufgabe 22.24.
Beweise Lemma 22.11.

Aufgabe 22.25.
Untersuche, ob die folgende zweistellige Operation die Menge $G := \mathbb{Q} \times \mathbb{Q}$ zu einer Gruppe macht:

$$G \times G \longrightarrow G : \big((a,b),(a',b')\big) \mapsto (a,b) \cdot (a',b') := (aa', bb').$$

Aufgabe 22.26.
Untersuche, ob die folgende zweistellige Operation die Menge $G := \mathbb{Q}_{>0} \times \mathbb{Q}_{>0}$ zu einer Gruppe macht:

$$G \times G \longrightarrow G : \big((a,b),(a',b')\big) \mapsto (a,b) \cdot (a',b') := (aa', bb').$$

Aufgabe 22.27.
Es seien $(G,\cdot)$ und $(H,*)$ zwei Gruppen. Wir definieren auf der Menge $G \times H = \{(x,y) \mid x \in G, y \in H\}$ eine zweistellige Operation durch

$$(x,y) \circ (x',y') := (x \cdot x', y * y')$$

für $(x,y),(x',y') \in G \times H$. Zeige, daß dann $(G \times H, \circ)$ eine Gruppe ist.

Aufgabe 22.28.
Untersuche, welche der folgenden zweistelligen Operationen Gruppen definieren:

a. $G = \big(\mathbb{Q}\setminus\{0\}\big) \times \big(\mathbb{Q}\setminus\{0\}\big)$ mit $(a,b)\cdot(a',b') = (ab', ba')$ für $a,a',b,b' \in \mathbb{Q}\setminus\{0\}$,

b. $G = \mathbb{R}\times\mathbb{R}\setminus\big\{(0,0)\big\}$ mit $(a,b)\cdot(a',b') = (aa'-bb', ab'+ba')$ für $a,a',b,b' \in \mathbb{R}$.

Aufgabe 22.29.
Finde alle möglichen zweistelligen Operationen auf der Menge $G = \{e,a,b\}$, bezüglich derer G eine Gruppe mit neutralem Element e wird.

Aufgabe 22.30.
Finde alle möglichen zweistelligen Operationen auf der Menge $G = \{e,a,b,c\}$, bezüglich derer G eine Gruppe mit neutralem Element e wird. Dabei sollten nur Möglichkeiten aufgelistet werden, die nicht durch Vertauschung der Buchstaben a, b und c ineinander überführt werden können.

Aufgabe 22.31.
Es sei $(G,\cdot)$ ein Gruppe mit neutralem Element e. Zeige, falls $g^2 = e$ für alle $g \in G$, so ist G abelsch.

Aufgabe 22.32.
Ein Schema der Form

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

mit $a, b, c, d \in \mathbb{R}$ wollen wir eine *relle* $2x2$*-Matrix* nennen, und $\mathrm{Mat}_2(\mathbb{R})$ soll die Menge solcher Matrizen sein. Für zwei relle $2x2$-Matrizen definieren wir ihr Produkt als

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \circ \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}.$$

Ferner bezeichnen wir

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \in \mathbb{R}$$

als *Determinante* der Matrix, und definieren

$$\mathrm{Gl}_2(\mathbb{R}) = \{A \in \mathrm{Mat}_2(\mathbb{R}) \mid \det(A) \neq 0\}.$$

Zeige:

a. Für $A, B \in \mathrm{Mat}_2(\mathbb{R})$ gilt $\det(A \cdot B) = \det(A) \cdot \det(B)$.

b. $\big(\mathrm{Gl}_2(\mathbb{R}), \circ\big)$ ist eine nicht-abelsche Gruppe.

Aufgabe 22.33 (Boolsche Gruppe).
Es sei M eine Menge.

a. Sind $X, Y, Z \subseteq M$, dann gelten

$$X \setminus \big((Y \setminus Z) \cup (Z \setminus Y)\big) = \big(X \setminus (Y \cup Z)\big) \cup (X \cap Y \cap Z)$$

und

$$\big((X \setminus Y) \cup (Y \setminus X)\big) \setminus Z = \big(X \setminus (Y \cup Z)\big) \cup \big(Y \setminus (X \cup Z)\big).$$

b. Wir definieren auf der Potenzmenge $G = \mathcal{P}(M) = \{A \mid A \subseteq M\}$ von M eine zweistellige Operation durch

$$A + B := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$$

für $A, B \in G$. Zeige, $(G, +)$ ist eine abelsche Gruppe.

Aufgabe 22.34.
Sei M eine Menge, $m \in M$, $k \in \mathbb{N}$ und $\sigma \in \mathrm{Sym}(M)$ mit $\sigma^k(m) = m$.
Zeige, dann ist auch $\sigma^{q \cdot k}(m) = m$ für alle $q \in \mathbb{Z}$.

Aufgabe 22.35.
Es sei $(G,\cdot)$ eine Gruppe und $a \in G$ sei fest gegeben. Wir definieren eine zweistellige Operation auf G durch

$$* : G \times G \longrightarrow G : (g,h) \mapsto g * h = g \cdot \big(a^{-1} \cdot h\big).$$

Überprüfe, ob $(G,*)$ eine Gruppe ist.

Aufgabe 22.36.
Zeige, die Menge

$$U = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \;\middle|\; a,b \in \mathbb{R}, (a,b) \neq (0,0) \right\}$$

ist eine Untergruppe von $\big(\mathrm{Gl}_2(\mathbb{R}), \circ\big)$.

Aufgabe 22.37.
Es sei $(G,\cdot)$ eine Gruppe und $g \in G$, so daß die Menge $\{g^n \mid n > 0\}$ endlich ist. Zeige, dann gibt es ein $n > 0$ mit $g^n = e_G$.

Aufgabe 22.38.
Es sei $(G,\cdot)$ eine Gruppe und $\emptyset \neq U \subseteq G$ eine endliche Teilmenge von G.
Zeige, genau dann ist U eine Untergruppe von G, wenn für alle $u,v \in U$ auch $u \cdot v \in U$.

Aufgabe 22.39.
Welche der folgenden Mengen sind Untergruppen von $\big(\mathrm{Sym}(\mathbb{R}), \circ\big)$?

a. $U = \{f \in \mathrm{Sym}(\mathbb{R}) \mid f(x) < f(y) \text{ falls } x > y\}$,

b. $V = \{f \in \mathrm{Sym}(\mathbb{R}) \mid |f(x)| = |x| \text{ für alle } x \in \mathbb{R}\}$.

Aufgabe 22.40.
Für zwei reelle Zahlen $a,b \in \mathbb{R}$ definieren wir die Abbildung

$$f_{a,b} : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto a \cdot x + b.$$

Welche der folgenden Mengen sind Untergruppen von $\big(\mathrm{Sym}(\mathbb{R}), \circ\big)$?

a. $U = \{f_{a,b} \mid a,b \in \mathbb{R}, a \neq 0\}$,

b. $V = \{f_{a,1} \mid a \in \mathbb{R}, a \neq 0\}$.

Aufgabe 22.41.
Sei $(G,\cdot)$ eine Gruppe. Zeige, daß die Menge

$$Z(G) := \{g \in G \mid g \cdot h = h \cdot g \ \ \forall\, h \in G\}$$

eine Untergruppe von G ist. Sie wird das Zentrum von G genannt.

Aufgabe 22.42.
Betrachte die Gruppe $(G,\cdot)$ aus Aufgabe 22.28 b. und die Gruppe $(U,\cdot)$ aus Aufgabe 22.36. Zeige, die Abbildung

$$\alpha : G \longrightarrow U : (a,b) \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

ist ein Gruppenisomorphismus.

Aufgabe 22.43.
Wir betrachten die Gruppe $U = \{f_{a,b} : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto ax+b \mid a,b \in \mathbb{R}, a \neq 0\}$ aus Aufgabe 22.40, wobei die Gruppenoperation die Verknüpfung von Abbildungen ist, sowie die Gruppe $(\mathbb{R} \setminus \{0\},\cdot)$. Zeige, daß die Abbildung

$$\alpha : U \longrightarrow \mathbb{R} \setminus \{0\} : f_{a,b} \mapsto a$$

ein Gruppenhomomorphismus ist.

Aufgabe 22.44.
Es sei $(G,\cdot)$ eine Gruppe und $g \in G$. Zeige:

a. Die Abbildung

$$\alpha : \mathbb{Z} \longrightarrow G : n \mapsto g^n$$

ist ein Gruppenhomomorphismus mit Bild $\mathrm{Im}(\alpha) = \langle g \rangle$.

b. α ist genau dann injektiv, wenn $g^k \neq g^l$ für alle $k,l \in \mathbb{Z}$ mit $k \neq l$.

c. Gibt es ganze Zahlen $k \neq l$ mit $g^k = g^l$, so existiert die Zahl

$$n = \min\{m \in \mathbb{N} \mid m > 0, g^m = e_G\}$$

und es gelten:

(i) $\mathrm{Ker}(\alpha) = \{m \in \mathbb{Z} \mid g^m = e\} = n\mathbb{Z}$,

(ii) $\langle g \rangle = \{e_G, g, g^2, \ldots, g^{n-1}\}$, und

(iii) $|\langle g \rangle| = n$.

Aufgabe 22.45.
Es sei $(G, \cdot)$ eine Gruppe und $h, k \in G$ fest gegeben. Prüfe, welche Bedingungen für h und k gelten müssen, damit die folgenden Abbildungen Gruppenhomomorphismen sind:

a. $\alpha : G \to G : g \mapsto h \cdot g$,

b. $\alpha : G \to G : g \mapsto h \cdot g \cdot h$,

c. $\beta : G \to G : g \mapsto h^{-1} \cdot g \cdot k$,

Aufgabe 22.46 (Satz von Cayley).
Es sei $(G, \cdot)$ eine Gruppe und $g \in G$.

a. Die Abbildung $L_g : G \longrightarrow G : h \mapsto g \cdot h$ ist bijektiv.

b. Die Abbildung $\alpha : G \longrightarrow \mathrm{Sym}(G) : g \mapsto L_g$ ist ein injektiver Homomorphismus.

Anmerkung: Man nennt die Aussage in Teil b. auch den *Satz von Cayley*. Er besagt letztlich, daß jede Gruppe isomorph zu einer Untergruppe einer symmetrischen Gruppe ist.

Aufgabe 22.47.
Da $(\mathbb{R} \setminus \{0\}, \cdot)$ und $(\mathbb{R}, +)$ Gruppen sind, ist nach Aufgabe 22.27 auch die Menge $G = (\mathbb{R} \setminus \{0\}) \times \mathbb{R}$ mit der Operation

$$(r, s) * (r', s') := (r \cdot r', s + s')$$

eine Gruppe. Zudem wissen wir, daß die komplexen Zahlen $(\mathbb{C} \setminus \{0\}, \cdot)$ ohne die Null bezüglich der Multiplikation eine Gruppe sind. Zeige, daß die Abbildung

$$\alpha : G \longrightarrow \mathbb{C} \setminus \{0\} : (r, s) \mapsto r \cdot \exp(s \cdot \pi \cdot i)$$

ein Gruppenhomomorphismus ist. Bestimme das Bild und den Kern von α. Ist α injektiv / surjektiv? Dabei ist π die Kreiszahl und i die imaginäre Einheit.

Aufgabe 22.48.
Da $(\mathbb{R}, +)$ eine Gruppe ist, ist nach Aufgabe 22.27 auch $(\mathbb{R}^2, +)$ mit der komponentenweisen Addition eine Gruppe. Zeige, daß folgende Abbildung

$$\alpha : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \mapsto 2x + 3y$$

ein Gruppenhomomorphismus ist. Bestimme das Bild und den Kern von α. Ist α injektiv / surjektiv?

Aufgabe 22.49.
Es sei $\alpha : (G, \cdot) \longrightarrow (H, *)$ ein Gruppenhomomorphismus, $g \in G$ und $g' \in \mathrm{Ker}(\alpha)$. Zeige, dann gilt $g^{-1} \cdot g' \cdot g \in \mathrm{Ker}(\alpha)$.

Aufgabe 22.50.
Bestimme alle Gruppenhomomorphismen von $(\mathbb{Z}, +)$ nach $(\mathbb{R}, +)$.

§ 23 Die symmetrische Gruppe

Im vorliegenden Abschnitt stellen wir die für uns relevanten Ergebnisse zur symmetrischen Gruppe $\mathbb{S}_n$ zusammen. Wir verzichten auf Beweise und erläutern sie stattdessen an Beispielen. Die Beweise der Aussagen finden sich in einer etwas ausführlicheren Version des Abschnitts im Anhang (siehe A2)

Definition 23.1 (Permutationen).
Eine bijektive Abbildung $\sigma : \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$ nennen wir eine *Permutation* der Menge $\{1, \dots, n\}$, und wir bezeichnen mit

$$\mathbb{S}_n = \mathrm{Sym}(\{1, \dots, n\}) = \{\sigma : \{1, \dots, n\} \longrightarrow \{1, \dots, n\} \mid \sigma \text{ ist bijektiv}\}$$

die Menge aller Permutation der Menge $\{1, \dots, n\}$.
Eine Permutation $\sigma \in \mathbb{S}_n$ kann durch eine *Wertetabelle* der folgenden Form beschrieben werden:

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

bzw.

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \\ \sigma(a_1) & \sigma(a_2) & \dots & \sigma(a_n) \end{pmatrix},$$

falls $a_1, \dots, a_n$ irgendeine Anordnung der Zahlen $1, \dots, n$ ist.

Bemerkung 23.2 (Die symmetrische Gruppe $\mathbb{S}_n$).
In Beispiel 22.3 haben wir gezeigt, daß $\mathbb{S}_n$ mit der Komposition von Abbildungen eine Gruppe ist. Wir nennen $(\mathbb{S}_n, \circ)$ die *symmetrische Gruppe* vom Grad n. Die $\mathbb{S}_n$ enthält genau $n!$ Elemente.

Beispiel 23.3 ($\mathbb{S}_n$ nicht abelsch für $n \geq 3$).
Die Gruppe $\mathbb{S}_n$ ist für $n \geq 3$ nicht abelsch. In $\mathbb{S}_3$ gilt für die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in \mathbb{S}_3$$

nämlich

$$(\sigma \circ \pi)(1) = \sigma(\pi(1)) = \sigma(2) = 1 \neq 3 = \pi(2) = \pi(\sigma(1)) = (\pi \circ \sigma)(1)$$

und deshalb

$$\pi \circ \sigma \neq \sigma \circ \pi.$$

Beachte, daß es bei dem Schema nicht darauf ankommt, in welcher Reihenfolge die Zahlen von 1 bis n in der ersten Zeile stehen. Es gilt etwa:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \end{pmatrix}.$$

Es empfiehlt sich aber der Übersichtlichkeit halber für gewöhnlich, die Ziffern in aufsteigender Reihenfolge anzuordnen.

Bemerkung 23.4 (Invertieren einer Permutation).
Die oben eingeführte Darstellung einer Permutation hat den angenehmen Nebeneffekt, daß man das Inverse der Permutation leicht angeben kann, indem man einfach die beiden Zeilen vertauscht. Sprich, für eine Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix} \in \mathbb{S}_n$$

ist das Inverse σ^{-1} gegeben durch

$$\sigma^{-1} = \begin{pmatrix} \sigma(1) & \sigma(2) & \dots & \sigma(n) \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Definition 23.5 (Zyklen und Transpositionen).

a. Sei $\{1, \dots, n\} = \{a_1, \dots, a_k\} \dot\cup \{b_1, \dots, b_{n-k}\}$, $k \geq 2$, und

$$\sigma = \begin{pmatrix} a_1 & a_2 & \dots & a_{k-1} & a_k & b_1 & \dots & b_{n-k} \\ a_2 & a_3 & \dots & a_k & a_1 & b_1 & \dots & b_{n-k} \end{pmatrix} \in \mathbb{S}_n,$$

so heißt σ ein **k-Zyklus**, und wir sagen, daß sie die Zahlen $a_1, \dots, a_k$ *zyklisch vertauscht*. Die Abbildungsvorschrift eines solchen k-Zyklus läßt sich deutlich kompakter durch das folgende einzeilige Schema repräsentieren:

(67) $$\sigma = (a_1 \dots a_k).$$

b. Ein 2-Zyklus wird auch eine **Transposition** genannt. Eine Transposition $\tau = (i\ j)$ ist mithin eine Permutation, die nur die zwei Zahlen i und j miteinander vertauscht, alle anderen aber fest läßt.

c. Das neutrale Element von $\mathbb{S}_n$, per definitionem $\mathrm{id}_{\{1,\dots,n\}}$, wollen wir der Einfachheit halber mit id bezeichnen.

Bemerkung 23.6 (Zykelschreibweise).

a. Die Interpretation der Schreibweise in Gleichung (67) ist offensichtlich, das erste Element a_1 wird auf das zweite a_2 abgebildet, das zweite auf das dritte, und so weiter, bis schließlich das letzte, nämlich a_k, auf das erste, das heißt auf a_1, abgebildet wird – der *Kreis* schließt sich. Beachte hierbei, daß die Zyklen $(a_1 \dots a_k)$, $(a_k a_1 \dots a_{k-1})$, etc. übereinstimmen! Um diese Mehrdeutigkeit zu vermeiden, empfiehlt es sich, einen Zyklus stets mit der kleinsten der Zahlen $a_1, \dots, a_k$ zu beginnen.
Bisher haben wir k-Zyklen nur für $k \geq 2$ definiert. Wir können nun auch 1-Zyklen, etwa (1) oder (3), zulassen und definieren diese in natürlicher Weise als die Identität.

b. Die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix} \in \mathbb{S}_4 \quad \text{und} \quad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 2 & 5 \end{pmatrix} \in \mathbb{S}_5$$

sind jeweils 3-Zyklen, die die Zahlen $1, 4, 2$ zyklisch vertauschen. In der oben eingeführten Zyklenschreibweise gilt

$$\sigma = (1\ 4\ 2) \quad \text{und} \quad \pi = (1\ 4\ 2).$$

Damit wird der Nachteil dieser Schreibweise gegenüber dem zweizeiligen Schema deutlich – weder der Definitionsbereich noch der Wertebereich lassen sich aus der Zyklenschreibweise eindeutig ablesen. Aber diesen Preis sind wir für die gewonnene *Übersichtlichkeit* gerne bereit zu zahlen. Denn einerseits ist in Anwendungen meist zweifelsfrei bekannt, was n ist, und andererseits ist die wesentliche Information für uns letztlich, welche Zahlen durch die Permutation vertauscht werden, und nicht, welche unbewegt bleiben.

c. Für eine Transposition $\tau \in \mathbb{S}_n$ gilt $\tau^{-1} = \tau$, also $\tau^2 = \mathrm{id}$.

d. Für kleine Werte n ist $\mathbb{S}_n$ sehr übersichtlich, für große Werte n wird $\mathbb{S}_n$ jedoch riesig. $\mathbb{S}_1 = \{\mathrm{id}\}$ und $\mathbb{S}_2 = \{\mathrm{id}, (1\ 2)\}$. $\mathbb{S}_3 = \{\mathrm{id}, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$ hat schon sechs Elemente, $\mathbb{S}_4$ gar 24 und $\mathbb{S}_{64}$ ungefähr 10^{89}. Letztere Zahl entspricht in etwa der angenommenen Anzahl der Atome im Universum.

Satz 23.7 (Zyklenzerlegung und Signum).

a. Jede Permutation läßt sich als Produkt von disjunkten Zyklen schreiben.

b. Jede Permutation läßt sich als Produkt von Transpositionen schreiben.

c. Jede Permutation läßt sich als Produkt von Transpositionen benachbarter Zahlen schreiben.

d. Es gibt genau einen Gruppenhomomorphismus, das *Signum* genannt,

$$\mathrm{sgn} : (\mathbb{S}_n, \circ) \longrightarrow \big(\{1, -1\}, \cdot\big)$$

mit $\mathrm{sgn}(\tau) = -1$ für jede Transposition $\tau \in \mathbb{S}_n$. Insbesondere gilt, ist $\sigma \in \mathbb{S}_n$ ein Produkt von k Transpositionen, dann gilt mithin

$$\mathrm{sgn}(\sigma) = (-1)^k.$$

e. Für $\sigma \in \mathbb{S}_n$ gilt $\mathrm{sgn}(\sigma) = \mathrm{sgn}\big(\sigma^{-1}\big)$.

f. Ist $\mathbb{A}_n = \{\sigma \in \mathbb{S}_n \mid \mathrm{sgn}(\sigma) = 1\}$ und ist $\tau = (i\ j)$ eine Transposition, so gilt

$$\mathbb{S}_n = \mathbb{A}_n \dot\cup \tau\mathbb{A}_n$$

wobei $\tau\mathbb{A}_n = \{\tau \circ \sigma \mid \sigma \in \mathbb{A}_n\}$.

Beweis: Für den formalen Beweis der Aussagen verweisen wir auf den Anhang A2. □

Bemerkung 23.8 (sgn als Gruppenhomomorphismus).

a. Daß die Abbildung sgn ein *Gruppenhomomorphismus* ist, heißt

$$\operatorname{sgn}(\sigma \cdot \pi) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\pi)$$

für alle $\sigma, \pi \in \mathbb{S}_n$. Ist also $\sigma = \tau_1 \circ \cdots \circ \tau_k \in \mathbb{S}_n$ ein Produkt von k Transpositionen, dann gilt

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau_1) \cdot \ldots \cdot \operatorname{sgn}(\tau_k) = (-1)^k.$$

b. Einen Beweis der Aussagen findet der interessierte Leser im Anhang (siehe A2). Wir wollen uns hier damit begnügen, an einem Beispiel zu zeigen, was die Aussagen bedeuten und wie man die Zerlegungen bzw. das Signum berechnen kann.

Beispiel 23.9 (Zyklenzerlegung).
Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix} \in \mathbb{S}_5$$

hat die Zyklenzerlegung

$$\sigma = (1\ 2\ 5) \circ (3\ 4) = (3\ 4) \circ (1\ 2\ 5). \tag{68}$$

Eine berechtigte Frage ist, wie wir die Zyklenzerlegung in (68) gefunden haben. Wir wollen versuchen, dies so in Worte zu fassen, daß dem Leser daraus die allgemeine Vorgehensweise ersichtlich wird. Man starte mit der kleinsten Zahl, 1, und suche ihr Bild unter σ, also $\sigma(1) = 2$. Das liefert den Startteil des ersten Zyklus:

$$(1\ 2$$

Sodann betrachte man das Bild von 2 unter σ, also $\sigma(2) = 5$, und erhält:

$$(1\ 2\ 5$$

Man fährt mit dem Bild von 5 unter σ, also $\sigma(5) = 1$, fort. Da dieses das erste Element des ersten Zyklus war, schließen wir den Zyklus,

$$(1\ 2\ 5),$$

und beginnen den zweiten Zyklus mit der kleinsten Zahl in $\{1, \ldots, 5\}$, die noch nicht in dem ersten Zyklus vorkommt, also mit 3:

$$(1\ 2\ 5) \circ (3$$

Dann betrachten wir deren Bild unter σ, also $\sigma(3) = 4$, und setzen so unseren zweiten Zyklus fort:

$$(1\ 2\ 5) \circ (3\ 4$$

Da bereits alle fünf Elemente von $\{1,\dots,5\}$ aufgebraucht sind, muß notwendig $\sigma(4)=3$ gelten, was es auch tut, und wir können damit auch den zweiten Zyklus schließen:

$$\sigma = (1\ 2\ 5)\circ(3\ 4).$$

Wie gesagt, da in $\{1,\dots,5\}$ keine Zahl mehr übrig ist, sind wir fertig und haben die Zyklenzerlegung von σ gefunden. □

Beispiel 23.10 (Zerlegung in Transpositionen).
Wir wollen nun zeigen, wie man die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 8 & 7 & 1 & 4 & 5 & 6 & 9 & 2 \end{pmatrix} \in \mathbb{S}_9$$

als Produkt von Transpositionen schreiben kann und wie man ihr Signum berechnet.

Dazu zerlegen wir sie zunächst in ein Produkt disjunkter Zyklen, und mit Hilfe des Verfahrens aus Beispiel 23.9 erhalten wir

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 8 & 7 & 1 & 4 & 5 & 6 & 9 & 2 \end{pmatrix} = (1\ 3\ 7\ 6\ 5\ 4)\circ(2\ 8\ 9).$$

Dann schreiben wir die Zyklen als Produkte von Transpositionen:

$$(1\ 3\ 7\ 6\ 5\ 4) = (1\ 3)\circ(3\ 7)\circ(7\ 6)\circ(6\ 5)\circ(5\ 4)$$

und

$$(2\ 8\ 9) = (2\ 8)\circ(8\ 9).$$

Die Ergebnisse können wir dann zusammensetzen und erhalten

$$\sigma = (1\ 3)\circ(3\ 7)\circ(7\ 6)\circ(6\ 5)\circ(5\ 4)\circ(2\ 8)\circ(8\ 9).$$

Aus dem Beispiel läßt sich leicht ein allgemeines Verfahren ableiten, um *eine* solche Zerlegung zu berechnen. Man sollte beachten, daß die Zerlegung in ein Produkt nicht eindeutig ist. Sie läßt sich auf viele Arten variieren. Wichtig ist sie allein, um das Signum zu berechnen, denn es gilt

$$\operatorname{sgn}(\sigma) = (-1)^7 = -1,$$

da σ Produkt von sieben Transpositionen ist.

Will man die Permutation gar als Produkt von Transpositionen benachbarter Zahlen schreiben, so reicht es, zu zeigen, wie man eine beliebige Transposition als Produkt solcher Transpositionen schreiben kann. Dann kann man das Verfahren auf jede Transposition in der obigen Zerlegung anwenden. Wir führen dies hier nur am Beispiel der Transposition $(3\ 7)$ vor. Das allgemeine Verfahren kann man daraus leicht ablesen:

$$(3\ 7) = (3\ 4)\circ(4\ 5)\circ(5\ 6)\circ(6\ 7)\circ(5\ 6)\circ(4\ 5)\circ(3\ 4).$$

Beispiel 23.11 (Symmetrische Gruppe $\mathbb{S}_3$).
Wir können alle $6 = 3!$ Elemente der $\mathbb{S}_3$ auflisten:

$$\mathbb{S}_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}$$
$$= \{\mathrm{id} = (1), (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}\,.$$

Wir können dann auch für alle Elemente das Signum ausrechnen:

σ	id	(1 2)	(1 3)	(2 3)	(1 2 3)	(1 3 2)
$\mathrm{sgn}(\sigma)$	1	-1	-1	-1	1	1

Damit gilt dann

$$\mathbb{A}_3 = \{\mathrm{id}, (1\ 2\ 3), (1\ 3\ 2)\}$$

und

$$(1\ 2)\mathbb{A}_3 = \{(1\ 2) \circ \mathrm{id}, (1\ 2) \circ (1\ 2\ 3), (1\ 2) \circ (1\ 3\ 2)\} = \{(1\ 2), (2\ 3), (1\ 3)\},$$

woraus sich dann unmittelbar die folgende Gleichung ergibt:

$$\mathbb{S}_3 = \mathbb{A}_3 \dot\cup (1\ 2)\mathbb{A}_3.$$

Aufgaben

Aufgabe 23.12.
Betrachte die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 6 & 3 & 7 & 5 & 1 & 4 \end{pmatrix}, \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 2 & 7 & 1 & 4 & 3 & 6 \end{pmatrix} \in \mathbb{S}_7.$$

a. Berechne $\sigma \circ \pi$, $\pi \circ \sigma$, σ^{-1}, π^{-1}.

b. Bestimme für jede der Permutationen in a. die Zyklenzerlegung.

c. Schreibe $\sigma \circ \pi$ als ein Produkt von Transpositionen.

d. Schreibe π^{-1} als ein Produkt von Transpositionen aufeinander folgender Zahlen.

e. Berechne für jede der Permutationen in a. das Signum.

§ 24 Der Satz von Lagrange und Faktorgruppen

Der Begriff der *Faktorgruppe* zählt ganz ohne Frage zu den Begriffen, die Studienanfängern in Mathematik die größten Probleme bereiten. Die ersten Gruppen, die wir kennen gelernt haben, sind $(\mathbb{Z}, +)$ und $(\mathbb{Q}, +)$ – die Elemente sind schlicht Zahlen, etwas sehr vertrautes. Dann haben wir die symmetrische Gruppe betrachtet, deren Elemente Abbildungen sind. Für eine beliebige Menge M ist $\mathrm{Sym}(M)$ sicher schon nicht so ganz einfach gewesen, mußte doch z.B. die Gleichheit im Assoziativgesetz durch Einsetzen der Elemente von M in die Funktionen getestet werden. Der Übergang zur $\mathbb{S}_n$ sollte das Leben dann wieder einfacher gemacht haben, da die Elemente zu matrixähnlichen Schemata mit festen Rechenregeln wurden. Der Schritt zur Faktorgruppe scheint noch einmal höhere Anforderungen an das Abstraktionsvermögen der Studierenden zu stellen, sind doch die Elemente jetzt plötzlich selbst eigentlich Mengen. Wie gesagt, damit tun sich die meisten Studienanfänger recht schwer, dabei ist es eigentlich sehr einfach. Man muß nur bereit sein, zu vergessen, was die Elemente eigentlich sind und sich (wie bei den Permutationen in der $\mathbb{S}_n$) nur die Rechenregeln für Faktorgruppen merken – und die sind so einfach wie sie nur sein können.

Wie alle Gruppen bestehen auch Faktorgruppen aus einer Menge zusammen mit einer Gruppenoperation. Die ersten beiden Abschnitte dieses Kapitels beschäftigen sich mehr oder weniger mit den Voraussetzungen für die einer Faktorgruppe zugrundeliegenden Menge. Im vorliegenden Abschnitt werden wir die Faktorgruppen dann nur für abelsche Gruppen einführen, weil das für unsere Belange reicht. Für den allgemeinen Fall, der den Begriff des Normalteilers benötigt, verweisen wir den interessierten Leser auf den Anhang.

A) Linksnebenklassen

In diesem Abschnitt betrachten wir einen bestimmten Typ von Äquivalenzrelation. Allerdings wollen wir jetzt voraussetzen, daß die zugrundeliegende Menge eine Gruppe ist. Zur Beschreibung der Äquivalenzklassen benötigen wir die folgende Notation, die auch im weiteren Verlauf der Vorlesung noch öfter von Nutzen sein wird.

Notation 24.1 (Produkte von Mengen).
Es sei $(G, \cdot)$ ein Gruppe und $A, B \subseteq G$ seien zwei Teilmengen. Wir definieren

$$A \cdot B = \{a \cdot b \mid a \in A, b \in B\}.$$

Manchmal schreiben wir abküzend AB für $A \cdot B$, und wenn $A = \{g\}$ einelementig ist, dann schreiben wir gB statt $\{g\}B$ und Bg statt $B\{g\}$.

Man beachte, daß sich die Assoziativität der Gruppenoperation auf das Produkt von Teilmengen überträgt, das heißt, für $A, B, C \subseteq G$ gilt

$$(A \cdot B) \cdot C = \{(a \cdot b) \cdot c \mid a \in A, b \in B, c \in C\}$$
$$= \{a \cdot (b \cdot c) \mid a \in A, b \in B, c \in C\} = A \cdot (B \cdot C).$$

Proposition 24.2 (Linksnebenklassen als Äquivalenzklassen).
Es sei G eine Gruppe und $U \leq G$. Für zwei Elemente $g, h \in G$ definieren wir

$$g \sim h \quad :\Longleftrightarrow \quad g^{-1}h \in U.$$

Dann ist $\sim$ eine Äquivalenzrelation auf der Menge G und die zu g gehörende Äquivalenzklasse ist

$$\overline{g} = gU = \{gu \mid u \in U\}.$$

Wir nennen gU die zu g gehörende *Linksnebenklasse* von U in G und g einen *Repräsentanten* der Linksnebenklasse. Außerdem bezeichnen wir mit

$$G/U = \{gU \mid g \in G\}$$

die Menge aller Linksnebenklassen von U in G und nennen die Mächtigkeit von G/U

$$|G : U| := |G/U|$$

den *Index* von U in G.

Beweis: Wir müssen zeigen, daß die durch $\sim$ definierte Relation auf G reflexiv, symmetrisch und transitiv ist. Seien dazu $g, h, k \in G$.

- **R1:** Da $g^{-1}g = e \in U$, gilt $g \sim g$ und $\sim$ reflexiv.
- **R2:** Es gelte $g \sim h$ und damit $g^{-1}h \in U$. Aus der Abgeschlossenheit von U bezüglich der Inversenbildung folgt dann aber $h^{-1}g = \left(g^{-1}h\right)^{-1} \in U$. Es ist also $h \sim g$, und $\sim$ ist symmetrisch.
- **R3:** Es gelte $g \sim h$ und $h \sim k$ und damit $g^{-1}h \in U$ und $h^{-1}k \in U$. Wegen der Abgeschlossenheit von U bezüglich der Gruppenoperation folgt daraus $g^{-1}k = \left(g^{-1}h\right)\left(h^{-1}k\right) \in U$ und somit $g \sim k$. $\sim$ ist also auch transitiv.

Mithin ist $\sim$ eine Äquivalenzrelation.

Es bleibt zu zeigen, daß die Menge der zu g äquivalenten Elemente gU ist. Ist $h \in G$ mit $g \sim h$, so gilt nach Definition $g^{-1}h \in U$ und damit $h = g \cdot (g^{-1}h) \in gU$. Ist umgekehrt $h = gu \in gU$ mit $u \in U$, so gilt $g^{-1}h = g^{-1}gu = u \in U$ und somit $g \sim h$. □

Da eine Äquivalenzrelation auf der Menge, auf der sie definiert ist, eine disjunkte Zerlegung in Äquivalenzklassen induziert (siehe Proposition 6.10), erhalten wir das folgende Korollar geschenkt.

Korollar 24.3 (Zerlegung von G in Linksnebenklassen).
Es sei G eine Gruppe und $U \leq G$. Für $g, h \in G$ gilt entweder $gU \cap hU = \emptyset$ oder $gU = hU$, und G ist die disjunkte Vereinigung der Linksnebenklassen von U:

$$G = \dot{\bigcup_{\lambda \in G/U}} g_\lambda U,$$

wobei $g_\lambda \in G$ irgendein Repräsentant der Linksnebenklasse λ ist, d.h. $\lambda = g_\lambda U$.[1]

Beispiel 24.4 (Linksnebenklassen der alternierenden Gruppe).
Betrachten wir die Gruppe $G = \mathbb{S}_3$ und die Untergruppe $U = \mathbb{A}_3$. Dann gibt es zwei Nebenklassen (siehe auch Beispiel 23.11):

$$\mathbb{A}_3 = \mathrm{id}\,\mathbb{A}_3 = (1\ 2\ 3)\mathbb{A}_3 = (1\ 3\ 2)\mathbb{A}_3 = \{\mathrm{id}, (1\ 2\ 3), (1\ 3\ 2)\}$$

und

$$(1\ 2)\mathbb{A}_3 = (1\ 3)\mathbb{A}_3 = (2\ 3)\mathbb{A}_3 = \{(1\ 2), (1\ 3), (2\ 3)\}.$$

Mithin ist der Index $|\mathbb{S}_3 : \mathbb{A}_3|$ von $\mathbb{A}_3$ in $\mathbb{S}_3$ zwei.

Bemerkung 24.5 (Die Linksnebenklasse U).
Eine Linksnebenklasse von U in G kennt man, ganz unabhängig davon, was U und G konkret sind, nämlich die Linksnebenklasse des neutralen Elementes:

$$eU = U$$

D.h. die Untergruppe selbst ist immer eine Linksnebenklasse.

Zudem sollte man beachten, daß die möglichen Repräsentanten einer Linksnebenklasse genau die Elemente in der Nebenklasse sind. Insbesondere gilt $uU = U$ für jedes $u \in U$. □

Das vielleicht wichtigste Beispiel für unsere Vorlesung ist die Menge $\mathbb{Z}/n\mathbb{Z}$ der Linksnebenklassen der Untergruppe $n\mathbb{Z}$ in der Gruppe $(\mathbb{Z}, +)$. Um in diesem Beispiel alle Linksnebenklassen beschreiben zu können und für jede einen möglichst *einfachen* Repräsentanten angeben zu können, benötigen wir das Prinzip der *Division mit Rest* auf den ganzen Zahlen: für zwei ganze Zahlen m und $n \neq 0$ gibt es eindeutig bestimmte Zahlen q und r, so daß

$$m = q \cdot n + r$$

mit $0 \leq r < n$. Die Zahl q gibt an, wie oft n in die Zahl m passt, und die Zahl r ist der *Rest*, der bei der Division von m durch n übrig bleibt.

[1] Man beachte, daß bei der Vereinigung $\dot{\bigcup}_{\lambda \in G/U} g_\lambda U$ jede Nebenklasse von U in G nur *genau einmal* in der Vereinigung aufgeführt wird, da für jede Nebenklasse nur ein Vertreter gewählt wurde.

Proposition 24.6 (Die Elemente von $\mathbb{Z}_n$).
Ist $(G,\cdot)=(\mathbb{Z},+)$ und $U=n\mathbb{Z}$ für eine natürliche Zahl $n\geq 1$, dann hat U genau n Linksnebenklassen in G, nämlich:[2]

$$\begin{array}{rcccc}
\overline{0} & = & 0+n\mathbb{Z} & = & n\mathbb{Z}, \\
\overline{1} & = & 1+n\mathbb{Z} & = & \{1+nz \mid z\in\mathbb{Z}\}, \\
\overline{2} & = & 2+n\mathbb{Z} & = & \{2+nz \mid z\in\mathbb{Z}\}, \\
\vdots & & & & \\
\overline{n-1} & = & (n-1)+n\mathbb{Z} & = & \{n-1+nz \mid z\in\mathbb{Z}\}.
\end{array}$$

Der Index $|\mathbb{Z}:n\mathbb{Z}|$ von $n\mathbb{Z}$ in $\mathbb{Z}$ ist mithin n.

Beweis: Wir müssen zeigen, daß jede ganze Zahl $m\in\mathbb{Z}$ in einer der oben angeführten n Äquivalenzklassen liegt, und daß diese paarweise verschieden sind.

Sei also $m\in\mathbb{Z}$ eine beliebige ganze Zahl, dann existieren aufgrund der Division mit Rest ganze Zahlen $q,r\in\mathbb{Z}$, so daß

$$m=qn+r \quad \text{mit} \quad 0\leq r\leq n-1.$$

Aber damit folgt[3]

$$m-r=qn=nq\in n\mathbb{Z}=U.$$

Damit ist m äquivalent zu r, und deshalb $m\in\overline{r}$, wobei $\overline{r}$ eine der oben aufgeführten n Äquivalenzklassen ist.

Es bleibt für $0\leq i<j\leq n-1$ zu zeigen, daß $\overline{i}\neq\overline{j}$. Würde $\overline{i}=\overline{j}$ gelten, dann wäre j äquivalent zu i und somit $j-i\in n\mathbb{Z}$ ein Vielfaches von n. Nach Voraussetzung wissen wir aber, daß $0<j-i<n$ kein Vielfaches von n sein kann. Also sind $\overline{i}$ und $\overline{j}$ nicht gleich. □

Notation 24.7 (Elemente von $\mathbb{Z}_n$).
Im folgenden werden wir meist $\mathbb{Z}_n$ statt $\mathbb{Z}/n\mathbb{Z}$ schreiben. Außerdem schreiben wir für eine Restklasse $\overline{a}$ aus $\mathbb{Z}_n$ manchmal $\overline{a}_n$, wenn wir verdeutlichen wollen, modulo welcher Zahl wir rechnen. Dies ist immer dann wichtig, wenn wir parallel modulo verschiedener Zahlen rechnen wollen.

[2] Man beachte, daß dadurch, daß die Gruppenoperation die Addition ist, eine Linksnebenklasse nicht "$g\cdot U$" ist, sondern "$g+U$". Das wäre an sich vielleicht noch nicht so verwirrend, wenn in diesem konkreten Beispiel nicht auch noch die Untergruppe $U=n\mathbb{Z}$ selbst einer multiplikativ geschriebenen Nebenklasse zum Verwechseln ähnlich sähe! Das ist einer der wesentlichen Gründe, weshalb wir im Folgenden für dieses konkrete Beispiel meist die Notation $\overline{k}$ der Notation $k+n\mathbb{Z}$ vorziehen.

[3] Man beachte wieder, daß die Gruppenoperation in $(\mathbb{Z},+)$ die Addition ist. Die Bedingung "$g^{-1}h\in U$" übersetzt sich hier deshalb zu "$-g+h\in U$". Und da die Addition zudem kommutativ ist, schreiben wir meist lieber "$h-g\in U$".

Die Menge $\mathbb{Z}_n$ wird für den Rest der Vorlesung von großer Bedeutung sein. Wir führen deshalb hier noch einige übliche Sprechweisen im Zusammenhang mit diesem Beispiel ein.

Bemerkung 24.8 (Kongruenz modulo n).
Sei $n \in \mathbb{Z}$ fest gewählt. $x, y \in \mathbb{Z}$ heißen *kongruent modulo* n, falls

$$x - y \in n\mathbb{Z} = \{nq \mid q \in \mathbb{Z}\}.$$

Die Kongurenz ist genau die in Proposition 24.6 betrachtete Äquivalenzrelation, aber statt des Zeichens "$\sim$" hat sich in dieser Situation folgende Notation dafür eingebürgert, daß x kongruent zu y modulo n ist:

$$x \equiv y \ (n) \quad \text{oder} \quad x \equiv y \ (\mathrm{mod}\ n).$$

□

B) Der Satz von Lagrange

Wir wollen diesen Abschnitt mit einem wichtigen Satz, dem Satz von Lagrange abschließen. Seine wichtigste Aussage ist, daß bei einer endlichen Gruppe die Mächtigkeit einer Untergruppe stets die Mächtigkeit der Gruppe teilen muß! Das folgende Lemma ist ein zentraler Baustein im Beweis des Satzes.

Lemma 24.9 (Mächtigkeit von Linksnebenklassen).
Es sei G eine Gruppe, $U \leq G$ und $g \in G$. Dann ist die Abbildung

$$L_g : U \longrightarrow gU : u \mapsto gu$$

eine Bijektion. Insbesondere haben also alle Linksnebenklassen von U in G die Mächtigkeit $|U|$.

Beweis: Wegen der Kürzungsregel folgt aus $L_g(u) = gu = gu' = L_g(u')$ für $u, u' \in U$ unmittelbar, daß $u = u'$. Also ist L_g injektiv. Ist nun $h \in gU$ ein beliebiges Element, so gibt es nach Definition von gU ein $u \in U$, so daß $h = gu$. Aber damit ist $h = gu = L_g(u)$, und mithin ist L_g surjektiv. Die Aussage zur Mächtigkeit von Linksnebenklassen folgt, da zwei Mengen nach Definition genau dann gleichmächtig sind, wenn es eine Bijektion zwischen ihnen gibt. □

Damit sind wir nun in der Lage, den Satz von Lagrange zu formulieren und zu beweisen.

Satz 24.10 (Satz von Lagrange).
Es sei G eine endliche Gruppe und $U \leq G$ eine Untergruppe von G. Dann gilt

$$|G| = |U| \cdot |G : U|.$$

Insbesondere gilt, $|U|$ und $|G/U| = |G : U|$ sind Teiler von $|G|$.

Beweis: Da G endlich ist notwendig auch G/U endlich. Sei also $G/U = \{g_1U, \ldots, g_kU\}$ und die g_iU seien paarweise verschieden, insbesondere also $|G : U| = |G/U| = k$. Aus Korollar 24.3 und Lemma 24.9 folgt dann:

$$|G| \overset{6.11}{=} \sum_{i=1}^{k} |g_iU| \overset{24.9}{=} \sum_{i=1}^{k} |U| = |U| \cdot k = |U| \cdot |G : U|.$$

□

Aus dem Satz von Lagrange erhalten wir unmittelbar folgendes Korollar.

Korollar 24.11 (Die Ordnung eines Elementes).
Ist G eine Gruppe und $g \in G$, so definieren wir die *Ordnung* von g als $o(g) := |\langle g \rangle|$, und wenn G endlich ist, dann ist $o(g)$ ein Teiler von $|G|$.

Bemerkung 24.12 (Die Ordnung eines Elementes).
Ist G eine Gruppe und $g \in G$, dann folgt aus Aufgabe 22.44

$$o(g) = \inf\{k > 0 \mid g^k = e\} \in \mathbb{N} \cup \{\infty\}.$$

Wir wollen die Nützlichkeit des Satzes von Lagrange an einem Beispiel verdeutlichen.

Beispiel 24.13 (Die Untergruppen der $\mathbb{S}_3$).
Sei $U \leq \mathbb{S}_3$, dann gilt $|U| \in \{1, 2, 3, 6\}$ wegen des Satzes von Lagrange und da $|\mathbb{S}_3| = 3! = 6$.

1. **Fall:** $|U| = 1$**:** Dann ist notwendig $U = \{\mathrm{id}\}$, da das neutrale Element von $\mathbb{S}_3$ in U liegt.
2. **Fall:** $|U| = 6$**:** Da U eine Teilmenge von $\mathbb{S}_3$ ist muß mithin $U = \mathbb{S}_3$ gelten.
3. **Fall:** $|U| = 2$**:** Es gibt ein Element $\mathrm{id} \neq \sigma \in U$ und damit gilt $o(\sigma) \neq 1$. Aus Korollar 24.11 wissen wir, daß $o(\sigma)$ ein Teiler von $|U| = 2$ ist. Da 2 eine Primzahl ist, folgt mithin $o(\sigma) = 2$ und $U = \langle \sigma \rangle$ ist von σ erzeugt. Also gilt $\sigma \in \{(1\ 2), (1\ 3), (2\ 3)\}$ und wir erhalten drei Untergruppen der Ordnung 2:

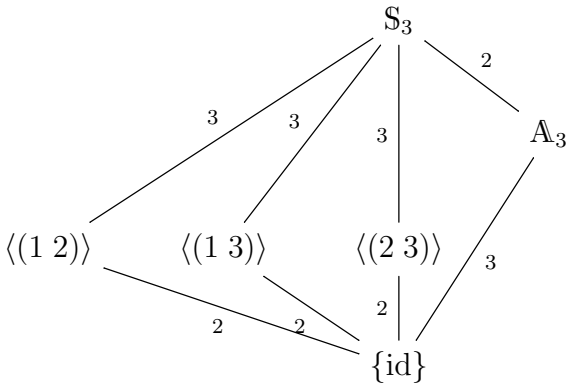
$$U = \{\mathrm{id}, (1\ 2)\} \text{ oder } U = \{\mathrm{id}, (1\ 3)\} \text{ oder } U = \{\mathrm{id}, (2\ 3)\}.$$

4. **Fall:** $|U| = 3$**:** Wie im dritten Fall gibt es ein $\mathrm{id} \neq \sigma \in U$ und $1 \neq o(\sigma) \mid |U| = 3$. Da auch 3 eine Primzahl ist gilt $o(\sigma) = 3$ und $U = \langle \sigma \rangle$. Dann ist aber $\sigma \in$

$\{(1\ 2\ 3), (1\ 3\ 2)\}$ und

$$U = \langle(1\ 2\ 3)\rangle = \langle(1\ 3\ 2)\rangle = \{\mathrm{id}, (1\ 2\ 3), (1\ 3\ 2)\} = \mathbb{A}_3.$$

Wir kennen mithin alle Untergruppen der $\mathbb{S}_3$ und können sie in folgendem *Untegruppendiagramm* festhalten:



Die Striche zwischen zwei Gruppen deuten an, daß die weiter oben stehende die weiter unten stehende enthält, und die Zahlen an den Strichen geben den Index der kleineren Gruppe in der größeren an.

C) Faktorgruppe

Jetzt sind wir in der Lage, eine Gruppenoperation auf G/U einzuführen. In der Hoffnung, daß die Notation $\overline{g}$ für die Linksnebenklasse gU einer Untergruppe das Rechnen mit den Elementen der Faktorgruppe G/U leichter macht, indem er verschleiert, daß das *Element* $\overline{g} = gU$ eigentlich eine *Menge* ist, werden wir den Satz mit dieser Notation formulieren. Wir beweisen den Satz zur Faktorgruppe nur für *abelsche* Gruppen.

Satz 24.14 (Die Faktorgruppe).
Sei $(G, \cdot)$ eine abelsche Gruppe und $U \leq G$ eine Untergruppe von G. Dann gilt[4]

$$\overline{g} \cdot \overline{h} = \overline{g \cdot h}, \quad \text{für } \overline{g}, \overline{h} \in G/U \tag{69}$$

und G/U ist bezüglich dieser Multiplikation eine abelsche Gruppe, die *Faktorgruppe* von G nach U.
Das *neutrale Element* von $(G/U, \cdot)$ ist die Linksnebenklasse $\overline{e} = U$, und das *Inverse* zu $\overline{g} = gU \in G/U$ ist die Linksnebenklasse $\overline{g^{-1}} = g^{-1}U$.
Außerdem ist die *Restklassenabbildung*

$$\pi : G \to G/U : g \mapsto \overline{g}$$

ein surjektiver Homomorphismus mit $\mathrm{Ker}(\pi) = U$.

[4]Dabei ist mit $\overline{g} \cdot \overline{h} = gU \cdot hU$ einfach das Produkt von Teilmengen von G gemeint, wie es in Notation 24.1 eingeführt wurde.

Beweis: Die Gleichheit in (69) folgt unmittelbar aus der Kommutativität der Gruppenoperation in G:

$$\overline{g} \cdot \overline{h} = g \cdot U \cdot h \cdot U = g \cdot h \cdot U \cdot U = g \cdot h \cdot U = \overline{gh},$$

da eine Untergruppe stets $U \cdot U = U$ gilt, wie man leicht sieht. Zeigen wir nun, daß G/U mit dieser Operation eine Gruppe ist.

Für $\overline{g}, \overline{h}, \overline{k} \in G/U$ folgt mittels der Assoziativität der Multiplikation in G:

$$\big(\overline{g} \cdot \overline{h}\big) \cdot \overline{k} = \overline{gh} \cdot \overline{k} = \overline{(gh)k} = \overline{g(hk)} = \overline{g} \cdot \overline{hk} = \overline{g} \cdot \big(\overline{h} \cdot \overline{k}\big).$$

Außerdem ist $\overline{e} \cdot \overline{g} = \overline{eg} = \overline{g}$, so daß $\overline{e}$ das Neutrale von G/U ist, und es gilt

$$\overline{g^{-1}} \cdot \overline{g} = \overline{g^{-1} \cdot g} = \overline{e},$$

und somit besitzt $\overline{g}$ ein Inverses, nämlich $\overline{g}^{-1} = \overline{g^{-1}}$. Also ist G/U eine Gruppe und die bezüglich des neutralen Elementes bzw. der Inversen getroffenen Aussagen sind ebenfalls gezeigt. Zudem ist G/U abelsch, da sich aus der Kommutativität der Multiplikation in G unmittelbar

$$\overline{g} \cdot \overline{h} = \overline{g \cdot h} = \overline{h \cdot g} = \overline{h} \cdot \overline{g}$$

für $\overline{g}, \overline{h} \in G/U$ ergibt.

Zudem folgt aus der Definition von π

$$\pi(gh) = \overline{gh} \overset{(69)}{=} \overline{g} \cdot \overline{h} = \pi(g) \cdot \pi(h)$$

und

$$\mathrm{Ker}(\pi) = \{g \in G \mid \pi(g) = \overline{e}\} = \{g \in G \mid \overline{g} = \overline{e}\} = \overline{e} = U,$$

so daß π ein Gruppenhomomorphismus mit $\mathrm{Ker}(\pi) = U$ ist. □

Bemerkung 24.15 (Rechenregeln in der Faktorgruppe).
Wir haben für den Beweis des Satzes ausgenutzt, daß das Produkt $\overline{g} \cdot \overline{h}$ ein Produkt von Teilmengen der Gruppe G ist. Diesen Umstand wollen wir nun tunlichst wieder *vergessen*! Wir merken uns nur: jedes Element $\overline{g}$ von G/U ist uns gegeben durch eine Repräsentanten und alle Operationen werden mittels dieser Repräsentanten ausgeführt. D.h. wir müssen uns nur folgende Regeln merken, um mit der Faktorgruppe rechnen zu können:

a. $\overline{g} \cdot \overline{h} = \overline{g \cdot h}$,
b. $\overline{g}^{-1} = \overline{g^{-1}}$,
c. $e_{G/U} = \overline{e} = \overline{u}$, wann immer $u \in U$.

Als unmittelbare Folgerung erhalten wir den Spezialfall $\mathbb{Z}_n$.

Korollar 24.16 ($\mathbb{Z}_n$ als Gruppe).
Für $n \in \mathbb{Z}$ ist $(\mathbb{Z}_n, +)$ eine abelsche Gruppe, wobei $\overline{x} + \overline{y} = \overline{x+y}$ für $x, y \in \mathbb{Z}$.

Bemerkung 24.17 (Rechnen in $\mathbb{Z}_n$ im Alltag).
Das Rechnen in $\mathbb{Z}_n$ für ausgewählte n ist uns seit unseren Kindertagen vertraut. Wenn unsere Uhr neun Uhr anzeigt, so wissen wir, daß es in fünf Stunden zwei Uhr ist. Wir rechnen dabei in $\mathbb{Z}_{12}$:

$$\overline{9} + \overline{5} = \overline{14} = \overline{2 + 1 \cdot 12} = \overline{2}.$$

Wer eine Uhr mit 24-Stundenrythmus bevorzugt, rechnet in $\mathbb{Z}_{24}$. Ist es jetzt neun Uhr, so war es vor 55 Stunden zwei Uhr:

$$\overline{9} - \overline{55} = \overline{-46} = \overline{2 - 2 \cdot 24} = \overline{2}.$$

Numeriert man die Wochentage von eins (Montag) bis sieben (Sonntag), so ist die Frage danach, welcher Wochentag in 51 Tagen ist, wenn heute Montag ist, eine Rechnung in $\mathbb{Z}_7$:

$$\overline{1} + \overline{51} = \overline{52} = \overline{3 + 7 \cdot 7} = \overline{3}.$$

In 51 Tagen ist also Mittwoch.

Um sich mit dem Rechnen in $\mathbb{Z}_n$ vertraut zu machen, empfehle ich, ein wenig mit Uhrzeiten und Wochentagen unter diesem Gesichtspunkt zu rechnen. □

Beispiel 24.18 (Verknüpfungstafeln).
Für Gruppen kleiner Ordnung, d. h. mit wenig Elementen, ist es sinnvoll sog. Verknüpfungstafeln aufzustellen, aus denen zu je zwei gegebenen Elementen die Verknüpfung der beiden Elemente abgelesen werden kann. Im Falle von $\mathbb{Z}_n$ erhalten wir für $n = 2, 3, 4$ die folgenden Verknüpfungstafeln:

$+$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{0}$	$\overline{1}$

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

□

Aufgaben

Aufgabe 24.19 (Produktformel).
Es seien $U, V \leq G$ Untergruppen der Gruppe $(G, \cdot)$.

a. Zeige, daß durch

$$(u, v) \sim (u', v') \iff u \cdot v = u' \cdot v'$$

eine Äquivalenzrelation auf der Menge $U \times V$ definiert wird.

b. Zeige, daß die Äquivalenzklasse von $(u, v) \in U \times V$ die Gestalt

$$\overline{(u, v)} = \left\{ \left(u \cdot y, y^{-1} \cdot v\right) \mid y \in U \cap V \right\}$$

besitzt und die Mächtigkeit $|U \cap V|$ hat.

c. Beweise, wenn U und V endlich sind, so gilt die Produktformel

$$|U \cdot V| = \frac{|U| \cdot |V|}{|U \cap V|}.$$

Bemerkung 24.20.
Die Formel in obiger Aufgabe ist besonders nützlich, wenn die Menge $U \cdot V$ eine Untergruppe von G ist. Das ist aber nicht immer der Fall, wie wir mit dem Satz von Lagrange leicht sehen können: das Produkt der Untergruppen $\langle(1\ 2)\rangle$ und $\langle(1\ 3)\rangle$ von $\mathbb{S}_3$ ist wegen der Aufgabe eine Teilmenge der Mächtigkeit 4 und kann nach dem Satz von Lagrange deshalb keine Untergruppe von $\mathbb{S}_3$ sein. Wir werden im folgenden Abschnitt (siehe Lemma A3.10) eine Bedingung kennen lernen, die V erfüllen muß, damit $U \cdot V$ eine Untergruppe von G wird.

Aufgabe 24.21.
Ist $(G, \cdot)$ eine Gruppe und $|G|$ eine Primzahl, so ist G zyklisch.

Aufgabe 24.22.
Bestimme alle Untergruppen der Gruppe $\mathbb{D}_8 = \langle(1\ 2\ 3\ 4), (2\ 4)\rangle$.

Aufgabe 24.23.
Bestimme alle Untergruppen der Gruppe $\mathbb{D}_{10} = \langle(1\ 2\ 3\ 4\ 5), (1\ 5) \circ (2\ 4)\rangle$.

Aufgabe 24.24.
Bestimme alle Untergruppen von $(\mathbb{Z}_{33}, +)$.

Aufgabe 24.25.
Betrachte für $m, n, a, b \in \mathbb{Z}_{>0}$ das Element $\big(\overline{a}_m, \overline{b}_n\big) \in \mathbb{Z}_m \times \mathbb{Z}_n$ in der Gruppe $(\mathbb{Z}_m \times \mathbb{Z}_n, +)$. Die Ordnung dieses Elementes läßt sich wie folgt berechnen

$$o\Big(\big(\overline{a}_m, \overline{b}_n\big)\Big) = \operatorname{kgv}\Big(o\big(\overline{a}_m\big), o\big(\overline{b}_n\big)\Big) = \operatorname{kgv}\left(\frac{\operatorname{kgv}(a,m)}{a}, \frac{\operatorname{kgv}(b,n)}{b}\right)$$

und ist ein Teiler von $\operatorname{kgv}(m,n)$. Insbesondere ist $\mathbb{Z}_m \times \mathbb{Z}_n$ nicht zyklisch, wenn m und n nicht teilerfremd sind.

Aufgabe 24.26.
Berechne die Ordnung von $\big(\overline{6}_{21}, \overline{9}_{33}\big) \in \mathbb{Z}_{21} \times \mathbb{Z}_{33}$.

Aufgabe 24.27.
Es sei σ und π zwei disjunkte Zyklen in $\mathbb{S}_n$ der Länge k bzw. l. Zeige, daß

$$o(\sigma \circ \pi) = \operatorname{kgv}(k,l).$$

Es ist ein wichtiges Prinzip, daß ein injektiver Homomorphismus alle guten Eigenschaften einer Gruppe bzw. ihrer Elemente erhält. Die Ordnung eines Elementes ist ein Beispiel für dieses Prinzip.

Aufgabe 24.28.
Es sei $\alpha : (G, \cdot) \longrightarrow (H, *)$ ein Gruppenhomomorphismus und $g \in G$. Wir nennen die Mächtigkeit des Erzeugnisses von g die *Ordnung* von g und bezeichnen sie mit $o(g) := |\langle g \rangle|$. Zeige:

a. Ist $o(g) < \infty$, so ist $o(g)$ ein Vielfaches von $o\big(\alpha(g)\big)$.

b. Zeige, α ist genau dann injektiv, wenn $o(g) = o\big(\alpha(g)\big)$ für alle $g \in G$ gilt.

§ 25 Ringe und Körper

In Kapitel 22 haben wir die mathematische Struktur der *Gruppe* eingeführt, und unsere ersten Beispiele waren die additiven Gruppen $(\mathbb{Z}, +)$ und $(\mathbb{Q}, +)$ der ganzen bzw. der rationalen Zahlen. Auf diesen Mengen haben wir aber neben der Addition jeweils noch eine zweite zweistellige Operation, die Multiplikation, bezüglich derer in beiden Mengen wiederum interessante Rechenregeln gelten. So ist die Menge $(\mathbb{Q} \setminus \{0\}, \cdot)$ wieder eine Gruppe, während $(\mathbb{Z} \setminus \{0\}, \cdot)$ zu dieser Eigenschaft (nur) die multiplikativen Inversen fehlen. Wir wollen diese Beispiele nun verallgemeinern und wiederholen dabei auch einige Begriffe und Ergebnisse aus Abschnitt 7.

A) Ringe

Definition 25.1 (Ring).

a. Ein *Ring mit Eins* ist ein Tripel $(R, +, \cdot)$ bestehend aus einer Menge R zusammen mit zwei zweistelligen Operationen

$$+ : R \times R \to R : (a, b) \mapsto a + b, \qquad (\text{“}\textit{Addition}\text{”})$$

und

$$\cdot : R \times R \to R : (a, b) \mapsto a \cdot b, \qquad (\text{“}\textit{Multiplikation}\text{”})$$

so, daß folgende Axiome erfüllt sind:

(i) $(R, +)$ ist eine abelsche Gruppe (mit neutralem Element 0_R).

(ii) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ für alle $a, b, c \in R$. (“*Assoziativität der Multiplikation*”)

(iii) Es gibt ein Element $1_R \in R$ mit $1_R \cdot a = a \cdot 1_R = a$ für alle $a \in R$. (“*Einselement*”)

(iv) $a \cdot (b + c) = a \cdot b + a \cdot c$ und $(b + c) \cdot a = b \cdot a + c \cdot a$ für alle $a, b, c \in R$. (“*Distributivität*”)

b. Ein Ring mit Eins $(R, +, \cdot)$ heißt *kommutativ*, falls $a \cdot b = b \cdot a$ für alle $a, b \in R$.

c. Ist $(R, +, \cdot)$ ein Ring mit Eins, dann heißt $a \in R$ eine *Einheit* oder *invertierbar* in R, falls es ein $a^{-1} \in R$ gibt mit $a \cdot a^{-1} = a^{-1} \cdot a = 1_R$. Wir bezeichnen mit

$$R^* = \{a \in R \mid a \text{ ist Einheit}\}$$

die Menge der Einheiten von R. Offenbar ist $(R^*, \cdot)$ eine Gruppe, die sogenannte *Einheitengruppe* von R.

Bemerkung 25.2.

Wir werden in Ringen für die Addition stets das Zeichen $+$ und für die Multiplikation das Zeichen $\cdot$ verwenden, auch wenn wir gleichzeitig verschiedene Ringe betrachten. Wir verzichten im Folgenden deshalb darauf, die Ringoperationen jeweils anzugeben und nennen verkürzend die dem Ring $(R, +, \cdot)$ zugrunde liegende Menge R einen Ring. Zudem werden wir statt $a \cdot b$ oft auch nur ab schreiben.

Das neutrale Element von $(R, +)$ werden wir mit 0_R oder einfach mit 0 bezeichnen und das *Nullelement* von R nennen; das Einselement bezeichnen wir mit 1_R oder 1.

Ist R ein Ring und sind $a, b \in R$, so schreiben wir statt $a + (-b)$ auch kurz $a - b$.

Da $(R^*, \cdot)$ eine Gruppe ist, sind das neutrale Element 1_R der Multiplikation und das Inverse eines Elementes $a \in R^*$ eindeutig bestimmt. □

Beispiel 25.3.

a. $(\mathbb{Z}, +, \cdot)$ mit der üblichen Addition und Multiplikation ist ein kommutativer Ring mit Eins.

b. Ist M eine beliebige Menge und $(R, +, \cdot)$ ein Ring mit Eins, so ist

$$R^M := \{f \mid f : M \to R \text{ ist eine Abbildung}\}$$

mit den punktweise definierten Operationen

$$+ : R^M \times R^M \to R^M : (f, g) \mapsto \big(f + g : M \to R : x \mapsto f(x) + g(x)\big),$$

und

$$\cdot : R^M \times R^M \to R^M : (f, g) \mapsto \big(f \cdot g : M \to R : x \mapsto f(x) \cdot g(x)\big),$$

ein Ring mit der Nullfunktion $0 : M \to R : x \mapsto 0_R$ als neutralem Element der Addition und der Einsfunktion $1 : M \to R : x \mapsto 1_R$ als Einselement, wie man mit etwas Fleiß nachprüft.

c. In Aufgabe 22.32 haben wir die Menge

$$\mathrm{Mat}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\middle|\; a, b, c, d \in \mathbb{R} \right\}$$

der reellen 2×2-Matrizen eingeführt und für zwei Matrizen

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$$

ihr Produkt als

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a \cdot a' + b \cdot c' & a \cdot b' + b \cdot d' \\ c \cdot a' + d \cdot c' & c \cdot b' + d \cdot d' \end{pmatrix}$$

definiert. Setzen wir zudem

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a+a' & b+b' \\ c+c' & d+d' \end{pmatrix},$$

so rechnet man mit etwas Geduld nach, daß $(\mathrm{Mat}_2(\mathbb{R}), +, \cdot)$ ein Ring mit Einselement

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

ist. Dieser Ring ist nicht-kommutativ, da

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Wir werden dieses Beispiel später auf Matrizen der Größe $n \times n$ über einem beliebigen Körper K verallgemeinern, und der Nachweis, daß auf diesem Wege ein nicht-kommutativer Ring mit Eins entsteht, wird dort auf weit geschickterem Weg als durch langatmiges Nachrechnen geführt.

d. Sei R ein kommutativer Ring mit Eins und t eine Veränderliche. Einen formalen Ausdruck der Form

$$\sum_{k=0}^{\infty} a_k \cdot t^k$$

mit $a_k \in R$ nennen wir eine *formale Potenzreihe* mit Koeffizienten in R und wir bezeichnen die Menge aller solcher Potenzreihen mit

$$R[[t]] := \left\{ \sum_{k=0}^{\infty} b_k \cdot t^k \;\middle|\; b_k \in R \right\}.$$

Für zwei formale Potenzreihen $\sum_{i=0}^{\infty} a_i \cdot t^i, \sum_{j=0}^{\infty} b_j \cdot t^j \in R[[t]]$ definieren wir ferner

$$\sum_{i=0}^{\infty} a_i \cdot t^i \;+\; \sum_{i=0}^{\infty} b_i \cdot t^i \;:=\; \sum_{i=0}^{\infty} (a_i + b_i) \cdot t^i \in R[[t]]$$

und angelehnt an das Cauchy-Produkt absolut konvergenter Reihen (siehe Satz 12.27)

$$\sum_{i=0}^{\infty} a_i \cdot t^i \;\cdot\; \sum_{j=0}^{\infty} b_j \cdot t^j \;:=\; \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i \cdot b_j \right) \cdot t^k \in R[[t]].$$

Dann ist $R[[t]]$ mit diesen beiden Operationen ein Ring mit Eins $1_{R[[t]]} = t^0$, der *Ring der formalen Potenzreihen* über R in der Unbestimmten t.
Man beachte, daß aus der Definition unmittelbar folgt, daß

$$\sum_{i=0}^{\infty} a_i \cdot t^i \;=\; \sum_{i=0}^{\infty} b_i \cdot t^i \quad\Longleftrightarrow\quad a_i = b_i \;\forall\, i \in \mathbb{N}.$$

Gilt $a_i = 0$ für $i \geq n$, so schreiben wir auch abkürzend

$$\sum_{i=0}^{n} a_i \cdot t^i = \sum_{i=0}^{\infty} a_i \cdot t^i.$$

Im obigen Beispiel ist der Nachweis des Assoziativgesetzes der Multiplikation im Potzenreihenring $R[[t]]$ nicht ganz offensichtlich. Wir führen für den interessierten Leser deshalb die Ringeigenschaften für $R[[t]]$ hier aus.

Beweis, daß der Potenzreihenring ein Ring ist: Nach Definition sind $+$ und $\cdot$ zwei zweistellige Operationen auf $R[[t]]$. Seien also $\sum_{i=0}^{\infty} a_i \cdot t^i, \sum_{i=0}^{\infty} b_i \cdot t^i, \sum_{i=0}^{\infty} c_i \cdot t^i \in R[[t]]$ gegeben. Dann gilt wegen der Assoziativität der Addition in R

$$\begin{aligned}\left(\sum_{i=0}^{\infty} a_i \cdot t^i + \sum_{i=0}^{\infty} b_i \cdot t^i\right) + \sum_{i=0}^{\infty} c_i \cdot t^i &= \sum_{i=0}^{\infty} \big((a_i + b_i) + c_i\big) \cdot t^i \\ &= \sum_{i=0}^{\infty} \big(a_i + (b_i + c_i)\big) \cdot t^i = \sum_{i=0}^{\infty} a_i \cdot t^i + \left(\sum_{i=0}^{\infty} b_i \cdot t^i + \sum_{i=0}^{\infty} c_i \cdot t^i\right)\end{aligned}$$

und wegen der Kommutativität der Addition in R

$$\begin{aligned}\sum_{i=0}^{\infty} a_i \cdot t^i + \sum_{i=0}^{\infty} b_i \cdot t^i = \sum_{i=0}^{\infty} (a_i + b_i) \cdot t^i \\ = \sum_{i=0}^{\infty} (b_i + a_i) \cdot t^i = \sum_{i=0}^{\infty} b_i \cdot t^i + \sum_{i=0}^{\infty} a_i \cdot t^i.\end{aligned}$$

Zudem gilt für die Nullfunktion $0_{R[[t]]} = \sum_{i=0}^{\infty} 0 \cdot t^i$

$$0_{R[[t]]} + \sum_{i=0}^{\infty} a_i \cdot t^i = \sum_{i=0}^{\infty} (0 + a_i) \cdot t^i = \sum_{i=0}^{\infty} a_i \cdot t^i,$$

und für $\sum_{i=0}^{\infty} (-a_i) \cdot t^i \in R[[t]]$ gilt

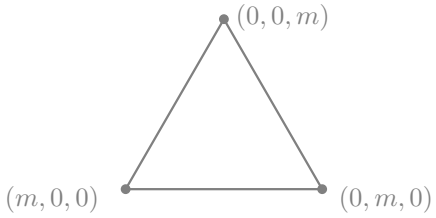
$$\sum_{i=0}^{\infty} (-a_i) \cdot t^i + \sum_{i=0}^{\infty} a_i \cdot t^i = \sum_{i=0}^{\infty} (-a_i + a_i) \cdot t^i = 0_{R[[t]]},$$

so daß $(R[[t]], +)$ eine abelsche Gruppe mit der Nullfunktion als neutralem Element ist. Man beachte nun, daß

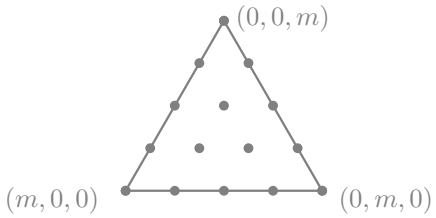
$$\sum_{k+l=m} \left(\sum_{i+j=k} a_i \cdot b_j\right) \cdot c_l = \sum_{i+j+l=m} a_i \cdot b_j \cdot c_l = \sum_{i+k=m} \left(a_i \cdot \sum_{j+l=k} b_j \cdot c_l\right), \tag{70}$$

da unter jeder der Summen jedes der Tripel (i, j, l) natürlicher Zahlen mit der Eigenschaft $i + j + l = m$ genau einmal vor kommt[5] und da in R das Assoziativgesetz der

[5] Man kann sich dies auch geometrisch veranschaulichen. Fassen wir (i, j, l) als Koordinaten des dreidimensionalen Raumes $\mathbb{R}^3$ auf, so bestimmt die Gleichung $i + j + l = m$ bei vorgegebenem m eine Ebene im Raum, nämlich die Ebene, die durch die drei Punkte $(m, 0, 0)$, $(0, m, 0)$ und $(0, 0, m)$ aufgespannt wird. Verbinden wir diese drei Punkte in dieser Ebene, so erhalten wir ein Dreieck:



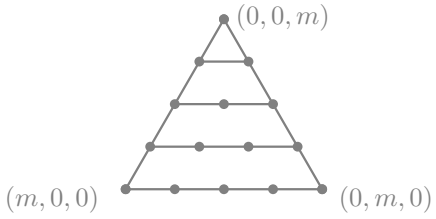
Die Punkte mit ganzzahligen Koordinaten in diesem Dreieck sind genau die Tripel nicht-negativer ganzer Zahlen, deren Summe m ist:



In der linken Seite von (70) zerlegen wir diese Menge wie folgt:

$$\bigcup_{l=0}^{m} \bigcup_{i+j=m-l} \{(i,j,l)\}.$$

In der inneren Summe werden also diejenigen ganzzahligen (i,j,l) in dem Dreieck zusammen gefaßt, für die die Koordinate l konstant ist und für die $i+j=m-l$ ist, d.h. die Punkte liegen auf einer Geraden parallel zur Geraden durch $(m,0,0)$ und $(0,m,0)$:



In der rechten Seite von (70) zerlegen wir diese Menge wie folgt:

$$\bigcup_{i=0}^{m} \bigcup_{j+l=m-i} \{(i,j,l)\}.$$

In der inneren Summe werden also diejenigen ganzzahligen (i,j,l) in dem Dreieck zusammen gefaßt, für die die Koordinate i konstant ist und für die $j+l=m-i$ ist, d.h. die Punkte liegen auf einer Geraden parallel zur Geraden durch $(0,m,0)$ und $(0,0,m)$:

Multiplikation und das Distributivgesetz gelten. Damit gilt aber

$$\begin{aligned}\left(\sum_{i=0}^{\infty} a_i \cdot t^i \;\cdot\; \sum_{j=0}^{\infty} b_j \cdot t^j\right) \;\cdot\; \sum_{l=0}^{\infty} c_l \cdot t^l &= \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i \cdot b_j\right) \cdot t^k \;\cdot\; \sum_{l=0}^{\infty} c_l \cdot t^l \\ &= \sum_{m=0}^{\infty} \left(\sum_{k+l=m} \left(\sum_{i+j=k} a_i \cdot b_j\right) \cdot c_l\right) \cdot t^m \\ &= \sum_{m=0}^{\infty} \left(\sum_{i+k=m} a_i \cdot \sum_{j+l=k} b_j \cdot c_l\right) \cdot t^m \\ &= \sum_{i=0}^{\infty} a_i \cdot t^i \;\cdot\; \sum_{k=0}^{\infty} \left(\sum_{j+l=k} b_j \cdot c_l\right) \cdot t^k \\ &= \sum_{i=0}^{\infty} a_i \cdot t^i \;\cdot\; \left(\sum_{j=0}^{\infty} b_j \cdot t^j \;\cdot\; \sum_{l=0}^{\infty} c_l \cdot t^l\right),\end{aligned}$$

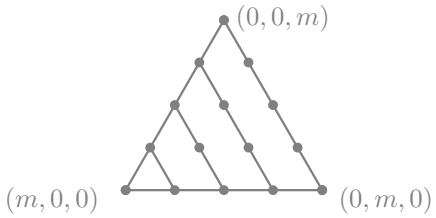
so daß die Multiplikation auf $R[[t]]$ assoziativ ist. Ferner folgt aus der Kommutativität der Multiplikation auf R unmittelbar

$$\begin{aligned}\sum_{i=0}^{\infty} a_i \cdot t^i \;\cdot\; \sum_{j=0}^{\infty} b_j \cdot t^j &= \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i \cdot b_j\right) \cdot t^k \\ &= \sum_{k=0}^{\infty} \left(\sum_{j+i=k} b_j \cdot a_i\right) \cdot t^k = \sum_{j=0}^{\infty} b_j \cdot t^j \;\cdot\; \sum_{i=0}^{\infty} a_i \cdot t^i.\end{aligned}$$

Und schließlich gilt für $1_{R[[t]]} = t^0 = \sum_{j=0}^{\infty} e_j \cdot t^j$ mit $e_0 = 1$ und $e_j = 0$ für $j \geq 1$:

$$1_{R[[t]]} \cdot \sum_{i=0}^{\infty} a_i \cdot t^i = \sum_{k=0}^{\infty} \left(\sum_{j+i=k} e_j \cdot a_i\right) \cdot t^k = \sum_{k=0}^{\infty} a_k \cdot t^k,$$

so daß t^0 unter Ausnutzung der Kommutativität der Multiplikation das Einselement von $R[[t]]$ ist.



In beiden Fällen wird jedes ganzzahlige Tripel (i, j, l) im Dreieck genau einmal betrachtet.

Es bleibt nur, die Distributivität zu zeigen:

$$\begin{aligned}\sum_{i=0}^{\infty} a_i \cdot t^i \cdot \left(\sum_{j=0}^{\infty} b_j \cdot t^j + \sum_{j=0}^{\infty} c_j \cdot t^j\right) &= \sum_{i=0}^{\infty} a_i \cdot t^i \cdot \sum_{j=0}^{\infty} (b_j + c_j) \cdot t^j \\ &= \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i \cdot (b_j + c_j)\right) \cdot t^k \\ &= \sum_{k=0}^{\infty} \left(\sum_{i+j=k} (a_i \cdot b_j + a_i \cdot c_j)\right) \cdot t^k \\ &= \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i \cdot b_j\right) \cdot t^k + \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i \cdot c_j\right) \cdot t^k \\ &= \sum_{i=0}^{\infty} a_i \cdot t^i \cdot \sum_{j=0}^{\infty} b_j \cdot t^j + \sum_{i=0}^{\infty} a_i \cdot t^i \cdot \sum_{j=0}^{\infty} c_j \cdot t^j.\end{aligned}$$

Das zweite Distributivgesetz folgt mittels der Kommutativität der Multiplikation.

Damit haben wir gezeigt, daß $(R[[t]], +, \cdot)$ ein kommutativer Ring mit Eins ist. □

B) Rechenregeln in Ringen

Wir wollen nun einige Rechenregeln für das Rechnen in Ringen aufstellen.

Lemma 25.4 (Rechenregeln).
Es sei R ein Ring mit Eins. Für $a, b, c \in R$ gelten:

a. $-(-a) = a$.

b. $0 \cdot a = a \cdot 0 = 0$.

c. $(-a) \cdot b = a \cdot (-b) = -(a \cdot b)$.

d. $a \cdot (b - c) = a \cdot b - a \cdot c$.

e. Für $a \in R^*$ ist $a^{-1} \in R^*$ und $\left(a^{-1}\right)^{-1} = a$.

f. Ist $1_R = 0_R$, so ist $R = \{0_R\}$ der *Nullring*.

Beweis: Die Aussage a. folgt unmittelbar aus Lemma 22.6.

b. Für $a \in R$ gilt $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$, also folgt $0 \cdot a = 0$ mittels der Kürzungsregeln in $(R, +)$. Analog sieht man $a \cdot 0 = 0$.

c. Für $a, b \in R$ gilt wegen b.:

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

also $-(a \cdot b) = (-a) \cdot b$. Die Gleichheit des Ausdrucks zu $a \cdot (-b)$ folgt analog.

d. Für $a, b \in R$ folgt unter Zuhilfenahme von a. und c.:

$$(-a) \cdot (-b) = -\big(a \cdot (-b)\big) = -\big(-(a \cdot b)\big) = a \cdot b.$$

e. Für $a, b, c \in R$ impliziert c.:

$$a \cdot (b - c) = a \cdot b + a \cdot (-c) = a \cdot b + \big(-(a \cdot c)\big) = a \cdot b - a \cdot c.$$

f. Ist $a \in R^*$ eine Einheit mit Inversem a^{-1}. Dann ist nach Definition a ein Inverses von a^{-1}, und insbesondere ist a^{-1} eine Einheit. Aus der Eindeutigkeit des Inversen (siehe Bemerkung 25.2) folgt dann, daß $a = (a^{-1})^{-1}$.

g. Ist $a \in R$, so gilt $a = 1_R \cdot a = 0_R \cdot a = 0_R$.

□

Die folgende Proposition ist ein praktisches Hilfsmittel, um neue Ringe oder Körper als Teilmengen bereits bekannter Ringe oder Körper zu finden.

Definition und Proposition 25.5 (Unterringe).
Sei R ein Ring mit Eins. Eine nicht-leere Teilmenge $S \subseteq R$ heißt ein *Unterring* oder *Teilring* von R, wenn folgende Bedingungen gelten:

- $1_R \in S$,
- $a + b \in S$ für alle $a, b \in S$,
- $-a \in S$ für alle $a \in S$ und
- $a \cdot b \in S$ für alle $a, b \in S$.

Ist S ein Unterring von R, so ist $(S, +, \cdot)$ selbst wieder ein Ring.

Beweis: Zunächst ist $(S, +)$ wegen des Untergruppenkriteriums eine Untergruppe von $(R, +)$. Die übrigen Axiome eines Ringes werden hier entweder gefordert oder übertragen sich unmittelbar, weil sie schon in der größeren Menge R gelten. □

Beispiel 25.6 (Unterring).
Die Menge der komplexen Zahlen $\mathbb{C}$ ist mit der üblichen Addition und Multiplikation ein kommutativer Ring mit Eins. Die Teilmenge der *ganzen Gaußschen Zahlen*

$$\mathbb{Z}[i] = \{m + n \cdot i \in \mathbb{C} \mid m, n \in \mathbb{Z}\}$$

ist dann ein Unterring von $\mathbb{C}$, denn für $m + n \cdot i, m' + n'\dot{i} \in \mathbb{Z}[i]$ gilt:

- $1 = 1 + 0 \cdot i \in \mathbb{Z}[i]$,
- $(m + n \cdot i) + (m' + n' \cdot i) = (m + m') + (n + n') \cdot i \in \mathbb{Z}[i]$,
- $-(m + n \cdot i) = (-m) + (-n) \cdot i \in \mathbb{Z}[i]$ und
- $(m + n \cdot i) \cdot (m' + n' \cdot i) = (m \cdot m' - n \cdot n') + (m \cdot n' + n \cdot m') \cdot i \in \mathbb{Z}[i]$.

Also ist $\mathbb{Z}[i]$ mit der üblichen Addition und Multiplikation komplexer Zahlen ein Ring.

C) Körper

Definition 25.7 (Körper).
Ein kommutativer Ring mit Eins $(R, +, \cdot)$ heißt *Körper*, falls $R^* = R \setminus \{0\}$.

Bemerkung 25.8 (Körper).
Unter Beachtung der Rechenregeln 25.4 kann man leicht zeigen, daß eine Tripel $(K, +, \cdot)$ genau dann ein *Körper* ist, wenn gilt (siehe auch Definition 7.5):

a. $(K, +)$ ist eine abelsche Gruppe.

b. $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe.

c. Für alle $x, y, z \in K$ gilt $x \cdot (y + z) = x \cdot y + x \cdot z$ und $(y + z) \cdot x = y \cdot x + z \cdot x$.

Beispiel 25.9 (Körper).
$(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ und $(\mathbb{C}, +, \cdot)$ mit der üblichen Addition und Multiplikation sind Körper, $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer Ring mit Eins, der kein Körper ist.

Lemma 25.10 (Nullteilerfreiheit von Körpern).
Ist $(K, +, \cdot)$ ein Körper und sind $a, b \in K \setminus \{0\}$, dann gilt $a \cdot b \neq 0$.

Beweis: Nach Bemerkung 25.8 ist $(K \setminus \{0\}, \cdot)$ eine Gruppe, woraus insbesondere die Abgeschlossenheit bezüglich der Multiplikation folgt. □

D) Der Ring $\mathbb{Z}_n$

Satz 25.11 ($\mathbb{Z}_n$ als kommutativer Ring mit Eins).
Die Menge

$$\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \ldots, \overline{n-1}\}$$

der Linksnebenklassen von $\mathbb{Z}$ modulo $n\mathbb{Z}$ wird durch die Operatioen

$$\overline{a} + \overline{b} := \overline{a + b}$$

und

$$\overline{a} \cdot \overline{b} := \overline{a \cdot b}$$

zu einem kommutativen Ring mit Eins $\overline{1}$.

Beweis: Wir haben in Proposition 24.6 schon gesehen, daß $(\mathbb{Z}_n, +)$ eine abelsche Gruppe ist.

Wir müssen nun als nächstes zeigen, daß die Multiplikation wohldefiniert ist. Seien dazu $\overline{a} = \overline{c}$ und $\overline{b} = \overline{d}$ gegeben. Dann gibt es zwei ganze Zahlen $x, y \in \mathbb{Z}$ mit

$$a = c + x \cdot n$$

und

$$b = d + y \cdot n.$$

Mithin gilt

$$a \cdot b = c \cdot d + n \cdot (cy + xd + xyn) \equiv c \cdot d \pmod{n}$$

und somit

$$\overline{a \cdot b} = \overline{c \cdot d}.$$

Die Multiplikation ist also wohldefiniert. Die restlichen Ringaxiome übertragen sich dann unmittelbar vom Ring der ganzen Zahlen auf $\mathbb{Z}_n$. □

Beispiel 25.12 ($\mathbb{Z}_6$ ist kein Körper).
Der Ring $\mathbb{Z}_6$ ist kein Körper, weil die Null als Produkt

$$\overline{2}_6 \cdot \overline{3}_6 = \overline{6}_6 = \overline{0}_6$$

von zwei Zahlen geschrieben werden kann, die beide nicht null sind (siehe Lemma 25.10). Man sagt auch, der Ring ist nicht nullteilerfrei.

Korollar 25.13 ($\mathbb{Z}_n$ als Körper).
$\mathbb{Z}_n$ ist genau dann ein Körper, wenn p eine Primzahl ist.

Beweis: Ist n keine Primzahl, so können wir n als Produkt

$$n = a \cdot b$$

von zwei Zahlen $1 < a, b < n$ schreiben. Wie im Beispiel oben gilt dann

$$\overline{a}_n \neq \overline{0}_n \neq \overline{b}_n$$

und die Null lässt sich wieder als Produkt

$$\overline{a}_n \cdot \overline{b}_n = \overline{a \cdot b}_n = \overline{n}_n = \overline{0}_n$$

von zwei Zahlen ungleich null schreiben. Also kann $\mathbb{Z}_n$ dann kein Körper sein (siehe Lemma 25.10).

Sei nun n eine Primzahl und $\overline{0}_n \neq \overline{a}_n \in \mathbb{Z}_n$ mit $1 \leq a < n$. Es bleibt zu zeigen, daß ein $\overline{b}_n \in \mathbb{Z}_n$ existiert mit

$$\overline{a}_n \cdot \overline{b}_n = \overline{1}_n.$$

Dazu betrachten wir die Abbildung

$$\alpha : \mathbb{Z}_n \longrightarrow \mathbb{Z}_n : \overline{x}_n \mapsto \overline{a}_n \cdot \overline{x}_n = \overline{a \cdot x}_n.$$

Angenommen, die Abbildung wäre nicht injektiv, dann würde es Zahlen

$$0 \leq x < y \leq n-1$$

geben mit

$$\overline{ax}_n = \alpha(\overline{x}_n) = \alpha(\overline{y}_n) = \overline{ay}_n$$

und damit

$$a \cdot (y-x) = ay - ax \equiv 0 \pmod{n}.$$

Das heißt, $a \cdot (y-x)$ wäre ein Vielfaches von n, und da n eine Primzahl ist, müsste n dann in der Primfaktorzerlegung von $a \cdot (y-x)$ vorkommen. Das ist wegen $1 \leq a \leq n-1$ und $1 \leq y - x \leq n-1$ aber nicht möglich. Also ist α injektiv.

Als injektive Selbstabbildung einer endlichen Menge ist α dann nach Korollar 5.6 auch surjektiv und mithin gibt es ein $\overline{b}_n$ mit

$$\overline{1}_n = \alpha(\overline{b}_n) = \overline{a}_n \cdot \overline{b}_n.$$

□

Beispiel 25.14 (Die Körper $\mathbb{Z}_p$ und $\mathbb{F}_p$ sind identisch.).
Der Körper $\mathbb{Z}_2$ enthält nur zwei Elemente, das Neutrale der Addition $\overline{0}_2$ und das Neutrale der Multiplikation $\overline{1}_2$. Der Körper mit zwei Elementen ist bis auf Isomorphie eindeutig bestimmt. Wir haben ihn bereits in Beispiel 7.6 kennengelernt und haben ihn dort $\mathbb{F}_2$ genannt.

Allgemeiner haben wir in Beispiel 7.6 für jede Primzahl p einen Körper $\mathbb{F}_p$ mit genau p Elmenten eingeführt und mit Hilfe von Division mit Rest durch p eine Addition und Multiplikation definiert. Ein nur etwas genauerer Blick zeigt, dass es sich dabei um den Körper $\mathbb{Z}_p$ handelt, wenn man die Zahl a in $\mathbb{F}_p$ mit der Restklasse $\overline{a}_p$ in $\mathbb{Z}_p$ identifiziert. Damit haben wir jetzt auch gezeigt, dass $\mathbb{F}_p$ in der Tat ein Körper ist.

Aufgaben

Aufgabe 25.15.
Es sei R ein kommutativer Ring mit Eins und $f = \sum_{k=0}^{\infty} a_k \cdot t^k \in R[[t]]$ eine formale Potenzreihe über R. Zeige, f ist genau dann eine Einheit in $R[[t]]$, wenn a_0 eine Einheit in R ist.

Hinweis, wenn a_0 eine Einheit in R ist, so ist eine Reihe $g = \sum_{k=0}^{\infty} b_k \cdot t^k$ mit $f \cdot g = t^0$ gesucht. Multipliziere die linke Seite der Gleichung aus und löse die Gleichungen, die sich für die Koeffizienten ergeben rekursiv.

Aufgabe 25.16.
Bestimme die Einheitengruppe $\mathbb{Z}[i]^*$ des Ringes $\mathbb{Z}[i]$.

Aufgabe 25.17.
Für $\omega \in \mathbb{Z}$, $\omega \geq 2$, bezeichnen wir mit $\sqrt{-\omega}$ die komplexe Zahl $i \cdot \sqrt{\omega}$.

a. Zeige, $\mathbb{Z}\left[\sqrt{-\omega}\right] := \{a + b \cdot \sqrt{-\omega} \in \mathbb{C} \mid a, b \in \mathbb{Z}\}$ ist ein kommutativer Ring mit Eins, wobei die Addition und die Multiplikation einfach die Addition und Multiplikation komplexer Zahlen sein sollen.

b. Zeige, $\mathbb{Z}\left[\sqrt{-\omega}\right]^* = \{1, -1\}$.

§ 26 Der Polynomring $K[t]$

In diesem Abschnitt sei K stets ein Körper.

A) Grundlegende Eigenschaften des Polynomrings

Wir verallgemeinern hier zunächst den Begriff des Polynoms, den wir in Definition 13.11 über den reellen Zahlen eingeführt haben. Auf Beweise werden wir weitgehend verzichten und verweisen für diese auf den Anhang.

Definition 26.1 (Der Polynomring).
Ein *Polyom* in der Unbestimmten t mit Kooeffienten in K ist eine formale Potenzreihe der Form

$$\sum_{k=0}^{n} a_k \cdot t^k = a_n \cdot t^n + a_{n-1} \cdot t^{n-1} + \ldots + a_1 \cdot t^1 + a_0 \cdot t^0 \in K[[t]],$$

d.h. eine Potenzreihe, bei der nur endlich viele Koeffitienten ungleich null sind. Die Menge

$$K[t] = \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; a_k \in K, n \in \mathbb{N} \right\}$$

aller Polynome in der Unbestimmten t mit Koeffizienten in K heißt der *Polynomring* in der Unbestimmten t über dem Körper K.

Bemerkung 26.2 (Rechenoperationen im Polynomring).
In Beispiel 25.3 haben wir eine Addition und eine Multiplikation für Potenzreihen eingeführt, die wir hier nochmal für Polynome formulieren wollen.

Für zwei Polynome $f = \sum_{k=0}^{n} a_k \cdot t^k \in K[t]$ und $g = \sum_{k=0}^{m} b_k \cdot t^k \in K[t]$ gelten

$$f + g = \sum_{k=0}^{\max\{m,n\}} (a_k + b_k) \cdot t^k \tag{71}$$

mit $a_k = 0$ für $k > n$ und $b_k = 0$ für $k > m$ sowie

$$f \cdot g = \sum_{k=0}^{m+n} \left(\sum_{i=0}^{k} a_i \cdot b_{k-i} \right) \cdot t^k = \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i \cdot b_j \right) \cdot t^k \tag{72}$$

wobei $a_i = 0$ für $i > n$ und $b_j = 0$ für $j > m$.

Ein Skalar $\lambda \in K$ ist ein konstantes Polynom und für das Produkt mit λ gilt dann

$$\lambda \cdot f = \sum_{k=0}^{n} (\lambda \cdot a_k) \cdot t^k.$$

Dies wird im Zusammenhang mit Vektorräumen interessant sein.

Wir erinnern zudem daran: für zwei Polynome $f = \sum_{k=0}^{n} a_k \cdot t^k$ und $g = \sum_{k=0}^{m} b_k \cdot t^k$ gilt

$$f = g \iff a_k = b_k \quad \forall\, k = 0, \ldots, \max\{m, n\}, \tag{73}$$

wobei $a_k = 0$ für $k > n$ und $b_k = 0$ für $k > m$. Wir sagen, zwei Polynome sind gleich, wenn ihre Koeffizienten übereinstimmen und sprechen dabei vom *Koeffizientenvergleich*.

Proposition 26.3 (Der Polynomring als kommutativer Ring mit Eins.).
Der Polynomring $K[t]$ ist ein Unterring von $K[[t]]$ und ist damit insbesondere selbst ein kommutativer Ring mit Eins $1_{K[t]} = t^0$.

Beweis: $K[t]$ ist eine Teilmenge des Potenzreihenrings $K[[t]]$ und erfüllt offensichtlich die Bedingungen von Proposition 25.5. Mithin ist $K[t]$ selbst ein kommutativer Ring mit Eins. □

Bemerkung 26.4 (t^0 wird nicht geschrieben.).
Wir lassen bei der Schreibweise eines Polynoms

$$a_n \cdot t^n + \ldots + a_1 \cdot t^1 + a_0 \cdot t^0 = a_n \cdot t^n + \ldots + a_1 \cdot t + a_0$$

beim konstanten Anteil meist das t^0 weg. Der Körper K ist dann als Teilmenge der konstanten Polynome ein Unterring $K[t]$.

Definition 26.5 (Der Grad eines Polynoms).
Sei $f = \sum_{k=0}^{n} a_k t^k \in K[t]$ mit $a_n \neq 0$, dann heißt

$$\deg(f) := n$$

der *Grad* von f und

$$\mathrm{lc}(f) := a_n$$

der *Leitkoeffizient* von f. Zudem setzen wir $\deg(0) := -\infty$ und $\mathrm{lc}(0) := 0$.
Ist $\mathrm{lc}(f) = 1$ oder $f = 0$, so nennen wir f *normiert*.

Bemerkung 26.6 (Konstante Polynome).
Beachte, ein Polynom f ist genau dann konstant, wenn $\deg(f) \leq 0$.

Lemma 26.7 (Gradformeln).
Seien $f, g \in K[t]$. Dann gelten:

a. $\deg(f + g) \leq \max\{\deg(f), \deg(g)\}$.

b. $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Beweis: Ist $f = 0$ oder $g = 0$, so sind die Aussagen offenbar korrekt und wir können deshalb $f \neq 0 \neq g$ annehmen. Dann folgt a. unmittelbar aus (71) und b. aus (72), wobei wir für b. beachten, daß der Koeffizient von t^{m+n} gerade $a_m \cdot b_n = \mathrm{lc}(f) \cdot \mathrm{lc}(g)$ ist. □

Beispiel 26.8 (Gradformeln).
Sei $f = 2t + 1, g = -2t + 1 \in \mathbb{Q}[t]$, dann gilt $f + g = 2$, also $\deg(f + g) < \max\{\deg(f), \deg(g)\}$, aber $f \cdot g = -4t^2 + 1$ und somit $\deg(f \cdot g) = \deg(f) + \deg(g)$.

B) Nullstellen und Division mit Rest

Definition 26.9 (Nullstellen von Polynomen).
Sei K ein Körper, $\lambda \in K$ und $f = \sum_{k=0}^{n} a_k \cdot t^k \in K[t]$. Wir setzen

$$f(\lambda) := \sum_{k=0}^{n} a_k \cdot \lambda^k \in K.$$

Gilt $f(\lambda) = 0$, so heißt λ eine *Nullstelle* von f in K.

Beispiel 26.10 (Nullstelle).
Sei $f = t^2 - 4 = t^2 - 4 \cdot t^0 \in \mathbb{R}[t]$ und $\lambda = 2$, dann gilt

$$f(\lambda) = 2^2 - 4 = 0.$$

Also ist λ eine Nullstelle von f.

Definition 26.11 (Ideale).
Eine nicht-leere Teilmenge $I \subseteq K[t]$ heißt ein *Ideal* von $K[t]$, wenn für $f, g \in I$ und $h \in K[t]$ stets $f + g \in I$ und $h \cdot f \in I$ gilt.

Beispiel 26.12 (Ideal).
Die Menge aller reellen Polynome, die 1 als Nullstelle haben, ist ein Ideal:

$$I = \{f \in \mathbb{R}[t] \mid f(1) = 0\}.$$

Denn für $f, g \in I$ und $h \in \mathbb{R}[t]$ gilt

$$(f + g)(1) = f(1) + g(1) = 0 + 0 = 0$$

und

$$(h \cdot f)(1) = h(1) \cdot f(1) = h(1) \cdot 0 = 0,$$

woraus $f + g \in I$ und $h \cdot f \in I$ folgt. Zudem ist I nicht die leere Menge, weil das Nullpolynom offenbar in I liegt.

Satz 26.13 (Division mit Rest).
Es seien $f, g \in K[t] \setminus \{0\}$ und $I \subseteq K[t]$ ein Ideal.

a. Es gibt eindeutige Polynome $q, r \in K[t]$ mit $f = q \cdot g + r$ und $\deg(r) < \deg(g)$.

b. Ist $\lambda \in K$ eine Nullstelle von f in K, dann gibt es ein Polynom $q \in K[t]$ mit $f = q \cdot (t - \lambda)$, d.h. wir können $t - \lambda$ als Linearfaktor abspalten.

c. Ist $\deg(f) = n$, so hat f höchstens n Nullstellen in K.

d. Es gibt genau ein normiertes Polynom $\mu \in K[t]$ mit $I \stackrel{!}{=} \{\mu \cdot p \mid p \in K[t]\}$.

Beweis: Für den Beweis von Teil a. verweisen wir den Leser auf den Anhang (siehe Proposition A6.27).

Für den Teil b. teilen wir f durch $t - \lambda$ mit Rest und erhalten so Polynome $q, r \in K[t]$ mit

$$f = q \cdot (t - \lambda) + r$$

und $\deg(r) < \deg(t - \lambda) = 1$. Mithin ist r ein konstantes Polynom, und wenn wir auf beiden Seiten λ einsetzen, erhalten wir

$$0 = f(\lambda) = q(\lambda) \cdot (\lambda - \lambda) + r = r.$$

Den Teil c. beweisen wir mit Induktion nach dem Grad n von f. Wenn f eine Nullstelle λ hat, können wir sie nach Teil b. als Linearfaktor abspalten und erhalten $f = q \cdot (t - \lambda)$. Da q nur nur noch den Grad $n - 1$ hat, hat q nach Indutionsvoraussetzung maximal $n - 1$ Nullstellen und f hat somit maximal n Stück.

Für den Teil d. können wir ohne Einschränkung annehmen, daß I nicht nur das Nullpolynom enthält. Wir wählen dann ein normiertes Polynom $\mu \neq 0$ von minimalem Grad in I. Ist nun $f \in I$ beliebig, so teilen wir f durch μ mit Rest und erhalten $q, r \in K[t]$ mit

$$f = q \cdot \mu + r$$

und $\deg(r) < \deg(\mu)$. Aber aus der Gleichung folgt

$$r = f - q \cdot \mu \in I,$$

und da μ in I von minimalem Grad unter den Nicht-Null-Polynomen ist, muss $r = 0$ gelten. Daraus folgt, dass jedes Polynom in I ein Vielfaches von μ ist. □

Beispiel 26.14.
Sei $f = t^3 - 1 \in \mathbb{R}[t]$, dann gilt offenbar $f(1) = 1^3 - 1 = 0$. Polynomdivision liefert:

$$\begin{array}{rl} (t^3 & -1) : (t-1) = t^2 + t + 1. \\ \underline{t^3 - t^2} & \\ t^2 & \\ \underline{t^2 - t} & \\ t - & 1 \\ \underline{t - } & \underline{1} \\ - & \end{array}$$

Also gilt $f = (t^2 + t + 1) \cdot (t - 1)$.

Bemerkung 26.15 (Polynome versus Polynomfunktionen).

a. Auch die Menge K^K aller Abbildungen von K nach K ist eine Ring und

$$\psi : K[t] \longrightarrow K^K : f \mapsto f$$

ist ein Ringhomomorphismus, der einem Polynom seine *Polynomfunktion* zuordnet.

b. Zwei verschiedene Polynome können dieselbe Polynomfunktion liefern!
Die beiden Polynome $f = t^2 - t \in \mathbb{F}_2[t]$ und $g = 0 \in \mathbb{F}_2[t]$ induzieren beide als Polynomfunktion die Funktion konstant Null, da

$$f(0) = 0^2 - 0 = 0 \quad \text{und} \quad f(1) = 1^2 - 1 = 0.$$

Die Polynome f und g sind aber verschieden. D.h., ψ ist nicht injektiv.

c. Enthält K *unendlich* viele Elemente, so ist die Abbildung ψ aus Teil a. injektiv, d.h. zwei Polynome sind genau dann verschieden, wenn die zugehörigen Polynomfunktionen verschieden sind. Liefern nämlich zwei Polynome f und g die gleichen Polynomfunktionen, so hat die Differenz $f - g$ unendlich viele Nullstellen und muß wegen Satz 26.13 somit das Nullpolynom sein.

C) Der Fundamentalsatz der Algebra

Definition 26.16 (Vielfachheiten von Nullstellen).

a. Ist $\lambda \in K$ und $f = (t - \lambda)^m \cdot g \in K[t]$ mit $m \geq 1$ und $g(\lambda) \neq 0$, so nennen wir λ eine Nullstelle von f mit *Vielfachheit* $\text{mult}(f, \lambda) = m$.

b. Es sei $K \subseteq L$ ein Teilkörper des Körpers L und $f \in K[t]$ mit $n = \deg(f) > 0$. Gibt es $b_1, \ldots, b_n \in L$ und ein $0 \neq c \in L$ mit $f = c \cdot (t - b_1) \cdot \ldots \cdot (t - b_n)$ so sagen wir, daß f über L in *Linearfaktoren* zerfällt.

c. Wir nennen einen Körper K *algebraisch abgeschlossen*, wenn jedes nichtkonstante Polynom in $K[t]$ über K in Linearfaktoren zerfällt.

Beispiel 26.17.
Betrachte das Polynom $f = t^4 - 2t^3 + 2t^2 - 2t + 1 = (t-1)^2 \cdot (t^2+1) \in \mathbb{R}[t]$. Dann ist $\lambda = 1$ keine Nullstelle von t^2+1. Mithin ist $\lambda = 1$ eine Nullstelle von f mit Vielfachheit $\mathrm{mult}(f,1) = 2$. Man beachte, daß f über $\mathbb{R}$ nicht in Linearfaktoren zerfällt, da t^2+1 keine Nullstelle besitzt. Über $\mathbb{C}$ zerfällt f hingegen in Linearfaktoren

$$f = (t-1)^2 \cdot (t-i) \cdot (t+i).$$

Satz 26.18 (Fundamentalsatz der Algebra).

a. Der Köper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen.

b. Jeder Körper K ist Teilkörper eines algebraisch abgeschlossenen Körpers. Der kleinste solche Oberkörper von K ist bis auf Isomorphie eindeutig bestimmt und wird der *algebraische Abschluß* $\overline{K}$ von K genannt.

Beweis: Die als Fundamentalsatz der Algebra bekannte Aussage in Teil a. wird in der Vorlesung Einführung in die Funktionentheorie mit Mitteln der Analysis bewiesen und unter Umständen auch in der Vorlesung Einführung in die Algebra mit algebraischen Mitteln. Die Aussage in Teil b. ist Bestandteil der Vorlesung Einführung in die Algebra. □

D) Irreduzible Polynome

Ist der Körper K nicht algebraisch abgeschlossen, so zerfällt nicht jedes Polynom in Linearfaktoren. Zumindest aber läßt sich jedes Polynom als Produkt von nicht mehr weiter zerlegbaren Polynomen schreiben.

Definition 26.19 (Irreduzible Polynome).
Ein nicht-konstantes Polynom $f \in K[t] \setminus K$ heißt *irreduzibel*, wenn aus $f = g \cdot h$ mit $g, h \in K[t]$ stets $\deg(g) = 0$ oder $\deg(h) = 0$ folgt.

Beispiel 26.20.
Aus $f = g \cdot h$ folgt mit der Gradformel

$$\deg(f) = \deg(g) + \deg(h).$$

Ist $\deg(f) = 1$, so folgt unmittelbar, daß f irreduzibel ist. Ist $\deg(f) \in \{2,3\}$, so ist f genau dann irreduzibel, wenn man von f keinen Faktor vom Grad 1 abspalten kann, d.h. wenn f keine Nullstelle in K hat.

Satz 26.21 (Primfaktorzerlegung im Polynomring).
Jedes nicht-konstante normierte Polynom in $K[t]$ läßt sich als Produkt von endlich vielen normierten irreduziblen Polynomen schreiben, und diese Faktoren sind bis auf die Reihenfolge eindeutig bestimmt.

Beweis: Für die Aussage verweisen wir auf den Anhang (siehe Satz A6.62). □

Beispiel 26.22.
Das Polynom $f = t^4 - 2t^3 + 2t^2 - 2t + 1$ aus Beispiel 26.17 hat in $\mathbb{R}[t]$ die Primfaktorzerlegung

$$f = (t-1)^2 \cdot \left(t^2+1\right)$$

und in $\mathbb{C}[t]$ die Primfaktorzerlegung

$$f = (t-1)^2 \cdot (t-i) \cdot (t+i).$$

Satz 26.23 (Bézout-Identität).
Seien $f, g \in K[t]$ zwei normierte teilerfremde Polynome, d.h. sie haben keinen Primfaktor gemeinsam, so gibt es Polynome $p, q \in K[t]$ mit

$$1 = p \cdot f + q \cdot g.$$

Allgemeiner gilt, sind $q_1, \ldots, q_r \in K[t]$ normierte Polynome und gibt es keinen Primfaktor, den alle gemeinsam haben, dann gibt es Polynome $p_1, \ldots, p_r \in K[t]$ mit

$$1 = p_1 \cdot q_1 + \ldots + p_r \cdot q_r.$$

Beweis: Der Beweis der allgemeinen Aussage folgt aus Satz A6.54 im Anhang mittels einfacher Induktion. Wir wollen die Aussage hier im Spezialfall $g = t - \lambda$ beweisen. Teilen wir f durch $t - \lambda$ mit Rest, so finden wir Polynome $q, r \in K[t]$ mit

$$f = q \cdot (t-\lambda) + r$$

und $\deg(r) < \deg(t-\lambda) = 1$. Damit ist r eine Konstante. Wäre $r = 0$, so wäre $t - \lambda$ ein gemeinsamer Primfaktor von f und g, also ist $r \neq 0$. Dann ist

$$1 = \frac{1}{r} \cdot f - \frac{q}{r} \cdot g$$

die gesuchte Darstellung. □

Bemerkung 26.24 (Rationale Funktionen).
Der Polynomring $K[t]$ über einem Körper hat sehr viele Eigenschaften mit dem Ring $\mathbb{Z}$ der ganzen Zahlen gemeinsam – in beiden Ringen gibt es eine Division mit Rest und in beiden Ringen hat man eine eindeutige Primfaktorzerlegung (siehe Abschnitt A6 im

Anhang). Deshalb kann man bei den Polynomen das Problem der fehlenden multiplikativen Inversen genauso lösen wie im Fall der ganzen Zahlen, man führt Brüche ein. Dies führt zum Körper

$$K(t) = \left\{ \frac{f}{g} \;\middle|\; f, g \in K[t], g \neq 0 \right\}$$

der *rationalen Funktionen*. Das Kürzen von Brüchen funktioniert wie in den rationalen Zahlen, und gleiches gilt für die Addition und die Multiplikation.

Aufgaben

Aufgabe 26.25 (Polynominterpolation).
Es seien $b_0, \ldots, b_n \in K$ paarweise verschieden und $c_0, \ldots, c_n \in K$ beliebig. Zeige, es gibt genau ein Polynom $f \in K[t]$ vom Grad $\deg(f) \leq n$ mit $f(b_i) = c_i$ für alle $i = 0, \ldots, n$.

Aufgabe 26.26.
Zerlege das Polynom $f = t^4 + t^3 + 2t - 4 \in \mathbb{C}[t]$ in Linearfaktoren.

Aufgabe 26.27.
Zeige, ist $f \in \mathbb{R}[t]$ irreduzibel, so ist $\deg(f) \in \{1, 2\}$.

Hinweis: Betrachte für eine komplexe Nullstelle λ von f die Fälle $\lambda \in \mathbb{R}$ und $\lambda \in \mathbb{C} \setminus \mathbb{R}$. In letzterem Fall zeige, daß auch das konjugiert Komplexe von λ eine Nullstelle von f ist und betrachte dann das Polynom $g = (t - \lambda) \cdot (t - \overline{\lambda})$.

Kapitel IV

Vektorräume und lineare Abbildungen

Im folgenden wollen wir die Theorie der Vektorräume und der linearen Abbildungen studieren, unter anderem mit dem Ziel, die Lösungsmengen von linearen Gleichungssystemen zu verstehen und berechnen zu können. Der Analysis liegen die Körper der reellen und komplexen Zahlen zugrunde. Wesentlichster Baustein neben der Addition und Multiplikation sind dort der Absolutbetrag mit Werten in $\mathbb{R}$ und die Ordnungsrelation auf $\mathbb{R}$, die $\mathbb{R}$ zu einem vollständigen angeordneten Körper machen. Für die lineare Algebra spielen der Absolutbetrag und die Ordnungsrelation keine Rolle mehr. Wir kommen ohne ε's und δ's und komplizierte Abschätzungen aus. Deshalb können wir unser Arsenal an Grundstrukturen, mit denen wir arbeiten wollen, auch erweitern.

> K wird im folgenden einen *beliebigen Körper* bezeichnen,

etwa $\mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$ oder auch einen endlichen *Körper* wie etwa $\mathbb{F}_2$ in Beispiel 7.6 oder allgemeiner $\mathbb{Z}_p$ in Korollar 25.13.

§ 27 Rechnen mit Matrizen

In der Vorlesung werden wir vor allem endlich-dimensionale Vektorräume betrachten. Ist n die Dimension des Vektorraums V, so ist dieser isomorph zu dem in diesem Abschnitt eingeführten Vektorraum K^n. Abbildungen zwischen endlich-dimensionalen Vektorräumen, die die Struktur respektieren, können wir stets mit Hilfe von Matrizen beschreiben, und die Darstellung über Matrizen erlaubt es, die Eigenschaften der Abbildungen effizient zu berechnen. Die in diesem Abschnitt eingeführten Strukturen bilden insofern die Grundlage für alles, was wir im weiteren Verlauf des Kapitels untersuchen wollen.

A) Grundlegende Begriffe und Operationen

Definition 27.1 (Matrizen und der K^n).
Es seien $m, n \geq 1$ zwei positive ganze Zahlen.

a. Eine $m \times n$-*Matrix* über K ist ein rechteckiges Schema A mit Einträgen aus K der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Wenn keine Unklarheiten zu befürchten sind, schreiben wir verkürzt auch

$$A = (a_{ij})_{i=1,\dots,m;j=1,\dots,n} = (a_{ij}).$$

b. Die Menge aller $m \times n$-Matrizen über K wird mit $\mathrm{Mat}(m \times n, K)$ bezeichnet, und falls $m = n$, dann auch kurz mit $\mathrm{Mat}_n(K) = \mathrm{Mat}(n, K)$ und man spricht von *quadratischen Matrizen*.

c. Ist $A = (a_{ij})$ eine $m \times n$-Matrix, dann bezeichnen wir

$$a_i := (a_{i1}, \dots, a_{in})$$

als den i-ten *Zeilenvektor* von A und

$$a^j := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

als den j-ten *Spaltenvektor* von A.

d. Ist $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$, so heißt die $n \times m$-Matrix

$$A^t := \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix},$$

d. h. für $A^t = (a'_{ij})$ gilt $a'_{ij} = a_{ji}$, die *Transponierte* von A.

e. Schließlich definieren wir

$$K^n := \mathrm{Mat}(n \times 1, K) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \,\middle|\, x_1, \dots, x_n \in K \right\}.$$

Die Elemente von K^n heißen *Vektoren* oder *Punkte* im K^n. x_i heißt die i-te *Komponente* des Vektors x.

Definition 27.2 (Operationen mit Matrizen).

a. Es seien $A = (a_{ij}), B = (b_{ij}) \in \mathrm{Mat}(m \times n, K)$ und $\lambda \in K$. Dann definiert man

$$A + B := (a_{ij} + b_{ij}) = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{pmatrix},$$

sowie

$$\lambda \cdot A := (\lambda a_{ij}) = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{pmatrix}.$$

b. Sind $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$ und $B = (b_{jk}) \in \mathrm{Mat}(n \times p, K)$ zwei Matrizen, wobei A genauso viele Spalten wie B Zeilen hat. Dann definieren wir das *Matrixprodukt* durch

$$A \circ B := C, \quad \text{mit } C = (c_{ik}) \in \mathrm{Mat}(m \times p, K) \text{ und } c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk}.$$

Beispiel 27.3.

Folgende Matrizen $A, B \in \mathrm{Mat}(2 \times 3, \mathbb{K})$ und $C \in \mathrm{Mat}(3 \times 2, \mathbb{K})$ seien gegeben:

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 2 \end{pmatrix}, B = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 4 & 5 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}.$$

Dann gilt:

$$A + B = \begin{pmatrix} 4 & 2 & 3 \\ 3 & 5 & 7 \end{pmatrix}, \; 3 \cdot A = \begin{pmatrix} 3 & 0 & 6 \\ 9 & 3 & 6 \end{pmatrix} \text{ und } A \circ C = \begin{pmatrix} 7 & 2 \\ 11 & 3 \end{pmatrix}.$$

Bemerkung 27.4.

a. Die in Definition 27.2 a. definierte Addition zweier Matrizen definiert auf $\mathrm{Mat}(m \times n, K)$ offensichtlich eine zweistellige Operation, bezüglich derer $\mathrm{Mat}(m \times n, K)$ eine abelsche Gruppe $(\mathrm{Mat}(m \times n, K), +)$ wird, wie man leicht nachprüft.

b. Wir werden meist kurz λA bzw. λx schreiben, statt $\lambda \cdot A$ bzw. $\lambda \cdot x$, wenn $\lambda \in K$, $A \in \mathrm{Mat}(m \times n, K)$ und $x \in K^n$.

c. Wir schreiben statt $A \circ B$ häufig kurz AB, insbesondere auch Ax statt $A \circ x$.

d. Spaltenvektoren nehmen im Skript sehr viel Raum ein. Um platzsparend arbeiten zu können, werden wir deshalb statt den Spaltenvektor $x \in K^n$ als

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

anzugeben, meist den *transponierten* Zeilenvektor

$$x = (x_1 \ \dots \ x_n)^t$$

betrachten, und um Mißverständnissen vorzubeugen, fügen wir zudem meist Kommata als Trennsymbole ein

$$x = (x_1, \dots, x_n)^t.$$

e. Man beachte, daß das Produkt nur dann definiert ist, wenn A soviele Spalten wie B Zeilen hat. Das Produkt $A \circ B$ hat dann soviele Zeilen wie A und soviele Spalten wie B.

B) Die Abbildung zu einer Matrix

Jede Matrix definiert wie folgt eine Abbildung.

Definition 27.5 (Die Abbildung f_A).
Ist $A \in \mathrm{Mat}(m \times n, K)$, so definieren wir

$$f_A : K^n \to K^m : x \mapsto Ax = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}.$$

f_A heißt die *zu A assoziierte* oder *zu A gehörige Abbildung.*

Beispiel 27.6.
Die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$$

definiert die Abbildung

$$f_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x_1, x_2)^t \mapsto (x_1 + 2x_2, 3x_2)^t.$$

Wir wollen uns nun mit der Frage beschäftigen, ob zwei verschiedene Matrizen dieselbe Abbildung definieren können.

Bemerkung 27.7 (Einheitsvektoren).
Um den Zusammenhang zwischen A und f_A besser zu verstehen, betrachten wir für $i = 1, \ldots, n$ den i-ten *Einheitsvektor* $e_i = (\delta_{1i}, \cdots, \delta_{ni})^t \in K^n$, wobei

$$\delta_{ji} := \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

das *Kronecker Symbol* ist, d. h.

$$e_i := \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

wobei die Eins in der i-ten Komponente steht.

Es ist dann

$$f_A(e_i) = \begin{pmatrix} \sum_{j=1}^n a_{1j}\delta_{ji} \\ \vdots \\ \sum_{j=1}^n a_{mj}\delta_{ji} \end{pmatrix} = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} = a^i,$$

d. h. die i-te Spalte von A ist das Bild des i-ten Einheitsvektors unter f_A.

Hieraus folgt insbesondere, daß A durch f_A eindeutig bestimmt ist.

C) Rechenregeln für Matrizen

Lemma 27.8 (Einfache Rechenregeln für Matrizen).
Für $x, y \in K^n$, $A, B \in \mathrm{Mat}(m \times n, K)$, $C \in \mathrm{Mat}(n \times p, K)$ und $\lambda \in K$ gelten:

a. $A(x + y) = Ax + Ay$ und $A(\lambda x) = \lambda Ax$,

b. $\lambda \cdot (A \circ C) = (\lambda \cdot A) \circ C = A \circ (\lambda \cdot C)$,

c. $f_{A+B} = f_A + f_B$, und

d. $f_{\lambda A} = \lambda f_A$.

Beweis: Der Beweis ist eine Übungsaufgabe. □

Wir wollen jetzt sehen, wie sich die Multiplikation von Matrizen mit den zugehörigen Abbildungen verträgt.

Satz 27.9 (Matrixmultiplikation und Komposition).
Für $A \in \mathrm{Mat}(m \times n, K)$ und $B \in \mathrm{Mat}(n \times p, K)$ gilt:

$$f_{A \circ B} = f_A \circ f_B.$$

Beweis: Da Definitionsbereich und Wertebereich von beiden Abbildungen übereinstimmen, reicht es zu zeigen:

$$(f_{A\circ B})(x) = (f_A \circ f_B)(x), \quad \text{für alle } x \in K^p.$$

Seien $A = (a_{ij})$ und $B = (b_{jk})$, und sei $x = (x_1, \ldots, x_p)^t \in K^p$ gegeben.

$$(f_{A\circ B})(x) = (A \circ B)x = \begin{pmatrix} \sum_{j=1}^n a_{1j}b_{j1} & \ldots & \sum_{j=1}^n a_{1j}b_{jp} \\ \vdots & & \vdots \\ \sum_{j=1}^n a_{mj}b_{j1} & \ldots & \sum_{j=1}^n a_{mj}b_{jp} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=1}^p \sum_{j=1}^n a_{1j}b_{jk}x_k \\ \vdots \\ \sum_{k=1}^p \sum_{j=1}^n a_{mj}b_{jk}x_k \end{pmatrix}.$$

Ferner gilt:

$$(f_A \circ f_B)(x) = f_A(Bx) = A(Bx) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} \sum_{k=1}^p b_{1k}x_k \\ \vdots \\ \sum_{k=1}^p b_{nk}x_k \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^n a_{1j} \sum_{k=1}^p b_{jk}x_k \\ \vdots \\ \sum_{j=1}^n a_{mj} \sum_{k=1}^p b_{jk}x_k \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n \sum_{k=1}^p a_{1j}b_{jk}x_k \\ \vdots \\ \sum_{j=1}^n \sum_{k=1}^p a_{mj}b_{jk}x_k \end{pmatrix}.$$

Beide Ausdrücke stimmen (bis auf die Reihenfolge der Summation) überein, was zu zeigen war. □

Korollar 27.10 (Die Matrixmultiplikation ist assoziativ.).
Für $A \in \mathrm{Mat}(m \times n, K)$, $B \in \mathrm{Mat}(n \times p, K)$ und $C \in \mathrm{Mat}(p \times q, K)$ gilt

$$A \circ (B \circ C) = (A \circ B) \circ C.$$

Beweis: Dies folgt aus Satz 27.9, da die Komposition von Abbildungen assoziativ ist und da eine Matrix A durch die Abbildung f_A eindeutig bestimmt ist. □

Man kann die Aussage des Korollars natürlich auch direkt nachweisen, was auf die gleiche Rechnung wie in 27.9 führt - statt des einen Vektors x hat man die q Spaltenvektoren von C zu multiplizieren, was etwas mehr Schreibarbeit bedeutet.

Lemma 27.11 (Distributivgesetze).
Sind $A, B \in \mathrm{Mat}(m \times n, K)$ und $C, D \in \mathrm{Mat}(n \times p, K)$, so gelten die Distributivgesetze:

$$A \circ (C + D) = A \circ C + A \circ D,$$

sowie

$$(A + B) \circ C = A \circ C + B \circ C.$$

Beweis: Die Aussage kann wie Korollar 27.10 aus Lemma 27.8 und Satz 27.9 abgeleitet werden und sei dem Leser als Übung anempfohlen. □

D) Invertierbare Matrizen und die allgemeine lineare Gruppe

Definition 27.12 (Invertierbare Matrizen).
Eine Matrix $A \in \mathrm{Mat}_n(K)$ heißt *invertierbar*, falls es eine Matrix $A^{-1} \in \mathrm{Mat}_n(K)$ gibt, so daß

$$A \circ A^{-1} = A^{-1} \circ A = \mathbb{1}_n,$$

wobei die Matrix $\mathbb{1}_n = (\delta_{ij})_{i,j=1,\dots,n} \in \mathrm{Mat}_n(K)$ die *Einheitsmatrix* ist, die auf der Diagonalen Einsen und außerhalb der Diagonalen Nullen als Einträge hat. Eine Matrix mit der Eigenschaft von A^{-1} nennt man eine *Inverse* zu A.

Satz 27.13 (Die allgemeine lineare Gruppe $\mathrm{Gl}_n(K)$).
Die Menge der invertierbaren $n \times n$-Matrizen

$$\mathrm{Gl}_n(K) = \{A \in \mathrm{Mat}_n(K) \mid A \text{ ist invertierbar}\}$$

ist eine Gruppe mit neutralem Element $\mathbb{1}_n$, die für $n > 1$ nicht kommutativ ist. Insbesondere ist die Inverse zu A eindeutig bestimmt und es gelten für $A, B \in \mathrm{Gl}_n(K)$

$$(AB)^{-1} = B^{-1}A^{-1} \quad \text{und} \quad (A^{-1})^{-1} = A.$$

Beweis: Die Assoziativität der Operation folgt aus Korollar 27.10. Daß $\mathbb{1}_n$ das neutrale Element ist, folgt unmittelbar aus der Definition der Multiplikation, und aus der Definition der Invertierbarkeit folgt dann, daß A^{-1} das zu A Inverse ist und in $\mathrm{Gl}_n(K)$ liegt. Also ist $(\mathrm{Gl}_n(K), \circ)$ eine Gruppe. Daß diese nicht kommutativ ist, überlassen wir dem Leser als Übungsaufgabe. □

Aufgaben

Aufgabe 27.14.
Zeige, daß die Matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_2(K)$$

genau dann invertierbar ist, wenn $ad - bc \neq 0$ ist. Die Inverse ist dann

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Aufgabe 27.15.
Es sei $A \in \mathrm{Mat}(m \times n, K)$ und $B \in \mathrm{Mat}(n \times p, K)$. Zeige, $(AB)^t = B^t A^t$.

Aufgabe 27.16 (Nilpotente Matrizen).
Es sei $N = (n_{ij}) \in \mathrm{Mat}_n(K)$ eine Matrix, für die die Einträge auf der oberen Nebendiagonale alle 1 sind und für die alle anderen Einträge 0 sind, d.h. $n_{ij} = \delta_{j-i,1}$.
Zeige für $k = 1, \ldots, n$, daß die Einträge der Matrix $N^k = \left(n_{ij}^{(k)}\right)$ auf der k-ten oberen Nebendiagonale alle 1 und alle anderen Einträge 0 sind, d.h. $n_{ij}^{(k)} = \delta_{j-i,k}$. Insbesondere ist $N^n = 0$ und $N^k \neq 0$ für $k < n$.

§ 28 Vektorräume und lineare Abbildungen

A) Vektorräume

Definition 28.1 (Vektorräume).
Ein K-*Vektorraum* (oder Vektorraum über K) besteht aus einer nicht-leeren Menge V sowie einer zweistelligen Operation

$$+ : V \times V \to V : (x, y) \mapsto x + y,$$

die *Vektoraddition* genannt wird, und einer zweistelligen Operation

$$\cdot : K \times V \to V : (\lambda, x) \mapsto \lambda \cdot x = \lambda x,$$

die *Skalarmultiplikation* genannt wird, so daß die folgenden Gesetze gelten:

a. $(V, +)$ ist eine abelsche Gruppe,

b. für $\lambda, \mu \in K$ und $x, y \in V$ gelten:
   (i) $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$, ("*verallgemeinertes Distributivgesetz*")
   (ii) $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$, und ("*verallgemeinertes Distributivgesetz*")
   (iii) $(\lambda \cdot \mu) \cdot x = \lambda \cdot (\mu \cdot x)$. ("*verallgemeinertes Assoziativgesetz*")
   (iv) $1 \cdot x = x$.

Die Elemente aus V nennt man *Vektoren* und die aus K *Skalare*. Der *Nullvektor*, d. h. das neutrale Element aus V bezüglich der Addition, wird mit 0 bzw. mit 0_V bezeichnet und das neutrale Element von $(K, +)$ ebenfalls mit 0 bzw. mit 0_K.

Beispiel 28.2.

a. Der *Nullraum* $V = \{0\}$ mit $\lambda \cdot 0 = 0$ für alle $\lambda \in K$ ist für jeden Körper K ein K-Vektorraum. Man bezeichnet den Nullraum auch mit K^0.

b. Der Körper K selbst mit der Körperaddition als Vektoraddition und der Körpermultiplikation als Skalarmultiplikation ist ein K-Vektorraum.

c. $\mathbb{R}$ ist ein $\mathbb{Q}$-Vektorraum und $\mathbb{C}$ ist ein $\mathbb{R}$-Vektorraum, jeweils mit der üblichen Addition und Multiplikation.

d. Die Menge $\mathrm{Mat}(m \times n, K)$ der $m \times n$-Matrizen über K mittels der in Definition 27.2 definierten Addition und Skalarmultiplikation ist ein K-Vektorraum mit der *Nullmatrix*

$$0 := \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

als Nullvektor.

e. Damit ist insbesondere K^n mit der komponentenweisen Addition und Skalarmultiplikation ein K-Vektorraum mit $0_{K^n} = (0, \ldots, 0)^t$.
Speziell sind $\mathbb{R}^n$, $\mathbb{C}^n$ und $\mathbb{F}_2^n$ Vektorräume über $\mathbb{R}$, $\mathbb{C}$ bzw. $\mathbb{F}_2$.

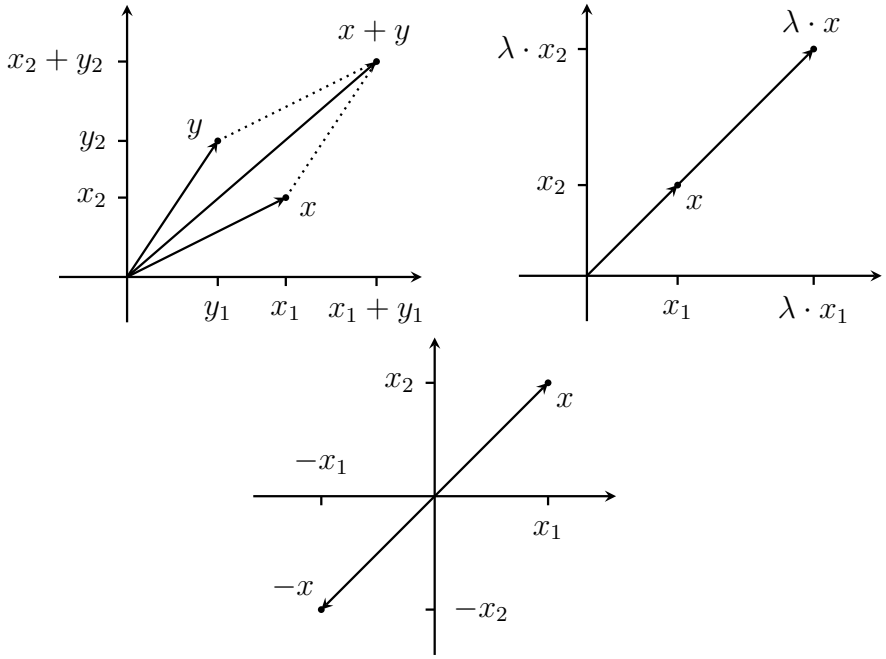


Abbildung 1: Addition und Skalarmultiplikation in $\mathbb{R}^2$

f. Ist M eine Menge und V ein K-Vektorraum, so wird die Menge

$$V^M = \{f : M \longrightarrow V \mid f \text{ ist eine Abbildung}\}$$

durch die Operationen

$$+ : V^M \times V^M \to V^M : (f, g) \mapsto \big(f + g : M \to V : x \mapsto f(x) + g(x)\big)$$

und

$$\cdot : K \times V^M \to V^M : (\lambda, f) \mapsto \big(\lambda \cdot f : M \to V : x \mapsto \lambda \cdot f(x)\big)$$

zu einem K-Vektorraum, wie man leicht nachrechnet.
Ist z.B. $M = \mathbb{N}$ und $K = V = \mathbb{R}$, so ist

$$\mathbb{R}^{\mathbb{N}} = \{a : \mathbb{N} \longrightarrow \mathbb{R} \mid a \text{ ist eine Abbildung}\} = \big\{(a_n)_{n\in\mathbb{N}} \;\big|\; a_n \in \mathbb{R}\big\}$$

der Vektorraum der Folgen in $\mathbb{R}$. Unsere Definitionen sorgen dafür, daß Folgen komponentenweise addiert und mit Skalaren multipliziert werden.

g. Der Polynomring

$$K[t] = \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; n \in \mathbb{N}, a_k \in K \right\}$$

wird mit der Addition aus Definition 26.1

$$\sum_{k=0}^{m} a_k \cdot t^k + \sum_{k=0}^{n} b_k \cdot t^k := \sum_{k=0}^{\max\{m,n\}} (a_k + b_k) \cdot t^k,$$

wobei $a_k = 0$ für $k > m$ und $b_k = 0$ für $k > n$ gelten soll, und der dort definierten Skalarmultiplikation

$$\lambda \cdot \sum_{k=0}^{n} a_k \cdot t^k := \sum_{k=0}^{n} (\lambda \cdot a_k) \cdot t^k$$

zu einem K-Vektorraum, wie man leicht nachprüft.

h. Da man für $M = \{1, \ldots, n\}$ eine Abbildung $f : M \to K$ in eindeutiger Weise durch das Tupel der Bildelemente $\big(f(1), \ldots, f(n)\big)$ beschreiben kann, sieht man leicht, daß die Zuordnung

$$K^M \to K^n : f \mapsto \big(f(1), \ldots, f(n)\big)^t$$

in diesem Falle eine Bijektion ist. Man prüft überdies leicht nach, daß diese Abbildung ein Vektorraumhomomorphismus im Sinne von Definition 28.19 ist. K^M und K^n sind dann also isomorph.

Lemma 28.3 (Einfache Rechenregeln für Vektoren).
In einem K-Vektorraum gelten folgende Rechenregeln:

a. $0_K \cdot x = 0_V$ und $\lambda \cdot 0_V = 0_V$ für alle $x \in V$, $\lambda \in K$.

b. Für $\lambda \in K$ und $x \in V$ gilt:

$$\lambda \cdot x = 0_V \quad \Longrightarrow \quad \lambda = 0 \quad \text{oder} \quad x = 0.$$

c. $(-1) \cdot x = -x$ für alle $x \in V$.

Beweis: Es seien $x \in V$ und $\lambda \in K$ gegeben.

a. Es gilt:

$$0_V + 0_K \cdot x = 0_K \cdot x = (0_K + 0_K) \cdot x = 0_K \cdot x + 0_K \cdot x,$$

also $0_V = 0_K \cdot x$, wie aus den Kürzungsregeln für $(V, +)$ folgt. Analog gilt:

$$0_V + \lambda \cdot 0_V = \lambda \cdot 0_V = \lambda \cdot (0_V + 0_V) = \lambda \cdot 0_V + \lambda \cdot 0_V,$$

und damit $0_V = \lambda \cdot 0_V$.

b. Ist $\lambda \in K$ mit $\lambda \neq 0$, dann gibt es ein Inverses $\lambda^{-1} \in K$. Aus $\lambda \cdot x = 0$ folgt dann aber wegen a. und den Vektorraumaxiomen

$$0_V = \lambda^{-1} \cdot 0_V = \lambda^{-1} \cdot (\lambda \cdot x) = \big(\lambda^{-1} \cdot \lambda\big) \cdot x = 1 \cdot x = x.$$

c. Für $x \in K$ gilt:

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

Also ist $(-1) \cdot x$ das (eindeutig bestimmte) additive Inverse zu x.

□

B) Unterräume

Definition 28.4 (Unterräume).
Es sei V ein Vektorraum über K. Eine nicht-leere Teilmenge $U \subseteq V$ von V heißt *Unterraum*, wenn für alle $\lambda \in K$ und $x, y \in U$ gilt

$$\lambda \cdot x \in U \quad \text{und} \quad x + y \in U. \tag{74}$$

Man sagt, U sei *abgeschlossen* bezüglich der Addition und der Skalarmultiplikation. Wir schreiben $U \leq V$, um auszudrücken, daß U ein Unterraum von V ist.

Proposition 28.5 (Unterräume sind Vektorräume.).
Jeder Unterraum eines K-Vektorraums ist selbst ein K-Vektorraum.

Beweis: Aus der Abgeschlossenheit bezüglich der Skalarmultiplikation folgt insbesondere

$$-x = (-1) \cdot x \in U$$

für alle $x \in U$, d.h. die Abgeschlossenheit bezüglich der Inversenbildung. Zusammen mit der Abgeschlossenheit bezüglich der Addtion ist U also eine Untergruppe von $(V, +)$ und damit selbst eine abelsche Gruppe nach Proposition 22.8.

Die verbleibenden Axiome aus Teil b. in Definition 28.1 vererben sich analog zum Assoziativgesetz unmittelbar von V auf U. Damit ist U also ein K-Vektorraum. □

Beispiel 28.6.

a. Ist V ein K-Vektorraum, so ist $\{0_V\}$ stets ein Unterraum von V. Ferner ist V selbst ein Unterraum. Man nennt diese beiden auch die *trivialen Unterräume*.

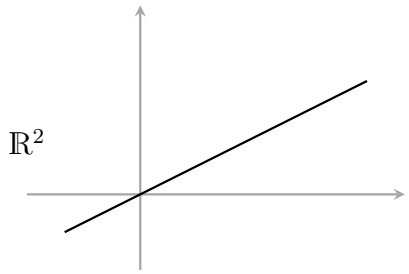
b. Eine Gerade G durch den Ursprung in der Ebene $\mathbb{R}^2$ mit Steigung m genügt der Gleichung $y = m \cdot x$ und ist deshalb

$$G = \left\{ (x, mx)^t \;\middle|\; x \in \mathbb{R} \right\}.$$

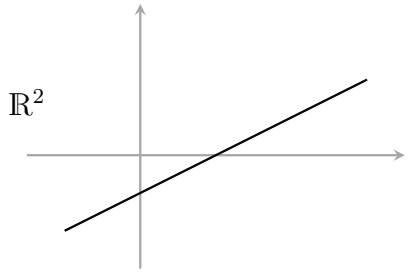
Für $v = (x, mx)^t, w = (x', mx')^t \in G$ und $\lambda \in \mathbb{R}$ gilt dann

$$v + w = \big(x + x', m \cdot (x + x')\big)^t, \; \lambda \cdot v = (\lambda x, m\lambda x)^t \in G.$$

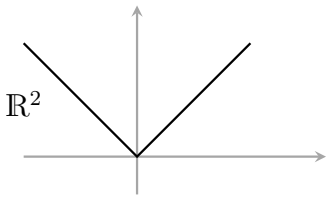
Mithin ist G ein Unterraum von $\mathbb{R}^2$.



c. Eine Gerade in $\mathbb{R}^2$, die nicht durch den Ursprung geht, kann kein Unterraum von $\mathbb{R}^2$ sein, da ein Unterraum den Nullvektor enthalten muß.



d. Der Graph der Betragsfunktion ist *kein* Unterraum von $\mathbb{R}^2$, da $(-1,1)^t$ und $(1,1)^t$ auf dem Graphen liegen, ihre Summe $(0,2)^t$ aber nicht.



e. Die Menge

$$U := \{(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid (a_n)_{n\in\mathbb{N}} \text{ ist konvergent}\}$$

der *konvergenten Folgen* in $\mathbb{R}$ ist ein Unterraum des $\mathbb{R}$-Vektorraums $\mathbb{R}^{\mathbb{N}}$ aller Folgen. Dies folgt aus den Grenzwertsätzen für Folgen 11.15.

f. Ist I ein Intervall in $\mathbb{R}$, so sind die Menge $\mathcal{C}(I,\mathbb{R})$ aller auf I stetigen Abbildungen sowie die Menge $\mathcal{C}^k(I,\mathbb{R})$ aller auf I k-fach stetig differenzierbaren Abbildungen Unterräume des Vektorraums $\mathbb{R}^I$ aller Abbildungen von I nach $\mathbb{R}$.
Solche Funktionenräume spielen in der Analysis eine große Rolle. Sie sind für kein n isomorph zu $\mathbb{R}^n$, und sie sind ein gutes Beispiel für den Wert der abstrakten Theorie der Vektorräume.

g. Ist $n \in \mathbb{N}$ fest vorgegeben, so bilden die Polynome vom Grad höchstens n

$$P_n := \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; a_k \in K \right\}$$

einen Unterraum des Vektorraums der Polynome $K[t]$.[1]

[1] Man beachte, obwohl $K[t]$ auch ein Ring ist, ist P_n für $n \geq 1$ kein Unterring von $K[t]$, weil $t^n \cdot t^n = t^{2n} \notin P_n$.

Lemma 28.7 (Durchschnitt von Unterräumen).
Der Durchschnitt beliebig vieler Unterräume ist wieder ein Unterraum.

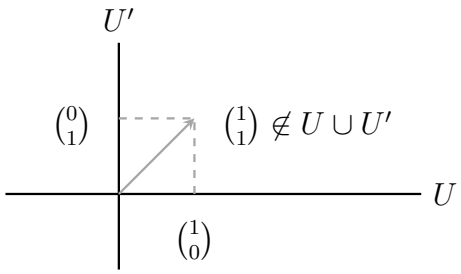
Beweis: Es seien U_i für $i \in I$ Unterräume eines K-Vektorraums V. Da $0_V \in U_i$ für alle $i \in I$, ist $U := \bigcap_{i\in I} U_i$ nicht die leere Menge. Es bleibt also zu zeigen, daß für $x, y \in U$ und $\lambda \in K$ gilt:

$$x + y \in U \quad \text{und} \quad \lambda x \in U.$$

Für ein beliebiges $i \in I$ gilt, da U_i ein Unterraum von V ist und da $x, y \in U \subseteq U_i$, daß $x + y \in U_i$ und $\lambda x \in U_i$. Also liegen die Vektoren im Durchschnitt U. □

Bemerkung 28.8.
Die Vereinigung von zwei Unterräumen ist i. a. kein Unterraum mehr!

Sei etwa U die x-Achse und U' die y-Achse im $\mathbb{R}^2$. Beides sind Unterräume von $\mathbb{R}^2$. Es gilt aber $(1,1)^t = e_1 + e_2 \notin U \cup U'$ und kann $U \cup U'$ kein Unterraum von $\mathbb{R}^2$ sein.



C) Die lineare Hülle und Summen von Unterräumen

Definition 28.9 (Linearkombination und lineare Hülle).
Es sei V ein K-Vektorraum.

a. Wir nennen $x \in V$ eine *Linearkombination* von $x_1, \ldots, x_r \in V$, falls es $\lambda_1, \ldots, \lambda_r \in K$ gibt mit

$$x = \lambda_1 x_1 + \ldots + \lambda_r x_r.$$

Ist eines der λ_i ungleich Null, so nennen wir die Linearkombination *nicht-trivial*.

b. Ist $M \subseteq V$, so nennen wir den Durchschnitt

$$\mathrm{Lin}\,(M) := \langle M \rangle := \bigcap_{M \subseteq U \leq V} U$$

aller Unterräume von V, die M enthalten, die *lineare Hülle* oder das *Erzeugnis* von M.

Bemerkung 28.10.

a. Man beachte, daß die lineare Hülle von M wegen Lemma 28.7 ein Unterraum von V ist. Aufgrund der Definition ist es *der kleinste* Unterraum, der M enthält.

b. Ist $M = \emptyset$, so ist $\mathrm{Lin}\,(M) = \mathrm{Lin}\,(\emptyset) = \{0\}$.

c. Eine Linearkombination ist immer eine *endliche* Summe von Vielfachen von Vektoren aus V. In der linearen Algebra wird es *nie* unendliche Summen geben.

d. Mit Induktion nach der Anzahl der Summanden folgt aus (74) unmittelbar, daß ein Unterraum U abgeschlossen bezüglich endlicher Linearkombinationen von Vektoren aus U ist.

Proposition 28.11 (Lineare Hülle = Menge der Linearkombinationen). Ist V ein K-Vektorraum und $\emptyset \neq M \subseteq V$, so ist die lineare Hülle von M

$$\mathrm{Lin}\,(M) = \{\lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r \mid r \in \mathbb{N}, x_1, \ldots, x_r \in M, \lambda_1, \ldots, \lambda_r \in K\} \leq V$$

die Menge aller Linearkombinationen von Elementen in M.

Beweis: Wir setzen

$$U := \{\lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r \mid r \in \mathbb{N}, x_1, \ldots, x_r \in M, \lambda_1, \ldots, \lambda_r \in K\}.$$

Als Unterraum, der M enthält, enthält $\mathrm{Lin}\,(M)$ auch jede endliche Linearkombination von Elementen in M, also auch U.

Wir wollen nun zeigen, daß U ein Unterraum von V ist, der M enthält, da er aufgrund der Definition des Erzeugnisses dann auch das Erzeugnis von M enthält. Dazu beachten wir zunächst, daß für $x \in M$ auch $x = 1 \cdot x \in U$ gilt. Also ist $M \subseteq U$ und somit ist U auch nicht leer. Seien nun $x = \sum_{i=1}^{r} \lambda_i \cdot x_i$ und $y = \sum_{j=1}^{s} \mu_j \cdot y_j$ mit $x_i, y_j \in M$ und $\lambda_i, \mu_i \in K$ sowie $\lambda \in K$ gegeben. Dann ist

$$x + y = \lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r + \mu_1 \cdot y_1 + \ldots + \mu_s \cdot y_s \in U,$$

weil es eine endliche Linearkombination von Elementen in M ist, und ebenso ist

$$\lambda \cdot x = \sum_{i=1}^{r} (\lambda \cdot \lambda_i) \cdot x_i \in U.$$

Also ist U ein Unterraum von V. □

Beispiel 28.12.

a. Ist $M = \{x_1, \ldots, x_n\} \subseteq V$ endlich, so ist das Erzeugnis von M

$$\mathrm{Lin}\,(x_1, \ldots, x_n) := \mathrm{Lin}\,(M) = \{\lambda_1 \cdot x_1 + \ldots + \lambda_n \cdot x_n \mid \lambda_1, \ldots, \lambda_n \in K\}.$$

Insbesondere gilt $\mathrm{Lin}\,(x) := \mathrm{Lin}\,(\{x\}) = \{\lambda \cdot x \mid \lambda \in K\}$.

b. Die Lineare Hülle der Vektoren $x_1 = (1,0)^t$ und $x_2 = (0,1)^t$ in $\mathbb{R}^2$ ist

$$\mathrm{Lin}\,(x_1, x_2) = \{\lambda_1 \cdot x_1 + \lambda_2 \cdot x_2 = (\lambda_1, \lambda_2)^t \mid \lambda_1, \lambda_2 \in \mathbb{R}\} = \mathbb{R}^2.$$

c. Die Lineare Hülle von $x = (1, m)^t \in \mathbb{R}^2$ ist die Gerade

$$\mathrm{Lin}\,(x) = \{(\lambda, \lambda m)^t \mid \lambda \in \mathbb{R}\}.$$

d. Es gilt offenbar $\mathrm{Lin}\,(t^0, t^1, \ldots, t^n) = P_n$.

Proposition 28.13 (Summe zweier Unterräume).
Es seien U und U' Unterräume des K-Vektorraums V. Dann gilt

$$U + U' := \{u + u' \mid u \in U, u' \in U'\} = \mathrm{Lin}\,(U \cup U') \leq V,$$

und wir nennen diesen Unterraum von V die *Summe* von U und U'.

Beweis: Wegen Proposition 28.11 ist $U + U'$ in $\mathrm{Lin}\,(U \cup U')$ enthalten, da die Elemente von $U + U'$ Linearkombinationen von Elementen in $U \cup U'$ sind.

Umgekehrt ist jede Linearkombination von Elementen in $U \cup U'$ von der Form $\sum_{i=1}^{r} \lambda_i \cdot u_i + \sum_{j=1}^{s} \mu_j \cdot u'_j$ mit $u_i \in U$, $u'_j \in U'$ und $\lambda_i, \mu_j \in K$. Da U und U' aber Unterräume sind, ist

$$u := \sum_{i=1}^{r} \lambda_i \cdot u_i \in U$$

und

$$u' := \sum_{j=1}^{s} \mu_j \cdot u'_j \in U'.$$

Deshalb ist die Linearkombination

$$\sum_{i=1}^{r} \lambda_i \cdot u_i + \sum_{j=1}^{s} \mu_j \cdot u'_j = u + u' \in U + U'$$

in $U + U'$, und mit Proposition 28.11 enthält $U + U'$ auch $\mathrm{Lin}\,(U \cup U')$. □

Bemerkung 28.14 (Summen von Unterräumen).

a. Die Summe zweier Unterräume ersetzt ihre Vereinigung in der Theorie der Vektorräume. Sie ist der kleinste Unterraum, der beide enthält. Im Beispiel aus Bemerkung 28.8 ist die Summe der beiden Unterräume ganz $\mathbb{R}^2$.

b. Analog zu Proposition 28.13 zeigt man allgemeiner: Sind $U_1, \ldots, U_n$ Unterräume des K-Vektorraums V, so gilt

$$U_1 + \ldots + U_n := \{u_1 + \ldots + u_n \mid u_i \in U_i\} = \mathrm{Lin}\,(U_1 \cup \ldots \cup U_n)\,.$$

Beispiel 28.15.
Jeder Vektor x in $U + U'$ läßt sich schreiben als $x = u + u'$ mit $u \in U$ und $u' \in U'$, diese Darstellung muß aber nicht eindeutig sein.

Sind z.B. $U = \mathrm{Lin}\,((1,0,0)^t, (0,1,1)^t)$ und $U' = \mathrm{Lin}\,((1,1,0)^t, (1,0,1)^t)$ als Unterräume von $\mathbb{R}^3$ gegeben, so können wir den Vektor $x = (1,0,-1)^t$ auf folgende beiden Weisen als Summe zweier Vektoren in U und U' schreiben:

$$x = (0,-1,-1)^t + (1,1,0)^t = (2,0,0)^t + (-1,0,-1)^t.$$

Definition 28.16 (Direkte Summe).
Es seien $U_1, \ldots, U_n$ Unterräume des K-Vektorraums V. Wir nennen die Summe $U = U_1 + \ldots + U_n$ eine *direkte Summe*, wenn sich jeder Vektor $x \in U_1 + \ldots + U_n$ auf eindeutige Weise als Summe $x = u_1 + \ldots + u_n$ mit $u_i \in U_i$ schreiben läßt. Wir schreiben dann $U = U_1 \oplus \ldots \oplus U_n$.

Proposition 28.17 (Direkte Summe zweier Unterräume).
Es seien U, U' und W Unterräume des K-Vektorraums V.
Genau dann gilt $W = U \oplus U'$, wenn $W = U + U'$ und $U \cap U' = \{0\}$.

Beweis: Ist die Summe $W = U \oplus U'$, so gilt insbesondere $W = U + U'$. Für $x \in U \cap U'$, gilt zudem

$$x = x + 0 = 0 + x \in U + U',$$

und wegen der Eindeutigkeit der Darstellung in $U + U'$ muß $x = 0$ sein.

Ist umgekehrt $W = U + U'$ und $U \cap U' = \{0\}$ und sind $x_1 + x_1' = x_2 + x_2' \in U + U' = W$ mit $x_i \in U$ und $x_i' \in U'$, $i = 1, 2$, so gilt:

$$x_1 - x_2 = x_2' - x_1' \in U \cap U' = \{0\}.$$

Also ist $x_1 = x_2$ und $x_1' = x_2'$, d. h. die Darstellung ist eindeutig. □

Beispiel 28.18.
Betrachte die Unterräume $U = \mathrm{Lin}\,((1,1,1)^t)$ und $U' = \mathrm{Lin}\,((1,0,1)^t)$ von $\mathbb{R}^3$. Ein Vektor x liegt genau dann im Durchschnitt $U \cap U'$, wenn es $\lambda, \mu \in K$ gibt mit

$$x = \lambda \cdot (1,1,1)^t = (\lambda, \lambda, \lambda)^t$$

und

$$x = \mu \cdot (1,0,1)^t = (\mu, 0, \mu)^t.$$

Gleichsetzen der beiden Ausdrücke liefert die Bedingungen $\lambda = \mu$ und $\lambda = 0$, also gilt $x = (0,0,0)^t$, d.h.

$$U \cap U' = \{(0,0,0)^t\}.$$

Damit ist die Summe $U + U'$ eine direkte Summe.

D) Lineare Abbildungen

Zu jeder Struktur gehören die strukturerhaltenden Abbildungen.

Definition 28.19 (Lineare Abbildungen).
Es seien V und W zwei K-Vektorräume.

a. Eine Abbildung $f : V \to W$ heißt *K-lineare Abbildung* oder *Vektorraumhomomorphismus*, wenn für alle $\lambda \in K$ und $x, y \in V$ gilt

$$f(x + y) = f(x) + f(y) \quad \text{und} \quad f(\lambda \cdot x) = \lambda \cdot f(x).$$

b. Eine injektive (bzw. surjektive bzw. bijektive) K-lineare Abbildung heißt auch *Monomorphismus* (bzw. *Epimorphismus* bzw. *Isomorphismus*). Gilt $V = W$, so nennen wir eine K-lineare Abbildung auch einen *Endomorphismus*, und ist sie zudem bijektiv, so sprechen wir von einem *Automorphismus*.

c. Existiert ein Isomorphismus von V nach W, so nennen wir V und W *isomorph* und schreiben $V \cong W$.

d. Die Menge aller K-linearen Abbildungen von V nach W bezeichnen wir mit $\mathrm{Hom}_K(V, W)$ und die Menge aller Endomorphismen von V mit $\mathrm{End}_K(V)$.

Bemerkung 28.20.
Die beiden Bedingungen in Definition 28.19 a. lassen sich zusammenfassen zu der Bedingung $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$ für alle $\lambda, \mu \in K$ und $x, y \in V$.

Beispiel 28.21.
Wir wollen uns nun einige Beispiele für lineare Abbildungen anschauen.

a. Die Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2)^t \mapsto x_1 - x_2$ ist $\mathbb{R}$-linear.
Denn für $x = (x_1, x_2)$ und $y = (y_1, y_2)$ sowie $\lambda \in \mathbb{R}$ gilt

$$\begin{aligned} f(x + y) &= f\big((x_1 + y_1, x_2 + y_2)\big) = x_1 + y_1 - x_2 - y_2 \\ &= x_1 - x_2 + y_1 - y_2 = f(x) + f(y) \end{aligned}$$

und

$$f(\lambda x) = f\big((\lambda x_1, \lambda x_2)\big) = \lambda x_1 - \lambda x_2 = \lambda \cdot (x_1 - x_2) = \lambda \cdot f(x).$$

b. Ist I ein Intervall, so ist die Abbildung

$$D : \mathcal{C}^1(I, \mathbb{R}) \longrightarrow \mathcal{C}(I, \mathbb{R}) : f \mapsto f'$$

$\mathbb{R}$-linear, da aus der Linearität der Ableitung 17.9 folgt

$$D(\lambda f + \mu g) = (\lambda f + \mu g)' = \lambda \cdot f' + \mu \cdot g' = \lambda \cdot D(f) + \mu \cdot D(g).$$

c. Die Abbildung

$$f : \mathbb{R}[t] \longrightarrow \mathbb{R}^{\mathbb{R}} : \sum_{k=0}^{n} a_k \cdot t^k \mapsto (a_k)_{k\in\mathbb{N}},$$

wobei $a_k = 0$ für $k > n$ gelten soll, ist eine injektive $\mathbb{R}$-lineare Abbildung, also ein Monomorphismus. Ihr Bild, der Unterraum

$$\mathrm{Im}(f) = \{(a_k)_{k\in\mathbb{N}} \mid \text{ nur endlich viele } a_k \text{ sind nicht } 0\}$$

der *abbrechenden Folgen* in $\mathbb{R}$, ist mithin isomorph zu $K[t]$.

d. Die formale Ableitung

$$d : K[t] \longrightarrow K[t] : \sum_{k=0}^{n} a_k \cdot t^k \mapsto \sum_{k=1}^{n} k \cdot a_k \cdot t^{k-1}$$

ist eine K-lineare Abbildung, wie man leicht nachrechnet.

Lemma 28.22 (Einfache Eigenschaften linearer Abbildungen).
Seien U, V und W K-Vektorräume und $f : U \longrightarrow V$ und $g : V \longrightarrow W$ seien K-linear. Ferner seien $x, x_1, \ldots, x_n \in U$ und $\lambda_1, \ldots, \lambda_n \in K$. Dann gelten:

a. $f(0_U) = 0_V$ und $f(-x) = -f(x)$.

b. $f(\lambda_1 x_1 + \ldots + \lambda_n x_n) = \lambda_1 f(x_1) + \ldots + \lambda_n f(x_n)$.

c. Ist f bijektiv, so ist $f^{-1} : V \longrightarrow U$ K-linear.

d. $g \circ f : U \longrightarrow W$ ist K-linear.

e. $\mathrm{Hom}_K(U, V)$ ist ein Unterraum von V^U.

Beweis: a. Aus der Verträglichkeit mit der Skalarmultiplikation folgen

$$f(0_U) = f(0_K \cdot 0_U) = 0_K \cdot f(0_U) = 0_V$$

und

$$f(-x) = f((-1) \cdot x) = (-1) \cdot f(x) = -f(x).$$

b. Die Aussage folgt mittels Induktion aus den beiden Bedingungen für Linearität.

c. Seien $y, y' \in V$ und $\lambda, \lambda' \in K$ sowie $x = f^{-1}(y)$ und $x' = f^{-1}(y')$. Wegen der Linearität von f gilt

$$f(\lambda \cdot x + \lambda' \cdot x') = \lambda \cdot f(x) + \lambda' \cdot f(x').$$

Wenden wir auf beiden Seiten f^{-1} an, so erhalten wir

$$\begin{aligned}\lambda \cdot f^{-1}(y) + \lambda' \cdot f^{-1}(y') &= \lambda \cdot x + \lambda' \cdot x' = f^{-1}\big(f(\lambda \cdot x + \lambda' \cdot x')\big) \\ &= f^{-1}\big(\lambda \cdot f(x) + \lambda' \cdot f(x')\big) = f^{-1}\big(\lambda \cdot y + \lambda' \cdot y'\big).\end{aligned}$$

Mithin ist f^{-1} eine lineare Abbildung.

d. Seien $\lambda, \mu \in K$ und $x, y \in U$, so gelten

$$\begin{aligned}(g \circ f)(\lambda x + \mu y) &= g\big(f(\lambda x + \mu y)\big) = g(\lambda f(x) + \mu f(y)\big) \\ &= \lambda g\big(f(x)\big) + \mu g(\big(f(y)\big) = \lambda(g \circ f)(x) + \mu(g \circ f)(y).\end{aligned}$$

e. Dies folgt aus Aufgabe 28.40.

□

Proposition 28.23 (f_A ist linear.).
Für $A \in \mathrm{Mat}(m \times n, K)$ ist $f_A : K^n \longrightarrow K^m$ eine K-lineare Abbildung.

Beweis: Aus Lemma 27.8 folgt für $x, y \in K^n$ und $\lambda \in K$

$$f_A(x + y) = A(x + y) = Ax + Ay = f_A(x) + f_A(y)$$

und

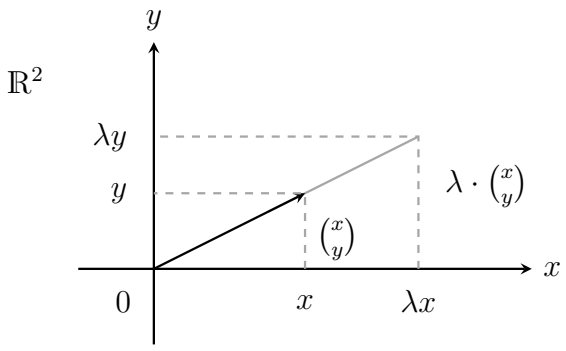
$$f_A(\lambda x) = A(\lambda x) = \lambda(Ax) = \lambda f_A(x).$$

□

Beispiel 28.24.

a. Im Fall $n = 1$ und $A = (a)$ ist die K-lineare Abbildung $f_A : K \to K : x \mapsto a \cdot x$ gerade die Multiplikation mit a.

b. Die lineare Abbildung $f_A : \mathbb{R}^2 \mapsto \mathbb{R}^2 : (x, y)^t \mapsto (\lambda x, \lambda y)^t$ zu $A = \lambda \mathbb{1}_2$ ist eine *Streckung* um den Faktor λ.



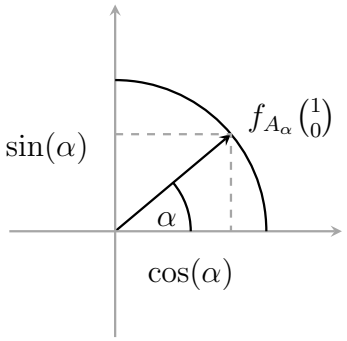
c. Für $\alpha \in \mathbb{R}$ setzen wir

$$A_\alpha := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Dann ist die lineare Abbildung $f_{A_\alpha} : \mathbb{R}^2 \to \mathbb{R}^2$ eine Drehung um den Winkel α. Beachte dazu, daß

$$A_\alpha e_1 = \big(\cos(\alpha), \sin(\alpha)\big)^t \quad \text{und} \quad A_\alpha e_2 = \big(-\sin(\alpha), \cos(\alpha)\big)^t,$$

woraus die Aussage für die *Einheitsvektoren* e_1 und e_2 unmittelbar folgt.



Daraus leitet sich die Aussage für einen beliebigen Vektor $(x, y)^t$ mittels der Linearität von f_{A_α} ab: $f_{A_\alpha}\big((x, y)^t\big) = x f_{A_\alpha}(e_1) + y f_{A_\alpha}(e_2)$.

d. Ist $n \geq m$, so ist die Abbildung

$$\mathrm{pr} : K^n \to K^m : (x_1, \ldots, x_n)^t \mapsto (x_1, \ldots, x_m)^t$$

eine K-lineare Abbildung, genannt die kanonische *Projektion*.
Ist $m \geq n$, dann ist die kanonische *Inklusion*

$$i_{K^n} : K^n \to K^m : (x_1, \ldots, x_n)^t \mapsto (x_1, \ldots, x_n, 0, \ldots, 0)^t$$

ebenfalls K-linear. Beides prüft man leicht nach.

Proposition 28.25 (Kern und Bild sind Unterräume).
Es seien V und W K-Vektorräume und $f : V \longrightarrow W$ sei K-linear.

a. Ist U ein Unterraum von V, so ist $f(U)$ ein Unterraum von W.

b. Ist U ein Unterraum von W, so ist $f^{-1}(U)$ ein Unterraum von V.

c. Das *Bild* $\mathrm{Im}(f) = f(V)$ von f ist ein Unterraum von W.

d. Der *Kern* von f, $\mathrm{Ker}(f) = \{x \in V \mid f(x) = 0\}$, ist ein Unterraum von V.

Beweis:

a. Es sei U ein Unterraum von V. Dann ist $0_V \in U$ und somit $0_W = f(0_V) \in f(U)$, so daß $f(U)$ nicht leer ist. Sind $\lambda \in K$ und $u = f(x), v = f(y) \in f(U)$ mit $x, y \in U$, so gilt

$$u + v = f(x) + f(y) = f(x + y) \in f(U)$$

und

$$\lambda u = \lambda f(x) = f(\lambda x) \in f(U).$$

Also ist $f(U)$ ein Unterraum von W.

b. Es sei U ein Unterraum von W. Dann ist $0_W \in U$ und wegen $f(0_V) = 0_W$ ist dann $0_V \in f^{-1}(U)$, so daß $f^{-1}(U)$ nicht leer ist. Sind $\lambda \in K$ und $x, y \in f^{-1}(U)$, so gilt $f(x), f(y) \in U$ und somit

$$f(x+y) = f(x) + f(y) \in U$$

und

$$f(\lambda x) = \lambda f(x) \in U.$$

Also auch $x + y \in f^{-1}(U)$ und $\lambda x \in f^{-1}(U)$, so daß $f^{-1}(U)$ ein Unterraum von V ist.

c. Dies folgt aus a. mit $U = V$.

d. Dies folgt aus b. mit $U = \{0_W\}$.

□

Beispiel 28.26.
Die Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2) \mapsto x_1 - x_2$ aus Beispiel 28.21 hat den Kern

$$\mathrm{Ker}(f) = \{x \in \mathbb{R}^2 \mid f(x) = x_1 - x_2 = 0\} = \left\{(\lambda, \lambda)^t \;\middle|\; \lambda \in K\right\}$$

und das Bild ist $\mathrm{Im}(f) = \mathbb{R}$.

Proposition 28.27 (Injektivität linearer Abbildungen).
Eine lineare Abbildung $f : V \longrightarrow W$ ist genau dann injektiv, wenn

$$\mathrm{Ker}(f) = \{0_V\}.$$

Beweis: Die Aussage folgt unmittelbar aus Lemma 22.23, da f auch ein Gruppenhomomorphismus der additiven Gruppen ist. □

E) Faktorräume

Da ein Unterraum U eines Vektorraums V zugleich eine Untergruppe der abelschen Gruppe $(V, +)$ ist, sind der Begriff der *Nebenklasse* von $x \in V$ modulo U und die Faktorgruppe $(V/U, +)$ bereits aus Abschnitt 24 bekannt. Wir wollen sie uns aber hier noch einmal ins Gedächtnis rufen und dann zeigen, daß V/U mit der vertreterweise definierten Skalarmultiplikation sogar ein Vektorraum ist.

Definition 28.28 (Faktorraum).
Es sei V ein K-Vektorraum und U ein Unterraum von V.

a. Für $x \in V$ heißt

$$\overline{x} := x + U := \{x + u \mid u \in U\}$$

die *Restklasse* oder *Nebenklasse* von x modulo U und x heißt ein *Vertreter* der Restklasse. Man nennt $x + U$ auch einen *affinen Raum* parallel zum Unterraum U mit Aufpunkt x.

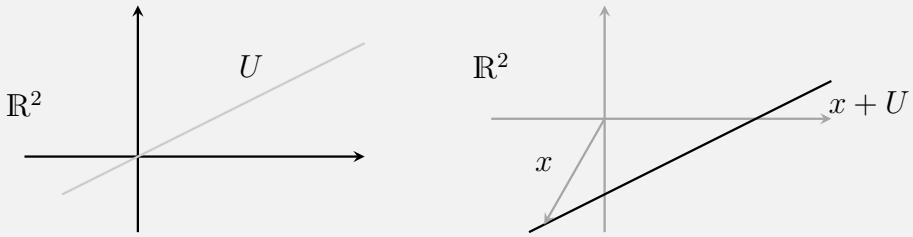


Abbildung 2: Ein affiner Raum $x + U$ zum Unterraum U.

Man beachte, daß aus der Notation $\overline{x}$ nicht mehr abzulesen ist, modulo welchem Unterraum man rechnet. Die Notation ist deshalb mit Vorsicht zu verwenden.

b. Wir nennen die Menge der Restklassen modulo U

$$V/U := \{x + U \mid x \in V\} = \{\overline{x} \mid x \in V\}$$

auch den *Faktorraum* von V modulo U.

Lemma 28.29 (Rechnen mit Restklassen).
Es sei V ein K-Vektorraum, U ein Unterraum, $x, x', y, y' \in V$ und $\lambda \in K$. In V/U gelten dann die folgenden Aussagen:

a. Entweder $\overline{x} = \overline{y}$ oder $\overline{x} \cap \overline{y} = \emptyset$.

b. Es gilt:

$$\overline{x} = \overline{y} \quad \Longleftrightarrow \quad x - y \in U.$$

Insbesondere, $\overline{x} = \overline{0} = U$ genau dann, wenn $x \in U$.

c. Gilt $\overline{x} = \overline{x'}$ und $\overline{y} = \overline{y'}$, so gelten auch

$$\overline{x + y} = \overline{x' + y'} \quad \text{und} \quad \overline{\lambda x} = \overline{\lambda x'}.$$

Beweis:

a. Dies folgt unmittelbar aus Proposition 6.10, da $\overline{x}$ und $\overline{y}$ nach Proposition 24.2 Äquivalenzklassen sind.

b. Dies folgt aus Proposition 24.2.

c. Wir wollen hier zur Verdeutlichung $x+U$ und $y+U$ statt $\overline{x}$ und $\overline{y}$ schreiben. Aus $x+U=x'+U$ sowie $y+U=y'+U$ folgt nach b.

$$x-x', y-y' \in U.$$

Damit gilt dann auch

$$(x+y)-(x'+y')=(x-x')+(y-y') \in U$$

und

$$(\lambda x-\lambda x')=\lambda\cdot(x-x') \in U.$$

Wegen b. gilt dann wieder $(x+y)+U=(x'+y')+U$ und $\lambda x+U=\lambda x'+U$.

□

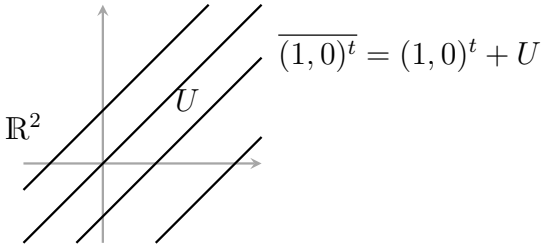
Beispiel 28.30.
Sei $V=\mathbb{R}^2$ und $U=\operatorname{Lin}((1,1)^t)$. Dann gilt

$$\overline{(1,0)^t}=(1,0)^t+U=(2,1)^t+U=\overline{(2,1)^t}$$

und

$$\overline{(1,0)^t}=(1,0)^t+U\neq(0,1)^t+U=\overline{(0,1)^t}.$$

Die Restklassen in V/U sind in diesem Fall genau die zu U parallelen Geraden.



Satz 28.31 (Der Faktorraum ist ein Vektorraum.).
Es sei V ein K-Vektorraum und U ein Unterraum. Dann definiert

(75) $$\overline{x}+\overline{y}:=\overline{x+y}$$

und

(76) $$\lambda\cdot\overline{x}:=\overline{\lambda x}$$

für $\overline{x},\overline{y}\in V/U$ und $\lambda\in K$ eine Addition und eine Skalarmultiplikation auf V/U bezüglich derer der Faktorraum V/U ein K-Vektorraum ist.
Zudem ist die Abbildung

$$\pi: V\longrightarrow V/U: x\mapsto\overline{x}$$

eine surjektive K-lineare Abbildung mit $\operatorname{Ker}(\pi)=U$, die wir die *Restklassenabbildung* nennen.

Beweis: Wir beachten zunächst, daß die Addition und Skalarmultiplikation wohldefiniert sind, weil sie nach Lemma 28.29 nicht von den gewählten Vertretern der Nebenklassen abhängen. Aus Satz 24.14 wissen wir zudem schon, daß V/U mit der angegebenen Addition eine abelsche Gruppe ist.

Es bleiben also nur die Axiome für die Skalarmultiplikation nachzurechnen. Dazu seien $\overline{x}, \overline{y} \in V/U$ und $\lambda, \mu \in K$ gegeben. Die gesuchten Axiome vererben sich von V auf V/U:

$$(\lambda + \mu) \cdot \overline{x} = \overline{(\lambda + \mu) \cdot x} = \overline{\lambda x + \mu x} = \overline{\lambda x} + \overline{\mu x} = \lambda \cdot \overline{x} + \mu \cdot \overline{x}$$

und

$$\lambda \cdot \overline{x + y} = \overline{\lambda \cdot (x + y)} = \overline{\lambda x + \lambda y} = \overline{\lambda x} + \overline{\lambda y} = \lambda \cdot \overline{x} + \lambda \cdot \overline{y}$$

und

$$(\lambda \cdot \mu) \cdot \overline{x} = \overline{(\lambda \cdot \mu) \cdot x} = \overline{\lambda \cdot (\mu \cdot x)} = \lambda \overline{\mu \cdot x} = \lambda \cdot \big(\mu \cdot \overline{x}\big)$$

und

$$1 \cdot \overline{x} = \overline{1 \cdot x} = \overline{x}.$$

Also ist V/U ein K-Vektorraum.

Es bleibt, die Aussagen zur Restklassenabbildung π zu zeigen. Die Linearität von π folgt aus der Definition der Operationen auf V/U. Sind $x, y \in V$ und $\lambda \in K$, dann gilt

$$\pi(x + y) = \overline{x + y} = \overline{x} + \overline{y} = \pi(x) + \pi(y)$$

und

$$\pi(\lambda x) = \overline{\lambda x} = \lambda \cdot \overline{x} = \lambda \cdot \pi(x).$$

Außerdem ist π surjektiv, da jedes $\overline{x} \in V/U$ sich schreiben läßt als $\overline{x} = \pi(x)$. Ds gilt

$$x \in \operatorname{Ker}(\pi) \iff \overline{x} = \pi(x) = \overline{0} \iff x \in U.$$

Damit sind alle Aussagen des Satzes bewiesen. □

Bemerkung 28.32 (Die vier Rechenregeln für den Faktorraum).
Um mit dem Faktorraum rechnen zu können, braucht man nur die Rechenregeln:

a. $\overline{0} = 0 + U = U$ ist der Nullvektor.

b. $\overline{x} + \overline{y} = \overline{x + y}$.

c. $\lambda \cdot \overline{x} = \overline{\lambda x}$.

d. $\overline{x} = \overline{y} \iff x - y \in U$.

Satz 28.33 (Homomorphiesatz).
Ist $f : V \longrightarrow W$ eine K-lineare Abbildung, so ist

$$\overline{f} : V/\operatorname{Ker}(f) \longrightarrow \operatorname{Im}(f) : \overline{x} \mapsto f(x)$$

ein Isomorphismus. Insbesondere gilt also $V/\operatorname{Ker}(f) \cong \operatorname{Im}(f)$.

Beweis: Da wir für die Definition von $\overline{f}(\overline{x})$ wieder den Restklassenvertreter x verwendet haben, müssen wir wieder zeigen, daß unsere Definition nicht von der speziellen Wahl des Vertreters abhängt. Man sagt wieder, wir müssen die *Wohldefiniertheit* von $\overline{f}$ zeigen.

Seien dazu $\overline{x} = x + \operatorname{Ker}(f) = y + \operatorname{Ker}(f) = \overline{y}$ gegeben. Dann gilt

$$x - y \in \operatorname{Ker}(f),$$

und mithin $0 = f(x-y) = f(x) - f(y)$, oder alternativ $f(x) = f(y)$. Die Abbildung $\overline{f}$ ist also wohldefiniert.

Die Linearität von $\overline{f}$ folgt dann aus der Linearität von f. Seien dazu $\overline{x}, \overline{y} \in V/\operatorname{Ker}(f)$ und $\lambda \in K$, dann gilt

$$\overline{f}\big(\overline{x} + \overline{y}\big) = \overline{f}\big(\overline{x+y}\big) = f(x+y) = f(x) + f(y) = \overline{f}\big(\overline{x}\big) + \overline{f}\big(\overline{y}\big)$$

und

$$\overline{f}\big(\lambda \cdot \overline{x}\big) = \overline{f}\big(\overline{\lambda \cdot x}\big) = f(\lambda \cdot x) = \lambda \cdot f(x) = \lambda \cdot \overline{f}\big(\overline{x}\big).$$

Es bleibt noch zu zeigen, daß $\overline{f}$ surjektiv und injektiv ist.

Ist $y \in \operatorname{Im}(f)$, so gibt es ein $x \in V$ mit $f(x) = y$, und somit gilt

$$y = f(x) = \overline{f}\big(\overline{x}\big).$$

Also ist $\overline{f}$ surjektiv.

Für die Injektivität nutzen wir Proposition 28.27. Es gilt

$$\overline{x} \in \operatorname{Ker}\big(\overline{f}\big) \iff 0 = \overline{f}\big(\overline{x}\big) = f(x) \iff x \in \operatorname{Ker}(f) \iff \overline{x} = \overline{0}.$$

Also enthält der Kern von $\overline{f}$ nur den Nullvektor, und somit ist $\overline{f}$ injektiv. □

F) Direkte Komplemente

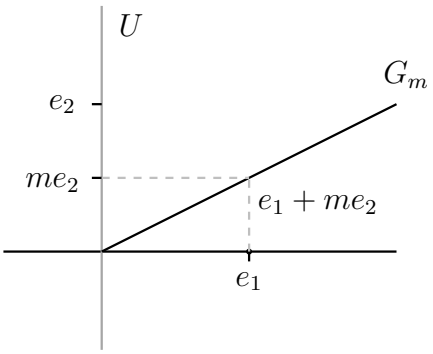
Definition 28.34 (Direkte Komplemente).
Es sei V ein K-Vektorraum und U und U' seien Unterräume von V. Dann heißt U' ein (direktes) *Komplement* von U, falls $V = U \oplus U'$.

Beispiel 28.35 (Komplemente sind nicht eindeutig.).
Ist $V = \mathbb{R}^2$ und $U = \operatorname{Lin}(e_2)$ die y-Achse, dann ist die Ursprungsgerade mit Steigung m

$$G_m := \operatorname{Lin}(e_1 + me_2)$$

für jedes $m \in \mathbb{R}$ ein Komplement von U. Beachte dazu nur, daß sich die Geraden U und G_m nur im Ursprung schneiden, d.h. $U \cap G_m = \{0\}$, und daß ein beliebiger Vektor $(x,y)^t \in \mathbb{R}^2$ sich schreiben läßt als

$$(x,y)^t = x \cdot (e_1 + me_2) + (-xm + y) \cdot e_2 \in G_m + U.$$



Proposition 28.36 (Der Faktorraum als Komplementersatz).
Sei V ein K-Vektorraum, U ein Unterraum von V und U' ein Komplement von U. Dann ist die Einschränkung der Restklassenabbildung

$$\pi_| : U' \to V/U : x \mapsto \overline{x}$$

auf U' ein Isomorphismus. Insbesondere sind je zwei Komplemente von U isomorph.

Beweis: Es ist klar, daß $\pi_|$ als Einschränkung einer K-linearen Abbildung wieder K-linear ist.

Wir zeigen zunächst, daß $\pi_|$ surjektiv ist. Sei dazu $\overline{x} \in V/U$ gegeben. Wegen $V = U \oplus U'$ läßt sich x als $x = y + z$ mit $y \in U$ und $z \in U'$ schreiben. Damit gilt:

$$\overline{x} = \overline{z} = \pi_|(z) \in \operatorname{Im}\left(\pi_|\right).$$

Also ist $\pi_|$ surjektiv.

Es bleibt zu zeigen, daß $\pi_|$ injektiv ist, d. h. $\operatorname{Ker}\left(\pi_|\right) = \{0\}$. Sei dazu $z \in \operatorname{Ker}\left(\pi_|\right)$, dann gilt

$$\overline{0} = \pi_|(z) = \overline{z}.$$

D. h. $z \in U$. Damit gilt aber $z \in U \cap U' = \{0\}$, also $z = 0$.

Seien schließlich U' und U'' zwei Komplemente von U, dann ist die Komposition

$$U' \xrightarrow{\pi_{|U'}} V/U \xrightarrow{\pi^{-1}_{|U''}} U''$$

ein Isomorphismus von U' nach U'' und die Komplemente sind isomorph zueinander. □

Bemerkung 28.37.
Da V/U isomorph zu jedem Komplement von U ist, kann man in vielen Situationen ein Komplement schlicht durch den Faktorraum ersetzen. Während es sehr viele Komplemente von U geben kann, gibt es nur *einen* Faktorraum. Dieser ist durch U eindeutig bestimmt. Das ist unter Umständen ein großer Vorteil!

Aufgaben

Aufgabe 28.38.
Welche der folgenden Teilmengen von K^3 sind Unterräume des K^3? Begründe Deine Anworten.

a. $\{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid x_1 \cdot x_2 = 2x_3\}$ für $K = \mathbb{R}$.

b. $\{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid ax_1 + x_2 + x_3 = a + 1\}$ für ein festes $a \in \mathbb{R}$ für $K = \mathbb{R}$.

c. $\{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid x_1 \leq 0\}$ für $K = \mathbb{R}$.

d. $\{(1, 0, 0)^t, (0, 1, 0)^t, (1, 1, 0)^t, (0, 0, 0)^t\}$ für $K = \mathbb{R}$ oder $K = \mathbb{F}_2$.

Aufgabe 28.39.
Gegeben seien die folgenden Teilmengen des $\mathbb{Q}$–Vektorraums $V = \mathbb{Q}^4$:

a. $U_1 = \{(x, x + 1, x + 2, x + 4)^t \mid x \in \mathbb{Q}\}$,

b. $U_2 = \{(x, 2x, 3x, 4x)^t \mid x \in \mathbb{Q}\}$,

c. $U_3 = \{(x_1, x_2, x_3, x_4)^t \mid x_1, x_2, x_3, x_4 \in \mathbb{Q},\ x_3 = x_1 + x_2,\ x_4 = x_2 + x_3\}$,

d. $U_4 = \{(x_1, x_2, x_3, x_4)^t \mid x_1, x_2, x_3, x_4 \in \mathbb{Q},\ x_2 \geq 0\}$,

e. $U_5 = \{(x_1, x_2, x_3, x_4)^t \mid x_1, x_2, x_3, x_4 \in \mathbb{Z}\}$.

Welche dieser Mengen sind Unterräume von V? Begründe Deine Aussage.

Aufgabe 28.40.
Seien U, V und W K-Vektorräume, $\lambda, \lambda' \in K$ und $f, f' \in \mathrm{Hom}_K(U, V)$ und $g, g' \in \mathrm{Hom}_K(V, W)$. Dann gelten:

a. $f + f', \lambda \cdot f \in \mathrm{Hom}_K(U, V)$, d.h. $\mathrm{Hom}_K(U, V)$ ist ein Unterraum von V^U.

b. $g \circ (\lambda f + \lambda' f') = \lambda(g \circ f) + \lambda'(g \circ f')$ und $(\lambda g + \lambda' g') \circ f = \lambda(g \circ f) + \lambda'(g' \circ f)$.

c. $\lambda(g \circ f) = (\lambda g) \circ f = g \circ (\lambda f)$.

Aufgabe 28.41.
Seien V, W K–Vektorräume und $f, g \in \mathrm{Hom}_K(V, W)$. Zeige

$$\mathrm{Ker}(f + g) \supseteq \mathrm{Ker}(f) \cap \mathrm{Ker}(g)$$

und

$$\mathrm{Im}(f + g) \subseteq \mathrm{Im}(f) + \mathrm{Im}(g).$$

Finde außerdem Beispiele, so dass die Inklusionen strikt sind.

Aufgabe 28.42 (f-invariante Unterräume).
Ist V ein K-Vektorraum, $f : V \longrightarrow V$ K-linear und $U \leq V$ ein Unterraum von V mit $f(U) \subseteq U$, so nennen wir U einen *f-invarianten* Unterraum von V.
Zeige, daß durch

$$f_U : U \longrightarrow U : x \mapsto f(x)$$

und

$$f_{V/U} : V/U \longrightarrow V/U : \overline{x} \mapsto \overline{f(x)}$$

K-lineare Abbildungen definiert werden.

Aufgabe 28.43.
Es sei $V = \mathbb{R}^{\mathbb{R}}$ der $\mathbb{R}$-Vektorraum aller Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$ und

$$U = \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid f(-x) = -f(x) \; \forall \, x \in \mathbb{R}\}$$

sowie

$$U' = \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid f(-x) = f(x) \; \forall \, x \in \mathbb{R}\}.$$

Zeige, daß U und U' Unterräume von V sind mit $V = U \oplus U'$.

Aufgabe 28.44 (Erster Isomorphiesatz).
Sei V ein K–Vektorraum und $U, U' \leq V$. Zeige

$$U/(U \cap U') \cong (U + U')/U'.$$

Aufgabe 28.45 (Projektionen).
Es sei V ein K-Vektorraum. $f \in \mathrm{End}_K(V)$ heißt *Projektion*, falls $f^2 = f$ gilt.
Zeige, die folgenden Aussagen sind äquivalent:

a. f ist eine Projektion,

b. $\mathrm{id}_V - f$ ist eine Projektion,

c. $\mathrm{Im}\left(\mathrm{id}_V - f\right) = \mathrm{Ker}(f)$,

d. $\mathrm{Ker}\left(\mathrm{id}_V - f\right) = \mathrm{Im}(f)$.

Zeige auch, sind obige Bedingungen erfüllt, so gilt zudem $V = \mathrm{Ker}(f) \oplus \mathrm{Im}(f)$.

§ 29 Basen von Vektorräumen

In diesem Abschnitt ist V stets ein K-Vektorraum.

Das wesentlichste Konzept im Zusammenhang mit Vektorräumen ist das der Basis. Mit Hilfe einer Basis können die Elemente eines Vektorraums effizient auf eindeutige Weise dargestellt werden. Wir führen in diesem Kapitel Basen als linear unabhängige Erzeugendensysteme ein.

A) Linear unabhängige Familien von Vektoren

Definition 29.1 (Familien).
Es seien I und X zwei Mengen.

a. Wir nennen ein Tupel der Form $F = (x_i)_{i\in I}$ mit $x_i \in X$ eine *Familie* von Elementen in X. Ist I endlich, so nennen wir die Familie *endlich* und setzen $|F| := |I|$.

b. Ist $F = (x_i)_{i\in I}$ eine Familie von Elementen in X, schreiben wir $x \in F$, wenn es ein $i \in I$ gibt mit $x = x_i$.

c. Sind $F = (x_i)_{i\in I}$ und $G = (y_i)_{i\in I}$ zwei Familien von Elementen in X, so schreiben wir $F = G$, falls $x_i = y_i$ für alle $i \in I$.

d. Ist $F = (x_i)_{i\in I}$ eine Familie von Elementen in X und $J \subseteq I$, so nennen wir $F' = (x_j)_{j\in J}$ eine *Teilfamilie* von F und F eine *Oberfamilie* von F', und wir schreiben $F' \subseteq F$. Ebenso schreiben wir $x \in F$, um auszudrücken, daß $x = x_i$ für ein $i \in I$ gilt.

e. Wir schreiben kurz $\operatorname{Lin}(F)$ für die lineare Hülle $\operatorname{Lin}(\{x_i \mid i \in I\})$ und nennen $\operatorname{Lin}(F)$ die *lineare Hülle* von F.

Beispiel 29.2.
Ist $I = \{1, 2, 3\}$ und $X = \mathbb{R}^2$, so wird durch $x_1 = (1, 0)^t$, $x_2 = (1, 1)^t$, $x_3 = (1, 0)^t$ eine endliche Familie $(x_1, x_2, x_3) = \big((1, 0)^t, (1, 1)^t, (1, 0)^t\big)$ definiert. $(x_1, x_3) = \big((1, 0)^t, (1, 0)^t\big)$ ist eine Teilfamilie.

Bemerkung 29.3 (Familien von Vektoren).

a. In einer Familie können Elemente auch *mehrfach* auftreten, in einer Menge geht das nicht. Z.B. $F = \big((1, 0)^t, (1, 1)^t, (1, 0)^t\big)$.

b. Ist die Menge I geordnet, so ordnen wir die Mitglieder der Familie F in der gleichen Weise. Z.B. $\big((1, 0)^t, (1, 1)^t, (1, 0)^t\big) \neq \big((1, 0)^t, (1, 0)^t, (1, 1)^t\big)$.

c. In unseren Anwendungen wird die Menge I meist $\{1, \ldots, n\}$ für eine positive natürliche Zahl n sein, und wir ordnen die Elemente dann in der naheliegenden Weise.

d. Formal korrekt sollte man die Familie $F = (x_i)_{i \in I}$ als Abbildung $F : I \longrightarrow X : i \mapsto x_i$ angeben. Die Tupelschreibweise ist aber suggestiver als die Schreibweise als Abbildung.

Definition 29.4 (Lineare Unabhängigkeit).
Es sei V ein K-Vektorraum und F eine Familie von Vektoren in V.

a. Eine endliche Familie $(x_1, \ldots, x_n)$ von Vektoren in V heißt *linear unabhängig*, wenn aus
$$\lambda_1 \cdot x_1 + \ldots + \lambda_n \cdot x_n = 0$$
stets
$$\lambda_1 = \ldots = \lambda_n = 0$$
folgt, d.h. wenn nur die triviale Linearkombination der x_i Null ergibt.
Wir sagen dann oft einfach, die Vektoren $x_1, \ldots, x_n$ seien linear unabhängig.

b. Eine endliche Familie $(x_1, \ldots, x_n)$ von Vektoren in V heißt *linear abhängig*, wenn es Skalare $\lambda_1, \ldots, \lambda_n \in K$ gibt, so daß
$$\lambda_1 \cdot x_1 + \ldots + \lambda_n \cdot x_n = 0,$$
aber nicht alle λ_i sind Null, d.h. wenn eine nicht-triviale Linearkombination der x_i Null ergibt.
Wir nennen oft einfach die Vektoren $x_1, \ldots, x_n$ linear abhängig.

c. F heißt *linear abhängig*, wenn es eine endliche linear abhängige Teilfamilie gibt.

d. F heißt *linear unabhängig*, wenn jede endliche Teilfamilie linear unabhängig ist.

Beispiel 29.5 (Lineare Unabhängigkeit).

a. Die Einheitsvektoren $e_1, \ldots, e_n \in K^n$ sind linear unabhängig. Denn aus
$$(0, \ldots, 0)^t = \lambda_1 \cdot e_1 + \ldots + \lambda_n \cdot e_n = (\lambda_1, \ldots, \lambda_n)^t$$
folgt unmittelbar $\lambda_1 = \ldots = \lambda_n = 0$.

b. Die Familie $\big((1,0)^t, (0,1)^t, (1,1)^t\big)$ ist linear abhängig, da
$$(1,0)^t + (0,1)^t - (1,1)^t = (0,0)^t.$$

c. Wir betrachten die Folge $e_k = (\delta_{n,k})_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, die als k-ten Eintrag eine Eins hat und ansonsten konstant Null ist. Dann ist die Familie $F = (e_k)_{k\in\mathbb{N}}$ linear unabhängig in $\mathbb{R}^{\mathbb{N}}$.
Um das zu sehen, betrachten wir die endliche Teilfamilie $(e_{k_1}, \ldots, e_{k_l})$ für $0 \leq k_1 < \ldots < k_l$. Dann folgt aus

$$\lambda_{k_1} \cdot e_{k_1} + \ldots + \lambda_{k_l} \cdot e_{k_l} = (0)_{n\in\mathbb{N}}$$

unmittelbar $\lambda_{k_1} = \ldots = \lambda_{k_l} = 0$, da die linke Folge als Folgenglied k_i den Wert λ_{k_i} hat. Also ist jede endliche Teilfamilie von F linear unabhängig, und somit ist auch F linear unabhängig.

Lemma 29.6 (Kriterien für lineare Abhängigkeit).
Es sei $F = (x_i)_{i\in I}$ eine Familie von Vektoren im K-Vektorraum V.

a. Ist $0 \in F$, so ist F linear abhängig.

b. Gibt es ein $i \neq j$ mit $x_i = x_j$, so ist F linear abhängig.

c. F ist genau dann linear abhängig, wenn es ein $x \in F$ gibt, das Linearkombination anderer Vektoren in F ist.

Beweis: Im ersten Fall ist $1 \cdot 0_V = 0_V$ eine nicht-triviale Linearkombination, im zweiten Fall ist $1 \cdot x_i - 1 \cdot x_j = 0_V$ eine solche. In jedem Fall ist F also linear abhängig, weil F eine endliche linear abhängige Teilfamilie enthält. Damit sind a. und b. gezeigt.

Ist F linear abhängig, so gibt es eine nicht-triviale Linearkombination

$$\sum_{j\in J} \lambda_j \cdot x_j = 0$$

mit $J \subseteq I$ endlich und nicht alle λ_j sind Null. Sei also $i \in J$ mit $\lambda_i \neq 0$, dann ist

$$x_i = \sum_{i\neq j\in J} -\frac{\lambda_j}{\lambda_i} \cdot x_j$$

Linearkombination anderer Vektoren in F.

Ist umgekehrt $x_i = \sum_{j\in J} \lambda_j \cdot x_j$ mit $J \subseteq I$ endlich und $i \in I \setminus J$, so ist

$$-x_i + \sum_{j\in J} \lambda_j \cdot x_j = 0$$

eine nicht-triviale Linearkombination, die Null ergibt. Mithin ist F linear abhängig. □

Beispiel 29.7.
In Beispiel 29.5 b. gilt

$$(1,0)^t = -(0,1)^t + (1,1)^t,$$

woraus ebenfalls die lineare Abhängigkeit der Familie folgt.

Notation 29.8 (Linearkombination).
Sei $F = (x_i)_{i\in I}$ eine Familie von Vektoren in V und I sei nicht notwendigerweise endlich. Wir werden des öfteren

$$x = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot x_i \tag{77}$$

schreiben, wenn wir sagen wollen, daß x eine Linearkombination von Vektoren in F ist. Formal korrekt müßte es lauten: es gibt eine endliche Teilfamilie $(x_j)_{j\in J}$ von F und Skalare $\lambda_j \in K$ für $j \in J$, so daß

$$x = \sum_{j\in J} \lambda_j \cdot x_j.$$

Wir interpretieren dies so, daß in (77) nur endlich viele der λ_i nicht Null sind, und daß somit die Summe auf der rechten Seite doch eine endliche Summe ist.

Mit dieser neuen Notation ist F genau dann linear unabhängig, wenn aus

$$\sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot x_i = 0$$

stets $\lambda_i = 0$ für alle $i \in I$ folgt; und analog ist F linear abhängig, wenn es eine Linearkombination

$$\sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot x_i = 0$$

gibt, bei der nicht alle λ_i Null sind.

Lemma 29.9 (Ergänzung linear unabhängiger Familien).
Ist $B = (x_i)_{i\in I}$ eine linear unabhängige Familie in V mit $\operatorname{Lin}(B) \subsetneq V$, so ist die Familie $(x, x_i \mid i \in I)$ für jedes $x \in V \setminus \operatorname{Lin}(B)$ linear unabhängig.

Beweis: Seien dazu $\lambda, \lambda_i \in K$, $i \in I$, mit

$$\lambda x + \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i x_i = 0.$$

Wäre $\lambda \neq 0$, so wäre

$$x = \sum_{\substack{i\in I\\ \text{endlich}}} -\frac{\lambda_i}{\lambda} \cdot x_i \in \operatorname{Lin}(B)$$

im Widerspruch zur Wahl von x. Also ist $\lambda = 0$, und somit folgt aus

$$\sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i x_i = \lambda x + \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i x_i = 0$$

und der linearen Unabhängigkeit von B, daß auch alle anderen λ_i Null sind. Also ist $(x, x_i \mid i \in I)$ linear unabhängig. □

B) Erzeugendensysteme und Basen

Definition 29.10 (Erzeugendensystem und Basis).
Es sei V ein K-Vektorraum und F eine Familie von Vektoren in V.

a. F heißt ein *Erzeugendensystem* von V, wenn $V = \mathrm{Lin}(F)$, d.h. wenn jeder Vektor in V eine Linearkombination von Vektoren in F ist.

b. F heißt eine *Basis* von V, wenn F ein linear unabhängiges Erzeugendensystem von V ist.

c. V heißt *endlich erzeugt*, wenn V ein endliches Erzeugendensystem besitzt.

Beispiel 29.11 (Erzeugendensystem und Basis).

a. Die Familie $B = (e_1, \ldots, e_n)$ der Einheitsvektoren im K^n ist eine Basis des K^n, die wir auch die *Standardbasis* oder die *kanonische Basis* des K^n nennen.
Denn nach Beispiel 29.5 ist B linear unabhängig und zudem ist ein beliebiger Vektor $x = (x_1, \ldots, x_n)^t \in K^n$ eine Linearkombination

$$x = x_1 \cdot e_1 + \ldots + x_n \cdot e_n$$

der Vektoren in B.

b. Analog sieht man, daß für $n, m \geq 1$ die Familie

$$\left(E_i^j \mid i = 1, \ldots, m; j = 1, \ldots, n\right),$$

wobei $E_i^j = (e_{lk})_{l=1,\ldots,m;k=1,\ldots,n}$ mit

$$e_{lk} = \delta_{il} \cdot \delta_{jk} = \begin{cases} 1, & \text{falls } l = i \text{ und } k = j, \\ 0, & \text{sonst} \end{cases}$$

die Matrix ist, die in Zeile i und Spalte j eine Eins als Eintrag hat und sonst nur Nullen, eine Basis des K-Vektorraums $\mathrm{Mat}(m \times n, K)$ ist.

c. Die Familie $\left((1,0)^t, (0,1)^t, (1,1)^t\right)$ in Beispiel 29.5 b. ist ein Erzeugendensystem von $\mathbb{R}^2$, aber keine Basis, da sie linear abhängig ist.

d. Die Familie $(e_k)_{k\in\mathbb{N}}$ im Vektorraum der Folgen $\mathbb{R}^{\mathbb{N}}$ aus Beispiel 29.5 c. ist *kein* Erzeugendensystem von $\mathbb{R}^{\mathbb{N}}$. Es scheint zwar, als gelte für eine beliebige Folge $(a_n)_{n\in\mathbb{N}}$

$$(a_n)_{n\in\mathbb{N}} = (a_0, a_1, a_2, \ldots) = \sum_{n\in\mathbb{N}} a_n \cdot e_n,$$

aber diese Summe ist *nicht* endlich und mithin keine zulässige Linearkombination! Die konstante Folge $(1)_{n\in\mathbb{N}}$ ist sicher keine endliche Linearkombination der e_k, da eine solche nur endlich viele Folgenglieder ungleich Null haben kann.

e. Die Familie $(1, i)$ ist eine Basis von $\mathbb{C}$ als $\mathbb{R}$-Vektorraum, da jede komplexe Zahl von der Gestalt $x + iy$ mit $x, y \in \mathbb{R}$ ist und da eine solche Zahl nur dann Null ist, wenn x und y beide Null sind.

f. Die Familie $B = (t^0, t^1, t^2, \ldots)$ ist eine Basis von $K[t]$.

Proposition 29.12 (Charakterisierung von Basen).
Für eine Familie B von Vektoren in V sind die folgenden Aussagen gleichwertig:

a. B ist eine Basis von V.

b. B ist ein minimales Erzeugendensystem von V.

c. B ist eine maximale linear unabhängige Familie in V.

Bemerkung 29.13.
Ein Erzeugendensystem B von V heißt *minimal*, wenn keine echte Teilfamilie von B ein Erzeugendensystem ist. Dies heißt *nicht*, daß sie in jedem anderen Erzeugendensystem enthalten ist! Es gibt nicht *das* minimale Erzeugendensystem.

Eine linear unabhängige Familie B in V heißt *maximal*, wenn keine echte Oberfamilie linear unabhängig ist. Dies heißt *nicht*, daß sie jede andere linear unabhänige Familie enthält! Es gibt nicht *die* maximale linear unabhängige Familie.

Beweis von Proposition 29.12: Es sei $B = (x_i)_{i \in I}$.

a. ⇒ b.: Ist B eine Basis, so erzeugt B den Vektorraum V per definitionem. Ist $(x_j \mid j \in J)$ eine echte Teilfamilie von B und ist $i \in I \setminus J$, so gibt es wegen der linearen Unabhängigkeit von B keine Darstellung

$$x_i - \sum_{\substack{j \in J \\ \text{endlich}}} \lambda_j x_j = 0$$

also ist $x_i \notin \operatorname{Lin}(x_j \mid j \in J)$.

b. ⇒ c.: Wir zeigen zunächst, daß B linear unabhängig ist. Angenommen, dies sei nicht der Fall, dann läßt sich nach Lemma 29.6 ein x_i als Linearkombination

$$x_i = \sum_{\substack{i \neq j \in I \\ \text{endlich}}} \lambda_j x_j$$

der übrigen Vektoren in B darstellen. Damit gilt dann aber

$$\operatorname{Lin}(x_j \mid j \in I \setminus \{i\}) = \operatorname{Lin}(x_j \mid j \in I) = V,$$

im Widerspruch zur Minimalität des Erzeugendensystems B.

Sei nun $(x_j \mid j \in J)$ mit $I \subsetneqq J$ eine echte Oberfamilie von B und $j \in J \setminus I$, so ist x_j eine Linearkombination

$$x_j = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i x_i$$

der Elemente in B, da B ein Erzeugendensystem ist. Folglich ist $(x_j \mid j \in J)$ linear abhängig nach Lemma 29.6.

c. $\Rightarrow$ a.: Da B linear unabhängig ist, bleibt zu zeigen, daß $\operatorname{Lin}(B) = V$. Gäbe es ein $x \in V \setminus \operatorname{Lin}(B)$, so wäre wegen Lemma 29.9 auch $(x, x_i \mid i \in I)$ linear unabhängig, im Widerspruch zur Maximalität von B.

□

Beispiel 29.14.
Wir betrachten die Familie $B = \big((1,1)^t, (1,-1)^t\big)$ in $\mathbb{R}^2$. Da sich ein beliebiger Vektor $(\lambda_1, \lambda_2)^t \in \mathbb{R}^2$ als

$$(\lambda_1, \lambda_2)^t = \frac{\lambda_1 + \lambda_2}{2} \cdot (1,1)^t + \frac{\lambda_1 - \lambda_2}{2} \cdot (1,-1)^t$$

schreiben läßt, ist B ein Erzeugendensystem von $\mathbb{R}^2$, und offenbar kann man weder $(1,1)^t$ noch $(1,-1)^t$ weglassen. B ist also ein minimales Erzeugendensystem und mithin eine Basis von $\mathbb{R}^2$.

Bemerkung 29.15 (Existenz von Basen).
Besitzt ein K-Vektorraum ein endliches Erzeugendensystem, so folgt aus Proposition 29.12 unmittelbar, daß er auch ein Basis besitzt, da wir aus dem Erzeugendensystem sukkzessive Elemente entfernen können, bis wir ein minimales Erzeugendensystem erhalten. Man kann in der Tat aber für einen beliebigen K-Vektorraum V die folgenden Aussagen zeigen:

- Jedes Erzeugendensystem von V enthält eine Basis.
- Jede linear unabhängige Familie in V läßt sich zu einer Basis ergänzen.
- Insbesondere besitzt V eine Basis.
- Jeder Unterraum U von V besitzt ein direktes Komplement.

Für die letzte Aussage wählt man eine Basis B von U ergänzt sie durch eine Familie B' zu einer Basis von V; dann erzeugt B' ein direktes Komplement.

Man beachte, daß Basen durchaus kompliziert sein können. Betrachten wir etwa $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum, so wird eine Basis von $\mathbb{R}$ notwendigerweise überabzählbar viele Elemente enthalten, da jede lineare Hülle von abzählbar vielen Elementen aus $\mathbb{R}$ mit Koeffizienten aus $\mathbb{Q}$ nur abzählbar viele Elemente enthalten kann.

C) Eindeutige Darstellbarkeit bezüglich einer Basis

Proposition 29.16 (Eindeutige Darstellbarkeit bezüglich einer Basis). Eine Familie B von Vektoren in V ist genau dann eine Basis von V, wenn jeder Vektor in V in eindeutiger Weise als Linearkombination von Elementen in B geschrieben werden kann.

Beweis: Sei zunächst $B = (x_i)_{i\in I}$ eine Basis von V und $x \in V$. Nach Voraussetzung ist B ein Erzeugendensystem von V und mithin ist

$$x = \sum_{\substack{i\in I \\ \text{endlich}}} \lambda_i \cdot x_i$$

eine Linearkombination von Vektoren in B. Ist nun

$$x = \sum_{\substack{i\in I \\ \text{endlich}}} \lambda'_i \cdot x_i$$

eine zweite Linearkombination von Vektoren in B, die x ergibt, so ist

$$0 = x - x = \sum_{\substack{i\in I \\ \text{endlich}}} \lambda_i \cdot x_i - \sum_{\substack{i\in I \\ \text{endlich}}} \lambda'_i \cdot x_i = \sum_{\substack{i\in I \\ \text{endlich}}} (\lambda_i - \lambda'_i) \cdot x_i$$

eine Linearkombination von Vektoren in B, die Null ergibt. Da B linear unabhängig ist, muß dann aber stets

$$\lambda_i - \lambda'_i = 0$$

gelten. Die Darstellung ist also eindeutig.

Sie nun umgekehrt jeder Vektor x in V auf eindeutige Weise als Linearkombination der Vektoren in B darstellbar. Dann ist offenbar B ein Erzeugendensystem von V, und 0_V kann nur auf die triviale Weise als Linearkombination von Vektoren in B dargestellt werden, so daß B auch linear unabhängig ist. □

Beispiel 29.17.
Kommen wir auf das Beispiel $B = \big((1,1)^t, (1,-1)^t\big)$ im $\mathbb{R}^2$ zurück. Wir haben bereits gesehen, dass sich ein beliebiger Vektor $(\lambda_1, \lambda_2)^t \in \mathbb{R}^2$ als

$$(\lambda_1, \lambda_2)^t = \frac{\lambda_1 + \lambda_2}{2} \cdot (1,1)^t + \frac{\lambda_1 - \lambda_2}{2} \cdot (1,-1)^t$$

schreiben läßt. Da B eine Basis ist, ist diese Darstellung auch eindeutig, wie man auch leicht direkt nachprüfen könnte.

Satz 29.18 (Existenz- und Eindeutigkeitssatz für lineare Abbildungen). Es seien V und W zwei K-Vektorräume, $B = (x_i)_{i\in I}$ eine Basis von V und $F = (y_i)_{i\in I}$ eine Familie von Vektoren in W.
Dann gibt es genau eine K-lineare Abbildung $f : V \longrightarrow W$, so daß für alle $i \in I$

$$f(x_i) = y_i.$$

Insbesondere, zwei lineare Abbildungen sind gleich, sobald sie auf einer Basis übereinstimmen.

Beweis: Jeder Vektor $x \in V$ läßt sich nach Proposition 29.16 in eindeutiger Weise als Linearkombination

$$x = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot x_i$$

schreiben. Wir definieren die Abbildung f dann durch

$$f(x) := \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot y_i. \tag{78}$$

Wir wollen nun zeigen, daß $f : V \longrightarrow W$ dann K-linear ist. Seien dazu

$$x = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot x_i, \quad x' = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda'_i \cdot x_i \in V$$

und $\lambda, \lambda' \in K$ gegeben, dann gilt für die eindeutige Darstellung von $\lambda x + \lambda' x'$ offenbar

$$\lambda x + \lambda' x' = \sum_{\substack{i\in I\\ \text{endlich}}} (\lambda \cdot \lambda_i + \lambda' \cdot \lambda'_i) \cdot x_i,$$

und mithin erhalten wir

$$\begin{aligned} f\big(\lambda x + \lambda' x'\big) &= f\left(\sum_{\substack{i\in I\\ \text{endlich}}} (\lambda \cdot \lambda_i + \lambda' \cdot \lambda'_i) \cdot x_i\right) \\ &= \sum_{\substack{i\in I\\ \text{endlich}}} (\lambda \cdot \lambda_i + \lambda' \cdot \lambda'_i) \cdot y_i \\ &= \lambda \cdot \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot y_i + \lambda' \cdot \sum_{\substack{i\in I\\ \text{endlich}}} \lambda'_i \cdot y_i \\ &= \lambda \cdot f(x) + \lambda' \cdot f(x'). \end{aligned}$$

Die Abbildung f ist also K-linear, und nach Definition gilt auch $f(x_i) = y_i$.

Es bleibt zu zeigen, daß es keine zweite K-lineare Abbildung geben kann, die diese Eigenschaft hat. Sei dazu $g : V \longrightarrow W$ eine K-lineare Abbildung mit $g(x_i) = y_i$ für alle

$i \in I$. Ein beliebiges $x \in V$ läßt sich wieder schreiben als

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot x_i$$

und dann gilt

$$f(x) = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot f(x_i) = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot y_i = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot g(x_i) = g(x).$$

Mithin stimmt f mit g überein. □

Bemerkung 29.19 (Existenz- und Eindeutigkeitssatz für lineare Abb.).
Satz 29.18 besagt, daß man die Werte einer linearen Abbildung auf einer Basis beliebig vorschreiben kann. Egal welche Vektoren im Zielbereich man als Bilder wählt, es gibt eine und nur eine lineare Abbildung, die den Basiselementen genau diese Vektoren zuordnet!

Wegen der Formel für $f(x)$ in (78) sagt man auch, daß sich f aus der Vorschrift $f(x_i) = y_i$, $i \in I$, durch *lineare Fortsetzung* ergibt.

Beispiel 29.20.
Setzen wir $x_1 = (1,1)^t$ und $x_2 = (1,-1)^t$, so ist $B = (x_1, x_2)$ eine Basis von $\mathbb{R}^2$. Wählen wir nun zudem $y_1 = (1,1)^t$ und $y_2 = (3,1)^t$, so muß es genau eine $\mathbb{R}$-lineare Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ geben mit

$$f\big((1,1)^t\big) = f(x_1) = y_1 = (1,1)^t \quad \text{und} \quad f\big((1,-1)^t\big) = f(x_2) = y_2 = (3,1)^t.$$

Diese besitzt die Abbildungsvorschrift

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x,y)^t \mapsto (2x - y, x)^t.$$

Wir werden später sehen, wie man die Abbildungsvorschrift systematisch bestimmen kann.

Korollar 29.21 (Alle linearen Abb. $K^n \to K^m$ sind von der Form f_A.).
Jede lineare Abbildung $f : K^n \longrightarrow K^m$ ist von der Form $f = f_A$ für eine eindeutig bestimmte Matrix $A \in \mathrm{Mat}(m \times n, K)$.

Beweis: Ist $f : K^n \longrightarrow K^m$ eine lineare Abbildung, so setzen wir $a^i := f(e_i) \in K^m$ für $i = 1, \ldots, n$ und bilden eine Matrix A mit den a^i als Spaltenvektoren. Dann ist f_A eine lineare Abbildung, mit

$$f_A(e_i) = Ae_i = a^i = f(e_i),$$

so daß aus der Eindeutigkeitsaussage in 29.18 unmittelbar $f_A = f$ folgt. Die Eindeutigkeit der Matrix A folgt aus der Tatsache, daß A die Abbildung f_A eindeutig festlegt (siehe Bemerkung 27.7). □

Aufgaben

Aufgabe 29.22.
Welche der folgenden Familien sind linear unabhängig / Erzeugendensysteme / Basen von $\mathbb{R}^2$?

a. $\big((1,0)^t,(0,1)^t,(1,1)^t\big)$.

b. $\big((1,1)^t,(2,2)^t\big)$.

c. $\big((1,3)^t\big)$.

d. $\big((1,1)^t,(1,-2)^t\big)$.

e. $\big((1,1)^t,(0,0)^t\big)$.

f. $\big((1,1)^t,(0,0)^t,(1,-2)^t\big)$.

g. $\big((1,2)^t,(2,1)^t\big)$.

Aufgabe 29.23.
Es sei V ein K-Vektorraum, $U \subset V$ ein Unterraum, $0 \neq x \in U$ und $y \in V \setminus U$. Zeige, daß (x,y) linear unabhängig ist.

Aufgabe 29.24.
Ist $f : V \longrightarrow W$ eine K-lineare Abbildung, F eine Familie von Vektoren in V, so ist

$$f\big(\operatorname{Lin}(F)\big) = \operatorname{Lin}\big(f(x) \mid x \in F\big).$$

Aufgabe 29.25.
Es seien V und W zwei K-Vektorräume, $f : V \longrightarrow W$ eine K-lineare Abbildung und B eine Basis von V.

a. Genau dann ist f surjektiv, wenn $f(B)$ ein Erzeugendensystem von W ist.

b. Genau dann ist f injektiv, wenn $f(B)$ linear unabhängig ist.

c. Genau dann ist f bijektiv, wenn $f(B)$ eine Basis von W ist.

Aufgabe 29.26.
Es seien $U_1,\ldots,U_k$ Unterräume eines K-Vektorraums V mit Basen $B_1,\ldots,B_k$. Zeige, genau dann ist $V = U_1 \oplus \ldots \oplus U_k$ die direkte Summe der U_i, wenn $B = B_1 \cup \ldots \cup B_k$ eine Basis von V ist.

Aufgabe 29.27.
Sei V ein $\mathbb{C}$–Vektorraum, dann ist V offensichtlich auch ein $\mathbb{R}$–Vektorraum, und seien $x_1, \ldots, x_n \in V$. Zeige, daß $(x_1, \ldots, x_n)$ genau dann linear unabhängig über $\mathbb{C}$ ist, wenn $(x_1, ix_1, \ldots, x_n, ix_n)$ linear unabhängig über $\mathbb{R}$ ist.

Aufgabe 29.28.
Es sei $(V, +, \cdot)$ ein K-Vektorraum, und $x_1, \ldots, x_n \in V$ seien linear abhängige Vektoren mit der Eigenschaft, daß je $n-1$ der Vektoren linear unabhängig sind. Zeige:

a. Es gibt $\lambda_1, \ldots, \lambda_n \in K \setminus \{0\}$ mit der Eigenschaft
$$\sum_{i=1}^{n} \lambda_i x_i = 0.$$

b. Gilt für $\mu_1, \ldots, \mu_n \in K$ ebenfalls $\sum_{i=1}^{n} \mu_i x_i = 0$, so gibt es ein $\nu \in K$ mit $\mu_i = \lambda_i \cdot \nu$ für alle $i = 1, \ldots, n$.

Aufgabe 29.29.
Es sei $A \in \mathrm{Mat}_n(K)$.

a. Genau dann ist f_A bijektiv, wenn $A \in \mathrm{Gl}_n(K)$.

b. Ist $A \in \mathrm{Gl}_n(K)$, so gilt $(f_A)^{-1} = f_{A^{-1}}$.

§ 30 Endlich-dimensionale Vektorräume

Wir betrachten jetzt *endlich erzeugte Vektorräume* V, d. h. Vektorräume, die ein endliches Erzeugendensystem besitzen. Nach Bemerikung 29.15 besitzt V dann auch eine endliche Basis. Für solche Vektorräume kann man die Sätze des vorigen Abschnitts teilweise verschärfen und vor allem kann man in diesen Vektorräumen mit Hilfe von Basen und Matrizen effizient rechnen.

In diesem Abschnitt ist V stets ein endlich erzeugter K-Vektorraum.

A) Austauschsatz von Steinitz

Lemma 30.1 (Austauschlemma).
Sei $B = (x_1, \ldots, x_n)$ eine Basis von V, $y = \sum_{i=1}^{n} \lambda_i x_i$ und $\lambda_j \neq 0$ für ein j.
Dann ist $(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n)$ eine Basis von V, d.h. man kann in der Basis B den Vektor x_j gegen y austauschen.

Beweis: Wegen $\lambda_j \neq 0$ gilt

$$x_j = \frac{1}{\lambda_j} \cdot y - \sum_{i \neq j} \frac{\lambda_i}{\lambda_j} \cdot x_i,$$

und somit

$$V = \operatorname{Lin}(x_1, \ldots, x_n) = \operatorname{Lin}(y, x_1, \ldots, x_n) = \operatorname{Lin}(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n).$$

Bleibt also zu zeigen, daß $(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n)$ linear unabhängig ist. Seien dazu $\mu_i \in K$, $i = 1, \ldots, n$, gegeben mit

$$0 = \mu_j y + \sum_{i \neq j} \mu_i x_i = \sum_{i=1}^{n} \mu_j \lambda_i x_i + \sum_{i \neq j} \mu_i x_i$$

$$= \mu_j \lambda_j x_j + \sum_{i \neq j} (\mu_j \lambda_i + \mu_i) x_i.$$

Dann folgt aus der linearen Unabhängigkeit von $x_1, \ldots, x_n$

$$\mu_j \lambda_j = 0 \quad \text{und} \quad \mu_i = -\mu_j \lambda_i, \text{ für } i \neq j.$$

Wegen $\lambda_j \neq 0$, ist also $\mu_j = 0$ und damit auch

$$\mu_i = 0 \quad \text{für } i \neq j.$$

Damit ist die lineare Unabhängigkeit von $(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n)$ gezeigt. □

Beispiel 30.2.
Ist zum Beispiel $E = (e_1, \ldots, e_n)$ die kanonische Basis des K^n und $x = (\lambda_1, \ldots, \lambda_n)^t \in K^n$ mit $\lambda_j \neq 0$, so können wir e_j gegen x austauschen und erhalten wieder eine Basis.

Konkret kann man in der Basis $E = (e_1, e_2, e_3)$ von $\mathbb{R}^3$ den Vektor $(1, 2, 0)^t$ gegen e_1 oder e_2 austauschen, nicht aber gegen e_3.

Das Austauschlemma wird benutzt, um den wichtigen Steinitzschen Austauschsatz zu beweisen.

Satz 30.3 (Austauschsatz von Steinitz).
Sei $(x_1, \ldots, x_n)$ eine Basis von V und $(y_1, \ldots, y_r)$ sei linear unabhängig in V. Dann lassen sich die $x_1, \ldots, x_n$ so umnumerieren, daß $(y_1, \ldots, y_r, x_{r+1}, \ldots, x_n)$ eine Basis von V ist. Insbesondere gilt $r \leq n$.

Beweis von Satz 30.3: Wir führen den Beweis mittels Induktion über r.

Für $r = 0$ ist die Behauptung offensichtlich richtig. Nehmen wir also an, daß $r > 0$ und daß die Behauptung bereits richtig ist für $r - 1$. D. h. nach evt. Umnumerieren ist $(y_1, \ldots, y_{r-1}, x_r, \ldots, x_n)$ eine Basis von V. Dann besitzt y_r eine Darstellung der Form

$$y_r = \lambda_1 y_1 + \ldots + \lambda_{r-1} y_{r-1} + \lambda_r x_r + \ldots + \lambda_n x_n,$$

mit $\lambda_i \in K$. Angenommen, $\lambda_r = \ldots = \lambda_n = 0$, dann wäre $(y_1, \ldots, y_r)$ linear abhängig, im Widerspruch zur Voraussetzung. Also gibt es ein $j \in \{r, \ldots, n\}$ mit $\lambda_j \neq 0$. Durch Umnumerieren können wir annehmen, daß $j = r$ gilt. Dann können wir aber nach dem Austauschlemma 30.1 y_r gegen x_r austauschen, und die Behauptung ist bewiesen. □

Bemerkung 30.4.

a. Der Austauschsatz von Steinitz besagt also, daß man - nach eventuellem Umnumerieren - die linear unabhängigen Vektoren $x_1, \ldots, x_r$ durch $y_1, \ldots, y_r$ ersetzen kann.

b. Im Austauschsatz tauschen wir nacheinander x_{i_1} durch y_1, x_{i_2} durch y_2, etc. und schließlich x_{i_r} durch y_r für geeignete $i_1, \ldots, i_r$ aus. Im j-ten Schritt wissen wir, daß wir eine Darstellung

$$y_j = \sum_{l=1}^{j-1} \lambda_{i_l} y_l + \sum_{l \notin \{i_1, \ldots, i_{j-1}\}} \lambda_l x_l$$

haben mit $\lambda_l \neq 0$ für ein $l \notin \{i_1, \ldots, i_{j-1}\}$, und setzen wir dann $i_j := l$, so können wir x_{i_j} durch y_j ersetzen.
Wie wir eine solche Darstellung von y_j mit Hilfe des Gauß'schen Algorithmus berechnen können, werden wir später sehen. Damit haben wir dann ein konstruktives Verfahren für die Anwendung des Steinitzschen Austauschsatzes.

Aus dem Satz von Steinitz folgen unmittelbar die folgenden beiden Korollare.

Korollar 30.5 (Basisergänzungssatz).
Ist $(y_1, \dots, y_r)$ eine linear unabhängige Familie im endlich erzeugten Vektorraum V, so kann diese zu einer Basis B von V ergänzt werden.

Beweis: Nach Voraussetzung besitzt V ein endliches Erzeugendensystem und dieses kann zu einer endlichen Basis $(x_1, \dots, x_n)$ von V ausgedünnt werden. Dann wenden wir auf $(y_1, \dots, y_r)$ und $(x_1, \dots, x_n)$ den Satz von Steinitz an. □

Korollar 30.6 (Existenz eines Komplementes).
Ist V ein endlich erzeugter Vektorraum und U ein Unterraum von V, dann besitzt U ein direktes Komplement.

Beweis: Wir wählen eine Basis B von U und ergänzen sie durch eine Familie B' zu einer Basis von V, dann erzeugt B' ein direktes Komplement von U in V. □

B) Die Dimension eines endlich-erzeugten Vektorraums

Als Folgerung des Steinitzschen Austauschsatzes erhalten wir den folgenden Satz.

Satz 30.7 (Alle Basen sind gleichmächtig.).

a. Ist V endlich erzeugt, so ist jede Basis von V endlich und alle Basen haben gleich viele Elemente.

b. Ist V nicht endlich erzeugt, so hat jede Basis unendlich viele Elemente.

Beweis: a. Nach Voraussetzung besitzt V ein endliches Erzeugendensystem E und nach Bemerkung 29.15 folgt dann auch, daß V eine endliche Basis $B = (x_1, \dots, x_n)$ besitzt. Dabei können wir o. E. annehmen, daß n die minimale Mächtigkeit einer Basis ist. Sei nun B' eine weitere Basis von V. Angenommen, $|B'| > n$. Dann gibt es eine linear unabhängige Teilfamilie $(y_1, \dots, y_{n+1})$ in B', im Widerspruch zum Austauschsatz von Steinitz, der verlangt $n + 1 \leq n$.

b. Dies ist offensichtlich, da jede Basis V erzeugt.

□

Satz 30.7 rechtfertigt die folgende Definition.

Definition 30.8 (Dimension eines Vektorraums).
Für einen (nicht notwendig endlich erzeugten) K-Vektorraum V definieren wir die *Dimension* von V durch

$$\dim_K(V) := \begin{cases} n, & \text{falls } V \text{ eine Basis mit } n < \infty \text{ Elementen besitzt,} \\ \infty, & \text{falls } V \text{ nicht endlich erzeugt ist.} \end{cases}$$

Ist $\dim_K(V) < \infty$, so nennen wir V einen *endlich-dimensionalen* K-Vektorraum.

Aus Satz 30.7 und Definition 30.8 folgt unmittelbar das folgende Korollar.

Korollar 30.9.
Sei $\dim_K(V) = n$, E ein Erzeugendensystem von V und F linear unabhängig in V. Dann gelten

$$|E| \geq n \quad \text{und} \quad |F| \leq n.$$

Zudem gilt die Gleichheit genau dann, wenn die jeweilige Familie eine Basis ist.

Beweis: Nach Bemerkung 29.15 ist F in einer Basis von V enthalten und E enthält eine Basis von V. Die Ungleichungen folgen dann aus Satz 30.7, und derselbe Satz liefert Gleichheit, wenn die Familien Basen sind. Gilt umgekehrt die Gleichheit, so muß E bzw. F ein minimales Erzeugendensystem bzw. eine maximal linear unabhängige Familie sein und somit nach Proposition 29.12 eine Basis. □

Beispiel 30.10.

a. Es gilt:
$$\dim_K(V) = 0 \Leftrightarrow V = \mathrm{Lin}(\emptyset) \Leftrightarrow V = \{0\}.$$

b. $\dim_K\left(K^n\right) = n$, da die kanonische Basis $E = (e_1, \ldots, e_n)$ genau n Elemente enthält.

c. $\dim_{\mathbb{Q}}(\mathbb{Q}) = \dim_{\mathbb{R}}(\mathbb{R}) = \dim_{\mathbb{C}}(\mathbb{C}) = 1$, aber $\dim_{\mathbb{Q}}(\mathbb{R}) = \infty$ und $\dim_{\mathbb{R}}(\mathbb{C}) = 2$. Für letzteres beachte man, daß die Familie $(1, i)$ eine $\mathbb{R}$-Basis von $\mathbb{C}$ ist.

d. Die Dimension des Vektorraums P_n der Polynome vom Grad höchstens n ist $\dim_K(P_n) = n + 1$, da $B = (t^0, t^1, \ldots, t^n)$ eine Basis ist.

Satz 30.11 (Karten eines Vektorraums).
Es sei $B = (x_1, \ldots, x_n)$ eine Basis von V und $E = (e_1, \ldots, e_n)$ die kanonische Basis des K^n. Dann bestimmt B einen Isomorphismus

$$\phi_B : V \to K^n : x_i \mapsto e_i, \text{ für } i = 1, \ldots, n,$$

durch *lineare Fortsetzung*. Man nennt ϕ_B die *Karte* von V zur Basis B.

Beweis: Nach Satz 29.18 bestimmen die Zuordnungen

$$x_i \mapsto e_i,\ i = 1, \ldots, n, \quad \text{und} \quad e_i \mapsto x_i,\ i = 1, \ldots, n,$$

zwei lineare Abbildungen $\phi_B : V \to K^n$ und $\phi^B : K^n \to V$. Es bleibt zu zeigen, daß

$$\phi_B \circ \phi^B = \mathrm{id}_{K^n} \quad \text{und} \quad \phi^B \circ \phi_B = \mathrm{id}_V\,.$$

Dazu reicht es wegen Satz 29.18, daß die beiden Seiten jeweils auf einer Basis übereinstimmen, und das tun sie offenbar. □

Daraus ergibt sich unmittelbar folgendes Korollar.

Korollar 30.12 (Dimension als einzige Invariante eines Vektorraums).
Für zwei endlich-dimensionale K-Vektorräume V und W sind gleichwertig:

a. $V \cong W$.

b. $\dim_K(V) = \dim_K(W)$.

Insbesondere ist jeder n-dimensionale K-Vektorraum isomorph zum K^n.

Beweis: Es seien $n = \dim_K(V)$ und $m = \dim_K(W)$.

Ist $f : V \to W$ ein Isomorphismus, so überführt er laut Aufgabe 29.25 eine Basis von V, die n Elemente enthält, in eine Basis von W, die m Elemente enthält. Mithin gilt $n = m$.

Ist umgekehrt $n = m$, dann sind nach Satz 30.11 beide Vektorräume isomorph zu K^n und mithin zueinander. □

Beispiel 30.13.
Die Abbildungen $\sin, \cos \in \mathbb{R}^{\mathbb{R}}$ sind linear unabhängig, da aus

$$\lambda \cdot \sin + \mu \cdot \cos = 0$$

insbesondere

$$0 = \lambda \cdot \sin(0) + \mu \cdot \cos(0) = \mu$$

und

$$0 = \lambda \cdot \sin\left(\frac{\pi}{2}\right) + \mu \cdot \cos\left(\frac{\pi}{2}\right) = \lambda$$

folgt. Aber dann hat der Vektorraum $\mathrm{Lin}\,(\sin, \cos)$ die Dimension zwei, da $(\sin, \cos)$ eine Basis ist, und deshalb gilt

$$\mathrm{Lin}\,(\sin, \cos) \cong \mathbb{R}^2.$$

C) Dimensionsformeln

Lemma 30.14.
Ist $\dim_K(V) < \infty$ und $U \subseteq V$ ein Unterraum. Dann gelten

$$\dim_K(U) \leq \dim_K(V),$$

und

$$U = V \iff \dim_K(U) = \dim_K(V).$$

Beweis: Ist U ein Unterraum, so kann eine Basis B von U zu einer Basis B' von V ergänzt werden, so daß $\dim_K(U) = |B| \leq |B'| = \dim_K(V)$ gelten muß.

Ist $U = V$, so ist offenbar auch $\dim_K(U) = \dim_K(V)$. Gilt umgekehrt $\dim_K(U) = \dim_K(V)$ und ist $B = (x_1, \ldots, x_n)$ eine Basis von U, so können wir sie nach dem Basisergänzungssatz 30.5 zu einer Basis B' von V ergänzen. Wegen $B \subseteq B'$ und $|B| = n = |B'|$ folgt dann aber notwendigerweise $B = B'$, und somit $U = \mathrm{Lin}\,(B) = \mathrm{Lin}\,(B') = V$. □

Satz 30.15 (Dimensionsformel für Unterräume).
Ist $\dim_K(V) < \infty$ und sind U und U' Unterräume von V, dann gilt:

$$\dim_K\left(U + U'\right) = \dim_K(U) + \dim_K\left(U'\right) - \dim_K\left(U \cap U'\right).$$

Beweis: Wir beweisen mehr, nämlich wie wir geeignete Basen von U, U' und $U \cap U'$ wählen können. Sei $B_{U\cap U'} := (x_1, \ldots, x_r)$ eine Basis von $U \cap U'$. Wir ergänzen $B_{U\cap U'}$ zu einer Basis $B_U := (x_1, \ldots, x_r, y_1, \ldots, y_s)$ von U, und zu einer Basis $B_{U'} := (x_1, \ldots, x_r, z_1, \ldots, z_t)$ von U'. Das geht nach dem Basisergänzungssatz 30.5.

Behauptung: $B_{U+U'} := (x_1, \ldots, x_r, y_1, \ldots, y_s, z_1, \ldots, z_t)$ ist Basis von $U + U'$.

Dazu zeigen wir zunächst, daß jedes Element von $U + U'$ eine Linearkombination von Elementen aus $B_{U+U'}$ ist. Sei also $x + x' \in U + U'$ mit $x \in U$ und $x' \in U'$. Dann gilt:

$$x = \sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j y_j \quad \text{und} \quad x' = \sum_{i=1}^{r} \lambda'_i x_i + \sum_{k=1}^{t} \mu'_k z_k,$$

mit $\lambda_i, \lambda'_i, \mu_j, \mu'_k \in K$, $i = 1, \ldots, r$, $j = 1, \ldots, s$, $k = 1, \ldots, t$. Daraus folgt:

$$x + x' = \sum_{i=1}^{r} \left(\lambda_i + \lambda'_i\right) x_i + \sum_{j=1}^{s} \mu_j y_j + \sum_{k=1}^{t} \mu'_k z_k \in \mathrm{Lin}\,(B_{U+U'})\,.$$

Dann müssen wir noch zeigen, daß $B_{U+U'}$ linear unabhängig ist. Sei dazu

$$\sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j y_j + \sum_{k=1}^{t} \nu_k z_k = 0 \tag{79}$$

eine Linearkombination der Null. Dann ist

$$\underbrace{\sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j y_j}_{\in U} = \underbrace{\sum_{k=1}^{t} -\nu_k z_k}_{\in U'} \in U \cap U'.$$

Da $B_{U\cap U'}$ eine Basis von $U \cap U'$ ist, gibt es also λ_i', so daß

$$\sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j y_j = \sum_{i=1}^{r} \lambda_i' x_i = \sum_{i=1}^{r} \lambda_i' x_i + \sum_{j=1}^{s} 0 \cdot y_j$$

gilt. Da B_U linear unabhängig ist, ergibt ein Koeffizientenvergleich auf beiden Seiten insbesondere $\mu_j = 0$ für alle $j = 1, \ldots, s$. Damit erhalten wir aus (79) dann

$$\sum_{i=1}^{r} \lambda_i x_i + \sum_{k=1}^{t} \nu_k z_k = 0,$$

und da $B_{U'}$ linear unabhängig ist, müssen notwendigerweise alle λ_i und ν_k Null sein. Damit haben wir dann auch gezeigt, daß $B_{U+U'}$ linear unabhängig ist.

Aus der Behauptung folgt,

$$\dim_K \left(U+U'\right) = r+s+t = (r+s)+(r+t)-r = \dim_K(U)+\dim_K\left(U'\right)-\dim_K\left(U\cap U'\right).$$

□

Beispiel 30.16.
Für die Unterräume $U = \mathrm{Lin}\left((1,0,0)^t, (1,1,1)^t\right)$ und $U' = \mathrm{Lin}\left((1,1,1)^t, (0,0,1)^t\right)$ von $\mathbb{R}^3$ sieht man leicht, daß

$$U \cap U' = \mathrm{Lin}\left((1,1,1)^t\right)$$

ein Vektorraum von Dimension eins ist, während U und U' jeweils Dimension zwei haben, da die angegebenen Erzeugendensysteme auch linear unabhängig sind. Mithin erhalten wir

$$\dim_{\mathbb{R}}(U + U') = 2 + 2 - 1 = 3 = \dim_{\mathbb{R}}(\mathbb{R}^3),$$

so daß

$$U + U' = \mathbb{R}^3$$

gelten muß. Da zudem

$$U + U' = \mathrm{Lin}\left((1,0,0)^t, (1,1,1)^t, (0,0,1)^t\right)$$

gilt, sehen wir, daß dieses Erzeugendensystem eine Basis des $\mathbb{R}^3$ ist.

Korollar 30.17 (Dimensionsformel für Komplemente).
Ist $\dim_K(V) < \infty$, dann sind für Unterräume U und U' von V die folgenden Aussagen äquivalent:

a. $V = U \oplus U'$.

b. $V = U + U'$ und $U \cap U' = \{0\}$.

c. $V = U + U'$ und $\dim_K(V) = \dim_K(U) + \dim_K\left(U'\right)$.

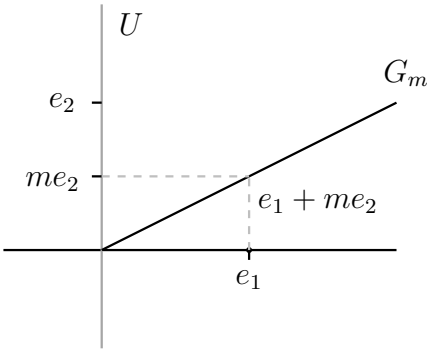
d. $U \cap U' = \{0\}$ und $\dim_K(V) = \dim_K(U) + \dim_K\left(U'\right)$.

Beweis: Dies ist eine direkte Folgerung aus Lemma 28.17, Lemma 30.14 und Satz 30.15. □

Beispiel 30.18.
Wir erinnern uns an Beispiel 28.35. Dort haben wir in $V = \mathbb{R}^2$ den Unterraum $U = \mathrm{Lin}\,(e_2)$, die y-Achse, betrachtet und gezeigt, daß jede Ursprungsgerade mit Steigung m

$$G_m := \mathrm{Lin}\,(e_1 + me_2)$$

ein Komplement von U ist. Dies können wir nun mit weniger Aufwand begründen, denn die beiden Geraden schneiden sich offenbar nur im Ursprung und ihre Dimensionen addieren sich zu zwei.



Korollar 30.19 (Dimensionsformel für Faktorräume).
Es sei V ein endlich-dimensionaler K-Vektorraum und U ein Unterraum von V. Dann gilt

$$\dim_K(V/U) = \dim_K(V) - \dim_K(U).$$

Beweis: Nach Bemerkung 29.15 besitzt U ein Komplement U', und nach Proposition 28.36 gilt $U' \cong V/U$. Aus Korollar 30.12 und Korollar 30.17 folgt dann

$$\dim_K(V) = \dim_K(U) + \dim_K(U') = \dim_K(U) + \dim_K(V/U).$$

□

Bemerkung 30.20 (Basis von V/U).
Es sei V ein K-Vektorraum und U ein Unterraum von V mit Basis $(x_1, \ldots, x_r)$. Dann sind die folgenden Aussagen für $y_1, \ldots, y_s \in V$ gleichwertig:

a. $(x_1, \ldots, x_r, y_1, \ldots, y_s)$ ist eine Basis von V.

b. $(y_1, \ldots, y_s)$ ist Basis eines Komplementes von U.

c. $\left(\overline{y_1}, \ldots, \overline{y_s}\right)$ ist eine Basis von V/U.

Beweis: Die Äquivalenz von b. und c. folgt aus Proposition 28.36 und Aufgabe 29.25. Die Äquivalenz von a. und b. folgt aus dem Beweis von Korollar 30.6 und Korollar 30.17. □

Satz 30.21 (Dimensionsformel für lineare Abbildungen).
Es sei $f : V \longrightarrow W$ eine K-lineare Abbildung und $\dim_K(V) < \infty$. Dann gilt

$$\dim_K(V) = \dim_K\big(\operatorname{Ker}(f)\big) + \dim_K\big(\operatorname{Im}(f)\big).$$

Beweis: Aus dem Homomorphiesatz 28.33 erhalten wir den Isomorphismus

$$V/\operatorname{Ker}(f) \cong \operatorname{Im}(f),$$

so daß die Formel dann aus Korollar 30.19 folgt. □

Beispiel 30.22.
Die lineare Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2)^t \mapsto x_1 - x_2$ hat den Kern $\operatorname{Lin}\left((1,1)^t\right)$ von Dimension eins und ist surjektiv. Wir erhalten also die Formel

$$\dim_{\mathbb{R}}(\mathbb{R}^2) = 2 = 1 + 1 = \dim_{\mathbb{R}}(\operatorname{Ker}(f)\big) + \dim_{\mathbb{R}}\big(\operatorname{Im}(f)\big).$$

D) Bijektivität linearer Abbildungen

Korollar 30.23 (Injektiv = surjektiv = bijektiv).
Es seien V und W zwei endlich-dimensionale K-Vektorräume gleicher Dimension und $f : V \longrightarrow W$ sei K-linear. Dann sind gleichwertig:

a. f ist bijektiv,

b. f ist injektiv,

c. f ist surjektiv.

Beweis: Ist f injektiv, so ist $\operatorname{Ker}(f) = \{0\}$, und wir erhalten aus der Dimensionsformel für lineare Abbildungen 30.21

$$\dim_K(W) = \dim_K(V) = \dim_K\big(\operatorname{Ker}(f)\big) + \dim_K\big(\operatorname{Im}(f)\big) = \dim_K\big(\operatorname{Im}(f)\big).$$

Wegen Lemma 30.14 gilt dann $W = \operatorname{Im}(f)$ und f ist surjektiv.

Ist f surjektiv, so ist $W = \mathrm{Im}(f)$ und wir erhalten aus der Dimensionsformel für lineare Abbildungen 30.21

$$\dim_K \big(\mathrm{Ker}(f)\big) = \dim_K(V) - \dim_K \big(\mathrm{Im}(f)\big) = \dim_K(V) - \dim_K(W) = 0.$$

Dann ist aber $\mathrm{Ker}(f) = \{0\}$ und somit ist f injektiv. □

Korollar 30.24 (Invertierbare Matrizen).
Sind $A, B \in \mathrm{Mat}_n(K)$ mit $AB = \mathbb{1}_n$, so gilt auch $BA = \mathbb{1}_n$ und $A \in \mathrm{Gl}_n(K)$.

Beweis: Aus $AB = \mathbb{1}_n$ folgt

$$f_A \circ f_B = f_{AB} = f_{\mathbb{1}_n} = \mathrm{id}_{K^n},$$

so daß f_B injektiv mit Linksinverser f_A ist. Nach Korollar 30.23 ist f_B dann aber schon bijektiv, und die Linksinverse ist die Inverse von f_B. Damit folgt dann auch

$$f_{\mathbb{1}_n} = \mathrm{id}_{K^n} = f_B \circ f_A = f_{BA},$$

und damit $BA = \mathbb{1}_n$. □

Aufgaben

Aufgabe 30.25.
Es sei V ein K-Vektorraum mit $\dim_K(V) = 5$, und U und U' Unterräume mit $\dim_K(U) = 3$ und $\dim_K(U') = 4$.

a. Welche Werte kann $\dim_K(U \cap U')$ annehmen?

b. Gib für jeden der Werte von $\dim_K(U \cap U')$ ein Beispiel (K, V, U, U') an.

Aufgabe 30.26.
Finde einen K-Vektorraum V sowie zwei K-lineare Abbildungen $f, g : V \longrightarrow V$, so daß folgendes gilt:

a. f ist injektiv, aber nicht surjektiv.

b. g ist surjektiv, aber nicht injektiv.

Aufgabe 30.27.
Es sei V ein K-Vektorraum mit $1 \leq \dim_K(V) = n < \infty$ und $g \in \mathrm{End}_K(V)$. Zeige, es gibt eine Zahl $0 \leq k \leq n$, so daß für alle $i \geq 1$ gilt:

$$\mathrm{Ker}(g^0) \subsetneqq \mathrm{Ker}(g^1) \subsetneqq \ldots \subsetneqq \mathrm{Ker}(g^k) = \mathrm{Ker}(g^{k+i}).$$

Aufgabe 30.28.
Es sei $B := \big((3,5,2)^t, (1,1,-1)^t, (2,4,1)^t\big)$.

a. Zeige, B ist eine Basis von $\mathbb{R}^3$.

b. Ersetze *mit Hilfe des Austauschsatzes von Steinitz* zwei Vektoren in B durch die Vektoren $(1,3,2)^t$ und $(-2,1,2)^t$.

Aufgabe 30.29.
Sei V ein K-Vektorraum und $F = (v_1, \ldots, v_5)$ eine linear unabhängige Familie in V. Welchen der Vektoren $v_1, \ldots, v_5$ kann man durch $v := v_2 - v_3 + v_4 - v_5$ ersetzen, so dass die daraus resultierende Familie wieder linear unabhängig ist? Begründe Deine Aussage.

Aufgabe 30.30.
Sei K ein Körper.

a. Begründe, weshalb die Mengen $U := \{(a_1, \ldots, a_n)^t \in K^n \mid a_1 = \ldots = a_n\}$ und $U' := \big\{(a_1, \ldots, a_n)^t \in K^n \;\big|\; a_1 + \ldots + a_n = 0\big\}$ Unterräume des K^n sind.

b. Bestimme $\dim_K(U)$, $\dim_K(U')$, $\dim_K(U \cap U')$ und $\dim_K(U + U')$.

§ 31 Lineare Abbildungen und Matrizen

In diesem Abschnitt sind V und W zwei endlich-dimensionale K-Vektorräume mit Basen $B = (b_1, \ldots, b_n)$ und $D = (d_1, \ldots, d_m)$, sofern nichts anderes gesagt wird.
Ferner bezeichnen wir mit $E = (e_1, \ldots, e_n)$ die kanonische Basis von K^n und mit $F = (f_1, \ldots, f_m)$ die kanonische Basis von K^m.

A) Matrixdarstellung linearer Abbildungen

Definition 31.1 (Matrixdarstellung einer linearen Abbildung).
Gegeben seien eine Basis $B = (b_1, \ldots, b_n)$ von V und eine Basis $D = (d_1, \ldots, d_m)$ von W sowie eine lineare Abbildung $f : V \longrightarrow W$.

a. Ist $x \in V$, so läßt sich x nach Proposition 29.16 auf eindeutige Weise darstellen als Linearkombination der Basis B

$$x = \lambda_1 \cdot b_1 + \ldots \lambda_n \cdot b_n.$$

Wir nennen den Vektor $M_B(x) = (\lambda_1, \ldots, \lambda_n)^t$ den *Koordinatenvektor* oder die *Koordinaten* von x bezüglich B.

b. Für jeden Basisvektor b_j in B läßt sich das Bild $f(b_j) \in W$ unter f nach Proposition 29.16 auf eindeutige Weise als Linearkombination der Basis D darstellen

$$f(b_j) = a_{1j} \cdot d_1 + \ldots + a_{mj} \cdot d_m.$$

Schreiben wir die Koeffizienten als Spalten in eine Matrix, so erhalten wir die Matrix

$$M_D^B(f) := \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \in \mathrm{Mat}(m \times n, K),$$

die sogenannte *Matrixdarstellung* von f bezüglich der Basen B und D.

Beispiel 31.2.

a. Der Koordinatenvektor eines Vektors $x = (x_1, \ldots, x_n)^t \in K^n$ bezüglich der kanonischen Basis $E = (e_1, \ldots, e_n)$ ist der Vektor $M_E(x) = x$ selbst.

b. Um den Koordinatenvektor von $x = (4, 0)^t \in \mathbb{R}^2$ bezüglich der Basis $B = \big((1, 1)^t, (1, -1)^t\big)$ zu bestimmen, muß man ihn als Linearkombination bezüglich

der Basis darstellen. Man sieht leicht, daß

$$x = 2 \cdot (1,1)^t + 2 \cdot (1,-1)^t$$

gilt. Damit erhalten wir

$$M_B(x) = (2,2)^t.$$

c. Sei $V = \mathbb{R}^2$ mit Basis $B = (b_1, b_2) = \big((1,2)^t, (1,1)^t\big)$ und $W = \mathbb{R}^3$ mit Basis $D = (d_1, d_2, d_3) = \big((1,1,0)^t, (0,1,1)^t, (0,0,1)^t\big)$, und sei $f : V \to W$ die lineare Abbildung, die definiert wird durch

$$\begin{aligned} b_1 &\mapsto 3d_1 - 4d_2 + 6d_3, \\ b_2 &\mapsto 3d_1 - 3d_2 + 4d_3. \end{aligned}$$

Dann gilt:

$$M_D^B(f) = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix}.$$

d. Ist $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$ eine Matrix und $f_A : K^n \longrightarrow K^m$ die zugehörige Abbildung, dann ist

$$f_A(e_j) = A \cdot e_j = j\text{-te Spalte von } A = a_{1j} \cdot f_1 + \ldots + a_{mj} \cdot f_m.$$

Die Matrixdarstellung f_A bezüglich der kanonischen Basen ist also

$$M_F^E(f_A) = A.$$

Bemerkung 31.3.

a. Mit der Notation aus Satz 30.11 gilt

$$M_B(x) = \phi_B(x),$$

d. h. der Koordinatenvektor von x unter B ist das Bild unter der Karte ϕ_B. Die Zuordnung

$$\phi_B : V \longrightarrow K^n : x \mapsto M_B(x)$$

ist ein Isomorphismus von Vektorräumen, der die Vektoren in einem festgelegten Koordinatensystem darstellt.

b. Nach Definition gilt zudem:

Die j-te Spalte von $M_D^B(f)$ ist der Koordinatenvektor von $f(b_j)$ bez. D.

Proposition 31.4 (Rechnen in Koordinaten).
Ist $f : V \longrightarrow W$ eine lineare Abbildung und $x \in V$, so gilt

$$M_D\big(f(x)\big) = M_D^B(f) \circ M_B(x),$$

d. h. der Koordinatenvektor $M_D\big(f(x)\big)$ von $f(x)$ bezüglich der Basis D ist das Matrixprodukt der Matrixdarstellung $M_D^B(f)$ von f bezüglich B und D mit dem Koordinatenvektor $M_B(x)$ von x bezüglich B.

Beweis: Für einen Vektor $x = \sum_{j=1}^{n} \lambda_j b_j$ und eine lineare Abbildung mit Matrixdarstellung $M_D^B(f) = (a_{ij})$ gilt

$$f(x) = \sum_{j=1}^{n} \lambda_j \cdot f(b_j) = \sum_{j=1}^{n} \lambda_j \cdot \sum_{i=1}^{m} a_{ij} \cdot d_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \cdot \lambda_j \right) \cdot d_i.$$

Daraus folgt dann

$$M_D\big(f(x)\big) = (a_{ij}) \circ (\lambda_1, \ldots, \lambda_n)^t = M_D^B(f) \circ M_B(x).$$

□

Beispiel 31.5.
Betrachten wir die lineare Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ aus Beispiel 31.2 und den Vektor $x = (0, 1)^t$, so gilt

$$x = 1 \cdot (1, 2)^t - 1 \cdot (1, 1)^t = 1 \cdot b_1 - 1 \cdot b_2$$

und damit

$$M_B(x) = (1, -1)^t.$$

Daraus leiten wir

$$M_D(f(x)) = M_D^B(f) \cdot M_B(x) = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}$$

ab, woraus

$$f(x) = 0 \cdot d_1 - 1 \cdot d_2 + 2 \cdot d_3 = 0 \cdot (1, 1, 0)^t - 1 \cdot (0, 1, 1)^t + 2 \cdot (0, 0, 1)^t = (0, -1, 1)^t$$

folgt. Wir können also die Bilder beliebiger Vektoren ausrechnen, obwohl wir die Abbildungsvorschrift nicht als geschlossene Formel in den Standardkoordinaten kennen. Das ist allerdings mühsam, und wir werden weiter unten sehen, wie man diese Abbildungsvorschrift aus der Matrixdarstellung gewinnen kann.

Satz 31.6 (Matrixdarstellung linearer Abbildungen).
Die Abbildung

$$M_D^B : \mathrm{Hom}_K(V,W) \longrightarrow \mathrm{Mat}(m \times n, K) : f \mapsto M_D^B(f)$$

ist ein Isomorphismus von K-Vektorräumen.

Beweis: Wir zeigen zunächst, daß die Abbildung linear ist. Sind $f, g \in \mathrm{Hom}_K(V,W)$ mit $M_D^B(f) = (a_{ij})$ und $M_D^B(g) = (b_{ij})$ und sind $\lambda, \mu \in K$, so gilt

$$(\lambda f + \mu g)(b_j) = \lambda \cdot f(b_j) + \mu \cdot g(b_j) = \lambda \cdot \sum_{i=1}^{m} a_{ij} \cdot d_i + \mu \cdot \sum_{i=1}^{m} b_{ij} \cdot d_i = \sum_{i=1}^{m} (\lambda \cdot a_{ij} + \mu \cdot b_{ij}) \cdot d_i,$$

woraus wir die Matrixdarstellung

$$M_D^B(\lambda f + \mu g) = (\lambda \cdot a_{ij} + \mu \cdot b_{ij}) = \lambda \cdot (a_{ij}) + \mu \cdot (b_{ij}) = \lambda \cdot M_D^B(f) + \mu \cdot M_D^B(g)$$

erhalten. Die Abbildung M_D^B ist also K-linear.

Es bleibt, zu zeigen, daß M_D^B bijektiv ist. Sei dazu $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$ eine beliebige $m \times n$-Matrix und setzen wir

$$y_j = \sum_{i=1}^{m} a_{ij} \cdot d_i \in W,$$

so gibt es nach dem Existenz- und Eindeutigkeitssatz für lineare Abbildungen 29.18 genau eine lineare Abbildung $f \in \mathrm{Hom}_K(V,W)$ mit

$$f(b_j) = y_j = \sum_{i=1}^{m} a_{ij} \cdot d_i$$

für $j = 1, \ldots, n$, d. h. es gibt genau eine lineare Abbildung f mit

$$M_D^B(f) = (a_{ij}).$$

Die Abbildung M_D^B ist also bijektiv. □

Bemerkung 31.7 (Matrixdarstellung bezüglich der kanonischen Basen).
Für $f \in \mathrm{Hom}_K\left(K^n, K^m\right)$ definieren wir eine Matrix $A_f \in \mathrm{Mat}(m \times n, K)$, deren j-te Spalte das Bild $f(e_j)$ des j-ten Einheitsvektors ist. Dann ist

$$M_F^E : \mathrm{Hom}_K\left(K^n, K^m\right) \longrightarrow \mathrm{Mat}(m \times n, K) : f \mapsto A_f$$

die Umkehrabbildung von

$$\mathrm{Mat}(m \times n, K) \longrightarrow \mathrm{Hom}_K\left(K^n, K^m\right) : A \mapsto f_A.$$

Diesen Spezialfall von Satz 31.6 hatten wir schon in Korollar 29.21 bewiesen, wobei die Linearität der Abbildung dabei unmittelbar aus Lemma 27.8 folgt.

Lemma 31.8 (Verträglichkeit von Matrixdarstellung und Komposition).
Sind $f \in \mathrm{Hom}_K(U,V)$ und $g \in \mathrm{Hom}_K(V,W)$ und sind B, C bzw. D Basen von U, V bzw. W, dann gilt

$$M_D^B(g \circ f) = M_D^C(g) \circ M_C^B(f).$$

Beweis: Ist $B = (b_1, \ldots, b_n)$, so gilt für den j-ten Einheitsvektor $e_j \in K^n$

$$M_B(b_j) = e_j$$

und mit Proposition 31.4 folgt dann

$$\begin{aligned} M_D^B(g \circ f) \circ e_j =& M_D^B(g \circ f) \circ M_B(b_j) = M_D\big((g \circ f)(b_j)\big) \\ =& M_D\big(g(f(b_j))\big) = M_D^C(g) \circ M_C\big(f(b_j)\big) \\ =& M_D^C(g) \circ \big(M_C^B(f) \circ M_B(b_j)\big) = \big(M_D^C(g) \circ M_C^B(f)\big) \circ e_j. \end{aligned}$$

Da die Multiplikation einer Matrix mit e_j die j-te Spalte dieser Matrix liefert, stimmen in $M_D^B(g \circ f)$ und in $M_D^C(g) \circ M_C^B(f)$ also die j-te Spalte überein und das für alle $j = 1, \ldots, n$. Die Matrizen sind also identisch. □

Bemerkung 31.9 (K-Algebren).
Ein K-Vektorraum $(B, +, \cdot)$, auf dem zusätzlich eine Multiplikation

$$\circ : B \times B \to B : (x, y) \mapsto x \circ y$$

definiert ist, so daß $(B, +, \circ)$ ein (nicht unbedingt kommutativer) Ring mit Eins 1_B ist, heißt eine *K-Algebra*, falls die Skalarmultiplikation mit der Ringmultiplikation verträglich ist, d. h. für $\lambda \in K$ und $x, y \in B$ gelten:

$$\lambda \cdot (x \circ y) = (\lambda \cdot x) \circ y = x \circ (\lambda \cdot y).$$

Ein *K-Algebrenisomorphismus* ist eine bijektive Abbildung $\varphi : A \to B$ zwischen zwei K-Algebren A und B, die mit allen drei Operationen verträglich ist und die 1_A auf 1_B abbildet.

$\big(\mathrm{End}_K(V), +, \cdot, \circ\big)$ und $\big(\mathrm{Mat}(n, K), +, \cdot, \circ\big)$ sind Beispiele für K-Algebren, die für unsere Vorlesung von Bedeutung sind, und das folgende Korollar besagt, daß diese isomorph zueinander sind.

Korollar 31.10 (M_B^B ist ein K-Algebrenisomorphismus).
Für zwei Endomorphismen $f \in \mathrm{End}_K(V)$ gilt

$$M_B^B(f \circ g) = M_B^B(f) \circ M_B^B(g).$$

Insbesondere, $M_B^B : \mathrm{End}_K(V) \longrightarrow \mathrm{Mat}_n(K)$ ist ein K-Algebrenisomorphismus.

Beweis: Die Aussage folgt unmittelbar aus Lemma 31.8. □

B) Basiswechsel

Unser Ziel ist es nun, in obigem Beispiel 31.5 aus $M_D^B(f)$ die Matrix $A_f = M_F^E(f)$ zu bestimmen. Dazu führen wir folgende allgemeine Begriffsbildung ein.

Definition 31.11 (Basiswechsel).
Sind $B = (b_1, \ldots, b_n)$ und $B' = (b'_1, \ldots, b'_n)$ zwei Basen von V, so besitzt jedes b_j eine eindeutige Darstellung als Linearkombination der Basis B'

$$b_j = a_{1j} \cdot b'_1 + \ldots + a_{nj} \cdot b'_n.$$

Schreiben wir die Koeffizienten als Spalten in eine Matrix, so erhalten wir die Matrix

$$T_{B'}^B := \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \in \mathrm{Mat}_n(K),$$

den *Basiswechsel* oder die *Koordinatentransformationsmatrix* bezüglich (B, B').

Es gilt also:

> Die j-te Spalte von $T_{B'}^B$ ist der Koordinatenvektor $M_{B'}(b_j)$ von b_j bez. B'.

Bemerkung 31.12 (Basiswechsel als Matrixdarstellung).
Offensichtlich ist der Basiswechsel ein Spezialfall einer Matrixdarstellung

$$T_{B'}^B = M_{B'}^B(\mathrm{id}_V),$$

nämlich die Matrixdarstellung der Identität bezüglich der Basen B und B'.

Lemma 31.13 (Basiswechselmatrizen sind invertierbar).
Sind B und B' zwei Basen von V, so ist $T_{B'}^B$ invertierbar mit

$$\left(T_{B'}^B\right)^{-1} = T_B^{B'}.$$

Beweis: Mit Lemma 31.8

$$T_B^{B'} \circ T_{B'}^B = M_B^{B'}(\mathrm{id}_V) \circ M_{B'}^B(\mathrm{id}_V) = M_B^B(\mathrm{id}_V \circ \mathrm{id}_V) = M_B^B(\mathrm{id}_V) = \mathbb{1}_n.$$

□

Satz 31.14 (Basiswechsel bei Matrixdarstellungen).
Seien B und B' Basen von V, D und D' Basen von W und $f \in \mathrm{Hom}_K(V,W)$. Dann gilt:

$$M_{D'}^{B'}(f) = T_{D'}^{D} \circ M_D^B(f) \circ T_B^{B'}.$$

Beweis: Die Aussage folgt unmittelbar aus Lemma 31.8

$$\begin{aligned} M_{D'}^{B'}(f) =& M_{D'}^{B'}(\mathrm{id}_W \circ f \circ \mathrm{id}_V) \\ =& M_{D'}^{D}(\mathrm{id}_W) \circ M_D^B(f) \circ M_B^{B'}(\mathrm{id}_V) \\ =& T_{D'}^{D} \circ M_D^B(f) \circ T_B^{B'}. \end{aligned}$$

□

Korollar 31.15 (Basiswechsel bei Endomorphismen).
Sind B und B' Basen von V, ist $T = T_B^{B'}$ und ist $f \in \mathrm{End}_K(V)$, so gilt

$$M_{B'}^{B'}(f) = T^{-1} \circ M_B^B(f) \circ T.$$

Beweis: Dies folgt aus Satz 31.14, weil nach Lemma 31.13 $\left(T_B^{B'}\right)^{-1} = T_{B'}^{B}$. □

Beispiel 31.16.
Wir wollen nun für die Abbildung in Beispiel 31.5 die Matrixdarstellung $M_F^E(f)$ bezüglich der kanonischen Basen berechnen. Nach Satz 31.14 gilt:

$$M_F^E(f) = T_F^D \circ M_D^B(f) \circ T_B^E.$$

Um T_F^D auszurechnen, müssen wir d_1, d_2 und d_3 in der kanonischen Basis ausdrücken und die Koeffizienten als Spaltenvektoren in die Matrix T_F^D übertragen:

$$T_F^D = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Um T_B^E zu ermitteln, müßten wir die Einheitsvektoren e_1 und e_2 als Linearkombination der Basis B darstellen, was auf das Lösen zweier Gleichungssysteme hinaus liefe. Stattdessen können wir aber auch T_E^B bestimmen und anschließend invertieren, was sich im Falle einer (2×2)-Matrix anbietet, da das Invertieren sehr einfach ist (vgl. Aufgabe 27.14),

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

sofern die Matrix invertierbar ist.

Analog zum Fall von T_F^D erhalten wir

$$T_E^B = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix},$$

und somit

$$T_B^E = \left(T_E^B\right)^{-1} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}.$$

Also gilt:

$$M_F^E(f) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \circ \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \circ \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

C) Äquivalenz von Matrizen und der Rang

Die Koordinatentransformationen in Vektorräumen mit Basen führen auf folgende Äquivalenzbegriffe für Matrizen.

Definition 31.17 (Äquivalenz von Matrizen).
Eine Matrix $A' \in \mathrm{Mat}(m \times n, K)$ heißt *äquivalent* zu $A \in \mathrm{Mat}(m \times n, K)$, falls es invertierbare Matrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$ gibt mit

$$A' = S \circ A \circ T.$$

Bemerkung 31.18 (Äquivalenz von Matrizen als Äquivalenzrelation).
Die Äquivalenz von Matrizen eine Äquivalenzrelation auf $\mathrm{Mat}(m \times n, K)$ ist.

Denn für $A, B, C \in \mathrm{Mat}(m \times n, K)$ gelten:

- $A = \mathbb{1}_m \circ A \circ \mathbb{1}_n$ und mithin ist A äquivalent zu A;
- wenn B äquivalent zu A ist, gibt es $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$ mit $B = S \circ A \circ T$ und mithin $S^{-1} \circ B \circ T^{-1} = A$, so daß auch A äquivalent zu B;
- wenn B äquivalent zu A und C äquivelent zu B ist, so gibt es Matrizen $S, U \in \mathrm{Gl}_m(K)$ und $T, V \in \mathrm{Gl}_n(K)$ mit $B = S \circ A \circ T$ und $C = U \circ B \circ V$ und mithin gilt auch $C = (U \circ S) \circ A \circ (T \circ V)$ mit $U \circ S \in \mathrm{Gl}_m(K)$ und $T \circ V \in \mathrm{Gl}_n(K)$, so daß C auch äquivalent zu A ist.

Beispiel 31.19.
Die Matrizen

$$A = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \quad \text{und} \quad A' = \begin{pmatrix} 3 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$$

sind äquivalent, da wir in Beispiel 31.16

$$A' = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \circ \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \circ \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$$

gezeigt haben.

Definition 31.20 (Rang).

a. Ist $f \in \mathrm{Hom}_K(V,W)$, so definieren wir den *Rang* von f als

$$\mathrm{rang}(f) := \dim_K \big(\mathrm{Im}(f)\big).$$

b. Ferner definieren wir für eine Matrix $A \in \mathrm{Mat}(m \times n, K)$ den *Rang* von A durch:

$$\mathrm{rang}(A) := \mathrm{rang}(f_A).$$

Bemerkung 31.21 (Rangabschätzung).

a. Für $f \in \mathrm{Hom}_K(V,W)$ gilt wegen der Dimensionsformel für lineare Abbildungen

$$\mathrm{rang}(f) = \dim_K(V) - \dim_K \big(\mathrm{Ker}(f)\big) \leq \dim_K(V)$$

und da $\mathrm{Im}(f)$ ein Unterraum von W ist, gilt auch $\mathrm{rang}(f) \leq \dim_K(W)$.

b. Man beachte, daß das Bild von f_A von den Spalten von A erzeugt wird, so daß der Rang von A die Anzahl linear unabhängiger Spalten von A ist.
Zudem folgt aus a. für $A \in \mathrm{Mat}(m \times n, K)$

$$\mathrm{rang}(A) \leq \min\{m,n\}.$$

Beispiel 31.22.
Da der Rang einer Matrix die Anzahl linear unabhängiger Spalten ist, gelten

$$\mathrm{rang}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0, \quad \mathrm{rang}\begin{pmatrix} 1 & 0 & 2 \\ 1 & 0 & 2 \end{pmatrix} = 1, \quad \mathrm{rang}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} = 2.$$

Satz 31.23 (Invertierbare Matrizen haben vollen Rang.).
Eine Matrix $A \in \mathrm{Mat}_n(K)$ ist genau dann invertierbar, wenn $\mathrm{rang}(A) = n$.
Damit sind die Spalten einer invertierbaren $n \times n$-Matrix eine Basis des K^n.

Beweis: Nach Aufgabe 29.29 ist A genau dann invertierbar, wenn $f_A : K^n \longrightarrow K^n$ bijektiv ist. Wegen Korollar 30.23 ist dies genau dann der Fall, wenn f_A surjektiv ist, d.h. wenn $\mathrm{Im}(f_A) = K^n$. Wegen $\mathrm{Im}(f_A) \subseteq K^n$ und Lemma 30.14 ist dies wiederum

genau dann der Fall, wenn

$$n = \dim_K(\mathrm{Im}(f_A)) = \mathrm{rang}(A).$$

Also ist A genau dann invertierbar, wenn $\mathrm{rang}(A) = n$ gilt. In diesem Fall sind die Spalten von A linear unabhängig in K^n und bilden mithin eine Basis des K^n. □

Bemerkung 31.24.
Sind die Matrizen $A, A' \in \mathrm{Mat}(m \times n, K)$ äquivalent, so gibt es Basen B von K^n und D von K^m, so daß

$$A' = M_D^B(f_A),$$

d.h. A und A' sind Matrixdarstellungen derselben linearen Abbildung f_A bezüglich verschiedener Basen!

Dazu betrachten wir einfach die Matrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$ mit $A' = S \circ A \circ T$. Die Spalten von S^{-1} sind linear unabhängig und bilden eine Basis D von K^m nach Satz 31.23. Ebenso bilden die Spalten von T eine Basis B von K^n. Für diese Basen gilt aber nach Konstruktion

$$T_F^D = S^{-1} \quad \text{und} \quad T_E^B = T.$$

Insgesamt erhalten wir also

$$M_D^B(f_A) = T_D^F \circ M_F^E(f_A) \circ T_E^B = S \circ A \circ T = A'.$$

Wir werden nun zeigen, daß der Rang der Matrixdarstellung einer linearen Abbildung *nicht* von der Wahl der Basen abhängt, bezüglich derer man die Matrixdarstellung bildet.

Proposition 31.25 (Rang einer Matrixdarstellung).
Ist $f \in \mathrm{Hom}_K(V, W)$, so gilt

$$\mathrm{rang}(f) = \mathrm{rang}\left(M_D^B(f)\right).$$

Insbesondere haben äquivalente Matrizen den gleichen Rang.

Beweis: Wir betrachten die Karten $\phi_B : V \longrightarrow K^n$ und $\phi_D : W \longrightarrow K^m$. Da ϕ_D ein Isomorphismus ist, erhält ϕ_D die Dimension eines Vektorraumes und es folgt

$$\begin{aligned}
\mathrm{rang}\left(M_D^B(f)\right) &= \mathrm{rang}\left(f_{M_D^B(f)}\right) = \mathrm{rang}\left(\phi_D \circ f \circ \phi_B^{-1}\right) \\
&= \dim_K\left(\phi_D\left(f\left(\phi_B^{-1}(K^n)\right)\right)\right) = \dim_K\left(\phi_D\left(f(V)\right)\right) \\
&= \dim_K\left(f(V)\right) = \mathrm{rang}(f)
\end{aligned}$$

Sind A und A' äquivalent, so sind sie nach Bemerkung 31.24 Matrixdarstellungen der gleichen Abbildung f_A bezüglich verschiedener Basen und wir haben gerade gesehen, daß der Rang der Matrixdarstellung nicht von der Wahl der Basen abhängt. □

Beispiel 31.26.
Die Abbildung f in Beispiel 31.2 hat den Rang zwei, wie man an ihrer Matrixdarstellung sieht:

$$M_D^B(f) = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix}.$$

Satz 31.27 (Normalform einer Matrixdarstellung bezüglich Äquivalenz).
Es sei $f \in \mathrm{Hom}_K(V,W)$ mit $\mathrm{rang}(f) = r$. Dann gibt es Basen B von V und D von W mit

$$M_D^B(f) = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right),$$

wobei hier 0 jeweils die Nullmatrix der entsprechenden Größe meint.[2] Wir bezeichnen die rechte Seite der obigen Gleichung auch als die *Normalform* von f bezüglich Äquivalenz.

Beweis: Wähle vermöge Lemma 30.6 ein Komplement U von $\mathrm{Ker}(f)$. Nach Satz 28.33 und Lemma 28.36 ist die folgende Abbildung ein Isomorphismus

$$f_{|U} : U \to \mathrm{Im}(f) : x \mapsto f(x).$$

Wähle eine Basis $(d_1, \ldots, d_r)$ von $\mathrm{Im}(f)$. Dann ist $(b_1, \ldots, b_r)$ mit $b_i := (f_{|U})^{-1}(d_i)$ eine Basis von U, nach Aufgabe 29.25. Wähle nun eine Basis $(b_{r+1}, \ldots, b_n)$ von $\mathrm{Ker}(f)$, dann ist $B = (b_1, \ldots, b_n)$ eine Basis von $V = U \oplus \mathrm{Ker}(f)$. Ergänze ferner $(d_1, \ldots, d_r)$ zu einer Basis $D = (d_1, \ldots, d_m)$ von W vermöge Korollar 30.5. Dann:

$$f(b_i) = \begin{cases} d_i, & i = 1, \ldots, r, \\ 0, & i = r+1, \ldots, n. \end{cases}$$

Also hat $M_D^B(f)$ die gewünschte Gestalt. □

So wie für eine lineare Abbildung eine Matrixdarstellung in Normalform existiert, existiert für jede Matrix auch eine Normalform bezüglich Äquivalenz.

[2] Man bezeichnet die vier Matrizen $\mathbb{1}_r \in \mathrm{Mat}_r(K)$, $0 \in \mathrm{Mat}\,(r \times (n-r), K)$, $0 \in \mathrm{Mat}\,((m-r) \times r, K)$ und $0 \in \mathrm{Mat}((m-r) \times (n-r), K)$ auch als *Blöcke* von $M_D^B(f)$ und die Matrix $M_D^B(f)$ als eine *Blockmatrix*.

Korollar 31.28 (Normalform einer Matrix bezüglich Äquivalenz).
Zu $A \in \mathrm{Mat}(m \times n, K)$ mit $r = \mathrm{rang}(A)$ existieren Matrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$ mit

$$S \circ A \circ T = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right). \tag{80}$$

Die rechte Seite heißt die *Normalform* von A bezüglich Äquivalenz.

Beweis: Anwendung des Satzes 31.27 auf $f_A : K^n \to K^m$ liefert B und D von K^n bzw. K^m mit

$$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right) = M_D^B(f_A) = T_D^F \circ M_F^E(f_A) \circ T_E^B = T_D^F \circ A \circ T_E^B.$$

Die Behauptung folgt also, da $S := T_D^F$ und $T := T_E^B$ invertierbar sind. □

Beispiel 31.29.
Die folgende Matrix $A \in \mathrm{Mat}(3 \times 4, \mathbb{R})$ hat Rang 2 und hat somit die Matrix B als Normalform:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 2 & 4 & 6 & 8 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Bemerkung 31.30 (Normalform als guter Repräsentant der Äquivalenzklasse).
Aus Korollar 31.28 folgt, daß zwei Matrizen genau dann äquivalent sind, wenn sie den gleichen Rang haben. $\mathrm{Mat}(m \times n, K)$ zerfällt also in $\min\{m, n\} + 1$ Äquivalenzklassen und jede Äquivalenzklasse ist durch den Rang einer ihrer Matrizen eindeutig bestimmt. Darüber hinaus besitzt jede Äquivalenzklasse $\overline{A}$, $A \in \mathrm{Mat}(m \times n, K)$, einen besonders schönen Repräsentanten, nämlich

$$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Diesen Repräsentanten der Äquivalenzklasse von A nennt man die *Normalform von A bezüglich Äquivalenz*.

Wir wollen die Normalform einer Matrix bezüglich Äquivalenz nun ausnutzen, um Eigenschaften wie den Rang oder die Invertierbarkeit der Matrix und ihrer Transponierten zueinander in Beziehung zu setzen.

Korollar 31.31 (Zeilen- und Spaltenrang).

a. $A \in \mathrm{Mat}_n(K)$ ist genau dann invertierbar, wenn A^t invertierbar ist. In dem Fall gilt
$$\left(A^t\right)^{-1} = \left(A^{-1}\right)^t.$$

b. Für eine Matrix $A \in \mathrm{Mat}(m \times n, K)$ gilt
$$\mathrm{rang}(A) = \mathrm{rang}\left(A^t\right).$$
Insbesondere ist die Anzahl linear unabhängiger Spalten in A gleich der Anzahl linear unabhängiger Zeilen!

Beweis: a. Es sei A invertierbar und $B = A^{-1}$. Dann gilt
$$B^t A^t = (AB)^t = \mathbb{1}_n^t = \mathbb{1}_n,$$
so daß A^t nach Korollar 30.24 invertierbar ist mit Inverser $\left(A^t\right)^{-1} = \left(A^{-1}\right)^t$. Ist umgekehrt A^t invertierbar, so ist nach dem eben gezeigten auch $A = (A^t)^t$ invertierbar.

b. Nach Korollar 31.28 finden wir invertierbare Matrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$, so daß
$$S \circ A \circ T = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right)$$
in Normalform mit $r = \mathrm{rang}(A)$ gegeben ist. Dann ist aber
$$T^t \circ A^t \circ S^t = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right)^t \in \mathrm{Mat}(n \times m, K)$$
ebenfalls eine Matrix in Normalform. Es gilt also
$$\mathrm{rang}\left(T^t \circ A^t \circ S^t\right) = r$$
und wegen Teil a. ist die Matrix $T^t \circ A^t \circ S^t$ äquivalent zu A^t, so daß sie nach Proposition 31.25 den gleichen Rang hat wie A^t.

□

Beispiel 31.32.
Die Matrix
$$A = \begin{pmatrix} 1 & 1 & 2 & 3 & 5 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & 3 & 1 \end{pmatrix}$$

hat offenbar den Rang 3, da die ersten drei Spalten schon linear unabhängig sind. Mithin hat auch die transponierte Matrix

$$A^t = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \\ 3 & 2 & 3 \\ 5 & 2 & 1 \end{pmatrix}$$

den Rang 3, d.h. die drei Spalten sind linear unabhängig.

Aufgaben

Aufgabe 31.33 (f-invarianter Unterraum).
Sei $f : V \longrightarrow V$ eine K-lineare Abbildung und $U \leq V$ ein Unterraum mit $f(U) \subseteq U$; U heißt dann ein *f-invarianter Unterraum*. Zeige, daß die Abbildungen

$$f_U : U \longrightarrow U : x \mapsto f(x)$$

und

$$f_{V/U} : V/U \longrightarrow V/U : \overline{x} \mapsto \overline{f(x)}$$

wohldefiniert und K-linear sind.

Aufgabe 31.34 (Matrixdarstellung in Blockform).
Es sei $f : V \longrightarrow V$ eine K-lineare Abbildung und $U \leq V$ ein f-invarianter Unterraum. Ferner sei $B' = (x_1, \ldots, x_k)$ eine Basis von U und $B = (x_1, \ldots, x_n)$ eine Ergänzung von B' zu einer Basis von V.
Dann ist $B'' = \left(\overline{x_{k+1}}, \ldots, \overline{x_n}\right)$ eine Basis von V/U und es gilt

$$M_B^B(f) = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right),$$

wobei $0 \in \mathrm{Mat}((n-k) \times k, K)$ die Nullmatrix ist und $* \in \mathrm{Mat}(k \times (n-k), K)$ eine geeignete Matrix ist.

Aufgabe 31.35 (Matrixdarstellung in Blockdiagonalgestalt).
Sei $f : V \longrightarrow V$ eine K-lineare Abbildung und $V = U_1 \oplus \ldots \oplus U_k$ die direkte Summe nicht-trivialer f-invarianter Unterräume $U_1, \ldots, U_k$ mit Basen $B_1, \ldots, B_k$.

Dann ist $B = B_1 \cup \ldots \cup B_k$ eine Basis von V und es gilt

$$M_B^B(f) = \left(\begin{array}{c|ccc} M_{B_1}^{B_1}(f_{U_1}) & 0 & \cdots & 0 \\ \hline 0 & \boxed{M_{B_2}^{B_2}(f_{U_2})} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{M_{B_k}^{B_k}(f_{U_k})} \end{array}\right).$$

Aufgabe 31.36 (Zyklische Unterräume).
Es sei $f \in \mathrm{End}_K(V)$, $0 \neq x \in V$ und $m \in \mathbb{N}$ mit $f^{m-1}(x) \neq 0$ und $f^m(x) = 0$.

a. Zeige, $B = \big(f^{m-1}(x), f^{m-2}(x), \ldots, f(x), x\big)$ ist eine Basis von $U = \mathrm{Lin}\,(B)$.

b. Zeige, U ist f-invariant.

c. Bestimme $M_B^B(f_U)$.

Wir nennen U einen *zyklischen Unterraum* von V.

Aufgabe 31.37.
Für Matrizen $A \in \mathrm{Mat}(n \times p, K)$ und $B \in \mathrm{Mat}(m \times n, K)$ gilt:

$$\mathrm{rang}(B \circ A) \leq \min\big\{\,\mathrm{rang}(A), \mathrm{rang}(B)\big\}.$$

Aufgabe 31.38.
Für eine Matrix $A \in \mathrm{Mat}(m \times n, K)$ bezeichne $\mathrm{ZR}(A)$ die lineare Hülle der Zeilen von A und $\mathrm{SR}(A)$ die lineare Hülle der Spalten von A.
Zeige für $A \in \mathrm{Mat}(m \times n, K)$, $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$

$$\mathrm{ZR}(A) = \mathrm{ZR}(SA) \quad \text{und} \quad \mathrm{SR}(A) = \mathrm{SR}(AT).$$

Aufgabe 31.39.
Betrachte den Vektorraum P_n der Polynome vom Grad höchstens n (siehe Beispiel 28.6) mit Basis $B = (t^0, t^1, \ldots, t^n)$ und die formale Ableitung

$$d : P_n \longrightarrow P_n : \sum_{k=0}^{n} a_k \cdot t^k \mapsto \sum_{k=1}^{n} k \cdot a_k \cdot t^{k-1},$$

von der wir aus Beispiel 28.21 bereits wissen, daß sie K-linear ist.

a. Berechne die Matrixdarstellung $M_B^B(d)$ und den Rang von d.

b. Zeige, daß im Fall $n = 3$ auch $D = (t^0, t^0 + t^1, t^1 + t^2, t^2 + t^3)$ eine Basis von P_3 ist und berechne die Basiswechsel T_B^D und T_D^B sowie die Matrixdarstellung $M_D^D(d)$.

§ 32 Der Gauß-Algorithmus

In diesem Abschnitt wollen wir zeigen, daß man jede Matrix durch elementare Zeilenoperationen in Zeilen-Stufen-Form transformieren kann, und einen Algorithmus angeben, der dies tut, den *Gauß-Algorithmus*.

Definition 32.1 (Zeilen-Stufen-Form).
Es sei $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$.

a. A ist in *Zeilen-Stufen-Form*, kurz ZSF, falls es ein r, mit $0 \leq r \leq m$ und Indizes $j_1, \ldots, j_r$ mit $1 \leq j_1 < j_2 < \ldots < j_r \leq m$ gibt, so daß folgendes gilt:
   (i) $a_{ij} = 0$ für $1 \leq i \leq r$ und $1 \leq j < j_i$,
   (ii) $a_{ij} = 0$ für $r < i \leq m$ und $j = 1, \ldots, n$, und
   (iii) $a_{ij_i} \neq 0$ für $i = 1, \ldots, r$.
   Die Körperelemente a_{ij_i} heißen die *Pivots* der Zeilen-Stufen-Form. Man beachte, daß A genau r linear unabhängige Zeilen hat und daß somit $r = \mathrm{rang}(A)$!

b. Eine Zeilen-Stufen-Form von A heißt *reduziert*, falls zusätzlich gilt:
   (iv) $a_{ij_i} = 1$ für $i = 1, \ldots, r$, und
   (v) $a_{kj_i} = 0$ für $k < i$ und $i = 1, \ldots, r$.

Bemerkung 32.2.
Eine Matrix A in Zeilen-Stufen-Form ist also von der folgenden Gestalt:

$$A = \begin{pmatrix}
0 & \ldots & 0 & a_{1j_1} & * & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & * \\
0 & \ldots & \ldots & 0 & \ldots & 0 & a_{2j_2} & * & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & * \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & \ldots & 0 & a_{3j_3} & * & \ldots & \ldots & \ldots & * \\
\vdots & & & & & & & & & & \ddots & & & & \vdots \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & a_{rj_r} & * & \ldots & * \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\
\vdots & & & & & & & & & & & & & & \vdots \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0
\end{pmatrix}$$

Hat A reduzierte Zeilen-Stufen-Form, so sind die Pivots alle Eins und die Einträge in der Spalte oberhalb der Pivots sind alle Null.

Beispiel 32.3.
Betrachte die Matrizen $A, B, C \in \mathrm{Mat}(4 \times 5, K)$ mit

$$A = \left(\begin{array}{c|cccc} 0 & 1 & 0 & 3 & 0 \\ \hline 0 & 0 & 1 & 2 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right), B = \left(\begin{array}{c|cccc} 1 & 2 & 3 & 4 & 5 \\ \hline 0 & 2 & 3 & 4 & 5 \\ \hline 0 & 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 0 & 4 & 1 \end{array}\right) \text{ und } C = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 3 & 4 & 0 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die Matrix A ist in reduzierter ZSF mit $\mathrm{rang}(A) = r = 3$, $j_1 = 2$, $j_2 = 3$ und $j_3 = 5$.
Die Matrix B ist in ZSF mit $\mathrm{rang}(B) = r = 4$ und $j_1 = 1$, $j_2 = 2$, $j_3 = 3$ und $j_4 = 4$.
Die ZSF ist aber nicht reduziert.
Die Matrix C ist nicht in ZSF. Aber durch Vertauschen der beiden ersten Zeilen entsteht eine Matrix, die ZSF hat.

Für die folgende Definition erinnern wir uns (siehe Beispiel 29.11) an die Matrizen

$$E_i^j = (e_{lk})_{l,k=1,\ldots,n} = (\delta_{il} \cdot \delta_{jk})_{l,k=1,\ldots,n} \in \mathrm{Mat}_n(K),$$

die an der Stelle (i, j) den Eintrag 1 und sonst nur Nullen als Einträge haben.

Definition 32.4 (Elementarmatrizen).
Es seien $0 \neq \lambda \in K$, $n > 0$ und $1 \leq i, j \leq n$ mit $i \neq j$. Wir definieren die folgenden quadratischen Matrizen in $\mathrm{Mat}(n, K)$, die auch als *Elementarmatrizen* bezeichnet werden:

a. $S_i(\lambda) := \mathbb{1}_n + (\lambda - 1) \cdot E_i^i$,

b. $Q_i^j(\lambda) := \mathbb{1}_n + \lambda \cdot E_i^j$, und

c. $P_i^j := \mathbb{1}_n - E_i^i - E_j^j + E_i^j + E_j^i$.

Die Matrizen P_i^j heißen zudem *Permutationsmatrizen*.

Bemerkung 32.5 (Elementare Zeilen- und Spaltenoperationen).
Es seien $0 \neq \lambda \in K$, $1 \leq i, j \leq n$ mit $i \neq j$ und $A \in \mathrm{Mat}(n \times m, K)$.

I $S_i(\lambda) \circ A$ geht aus A hervor, indem man die i-te Zeile von A mit λ multipliziert.

II $Q_i^j(\lambda) \circ A$ geht aus A hervor, indem man zur i-ten Zeile das λ-fache der j-ten Zeile addiert.

III $P_i^j \circ A$ geht aus A hervor, indem man die i-te und j-te Zeile vertauscht.

Man nennt die Multiplikation von links mit diesen Matrizen auch *elementare Zeilenoperationen*. Analog erhält man *elementare Spaltenoperationen*, indem man mit den Matrizen von rechts multipliziert.

I' $A \circ S_j(\lambda)$ geht aus A hervor, indem man die j-te Spalte von A mit λ multipliziert.

II' $A \circ Q_i^j(\lambda)$ geht aus A hervor, indem man zur j-ten Spalte das λ-fache der i-ten Spalte addiert.

III' $A \circ P_i^j$ geht aus A hervor, indem man die i-te und j-te Spalte vertauscht.

Proposition 32.6 (Elementarmatrizen sind invertierbar.).
Es seien $0 \neq \lambda \in K$, $1 \leq i, j \leq n$ mit $i \neq j$ und $A \in \mathrm{Mat}(n \times m, K)$. Dann gelten:

a. $S_i\big(\lambda^{-1}\big) \circ S_i(\lambda) = \mathbb{1}_n$,

b. $Q_i^j(-\lambda) \circ Q_i^j(\lambda) = \mathbb{1}_n$, und

c. $P_i^j \circ P_i^j = \mathbb{1}_n$.

Insbesondere sind die Elementarmatrizen invertierbar und die Inversen sind wiederum Elementarmatrizen vom gleichen Typ.

Beweis: Wir führen den Beweis für b. vor. Die übrigen Teile lassen sich dann analog zeigen. Für $0 \neq \lambda \in K$ gilt, vermittels der Distributivität der Matrixmultiplikation:

$$Q_i^j(-\lambda) \circ Q_i^j(\lambda) = \big(\mathbb{1}_n - \lambda \cdot E_i^j\big) \circ \big(\mathbb{1}_n + \lambda \cdot E_i^j\big) = \mathbb{1}_n - \lambda^2 \cdot E_i^j \circ E_i^j = \mathbb{1}_n,$$

da $E_i^j \circ E_i^j = 0$ wegen $i \neq j$. Beachte dazu, daß für $E_i^j \circ E_i^j = (c_{lk})$ gilt:

$$c_{lk} = \sum_{p=1}^{n} \delta_{il}\delta_{jp}\delta_{ip}\delta_{jk},$$

und daß für $i \neq j$ und p beliebig gilt $\delta_{jp}\delta_{ip} = 0$. □

Satz 32.7 (Existenz der reduzierten Zeilen-Stufen-Form).
Jede Matrix $A \in \mathrm{Mat}(m \times n, K)$ läßt sich mittels endlich vieler elementarer Zeilenoperationen in reduzierte Zeilen-Stufen-Form $\mathrm{rZSF}(A)$ überführen, d.h. es gibt Elementarmatrizen $T_1, \ldots, T_k$, so daß

$$\mathrm{rZSF}(A) = T_1 \circ \ldots \circ T_k \circ A.$$

Beweis: Sei $A \in \mathrm{Mat}(m \times n, K)$. Ist $A = 0$, so hat A bereits ZSF mit $r = 0$ und wir sind fertig.

Ist $A \neq 0$, so führe folgende Schritte durch:

1. Schritt: Durchlaufe die Spalten von oben nach unten, mit der ersten Spalte beginnend, bis der erste Eintrag $a_{i_1 j_1} \neq 0$ gefunden ist:

$$\begin{pmatrix} 0 & \dots & 0 & 0 & * & \dots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & 0 & * & \dots & * \\ \vdots & & \vdots & a_{i_1 j_1} & * & \dots & * \\ \vdots & & \vdots & * & * & \dots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & * & * & \dots & * \end{pmatrix}$$

2. Schritt: Steht $a_{i_1 j_1}$ nicht in der ersten Zeile, d. h. $i_1 \neq 1$, dann vertausche die Zeilen a_1 und a_{i_1} - Zeilenoperation vom Typ III. Die so entstandene Matrix heiße $\widetilde{A}_1 = (\tilde{a}_{ij})$. Dann ist $\tilde{a}_{1 j_1}$ unser erstes Pivot.

3. Schritt: Erzeuge in der Spalte $\tilde{a}^{j_1}$ von $\widetilde{A}_1$ unterhalb von $\tilde{a}_{1 j_1}$ Nullen durch elementare Operationen vom Typ II, d. h. addiere für $k = 2, \dots, m$ zur k-ten Zeile das $-\frac{\tilde{a}_{k j_1}}{\tilde{a}_{1 j_1}}$-fache der ersten Zeile. Die Spalten mit Index kleiner als j_1 werden dadurch nicht geändert. Das Ergebnis ist dann eine Matrix von der Form:

$$A^{(1)} := \left(\begin{array}{cccc|ccc} 0 & \dots & 0 & a^{(1)}_{1 j_1} & * & \dots & * \\ \hline 0 & \dots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & A_2 & \\ 0 & \dots & 0 & 0 & & & \end{array}\right),$$

wobei A_2 eine $(m-1) \times (n-j_1)$-Matrix ist, sofern $j_1 \neq n$.

Ist $n - j_1 = 0$ oder $m - 1 = 0$ oder $A^{(2)} = 0$, so sind wir fertig.

Andernfalls ist $A_2 \neq 0$, und wir führen Schritt 1-3 mit A_2 durch. Dabei kann man alle Zeilenoperationen auf die Matrix $A^{(1)}$ ausdehnen, ohne daß sich in den ersten j_1 Spalten etwas ändert, da dort nur Nullen stehen. Ist A_2 umgeformt, so erhält man eine Matrix $A^{(2)}$ der Form:

$$A^{(2)} = \left(\begin{array}{ccccccc|ccc} 0 & \dots & 0 & a^{(2)}_{1 j_1} & * & \dots & \dots & \dots & \dots & * \\ 0 & \dots & 0 & 0 & \dots & 0 & a^{(2)}_{2 j_2} & * & \dots & * \\ \hline 0 & \dots & 0 & 0 & \dots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & A_3 & \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & & & \end{array}\right)$$

mit einem Pivot $a^{(2)}_{2 j_2}$ und, sofern nicht $m - 2 = 0$ oder $n - j_2 = 0$, einer Matrix A_3, die eine Zeile und mindestens eine Spalte weniger als A_2 hat.

Ist $A_3 = 0$, so sind wir fertig. ansonsten fahren wir fort wie bisher und erhalten Matrizen $A^{(3)}$, A_4, $A^{(4)}$, etc.. Das Verfahren stoppt, falls nach r-maligem Durchlaufen der Schritte

1-3 entweder $r = m$ oder $r = n$ oder $A_{r+1} = 0$. In jedem der drei Fälle ist die Matrix $A^{(r)}$ in ZSF.

Um die Matrix $A^{(r)} = \left(a_{ij}^{(r)}\right)$ in reduzierte ZSF zu bringen, multiplizieren wir zunächst die Zeilen $a_i^{(r)}$, für $i = 1, \ldots, r$, mit $\frac{1}{a_{ij_i}^{(r)}}$, was einer elementaren Zeilenoperation vom Typ I entspricht. Die so entstehende Matrix heiße $A' = (a'_{ij})$. Sodann addiert man für $i = 1, \ldots, r$ und $k = 1, \ldots, i-1$ zur k-ten Zeile das $-a'_{kj_i}$-fache der i-ten Zeile – elementare Operationen vom Typ II – und nennt in jedem Schritt i die neue Matrix wieder A'. Man sieht unmittelbar, daß die entstehende Matrix $A'' = \left(a''_{ij}\right)$ reduzierte ZSF hat, da in Spalte j_i die Elemente a'_{kj_i} in $a''_{kj_i} = 0$, für $k < i$, übergegangen sind. □

Bemerkung 32.8 (Eindeutigkeit der reduzierten Zeilen-Stufen-Form).

a. Der Beweis des Satzes ist konstruktiv, das heißt, aus dem Beweis läßt sich ein Algorithmus zur Berechnung einer ZSF von A herleiten, der sogenannte *Gauß-Algorithmus*.

b. Die reduzierte Zeilenstufenform einer Matrix A ist eindeutig bestimmt, was die Bezeichnung $\mathrm{rZSF}(A)$ rechtfertigt.

Beweis der Eindeutigkeit der Zeilenstufenform: Es sei also $A \in \mathrm{Mat}(m \times n, K)$ eine $m \times n$-Matrix.

Da elementare Zeilenoperationen durch Multiplikation mit Elementarmatrizen von links realisiert werden, gilt für eine ZSF B von A, daß es eine invertierbare Matrix $S \in \mathrm{Gl}_m(K)$ gibt mit $B = S \circ A$ (vgl. auch Satz 32.7). Mit Aufgabe 31.38 folgt dann $\mathrm{ZR}(A) = \mathrm{ZR}(B)$, insbesondere gilt mit Korollar 30.9 also, daß die Nicht-Null-Zeilen von B eine Basis von $\mathrm{ZR}(A)$ bilden, da

$$r := \dim_K \big(\mathrm{ZR}(A)\big) = \mathrm{rang}(A) = \mathrm{rang}(B). \tag{81}$$

Seien nun $B = (b_{ij})$ und $B' = \left(b'_{ij}\right)$ zwei reduzierte ZSF von A mit Zeilenvektoren $b_1, \ldots, b_m$ bzw. $b'_1, \ldots, b'_m$ und Pivotspalten $\{j_1, \ldots, j_r\}$ bzw. $\{j'_1, \ldots, j'_r\}$ - beachte, daß die Anzahl $r = \mathrm{rang}(A)$ nach (81) für beide gleich ist. Wir zeigen nun per Induktion, daß die Zeilen der Matrizen B und B' übereinstimmen.

Induktionsbehauptung: Für $i \in \mathbb{N}$ gilt entweder $i \geq r$ oder $b_{r-i} = b'_{r-i}$, insbesondere also $j_{r-i} = j'_{r-i}$.

Induktionsanfang: $i = 0$. O. E. gelte $j_r \geq j'_r$. Da $b_r \in \mathrm{ZR}(A) = \mathrm{Lin}\,(b'_1, \ldots, b'_r)$, gibt es $\lambda_1, \ldots, \lambda_r \in K$ mit

$$b_r = \sum_{i=1}^{r} \lambda_i b'_i.$$

Insbesondere gilt für $i = 1, \ldots, r-1$

$$0 = b_{rj'_i} = \lambda_i \quad \text{und} \quad b_{rj'_r} = \lambda_r,$$

nach (iv) und (v) in Definition 32.1 angewandt auf die reduzierte ZSF B' mit Pivotspalten $j'_1, \ldots, j'_r$ sowie (i) angewandt auf die ZSF B. Also folgt $b_r = \lambda_r \cdot b'_r$. Da $b_r \neq 0$, muß $\lambda_r \neq 0$ gelten und somit $j'_r = j_r$ wegen (i) in 32.1. Aber dann gilt nach (iv) in 32.1 $1 = b_{rj_r} = \lambda_r$ und somit $b_r = b'_r$.

Induktionsschritt: $0 < i < r-1$ und die Behauptung gelte schon für $0, \ldots, i-1$. O. E. gelte $j_{r-i} \geq j'_{r-i}$. Nach Induktionsvoraussetzung gilt nun $b_{r-i} \in \mathrm{ZR}(A) = \mathrm{Lin}\left(b'_1, \ldots, b'_{r-i}, b_{r-i+1}, \ldots, b_r\right)$ also gibt es $\lambda_1, \ldots, \lambda_r \in K$ mit

$$b_{r-i} = \sum_{k=1}^{r-i} \lambda_k b'_k + \sum_{k=r-i+1}^{r} \lambda_k b_k.$$

Insbesondere gilt nach (v) in Definition 32.1, angewandt auf die reduzierte ZSF B, für $k = r-i+1, \ldots, r$

$$0 = b_{r-i\, j_k} = \lambda_k,$$

da $r - i < k$, und (i) angewandt auf B sowie (v) auf B' liefert für $k = 1, \ldots, r-i-1$

$$0 = b_{r-i\, j'_k} = \lambda_k,$$

da $j'_k < j'_{r-i} \leq j_{r-i}$. Insgesamt erhalten wir also wieder

$$b_{r-i} = \lambda_{r-i} b'_{r-i}. \tag{82}$$

Wäre $j_{r-i} > j'_{r-i}$, dann wäre wieder mit (i) $0 = b_{r-i\, j'_{r-i}} = \lambda_{r-i}$ im Widerspruch zu (82) und $b_{r-i} \neq 0$. Also ist $j_{r-i} = j'_{r-i}$ und dann folgt mit (iv) aus 32.1, daß $\lambda_{r-i} = b_{r-i\, j_{r-i}} = 1$, und damit aus (82)) $b_{r-i} = b'_{r-i}$.

Also haben wir mit Induktion gezeigt, daß die Zeilen von B und B' übereinstimmen, d. h. daß die reduzierte Zeilenstufenform von A eindeutig bestimmt ist. □

Beispiel 32.9.
Wir überführen nun die folgende Matrix in reduzierte ZSF.

$$\begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ \underline{-1} & 1 & -3 & 0 & 2 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} \xmapsto{I \leftrightarrow II} \begin{pmatrix} \underline{-1} & 1 & -3 & 0 & 2 \\ 0 & 0 & -1 & 2 & 3 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} \xmapsto{III \mapsto III + I}$$

$$\left(\begin{array}{cc|ccc} -1 & 1 & -3 & 0 & 2 \\ \hline 0 & 0 & \underline{-1} & 2 & 3 \\ 0 & 0 & -2 & 4 & 5 \end{array}\right) \xmapsto{III \mapsto III - 2 \cdot II} \begin{pmatrix} \underline{-1} & 1 & -3 & 0 & 2 \\ 0 & 0 & \underline{-1} & 2 & 3 \\ 0 & 0 & 0 & 0 & \underline{-1} \end{pmatrix} \xmapsto[III \mapsto -III]{I \mapsto -I, II \mapsto -II}$$

$$\begin{pmatrix} 1 & -1 & 3 & 0 & -2 \\ 0 & 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xmapsto[II \mapsto II + 3 \cdot III]{I \mapsto I + 2 \cdot III} \begin{pmatrix} 1 & -1 & 3 & 0 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xmapsto{I \mapsto I - 3 \cdot II}$$

$$\begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Die vierte Matrix besitzt bereits ZSF mit unterstrichenen Pivots, die letzte ist in reduzierter ZSF.

Wir bemerken, daß wir auch auf anderem Weg zum Ziel gekommen wären, und zwar durch andere Wahl der Pivots.

$$\begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ \underline{1} & -1 & 1 & 4 & 3 \end{pmatrix} \xmapsto{I \leftrightarrow III} \begin{pmatrix} 1 & -1 & 1 & 4 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ 0 & 0 & -1 & 2 & 3 \end{pmatrix} \xmapsto{II \mapsto II + I}$$

$$\left(\begin{array}{cc|ccc} 1 & -1 & 1 & 4 & 3 \\ \hline 0 & 0 & \underline{-2} & 4 & 5 \\ 0 & 0 & -1 & 2 & 3 \end{array}\right) \xmapsto{III \mapsto III - \frac{1}{2} \cdot II} \begin{pmatrix} \underline{1} & -1 & 1 & 4 & 3 \\ 0 & 0 & \underline{-2} & 4 & 5 \\ 0 & 0 & 0 & 0 & \underline{\frac{1}{2}} \end{pmatrix} \xmapsto[III \mapsto 2 \cdot III]{II \mapsto -\frac{1}{2} \cdot II}$$

$$\begin{pmatrix} 1 & -1 & 1 & 4 & 3 \\ 0 & 0 & 1 & -2 & -\frac{5}{2} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xmapsto[II \mapsto II + \frac{5}{2} \cdot III]{I \mapsto I - 3 \cdot III} \begin{pmatrix} 1 & -1 & 1 & 4 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xmapsto{I \mapsto I - II}$$

$$\begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

In der Praxis sind 1000×1000-Matrizen keine Seltenheit. Dort wird mit einer festen Stellenzahl gerechnet und deshalb treten bei großen Matrizen unter Umständen erhebliche Rundungsfehler auf. Es kommt der Wahl der richtigen Pivots eine große Bedeutung zu. Ist das gewählte Pivot zu klein, so kann bei Division durch dieses Pivot im dritten Schritt der Rundungsfehler riesig werden - für den Computer bedeutet dies in etwa, als ob man durch Null zu dividieren versuche. Deshalb wählt man in der Praxis das betragsmäßig größte Element als Pivot.

Rechnet man allerdings in Computeralgebrasystemen mit exakter Arithmetik, so spielt die Auslöschung durch Rundungsfehler keine Rolle. Dort muß man eher dafür sorgen, daß die Zahlen, d. h. die Zähler und Nenner, nicht zu groß werden, da dies zu erheblichen Geschwindigkeitsverlusten führen würde.

Wir wollen abschließend den Gauß-Algorithmus in leicht abgewandelter Form als rekursiven Algorithmus zur Bestimmung der reduzierten ZSF einer Matrix formulieren.

Algorithmus 32.10 (Gauß-Algorithmus).
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.
OUTPUT: $\mathrm{rZSF}(A)$, die reduzierte Zeilen-Stufen-Form von A.

0. Schritt: Falls $A = 0$, gehe zu Schritt 8.
1. Schritt: Falls $m = 1$, gehe zu Schritt 7.
2. Schritt: Durchlaufe die erste Spalte von oben nach unten, bis ein Element ungleich Null a_{i1} gefunden wurde oder das Ende der Spalte erreicht ist.
3. Schritt: Wurde kein $a_{i1} \neq 0$ gefunden, bilde eine Untermatrix B von A durch Streichen der ersten Spalte von A und gehe zu Schritt 6. Andernfalls, vertausche die Zeilen a_1 und a_i.
4. Schritt: Für $k = 2, \ldots, m$ addiere zur k-ten Zeile das $-\frac{a_{k1}}{a_{11}}$-fache der ersten.
5. Schritt: Falls $n = 1$, gehe zu Schritt 7. Andernfalls bilde eine Untermatrix B von A, durch Streichen der ersten Zeile und der ersten Spalte von A.
6. Schritt: Wende den Algorithmus auf die Untermatrix B an.[3]
7. Schritt: Die Matrix A ist nun in ZSF. Für $i = m$ bis $i = 1$, d. h. rückwärts zählend, durchlaufe die Zeile a_i, beginnend mit der ersten Spalte, bis ein Element $a_{ij} \neq 0$ gefunden wurde oder das Ende der Zeile erreicht ist.
In letzterem Fall tue nichts, in ersterem multipliziere die Zeile a_i mit $\frac{1}{a_{ij}}$ und addiere für $k = 1, \ldots, i-1$ zur k-ten Zeile das $-a_{kj}$-fache der i-ten Zeile.
8. Schritt: Gib die (veränderte) Matrix A zurück.

[3] Dies ist der Rekursionsschritt, indem der Algorithmus mit einer kleineren Untermatrix aufgerufen wird. Das Ergebnis, das man dabei zurück erhält, wird wieder in die Matrix A eingefügt.

A) Algorithmus zur Berechnung des Rangs einer Matrix

Lemma 32.11.
Elementare Zeilen- oder Spaltenoperationen ändern den Rang einer Matrix nicht.

Beweis: Multipliziert man eine Matrix A mit einer invertierbaren Matrix, so erhält man eine äquivalente Matrix. Wegen Proposition 31.25 ändert dies den Rang der Matrix nicht. Die Aussage folgt also mit Proposition 32.6. □

Algorithmus 32.12 (zur Bestimmung des Rangs).
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.
OUTPUT: $\mathrm{rang}(A)$

1. **Schritt:** Überführe A in ZSF.
2. **Schritt:** Zähle die Anzahl r der Nicht-Nullzeilen in der ZSF.
3. **Schritt:** Gib r zurück.

Beispiel 32.13.
In Beispiel 32.9 haben wir eine ZSF berechnet:

$$A := \begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} \longmapsto \quad \ldots \quad \longmapsto \begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Die Matrix A hat also Rang 3.

B) Algorithmus zur Berechnung der Inversen einer Matrix

Satz 32.14 (Kriterium für die Invertierbarkeit einer Matrix).
Es sei $A \in \mathrm{Mat}(n, K)$. Dann sind gleichwertig:

a. A ist invertierbar.

b. $\mathrm{rZSF}(A) = \mathbb{1}_n$.

c. Es gibt Elementarmatrizen $T_1, \ldots, T_k \in \mathrm{Mat}(n, K)$ mit:
$$T_k \circ \ldots \circ T_1 \circ A = \mathbb{1}_n.$$

d. Es gibt Elementarmatrizen $T'_1, \ldots, T'_k \in \mathrm{Mat}(n, K)$ mit:
$$A = T'_1 \circ \ldots \circ T'_k.$$

Insbesondere wird die Gruppe $\mathrm{Gl}_n(K)$ also von den Elementarmatrizen erzeugt.

Beweis: Nach Korollar 31.23 gilt, daß A genau dann invertierbar ist, wenn $\operatorname{rang}(A) = n$. Also folgt die Äquivalenz von a.-d. aus Satz 32.7 unter Berücksichtigung von Proposition 32.6. □

Aus Satz 32.14 leitet sich folgendes Verfahren zur Bestimmung der Inversen einer invertierbaren Matrix ab. Hierzu beachte man, daß für Elementarmatrizen $T_1, \ldots, T_k$, für die gilt, daß $T_k \circ \ldots \circ T_1 \circ A = \mathbb{1}_n$, auch gilt, daß

$$T_k \circ \ldots \circ T_1 \circ (A, \mathbb{1}_n) = (\mathbb{1}_n, T_k \circ \ldots \circ T_1) = \left(\mathbb{1}_n, A^{-1}\right).$$

Algorithmus 32.15 (zur Bestimmung der Inversen).
INPUT: $A \in \operatorname{Mat}(n, K)$.
OUTPUT: Inverse von A, falls sie existiert, eine Fehlermeldung sonst.

1. **Schritt:** Erweitere die Matrix A um $\mathbb{1}_n$ zu $C = (A, \mathbb{1}_n) \in \operatorname{Mat}(n \times 2n, K)$.
2. **Schritt:** Überführe C in reduzierte ZSF $C' = \left(A', B\right)$.
3. **Schritt:** Falls $\operatorname{rang}\left(A'\right) = n$, dann gib B zurück, sonst eine Fehlermeldung.

Beispiel 32.16.
Wir betrachten die 3×3-Matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in \operatorname{Mat}(3 \times 3, K)$$

und versuchen die Inverse mittels des Algorithmus 32.15 zu bestimmen.

A			$\mathbb{1}_n$			
1	1	1	1	0	0	
0	1	1	0	1	0	
1	0	1	0	0	1	$III \mapsto III - I$
1	1	1	1	0	0	
0	1	1	0	1	0	
0	−1	0	−1	0	1	$III \mapsto III + II$
1	1	1	1	0	0	$I \mapsto I - III$
0	1	1	0	1	0	$II \mapsto II - III$
0	0	1	−1	1	1	
1	1	0	2	−1	−1	$I \mapsto I - II$
0	1	0	1	0	−1	
0	0	1	−1	1	1	
1	0	0	1	−1	0	
0	1	0	1	0	−1	
0	0	1	−1	1	1	

Hieraus ergibt sich gemäß obigem Algorithmus zunächst, daß A invertierbar ist, und ferner, daß

$$A^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

C) Algorithmus zur Berechnung der Normalform einer Matrix

Korollar 32.17 (Normalform einer Matrix).
Sei $A \in \mathrm{Mat}(m \times n, K)$ mit $r = \mathrm{rang}(A)$, so läßt sich A durch endlich viele elementare Zeilen- und Spaltenoperationen auf die folgende Form bringen:

(83) $$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Beweis: Die Aussage folgt aus Korollar 31.28 und Korollar 32.14, da elementare Operationen nach Bemerkung 32.5 durch Multiplikation mit Elementarmatrizen realisierbar sind. □

Wir wollen nun noch an einem Beispiel zeigen, wie man eine Matrix mittels des gaußschen Verfahrens auf Normalform (83) bringt.

Algorithmus 32.18 (Normalform-Algorithmus).
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.
OUTPUT: Normalform $\mathrm{NF}(A)$ von A bezüglich Äquivalenz sowie die zugehörigen Transformationsmatrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$

1. **Schritt:** Überführe A durch elementare Zeilenoperationen in (reduzierte) ZSF und überführe $\mathbb{1}_m$ durch die selben Zeilenoperationen in eine Matrix S.
2. **Schritt:** Überführe A durch elementare Spaltenoperationen in Normalform und überführe $\mathbb{1}_n$ durch die selben Spaltenoperationen in eine Matrix T.
3. **Schritt:** Gib die Normalform von A sowie die Matrizen S und T zurück.

Beispiel 32.19.
Durch elementare Zeilen und Spaltenoperationen überführt man A_λ, $\lambda \in K$, in Normalform:

$\mathbb{1}_m$	A_λ	$\mathbb{1}_n$	
$1\ 0\ 0$	$1\ 0\ \lambda$	$1\ 0\ 0$	
$0\ 1\ 0$	$0\ 1\ 0$	$0\ 1\ 0$	$ZIII \mapsto ZIII - \lambda \cdot ZI$
$0\ 0\ 1$	$\lambda\ 0\ 1$	$0\ 0\ 1$	
$1\ 0\ 0$	$1\ 0\ \lambda$	$1\ 0\ 0$	
$0\ 1\ 0$	$0\ 1\ 0$	$0\ 1\ 0$	$SIII \mapsto SIII - \lambda \cdot SI$
$-\lambda\ 0\ 1$	$0\ 0\ 1-\lambda^2$	$0\ 0\ 1$	
$1\ 0\ 0$	$1\ 0\ 0$	$1\ 0\ -\lambda$	falls $\lambda = \pm 1$ fertig,
$0\ 1\ 0$	$0\ 1\ 0$	$0\ 1\ 0$	sonst $SIII \mapsto \frac{1}{1-\lambda^2} \cdot SIII$
$-\lambda\ 0\ 1$	$0\ 0\ 1-\lambda^2$	$0\ 0\ 1$	
$1\ 0\ 0$	$1\ 0\ 0$	$1\ 0\ -\frac{\lambda}{1-\lambda^2}$	
$0\ 1\ 0$	$0\ 1\ 0$	$0\ 1\ 0$	
$-\lambda\ 0\ 1$	$0\ 0\ 1$	$0\ 0\ \frac{1}{1-\lambda^2}$	
S	$\mathrm{NF}(A_\lambda)$	T	

Für die Normalform $\mathrm{NF}(A) = SAT$ erhalten wir also

$$\mathrm{NF}(A_\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\lambda & 0 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 & -\frac{\lambda}{1-\lambda^2} \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{1-\lambda^2} \end{pmatrix},$$

falls $\lambda \neq \pm 1$, und sonst

$$\mathrm{NF}(A_\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\lambda & 0 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 & -\lambda \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Insbesondere gilt, $\mathrm{rang}(A) = 3$ für $\lambda^2 \neq 1$ und $\mathrm{rang}(A) = 2$ sonst.

D) Algorithmus zur Berechnung einer Basis

Der folgende Algorithmus beruht auf der Tatsache, daß elementare Zeilenoperationen den Zeilenraum nicht verändern - vgl. Aufgabe 31.38.

Algorithmus 32.20 (Basisberechnung).
INPUT: Ein Erzeugendensystem F des Unterraums $U \subseteq K^n$.
OUTPUT: Eine Basis von U.

1. **Schritt:** Schreibe die Vektoren von F als Zeilen in eine Matrix A und überführe A in Zeilen-Stufen-Form.
2. **Schritt:** Gib die ersten $\mathrm{rang}(A)$ Zeilen als Vektoren zurück.

Beispiel 32.21.
Betrachte $U = \operatorname{Lin}\left((1,0,-1,2,3)^t, (1,-1,1,4,3)^t, (0,2,-4,-4,0)^t\right) \leq \mathbb{R}^5$:

$$\begin{pmatrix} 1 & 0 & -1 & 2 & 3 \\ 1 & -1 & 1 & 4 & 3 \\ 0 & 2 & -4 & -4 & 0 \end{pmatrix} \longmapsto \dots \longmapsto \begin{pmatrix} 1 & 0 & -1 & 2 & 3 \\ 0 & -1 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Also ist $B = \left((1,0,-1,2,3)^t, (0,-1,2,2,0)^t\right)$ eine Basis von U.

E) Algorithmus zum Test auf Injektivität / Surjektivität / Bijektivität

Bemerkung 32.22.
Sei $A \in \operatorname{Mat}(m \times n, K)$. Aus der Dimensionsformel für lineare Abbildungen 30.21

$$\dim_K \left(\operatorname{Ker}(f_A)\right) = n - \operatorname{rang}(A)$$

folgt unmittelbar:

- f_A ist injektiv $\iff \operatorname{rang}(A) = n$.
- f_A ist surjektiv $\iff \operatorname{rang}(A) = m$.
- f_A ist bijektiv $\iff \operatorname{rang}(A) = n = m$.

Algorithmus 32.23 (Test auf Injektivität / Surjektivität / Bijektivität).
INPUT: $A \in \operatorname{Mat}(m \times n, K)$.
OUTPUT: Meldung, ob f_A injektiv, surjektiv oder bijektiv ist.

1. **Schritt:** Bestimme den Rang r von A.
2. **Schritt:** Ist $r = m = n$, gib "f_A ist bijektiv" zurück. Ist $r = m < n$, gib "f_A ist surjektiv" zurück. Ist $r = n < m$, gib "f_A ist injektiv" zurück.

Beispiel 32.24.
Die zur folgenden Matrix $A \in \operatorname{Mat}(3 \times 5, \mathbb{R})$ gehörende Abbildung $f_A : \mathbb{R}^5 \longrightarrow \mathbb{R}^3$ ist weder injektiv noch surjektiv, da $\operatorname{rang}(A) = 2 < 3 = m$ und $\operatorname{rang}(A) = 2 < 5 = n$:

$$A = \begin{pmatrix} 1 & 0 & -1 & 2 & 3 \\ 1 & -1 & 1 & 4 & 3 \\ 0 & 2 & -4 & -4 & 0 \end{pmatrix} \longmapsto \dots \longmapsto \begin{pmatrix} 1 & 0 & -1 & 2 & 3 \\ 0 & -1 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

F) Algorithmus zur Berechnung der Summe zweier Unterräume

Die Berechnung der Summe zweier Unterräume, die durch Erzeuger gegeben sind, ist einfach, da man nur die Erzeuger der beiden Unterräume vereinigen muß.

Algorithmus 32.25 (Summe zweier Unterräume).
INPUT: Erzeugendensysteme F und G von Unterräumen U und U' des K^n.
OUTPUT: Eine Basis von $U + U'$.

1. **Schritt:** Bilde aus F und G ein Erzeugendensystem und berechne mittels 32.20 eine Basis von $U + U' = \langle F \cup G \rangle$.
2. **Schritt:** Gib diese Basis zurück.

G) Algorithmus zum Testen auf lineare Unabhängigkeit

Da eine endliche Familie von Vektoren genau dann linear unabhängig ist, wenn sie eine Basis ihres Erzeugnisses ist, und da die Dimension des Erzeugnisses einer solchen Familie gerade der Rang der Matrix ist, deren Spalten die Erzeuger sind, liefert Korollar 30.9 den folgenden Algorithmus.

Algorithmus 32.26 (Test auf lineare Unabhängigkeit).
INPUT: Eine Familie F von m Vektoren in K^n.
OUTPUT: Eins, falls F linear unabhängig ist, Null sonst.

1. **Schritt:** Ist F leer, gib Eins zurück, sonst schreibe die Vektoren in F als Spalten in eine Matrix A.
2. **Schritt:** Ist $\operatorname{rang}(A) = m$, so gib Eins zurück, sonst Null.

Ist $f = f_A$ für eine $m \times n$-Matrix A, dann wird das Bild von f von den Spalten von A erzeugt. Wir können eine Basis des Bildes also wie folgt bestimmen.

H) Algorithmus zur Berechnung des Bildes von f_A

Algorithmus 32.27 (Bild von f_A).
INPUT: $A \in \operatorname{Mat}(m \times n, K)$.
OUTPUT: Eine Basis von $\operatorname{Im}(f_A)$.

1. **Schritt:** Transponiere A und überführe die Transponierte in ZSF.
2. **Schritt:** Transponiere das Ergebnis wieder und gib die ersten $\operatorname{rang}(A)$ Spaltenvektoren zurück.

Aufgaben

Aufgabe 32.28.
Es seien $0 \neq \lambda \in K$, $1 \leq i, j \leq n$ mit $i \neq j$. Dann gelten:

$$Q_i^j(\lambda) = S_j\big(\lambda^{-1}\big) \circ Q_i^j(1) \circ S_j(\lambda),$$

und

$$P_i^j = Q_j^i(1) \circ Q_i^j(-1) \circ Q_j^i(1) \circ S_j(-1).$$

Aufgabe 32.29.
Berechne den Rang der folgenden Matrix in Abhängigkeit von a und b:

$$\begin{pmatrix} 0 & b & b & b \\ a & 0 & b & b \\ a & a & 0 & b \end{pmatrix} \in \mathrm{Mat}(3 \times 4, \mathbb{R}).$$

Aufgabe 32.30.
Bestimme die Inverse der Matrix

$$\begin{pmatrix} 1 & 3 & -1 & 4 \\ 2 & 5 & -1 & 3 \\ 0 & 4 & -3 & 1 \\ -3 & 1 & -5 & -2 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{R}).$$

Aufgabe 32.31.
Transformiere die folgende Matrix A in Normalform bezüglich Äquivalenz und gib auch die Transformationsmatrizen S und T an:

$$A = \begin{pmatrix} 1 & -2 & 3 & 0 \\ 2 & -7 & 10 & -1 \\ -2 & 4 & -7 & 2 \\ 3 & -5 & 7 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{R}).$$

Aufgabe 32.32.
Überprüfe die folgende Abbildung auf Injektivität und Surjektivität:

$$g : \mathbb{R}^4 \to \mathbb{R}^4 \ : \ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} 2x_1 + x_2 + x_4 \\ 3x_1 + x_2 + x_3 + x_4 \\ 2x_1 + x_3 + x_4 \\ 2x_1 + x_2 + x_3 \end{pmatrix}.$$

Aufgabe 32.33.
Es sei $U = \{(a_1, \dots, a_5)^t \in \mathbb{R}^5 \mid a_1 - 2a_2 = 0 = 2a_4 + a_5\} \leq \mathbb{R}^5$. Bestimme die Dimension von U sowie eine Basis von U, die den Vektor $(2, 1, 1, -1, 2)^t$ enthält.

Aufgabe 32.34.
Seien $U = \langle (1, 0, 1, 1)^t, (-1, 1, 0, 0)^t \rangle \leq \mathbb{R}^4$ und $U' = \langle (1, 0, 1, 0)^t, (1, 1, 1, 1)^t \rangle \leq \mathbb{R}^4$. Zeige, $\mathbb{R}^4 = U \oplus U'$.

§ 33 Lineare Gleichungssysteme

Definition 33.1 (Lineare Gleichungssysteme).

a. Ein *lineares Gleichungssystem* über K

$$\text{(LGS)} \qquad \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ & & & & \vdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m \end{array}$$

besteht aus m Gleichungen in n *Unbestimmten* oder *Variablen* $x_1, \dots, x_n$ mit $a_{ij}, b_i \in K$ für $1 \le i \le m$ und $1 \le j \le n$.
Da sich (LGS) mit $A = (a_{ij})$, $b = (b_1, \dots, b_m)^t$ und $x = (x_1, \dots, x_n)^t$ vermittels Matrixmultiplikation auch kurz schreiben läßt als

$$Ax = b,$$

sprechen wir meist von dem *linearen Gleichungssystem* $Ax = b$.

b. Die Matrix

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in \operatorname{Mat}(m \times n, K)$$

heißt *Koeffizientenmatrix* und der Vektor $b = (b_1, \dots, b_m)^t \in K^m$ die *Inhomogenität* des Gleichungssystems (LGS). Ferner heißt die Matrix

$$(A \mid b) := \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right) \in \operatorname{Mat}\big(m \times (n+1), K\big)$$

die *erweiterte Koeffizientenmatrix* von (LGS).

c. Das lineare Gleichungssystem (LGS) heißt *homogen*, falls $b = 0_{K^m}$ der Nullvektor in K^m ist. Ansonsten heißt das System *inhomogen*.

d. Ist ein lineares Gleichungssystem $Ax = b$ gegeben, so heißt das Gleichungssystem $Ax = 0$ (mit $0 = 0_{K^m}$) das *zugehörige homogene Gleichungssystem*.

e. Ein Vektor $c = (c_1, \dots, c_n)^t \in K^n$ heißt *Lösung* von (LGS), wenn die Gleichung $Ac = b$ erfüllt ist. Die Menge aller Lösungen von (LGS) wird mit

$$\operatorname{Lös}(A, b) := \big\{c \in K^n \mid Ac = b\big\}.$$

bezeichnet.

Bei einem linearen Gleichungssystem sind also Körperelemente a_{ij} und b_i fest vorgegeben, während für die Unbestimmten x_j Körperelemente c_j gesucht werden, die das Gleichungssystem lösen.

Falls $K = \mathbb{R}$, so kann ein lineares Gleichungssystem entweder gar keine, genau eine oder unendlich viele Lösungen haben. Wir werden im Folgenden mehrere Verfahren zur Lösung kennenlernen und uns, im Fall von mehr als einer Lösung, mit der Struktur der Lösungsmenge $\mathrm{Lös}(A, b)$ beschäftigen. Eine wichtige Rolle spielt dabei die lineare Abbildung $f_A : K^n \to K^m$.

Bemerkung 33.2 (Struktur des Lösungraums).
Es sei $A \in \mathrm{Mat}(m \times n, K)$ und $b \in K^m$.

a. Aus den Definitionen folgt unmittelbar

$$\mathrm{Lös}(A, 0) = \{c \in K^n \mid Ac = 0\} = \mathrm{Ker}(f_A),$$

so daß $\mathrm{Lös}(A, 0)$ ein Unterraum des K^n ist mit Dimension

$$\dim_K \big(\mathrm{Lös}(A, 0)\big) = \dim_K \big(\mathrm{Ker}(f_A)\big) = n - \mathrm{rang}(A).$$

Insbesondere ist $Ax = 0$ genau dann eindeutig lösbar, wenn $\mathrm{rang}(A) = n$.

b. Ebenfalls anhand der Definitionen sieht man, daß das lineare Gleichungssystem $Ax = b$ genau dann eine Lösung besitzt, wenn $b \in \mathrm{Im}(f_A) = \{Ac \mid c \in K^n\}$.

Beispiel 33.3.
Das lineare Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ 2x_1 + 3x_2 &= 1 \\ x_2 - x_3 &= 0 \end{aligned}$$

ist inhomogen, hat als Koeffizientenmatrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 0 \\ 0 & 1 & -1 \end{pmatrix} \in \mathrm{Mat}(3 \times 3, \mathbb{R}),$$

und als erweiterte Koeffizientenmatrix

$$(A \mid b) = \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 3 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{array}\right) \in \mathrm{Mat}(3 \times 4, \mathbb{R}).$$

Die Lösung ist in diesem Fall ganz einfach. Wir erhalten $x_3 = x_2$ aus der 3. Gleichung, $3x_2 = 1 - 2x_1$ aus der 2. und, wenn wir das in die erste Gleichung einsetzen, $x_1 + (1 - 2x_1) = 1$, also $x_1 = 0$. Einsetzen von $x_1 = 0$ in die 2. und 3. Gleichung liefert, daß $\big(0, \frac{1}{3}, \frac{1}{3}\big)^t$ die einzige Lösung ist.

Wir geben zunächst ein Kriterium für die Lösbarkeit eines Gleichungssystems.

Satz 33.4 (Kriterium für die Lösbarkeit eines LGS).
Ein Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn $\operatorname{rang}(A) = \operatorname{rang}(A \mid b)$.

Beweis: Wir beachten, daß $\operatorname{Im}(f_A)$ von den Spaltenvektoren $a^1, \ldots, a^n$ von A erzeugt wird, und erhalten deshalb:

$$\begin{aligned} Ax = b \text{ lösbar} &\iff b \in \operatorname{Im}(f_A) = \operatorname{Lin}\left(a^1, \ldots, a^n\right) \\ &\iff b \text{ ist Linearkombination von } a^1, \ldots, a^n \\ &\iff \operatorname{Im}(f_A) = \operatorname{Lin}\left(a^1, \ldots, a^n\right) = \operatorname{Lin}\left(a^1, \ldots, a^n, b\right) = \operatorname{Im}(f_{(A|b)}) \\ &\iff \operatorname{rang}(A) = \dim_K\left(\operatorname{Im}(f_A)\right) = \dim_K\left(\operatorname{Im}(f_{(A|b)})\right) = \operatorname{rang}(A \mid b), \end{aligned}$$

wobei wir für die letzte Äquivalenz berücksichtigen, daß $\operatorname{Im}(f_A) \subseteq \operatorname{Im}(f_{(A|b)})$ gilt. □

Der folgende Satz gibt Auskunft über die Struktur der Lösungsmenge eines linearen Gleichungssystems. Wir haben bereits gesehen, daß diese ein Unterraum ist, wenn das Gleichungssystem homogen ist, und wir werden nun zeigen, daß sie ein affiner Unterraum ist, wenn das Gleichungssystem inhomogen ist.

Satz 33.5 (Struktur von $\operatorname{Lös}(A, b)$ als affiner Raum).
Seien $A \in \operatorname{Mat}(m \times n, K)$, $b \in K^m$ und sei $c \in K^n$ eine Lösung des linearen Gleichungssystems $Ax = b$. Dann gilt:

$$\operatorname{Lös}(A, b) = c + \operatorname{Lös}(A, 0).$$

Beweis: Sei zunächst $y \in \operatorname{Lös}(A, 0)$. Dann gilt:

$$A\big(c + y\big) = Ac + Ay = b + 0 = b,$$

also ist $c + y \in \operatorname{Lös}(A, b)$.

Ist umgekehrt $x \in \operatorname{Lös}(A, b)$. Dann gilt für $y := x - c$

$$Ay = A\big(x - c\big) = Ax - Ac = b - b = 0,$$

also ist $y \in \operatorname{Lös}(A, 0)$. Aber damit ist $x = c + y \in c + \operatorname{Lös}(A, 0)$. □

Wir wollen nun einen Algorithmus kennenlernen, der es uns erlaubt, die Lösung eines linearen Gleichungssystems $Ax = b$ in parametrisierter Form zu bestimmen, d. h. eine spezielle Lösung und eine Basis des Lösungsraumes des zugehörigen homogenen Gleichungssystems zu berechnen. Der wichtigste Schritt ist hierbei die Überführung der erweiterten Koeffizientenmatrix $(A \mid b)$ in reduzierte Zeilen-Stufen-Form.

Lemma 33.6 (Element. Zeilenoperationen ändern Lösungsraum nicht.). Sind $A, A' \in \mathrm{Mat}(m \times n, K)$ und $b, b' \in K^m$ und entsteht die Matrix $(A' \mid b')$ aus $(A \mid b)$ durch elementare Zeilenoperationen, so gilt

$$\mathrm{Lös}(A, b) = \mathrm{Lös}\left(A', b'\right).$$

Beweis: Daß $(A' \mid b')$ aus $(A \mid b)$ durch elementare Zeilenoperationen hervorgeht, bedeutet, daß es eine invertierbare Matrix $S \in \mathrm{Gl}_m(K)$ gibt mit $A' = SA$ und $b' = Sb$.

Ist nun $c \in \mathrm{Lös}(A, b)$, dann gilt $Ac = b$ und damit

$$b' = Sb = SAc = A'c.$$

Also ist $c \in \mathrm{Lös}\left(A', b'\right)$.

Ist andererseits $c \in \mathrm{Lös}\left(A', b'\right)$, dann gilt $A'c = b'$ und damit

$$b = S^{-1}b' = S^{-1}A'c = Ac.$$

Also ist $c \in \mathrm{Lös}(A, b)$. □

Bemerkung 33.7.
Aus Lemma 33.6 und Satz 32.7 folgt, daß wir die erweiterte Koeffizientenmatrix eines Gleichungssystems $Ax = b$ mittels Gauß-Algorithmus in (reduzierte) ZSF überführen können, ohne daß sich die Lösungsmenge ändert.

Wir betrachten deshalb den Fall, daß die Matrix A in ZSF gegeben ist, näher.

Satz 33.8 (Lösbarkeitskriterium für LGS mittels Gauß-Algorithmus). Sei $A \in \mathrm{Mat}(m \times n, K)$ eine Matrix in Zeilen-Stufen-Form und $b \in K^m$. Die erweiterte Koeffizientenmatrix habe die Gestalt

(84)

$$(A \mid b) = \left(\begin{array}{ccccccccccc|c} 0 & \dots & 0 & a_{1j_1} & * & \dots & \dots & \dots & \dots & \dots & * & b_1 \\ 0 & \dots & \dots & 0 & \dots & 0 & a_{2j_2} & * & \dots & \dots & * & b_2 \\ \vdots & & & & & & & \ddots & & & \vdots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & a_{rj_r} & * & * & b_r \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & b_{r+1} \\ \vdots & & & & & & & & & & \vdots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & b_m \end{array}\right)$$

mit Pivots $a_{ij_i} \neq 0$ für $i = 1, \dots, r$. Dann gilt:

a. $Ax = b$ ist genau dann lösbar, wenn $b_{r+1} = \dots = b_m = 0$.

b. Sind $b_{r+1} = \dots = b_m = 0$ und $r = n$, so hat $Ax = b$ *genau eine* Lösung.

c. Sind $b_{r+1} = \ldots = b_m = 0$ und ist $r < n$, so hat $Ax = b$ mehr als eine Lösung. Genauer $\text{Lös}(A, b) = c + \text{Lös}(A, 0)$, wobei c eine spezielle Lösung ist und $\text{Lös}(A, 0)$ die Dimension $n - r$ hat.

Beweis: Die Aussagen folgen aus Satz 33.4, Satz 33.5 und Bemerkung 33.2. □

A) Der Gauß-Algorithmus zur Lösung eines (LGS)

Algorithmus 33.9 (Algorithmus zur Lösung eines LGS).
INPUT: Die erweiterte Koeffizientenmatrix $(A \mid b)$ eines LGS $Ax = b$.
OUTPUT: Eine spezielle Lösung c von $Ax = b$ und eine Basis B von $\text{Lös}(A, 0)$, sofern das Gleichungssystem lösbar ist.

1. **Schritt:** Berechne eine reduzierte Zeilen-Stufen-Form $(A' \mid b')$ von $(A \mid b)$ mit $r = \text{rang}(A')$.
2. **Schritt:** Ist $b'_{r+1} \neq 0$, dann ist das LGS nicht lösbar.
3. **Schritt:** Überführe $(A' \mid b')$ in eine $n \times (n+1)$-Matrix $(A'' \mid b'')$ durch Einfügen und Streichen von Nullzeilen, so daß die Pivotelemente anschließend auf der Diagonale der Matrix A'' stehen.
4. **Schritt:** Ersetze jede Null auf der Diagonale von A'' durch -1.
5. **Schritt:** Die spezielle Lösung ist $c := b''$ und die Spalten von A'', die eine -1 auf der Diagonale haben, sind eine Basis von $\text{Lös}(A, 0)$.

Beispiel 33.10.
Wir betrachten das Gleichungssystem:

$$\begin{aligned} x_1 + x_2 + x_3 - 2x_4 &= 1 \\ x_1 + x_2 - x_3 &= -1 \\ 3x_1 + 3x_2 + x_3 - 4x_4 &= 1 \end{aligned} \tag{85}$$

In Matrixschreibweise lautet das Gleichungssystem:

$$\begin{pmatrix} 1 & 1 & 1 & -2 \\ 1 & 1 & -1 & 0 \\ 3 & 3 & 1 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Durch den Gauß-Algorithmus überführen wir die erweiterte Koeffizientenmatrix in reduzierte Zeilen-Stufen-Form:

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & -2 & 1 \\ 1 & 1 & -1 & 0 & -1 \\ 3 & 3 & 1 & -4 & 1 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Wir sehen, daß $\operatorname{rang}(A) = \operatorname{rang}(A \mid b) = 2$, so daß das Gleichungssystem lösbar ist.

Um die Lösung zu berechnen, fügen wir als zweite Zeile eine Nullzeile ein, um eine 4×5-Matrix zu erzeugen und die Pivotelemente auf der Diagonalen zu haben, und ersetzen die Nullen auf der Diagonalen anschließend durch -1:

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 \end{array}\right).$$

Damit erhalten wir die letzte Spalte

$$c = (0, 0, 1, 0)^t$$

als spezielle Lösung von (85) und die Spalten 2 und 4 als Basis

$$B = \big((1, -1, 0, 0)^t, (-1, 0, -1, -1)^t\big)$$

des Lösungsraums $\text{Lös}(A, 0)$ des homogenen Gleichungsystems $Ax = 0$. Insgesamt gilt damit

$$\text{Lös}(A, b) = c + \text{Lös}(A, 0) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} -1 \\ 0 \\ -1 \\ -1 \end{pmatrix} \;\middle|\; s, t \in K \right\}.$$

Wir wollen nun einige Algorithmen angeben, denen der Algorithmus zur Lösung eines linearen Gleichungssystems zugrunde liegt.

B) Algorithmus zur Berechnung des Kerns von f_A

Ist $f = f_A$ für eine $m \times n$-Matrix A, dann ist der Kern von f gerade die Lösungsmenge $\text{Lös}(A, 0)$ des homogenen Gleichungssystems $Ax = 0$.

Algorithmus 33.11 (Kern von f_A).
INPUT: $A \in \text{Mat}(m \times n, K)$.
OUTPUT: Eine Basis von $\text{Ker}(f_A)$.

1. **Schritt:** Bestimme eine Lösung $\big(c, B\big)$ von $Ax = 0$ gemäß 33.9.
2. **Schritt:** Gib B als Basis zurück.

Beispiel 33.12.
Wir wollen den Kern der Linearen Abbildung $f_A : K^4 \longrightarrow K^3$ berechnen, die durch die Koeffizientenmatrix

$$A = \begin{pmatrix} 1 & 1 & 1 & -2 \\ 1 & 1 & -1 & 0 \\ 3 & 3 & 1 & -4 \end{pmatrix}$$

in Beispiel 33.10 gegeben ist. Dann gehen wir wie in Beispiel 33.10 vor, wobei wir die Inhomogenität durch den Nullvektor ersetzen oder einfach gänzlich ignorieren können. Die Rechnungen ändern sich nicht und wir erhalten wie dort

$$B = \big((1,-1,0,0)^t, (-1,0,-1,-1)^t\big)$$

als Basis von $\mathrm{Ker}(f_A) = \mathrm{Lös}(A, 0)$.

C) Algorithmus zur Berechnung der Transformationsmatrix $T^B_{B'}$

Sind $B = (b_1, \ldots, b_n)$ und $B' = (b'_1, \ldots, b'_n)$ zwei Basen des K^n und wollen wir die Transformationsmatrix $T^B_{B'}$ bestimmen, so müssen wir die Basisvektoren in B als Linearkombination der Basisvektoren in B' darstellen und die so erhaltenen Koeffizienten liefern die Spalten von $T^B_{B'}$. Wir müssen also n Gleichungsysteme Lösen, bei denen die Koeffizientenmatrix stets $b'_1, \ldots, b'_n$ als Spaltenvektoren hat und bei denen die Inhomogenitäten durch die Vektoren $b_1, \ldots, b_n$ gegeben werden. Da die Koeffizientenmatrix sich nicht ändert, können wir die n Gleichungssysteme simultan lösen, indem wir der erweiterten Koeffizientenmatrix gleich alle Vektoren $b_1, \ldots, b_n$ als zusätzliche Spalten anhängen.

Algorithmus 33.13 (Transformationsmatrix $T^B_{B'}$).
INPUT: Zwei Basen $B = (b_1, \ldots, b_n)$ und $B' = (b'_1, \ldots, b'_n)$ im K^n.
OUTPUT: Die Transformationsmatrix $T^B_{B'}$.

1. **Schritt:** Schreibe die Vektoren $b'_1, \ldots, b'_n, b_1, \ldots, b_n$ in dieser Reihenfolge als Spalten in eine Matrix A.
2. **Schritt:** Bringe A auf reduzierte ZSF.
3. **Schritt:** Die letzten n Spalten von $\mathrm{rZSF}(A)$ sind $T^B_{B'}$.

Beispiel 33.14.
Seien die zwei Basen $B = \big((1,1)^t, (1,-1)^t\big)$ und $B' = \big((1,2)^t, (-1,0)^t\big)$ des $\mathbb{R}^2$ gegeben.

B'		B	
1	-1	1	1
2	0	1	-1
1	-1	1	1
0	2	-1	-3
1	-1	1	1
0	1	$-\frac{1}{2}$	$-\frac{3}{2}$
1	0	$\frac{1}{2}$	$-\frac{1}{2}$
0	1	$-\frac{1}{2}$	$-\frac{3}{2}$
$\mathbb{1}_2$		$T^B_{B'}$	

D) Algorithmus zur Berechnung einer Matrixdarstellung $M^B_D(f)$

Wir wollen hier angeben, wie man die Matrixdarstellung einer linearen Abbildung $f : K^n \longrightarrow K^m$ bezüglich zweier Basen $B = (b_1, \ldots, b_n)$ von K^n und $D = (d_1, \ldots, d_m)$ von K^m berechnet. Die Grundidee ist ähnlich wie beim Algorithmus zur Berechnung der Transformationsmatrix.

Algorithmus 33.15 (Matrixdarstellung $M^B_D(f)$).
INPUT: Eine lineare Abbildung $f : K^n \longrightarrow K^m$, eine Basis $B = (b_1, \ldots, b_n)$ von K^n und eine Basis $D = (d_1, \ldots, d_m)$ im K^m.
OUTPUT: Die Matrixdarstellung $M^B_D(f)$.

1. **Schritt:** Schreibe die Vektoren $d_1, \ldots, d_m, f(b_1), \ldots, f(b_n)$ in dieser Reihenfolge als Spalten in eine Matrix A.
2. **Schritt:** Bringe A auf reduzierte ZSF.
3. **Schritt:** Die letzten n Spalten von rZSF(A) sind $M^B_D(f)$.

Beispiel 33.16.
Für die Basen $B = \big((1,0,1)^t, (1,1,0)^t, (0,0,1)^t\big)$ des K^3 und $D = \big((1,1)^t, (1,-1)^t\big)$ des K^2 sowie die lineare Abbildung

$$f : K^3 \longrightarrow K^2 : (x,y,z)^t \mapsto (x+y+z, x-z)^t$$

wollen wir die Matrixdarstellung $M^B_D(f)$ berechnen.

D		$f(B)$		
1	1	2	2	1
1	−1	0	1	−1
1	1	2	2	1
0	−2	−2	−1	−2
1	1	2	2	1
0	1	1	$\frac{1}{2}$	1
1	0	1	$\frac{3}{2}$	0
0	1	1	$\frac{1}{2}$	1
$\mathbb{1}_2$		$M_D^B(f)$		

Bemerkung 33.17.
Natürlich könnte man auch zunächst die Matrixdarstellung $M_F^E(f)$ bezüglich der kanonischen Basen berechnen, da man dazu einfach die Vektoren $f(e_i)$ in die Spalten der Matrix schreiben muß. Analog erhält man T_E^B, indem man die Vektoren von B in die Spalten der Matrix schreibt. Dann muß man nur noch T_D^F mit Hilfe des Algorithmus' zur Berechnung einer Transformationsmatrix bestimmen und kann die Matrizen multiplizieren, um $M_D^B(f)$ zu erhalten.

E) Algorithmus zum Austauschverfahren von Steinitz

Beim Austauschsatz von Steinitz müssen wir die Vektoren in $F = (y_1, \ldots, y_r)$, die wir in die Basis $B = (x_1, \ldots, x_n)$ hineintauschen wollen, sukzessive als Linearkombination der Basisvektoren in (dem veränderten) B darstellen, d.h. wir müssen immer wieder lineare Gleichungssysteme lösen.

Algorithmus 33.18 (Austauschverfahren von Steinitz).
INPUT: Eine Basis $B = (x_1, \ldots, x_n)$ und eine linear unabhängige Familie $F = (y_1, \ldots, y_r)$ von Vektoren in $V = \operatorname{Lin}(B) \subseteq K^n$.
OUTPUT: Eine Basis B' von V, die F enthält.

1. Schritt: Für $i = 1, \ldots, r$ tue:
- Schreibe die Vektoren in B als Spalten in eine Matrix A.
- Bilde die erweiterte Matrix (A, y_i).
- Überführe (A, y_i) in reduzierte Zeilen-Stufen-Form und suche in der letzten Spalte den ersten Eintrag ungleich Null.
- Streiche den entsprechenden Vektor aus B und füge y_i als letzten Vektor in B ein.

2. Schritt: Gib B zurück.

Beispiel 33.19.
Betrachte die linear unabhängige Familie $F = (y_1, y_2) = \big((1,2,1)^t, (1,2,2)^t\big)$ und die Basis $B = (x_1, x_2, x_3) = \big((1,1,0)^t, (1,0,1)^t, (0,0,1)^t\big)$. Wir wollen nun F in B hineintauschen.

Wir bilden die erweiterte Matrix (A, y_1) und überführen sie in reduzierte ZSF:

$$(A, y_1) = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

Da der erste Eintrag in der letzten Spalte nicht Null ist, streichen wir aus B den Vektor x_1 und fügen y_1 als letzten Vektor ein. Wir erhalten die neue Basis

$$B = (x_2, x_3, y_1).$$

Dann bilden wir wieder die erweiterte Matrix (A, y_2) und überführen sie in reduzierte ZSF:

$$(A, y_2) = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 1 & 1 & 1 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Der erste Eintrag der letzten Spalte, der nicht Null ist, ist der zweite, mithin müssen wir den zweiten Vektor in B streichen, das ist x_3, und fügen y_2 am Ende ein. Wir erhalten die Basis

$$B = (x_2, y_1, y_2).$$

Bemerkung 33.20 (Berechnung eines Komplementes / einer Basis für K^n/U).
Will man ein Komplement eines Unterraums U in K^n berechnen, so berechnet man zunächst eine Basis von U und tauscht diese anschließend mit Steinitz in die kanonische Basis von K^n. Die verbleibenden Vektoren der kanonischen Basis sind dann eine Basis für ein Komplement und zugleich sind deren Restklassen eine Basis für den Faktorraum K^n/U (siehe Bemerkung 30.20). Der obige Algorithmus erlaubt also auch die Berechnung eine Komplementes und einer Basis eines Faktorraums.

F) Algorithmus zur Berechnung von Gleichungen eines Unterraumes

Wir haben gesehen, daß Unterräume des K^n als Lösungsmengen von homogenen linearen Gleichungssystemen auftauchen. Um etwa den Schnitt zweier Unterräume des K^n zu bestimmen, ist es nützlich, aus dem Erzeugendensystem eines Unterraumes ein Gleichungssystem bestimmen zu können, das den Unterraum beschreibt.

Algorithmus 33.21 (Gleichungen eines Unterraumes).
INPUT: Eine Familie $F = (x_1, \ldots, x_m)$ von Vektoren im K^n.
OUTPUT: Eine Matrix $A \in \mathrm{Mat}(k \times n, K)$ mit $\mathrm{Lös}(A, 0) = \mathrm{Lin}\,(F)$.

1. **Schritt:** Schreibe die Vektoren aus F als Zeilen in eine Matrix $B \in \mathrm{Mat}(m \times n, K)$ und bestimme eine Basis $(y_1, \ldots, y_k)$ von $\mathrm{Ker}(f_B) = \mathrm{Lös}(B, 0)$.
2. **Schritt:** Schreibe die $y_1, \ldots, y_k$ als Zeilenvektoren in eine Matrix A.
3. **Schritt:** Gib A zurück.

Beispiel 33.22.
Finde ein lineares Gleichungssystem $Ax = 0$ mit Lösungsmenge

$$\mathrm{Lös}(A, 0) = \mathrm{Lin}\left((1, 2, 1)^t, (0, 1, 0)^t\right) \leq \mathbb{R}^3.$$

Dazu Bilden wir die 2×3-Matrix

$$B = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

und berechnen ihren Kern:

$$\mathrm{rZSF}(B, 0) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Somit ist der Vektor in der dritten Spalte eine Basis von $\mathrm{Lös}(B, 0)$ und wir erhalten

$$A = \begin{pmatrix} 1 & 0 & -1 \end{pmatrix}.$$

G) Algorithmus zur Berechnung des Durchschnitts zweier Unterräume

Abschließend sind wir nun in der Lage, einen Algorithmus anzugeben, mittels dessen sich eine Basis des Schnitts zweier Unterräume des K^n ermitteln läßt.

Algorithmus 33.23 (Durchschnitt zweier Unterräume).
INPUT: Zwei Familien F und G von Vektoren in K^n.
OUTPUT: Eine Basis des Schnitts von $\mathrm{Lin}\,(F)$ und $\mathrm{Lin}\,(G)$.

1. **Schritt:** Bestimme Matrizen A und A' gemäß 33.21, so daß $\mathrm{Lin}\,(F) = \mathrm{Lös}(A, 0)$ und $\mathrm{Lin}\,(G) = \mathrm{Lös}\left(A', 0\right)$.
2. **Schritt:** Bilde aus den Zeilen von A und A' eine gemeinsame Matrix A''.
3. **Schritt:** Bestimme eine Basis B von $\mathrm{Ker}(f_{A''}) = \mathrm{Lös}\left(A'', 0\right)$ gemäß 33.11 und gib B zurück.

Beispiel 33.24.
Wir wollen den Durchschnitt der Unterräume

$$U = \operatorname{Lin}\big((1,2,1)^t, (0,1,0)^t\big)$$

und

$$U' = \{(x,y,z)^t \in \mathbb{R}^3 \mid x+y+z = 0\}$$

berechnen. Der zweite Unterraum ist bereits als Lösungsmenge eines Gleichungssystems mit Koeffizientenmatrix $A' = (1\ 1\ 1)$ gegeben. Für den ersten Unterraum haben wir eine solche Darstellung $\text{Lös}(A,0)$ bereits in Beispiel 33.22 berechnet. Wir bilden eine neue Matrix A''

$$A'' = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

aus A und A' und lösen das zugehörige homogene Gleichungssystem

$$(A'',0) = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Damit ist die dritte Spalte

$$B = \big((-1,2,-1)^t\big)$$

eine Basis von $U \cap U'$.

H) Beispiele linearer Gleichungssysteme in Anwendung

Wir geben jetzt einige Beispiele von Gleichungssystemen, die zum Teil aus Anwendungen kommen. Wir werden diese nicht in der Vorlesung besprechen. Sie sollen dem interessierten Leser die große praktische Bedeutung linearer Gleichungssysteme illustrieren.

Beispiel 33.25 (Wie alt ist der Vater?).
Ein Vater hat einen Sohn und eine Tochter. Der Vater ist viermal so alt wie sein Sohn und der Sohn ist fünf Jahre älter als seine Schwester. In fünf Jahren sind Vater und Sohn zusammen sechsmal so alt wie die Tochter.

Wie alt sind Vater, Sohn und Tochter?

Das lineare Gleichungssystem mit v = Alter des Vaters, s = Alter des Sohnes, und t = Alter der Tochter lautet:

$$v = 4s,\ s = t+5,\ (v+5)+(s+5) = 6(t+5).$$

Das Gleichungssystem schreiben wir systematisch folgendermaßen auf:

$$\begin{aligned} v - 4s + 0 \cdot t &= 0, \\ 0 \cdot v + s - t &= 5, \\ v + s - 6t &= 20. \end{aligned}$$

Dies ist ein lineares Gleichungssystem in den Unbestimmten v, s, t.

Die Lösung mit Hilfe des Gaußschen Algorithmus geht wie folgt:

$$\left(\begin{array}{ccc|c} 1 & -4 & 0 & 0 \\ 0 & 1 & -1 & 5 \\ 1 & 1 & -6 & 20 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & -4 & 0 & 0 \\ 0 & 1 & -1 & 5 \\ 0 & 5 & -6 & 20 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & -4 & 0 & 0 \\ 0 & 1 & -1 & 5 \\ 0 & 0 & 1 & 5 \end{array}\right)$$
$$\mapsto \left(\begin{array}{ccc|c} 1 & -4 & 0 & 0 \\ 0 & 1 & 0 & 10 \\ 0 & 0 & 1 & 5 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 0 & 0 & 40 \\ 0 & 1 & 0 & 10 \\ 0 & 0 & 1 & 5 \end{array}\right)$$

Als Lösung erhalten wir also: $t = 5,\ s = 10,\ v = 40$, d. h. der Vater ist 40 Jahre alt, sein Sohn zehn und seine Tochter fünf.

Beispiel 33.26 (Schnitt zweier Ebenen).
Wir definieren eine *Ebene* im $\mathbb{R}^3$ als Lösungsmenge einer linearen Gleichung

$$E : a_1x_1 + a_2x_2 + a_3x_3 = b$$

mit $a_1, a_2, a_3, b \in \mathbb{R}$ und $a_i \neq 0$ für mindestens ein i.

Dies stimmt mit der Anschauung überein (sind alle a_i und b gleich 0, so erhalten wir als Lösungsmenge den ganzen $\mathbb{R}^3$, sind alle $a_i = 0$ und $b \neq 0$, so ist die Lösungsmenge leer).

Um den Schnitt der beiden Ebenen, die durch die Gleichungen $E_1 : x_1 + x_2 + 2x_3 = 2$ und $E_2 : x_1 + x_3 = 4$ gegeben sind, zu bestimmen, müssen wir also das Gleichungssystem aus diesen beiden Gleichungen lösen, wobei wir wie in Abschnitt A) beschrieben vorgehen:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 2 \\ 1 & 0 & 1 & 4 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 1 & -2 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & -1 & 0 \end{array}\right).$$

Wir erhalten als Lösungsmenge

$$E_1 \cap E_2 = \begin{pmatrix} 4 \\ -2 \\ 0 \end{pmatrix} + \mathbb{R} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Dies ist offensichtlich die Parameterdarstellung einer Geraden im $\mathbb{R}^3$ durch die Punkte $(4, -2, 0)^t$ und $(5, -1, -1)^t$.

Beispiel 33.27 (Schnitt zweier Ebenen).
Im allgemeinen werden sich zwei Ebenen, E_1, E_2, im $\mathbb{R}^3$ in einer Geraden schneiden, in Spezialfällen können die Ebenen aber parallel sein ($E_1 \cap E_2 = \emptyset$) oder übereinstimmen ($E_1 = E_2$).

Sei E_1 die Ebene

$$E_1 : x_1 + x_2 + 2x_2 = 3$$

und E_2 eine beliebige Ebene

$$E_2 : a_1x_1 + a_2x_2 + a_3x_3 = b.$$

Wir wollen feststellen für welche a_1, a_2, a_3, b entweder $E_1 \cap E_2$ eine Gerade, leer oder E_1 ist:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ a_1 & a_2 & a_3 & b \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & a_2 - a_1 & a_3 - 2a_1 & b - 3a_1 \end{array}\right).$$

Die letzte Gleichung lautet

$$(a_2 - a_1)x_2 + (a_3 - 2a_1)x_3 = b - 3a_1.$$

Ein wenig Überlegung liefert (da die Lösungsmenge der ersten Gleichung E_1 ist, und da die Lösungsmenge der zweiten Gleichung unabhängig von x_1 ist):

$$E_1 \cap E_2 = \emptyset \quad \Leftrightarrow \quad a_2 - a_1 = a_3 - 2a_1 = 0, (b - 3a_1) \neq 0, \tag{86}$$

$$E_1 = E_2 \quad \Leftrightarrow \quad a_2 - a_1 = a_3 - 2a_1 = b - 3a_1 = 0. \tag{87}$$

In allen anderen Fällen ist $E_1 \cap E_2$ eine Gerade.

Im Fall $E_1 = E_2$ haben wir wieder ein Gleichungssystem (87) mit drei Gleichungen in den vier Unbestimmten a_1, a_2, a_3, b zu lösen:

$$\left(\begin{array}{cccc|c} -1 & 1 & 0 & 0 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ -3 & 0 & 0 & 1 & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} -1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & -3 & 0 & 1 & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} -1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & -\frac{3}{2} & 1 & 0 \end{array}\right)$$

$$\mapsto \left(\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 1 & 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \end{array}\right)$$

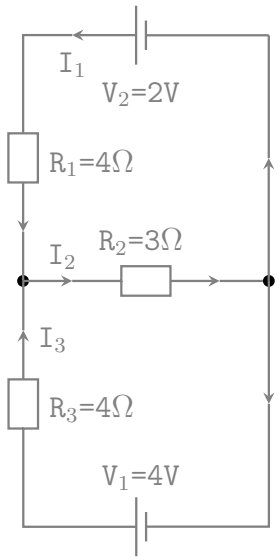
Als Lösung ergibt sich $a_1 = -\frac{b}{3}$, $a_2 = \frac{b}{3}$ und $a_3 = \frac{2b}{3}$, oder kurz

$$(a_1, a_2, a_3, b) = t \cdot (-1, 1, 2, 3)$$

mit $t \in \mathbb{R}$ beliebig. Daraus können wir aber alle drei Fälle ablesen:

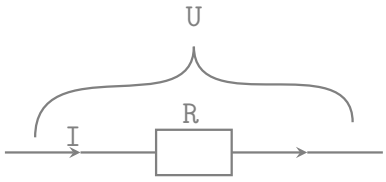
$E_1 = E_2$ genau dann, wenn die Gleichung von E_2 ein Vielfaches $\neq 0$ der Gleichung von E_1 ist; $E_1 \cap E_2 = \emptyset$ genau dann, wenn der Koeffizientenvektor (a_1, a_2, a_3) ein Vielfaches $\neq 0$ des Koeffizientenvektors von E_1 ist, aber die rechte Seite b von E_2 nicht das gleiche Vielfache der rechten Seite von E_1 ist; und $E_1 \cap E_2$ ist eine Gerade in allen anderen Fällen.

Beispiel 33.28 (Elektrische Netzwerke).
In einem einfachen elektrischen Netzwerk, wie z. B.

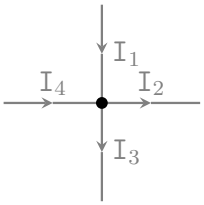


bezeichnet man mit U die Spannung, mit I den Strom und mit R den Widerstand, gemessen in Volt (V), Ampere (A) und Ohm (Ω) respektive. Dabei gelten folgende Gesetze:

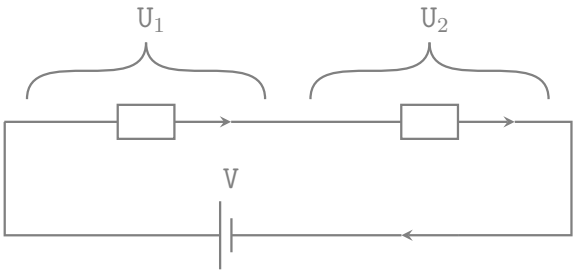
- *Ohmsches Gesetz*: Der Spannungsabfall über einen Widerstand ist das Produkt von Widerstand und Strom, $U=R\cdot I$.



- *1. Kirchhoffsches Gesetz (Knotengleichung):* Die Summe der in einen Knoten hineinfließenden Ströme ist gleich der Summe der hinausfließenden Ströme. Beispiel: $I_1+I_4=I_2+I_3$



- *2. Kirchhoffsches Gesetz (Maschengleichung):* Die Summe der Spannungsverluste in einem geschlossenen Kreis ist gleich der Gesamtspannung in einem Kreis. Beispiel: $V=U_1+U_2$



Im obigen Beispiel stellt man mit Hilfe der drei Gesetze das folgende lineare Gleichungssystem auf:

$$\begin{aligned} I_1 + I_3 &= I_2, && \text{(Knotengleichung)} \\ 4I_1 + 3I_2 &= 2, && \text{(1. Maschengleichung)} \\ 4I_3 + 3I_2 &= 4. && \text{(2. Maschengleichung)} \end{aligned}$$

Wir erhalten das folgende Gleichungssystem:

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 4 & 3 & 0 & 2 \\ 0 & 3 & 4 & 4 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 7 & -4 & 2 \\ 0 & 3 & 4 & 4 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 7 & -4 & 2 \\ 0 & 0 & 40 & 22 \end{array}\right)$$

$$\mapsto \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & -\frac{4}{7} & \frac{2}{7} \\ 0 & 0 & 1 & \frac{11}{20} \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{20} \\ 0 & 1 & 0 & \frac{3}{5} \\ 0 & 0 & 1 & \frac{11}{20} \end{array}\right)$$

woraus sich die folgende Lösung ergibt:

$$I_3 = \tfrac{11}{20}, I_2 = \tfrac{3}{5} \text{ und } I_1 = \tfrac{1}{20}.$$

Beispiel 33.29 (Kubische Splines).
Im "Computer aided geometric design" (CAGD) werden zum Design von Flächen und Kurven (z. B. im Automobil- oder Flugzeugbau) Flächen- und Kurvenstücke verwendet (meist durch sogenannte kubische Splines realisiert), die dann an den Endpunkten oder Randkurven glatt zusammenpassen müssen. Am bekanntesten sind die Bézier-Kubiken, die von dem französischen Auto-Designer bei Renault, P. Bézier, eingeführt wurden (diese werden heute z. B. auch in der Text-Beschreibungssprache PostScript verwendet).

Ein typisches Problem ist z.B. die Bestimmung einer kubischen Parabel

$$f(x) = ax^3 + bx^2 + cx + d$$

durch zwei Punkte (x_1, y_1), (x_2, y_2) in der Ebene mit vorgegebener Steigung m_1 in (x_1, y_1) und m_2 in (x_2, y_2).

Für $(x_1, y_1) = (0, 2)$, $(x_2, y_2) = (4, 0)$, $m_1 = -3$, $m_2 = -3$ ergibt sich aus

$$f'(x) = 3ax^2 + 2bx + c$$

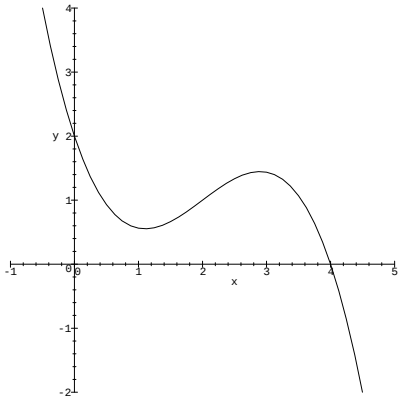
und

$$f(0) = 2, f(4) = 0, f'(0) = -3 \text{ und } f'(4) = -3$$

das lineare Gleichungssystem

$$\begin{aligned} d &= 2, \\ 64a + 16b + 4c + d &= 0, \\ c &= -3, \\ 48a + 8b + c &= -3, \end{aligned}$$

also $d = 2$, $c = -3$, $6a + b = 0$, $32a + 8b = 5$, und damit $a = -\frac{5}{16}$ und $b = \frac{15}{8}$. Die Kurve $y = -\frac{5}{16}x^3 + \frac{15}{8}x^2 - 3x + 2$ hat etwa die folgende Gestalt



Die Aufgabe ist, wie leicht zu sehen ist, stets lösbar und daher können kubische Splines stückweise definiert und glatt aneinander gesetzt werden.

Beispiel 33.30 (Leontieff-Modell).
Die folgende Planungsaufgabe zeigt, daß durchaus Gleichungen mit vielen Veränderlichen in der Praxis auftauchen.

Ein Konzern besitzt n Fabriken $F_1, \ldots, F_n$, in der Fabrik F_i wird das Produkt P_i hergestellt.

Zur Produktion einer Einheit von P_k werden a_{jk} Einheiten von P_j benötigt; wir nehmen an $a_{ii} = 0$.

Am Ende eines Produktionszyklus sind x_k Einheiten von P_k hergestellt, $k = 1, \ldots, n$; wir haben also einen Produktionsvektor $x = (x_1, \ldots, x_n)$.

Zur Herstellung von $x = (x_1, \ldots, x_n)$ werden

$$\sum_{k=1}^{n} a_{jk}x_k = a_{j1}x_1 + \cdots + a_{jn}x_n$$

Einheiten von P_j verbraucht.

Für den Markt verbleiben damit

$$y_j = x_j - \sum_{k=1}^{n} a_{jk}x_k$$

Einheiten von P_j.

Die Planungsaufgabe lautet nun:

Der Mehrbedarf $y = (y_1, \dots, y_n)$ ist vorgegeben. Gesucht ist ein Produktionsvektor $x = (x_1, \dots, x_n)$ mit

$$\begin{aligned} x_1 - (a_{11}x_1 + \cdots + a_{1n}x_n) &= y_1 \\ \vdots & \quad \vdots \\ x_n - (a_{n1}x_1 + \cdots + a_{nn}x_n) &= y_n. \end{aligned}$$

Also ist ein lineares Gleichungssystem zu lösen. Allerdings, und das macht das Problem schwerer, ist zu beachten, daß alle $x_i \geq 0$ sein müssen (natürlich sind auch die y_j und die $a_{j_k} \geq 0$).

(Das Modell heißt Leontieff-Modell und ist nach Vassili Leontieff benannt, der 1973 den Nobelpreis für Wirtschaftswissenschaften erhielt.)

Ein einfaches Beispiel mit zwei Fabriken, Verbrauchsmatrix

$$\begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{3} & 0 \end{pmatrix}$$

und zunächst unbestimmtem Mehrbedarf (y_1, y_2) liefert das Gleichungssystem

$$\left(\begin{array}{cc|c} 1 & -\frac{1}{2} & y_1 \\ -\frac{1}{3} & 1 & y_2 \end{array}\right) \mapsto \left(\begin{array}{cc|c} 1 & -\frac{1}{2} & y_1 \\ 0 & \frac{5}{6} & \frac{1}{3}y_1 + y_2 \end{array}\right) \mapsto \left(\begin{array}{cc|c} 1 & -\frac{1}{2} & y_1 \\ 0 & 1 & \frac{2}{5}y_1 + \frac{6}{5}y_2 \end{array}\right)$$
$$\mapsto \left(\begin{array}{cc|c} 1 & 0 & \frac{6}{5}y_1 + \frac{3}{5}y_2 \\ 0 & 1 & \frac{2}{5}y_1 + \frac{6}{5}y_2 \end{array}\right)$$

und damit $x_1 = \frac{6}{5}y_1 + \frac{3}{5}y_2$, $x_2 = \frac{2}{5}y_1 + \frac{6}{5}y_2$.

Beispiel 33.31 (Finde ein Gleichungssystem zu gegebener Lösung.).
Ein Gleichungssystem besitze die spezielle Lösung $(1, 0, 1)^t$ und das zugehörige homogene System besitze $(1, 1, 1)^t$ als Lösung und habe den Rang zwei. Finde ein Gleichungssystem, das diese Bedingungen erfüllt.

Da die Lösungen Vektoren im $\mathbb{R}^3$ sind, ist es ein System in drei Variablen.

Da der Rang zwei ist, hat die Zeilen-Stufen-Form zwei Zeilen. Da die Lösungsmenge nicht von der Form abhängt, können wir das System in Zeilen-Stufen-Form annehmen:

Problem: Finde eine Gerade im $\mathbb{R}^3$, die selbst durch $(1, 0, 1)^t$ geht und für die die in den Nullpunkt verschobene Gerade durch $(1, 1, 1)^t$ geht.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a_{22}x_2 + a_{23}x_3 &= b_2. \end{aligned}$$

$(1,0,1)^t$ ist Lösung:

$$a_{11} + a_{13} = b_1, \tag{1}$$

$$a_{23} = b_2. \tag{2}$$

$(1,1,1)^t$ ist Lösung des homogenen Systems:

$$a_{11} + a_{12} + a_{13} = 0, \tag{3}$$

$$a_{22} + a_{23} = 0. \tag{4}$$

Das zugehörige lineare Gleichungssystem in $a_{11}, a_{12}, , a_{13}, a_{22}, a_{23}, b_1, b_2$ lautet:

$$\begin{array}{c} \\ (1) \\ (3) \\ (4) \\ (2) \end{array} \begin{array}{c} \begin{array}{ccccccc} a_{11} & a_{12} & a_{13} & a_{22} & a_{23} & b_1 & b_2 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix} \end{array} \mapsto \begin{array}{c} \begin{array}{ccccccc} a_{11} & a_{12} & a_{13} & a_{22} & a_{23} & b_1 & b_2 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix} \end{array}.$$

Das System hat unendlich viele Lösungen, und da der Rang 2 sein soll, muß $a_{22} \neq 0$ und damit auch $a_{23} = -a_{22} \neq 0$ sein.

Wir wählen

$$a_{22} = 1 \Rightarrow a_{23} = b_2 = -1,$$
$$a_{12} = 1 \Rightarrow b_1 = -1,$$
$$a_{11} = 1 \Rightarrow a_{13} = -2.$$

Also ist

$$x_1 + x_2 - 2x_3 = -1,$$
$$x_2 - x_3 = -1$$

ein geeignetes Gleichungssystem.

Aufgaben

Aufgabe 33.32.
Berechne die Lösungsmenge des folgenden Gleichungssystems, sofern es lösbar ist:

$$\begin{array}{rcrcrcr} -x & + & 6y & + & 2z & = & 4 \\ 2x & - & 2y & - & z & = & 2 \\ 3x & - & 4y & - & 2z & = & 1 \end{array}$$

Aufgabe 33.33.
Für welche $a, b \in \mathbb{R}$ besitzt das lineare Gleichungssystem

$$\begin{array}{rcrcrcr} ax & + & & & z & = & ab \\ -2x & + & by & + & az & = & -b \\ & & by & + & (a+1)z & = & b \end{array}$$

außer $(b, 1, 0)$ noch weitere Lösungen. Bestimme sie.

Aufgabe 33.34.
Bestimme die Lösungsmenge des folgenden linearen Gleichungssystems über $\mathbb{R}$ in Abhängigkeit vom Parameter $t \in \mathbb{R}$:

$$\begin{array}{rcrcrcr} x & + & y & + & z & = & 1 \\ & & ty & + & z & = & 1 \\ tx & + & ty & + & z & = & 1+t \end{array}$$

Aufgabe 33.35.
Bestimme eine Basis des Kerns und des Bildes von f_A mit

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ -1 & -1 & 0 & 1 & 1 \\ -1 & -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & -1 & 0 \end{pmatrix} \in \mathrm{Mat}_5(\mathbb{R}).$$

Aufgabe 33.36.
Es sei $U = \langle (1,2,3,4)^t, (1,1,1,1)^t \rangle \leq \mathbb{R}^4$. Bestimme mit Hilfe des Austauschsatzes von Steinitz eine Basis von $\mathbb{R}^4/U$.

Aufgabe 33.37.
Es sei $U = \{(x+y, y, y-x)^t \mid x, y \in \mathbb{R}\}$ und $U' = \{(x, y, z)^t \in \mathbb{R}^3 \mid z = 2x + y\}$. Bestimme Basen von $U + U'$, $U \cap U'$, $\mathbb{R}^3/U$ und $\mathbb{R}^3/U'$.

Aufgabe 33.38.
Bestimme eine Basis für $U \cap U'$ mit

$$U = \langle (2,-1,1,-1)^t, (1,-2,2,1)^t, (3,-1,0,2)^t \rangle \leq \mathbb{R}^4$$

und

$$U' = \langle (3,-2,3,8)^t, (2,1,-5,3)^t \rangle \leq \mathbb{R}^4.$$

Aufgabe 33.39.
Wir betrachten

$$B = \big((1,1,1,1)^t, (-1,0,0,1)^t, (0,-1,0,1)^t, (0,0,-1,1)^t\big)$$

und

$$D = \big((1,1,0)^t, (0,1,1)^t, (0,0,1)^t\big).$$

a. Zeige, dass B eine Basis des $\mathbb{R}^4$ und D eine Basis des $\mathbb{R}^3$ ist.

b. Bestimme $M_D^B(f)$ für $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3 : (x_1, x_2, x_3, x_4)^t \mapsto (x_1 - x_2, x_3, x_2 + x_4)^t$.

c. Bestimme umgekehrt die Funktionsvorschrift für $g \in \mathrm{Hom}_{\mathbb{R}}\left(\mathbb{R}^4, \mathbb{R}^3\right)$ mit

$$M_D^B(g) = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 2 & 1 & 3 & 1 \\ 0 & -1 & 2 & 0 \end{pmatrix}.$$

Aufgabe 33.40.
Es sei V ein $\mathbb{R}$-Vektorraum, $B = (x_1, x_2, x_3)$ eine Basis von V und $B' = (y_1, y_2, y_3)$ mit $y_1 = x_1 + x_3$, $y_2 = x_1 + x_2$ und $y_3 = x_1 + x_2 + x_3$.

a. Zeige, dass B' eine Basis von V ist.

b. Bestimme $M_{B'}^{B'}(f)$, wobei $f \in \mathrm{End}_{\mathbb{R}}(V)$ gegeben ist durch

$$M_B^B(f) = \begin{pmatrix} a & 0 & b \\ -b & a & a \\ a & b & b \end{pmatrix} \text{ mit } a, b \in \mathbb{R}.$$

Aufgabe 33.41.
Seien $B = \big((1,1,1)^t, (1,1,0)^t, (1,0,-1)^t\big)$ und $B' = \big((2,1)^t, (1,1)^t\big)$. E bzw. E'

seien die kanonischen Basen des $\mathbb{R}^3$ bzw. des $\mathbb{R}^2$. Ferner sei $f \in \mathrm{Hom}_K\left(\mathbb{R}^3, \mathbb{R}^2\right)$ gegeben durch $f\left((x,y,z)^t\right) = (x-y+z, 2x+y)^t$.

a. Zeige, dass B und B' Basen des $\mathbb{R}^3$ bzw. des $\mathbb{R}^2$ sind.

b. Bestimme $M^E_{E'}(f)$.

c. Bestimme $M^B_{B'}(f)$ sowie die Transformationsmatrizen T^B_E und $T^{E'}_{B'}$ mit $T^{E'}_{B'} \cdot M^E_{E'}(f) \cdot T^B_E = M^B_{B'}(f)$.

§ 34 Die Determinante

Wir werden jetzt eine ganz neue Möglichkeit kennenlernen, um quadratische lineare Gleichungssysteme zu lösen, nämlich mit Hilfe von Determinanten. Die Determinante ordnet einer quadratischen Matrix über einem Körper ein Element des Körpers zu, das genau dann ungleich Null ist, wenn die Matrix invertierbar ist. Die Determinante liefert aber nicht nur ein nützliches Kriterium für die Invertierbarkeit, sie ist vor allem aus theoretischen Gründen von unschätzbarem Wert. Z. B. liefert die Cramersche Regel mit Hilfe der Determinante eine geschlossene Formel für die Lösung eines linearen Gleichungssystems. Aus dieser Formel lassen sich Eigenschaften der Lösungen als Funktionen der Koeffizienten der Matrix bestimmen.

Die Determinante einer Matrix ist eine *polynomiale Funktion* in den Einträgen der Matrix. Sind diese Einträge etwa reelle oder komplexe Zahlen, so hängt die Determinante stetig von den Einträgen ab. Daraus folgt z. B. die wichtige Tatsache, daß eine invertierbare Matrix bei kleiner Störung der Einträge invertierbar bleibt. Damit wird eine Verbindung zur Analysis hergestellt. Eine weitere wichtige Bedeutung in der Analysis hat die Determinante für die Volumenberechnung (siehe auch Bemerkung 36.30).

Wir werden die Eigenschaften der Determinante soweit entwickeln, wie sie in der linearen Algebra wichtig sind. Allerdings führt uns die Determinante auch hier schon auf eine höhere Stufe: die Determinante ist nicht nur linear, sie ist *multilinear*, wie wir gleich sehen werden.

A) Die Leibniz-Formel für die Determinante

Definition 34.1 (Determinante).
Wir definieren für $A \in \mathrm{Mat}_n(K)$ die *Determinante* von A durch die *Leibniz-Formel*

$$\det(A) := |A| := \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}. \tag{88}$$

Beispiel 34.2 (Determinanten für $n = 1, 2, 3$).

a. Ist $n = 1$, dann ist $A = (a) \in \mathrm{Mat}(1, K)$ und $\det(A) = a$.

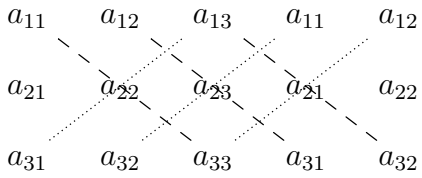
b. Ist $n = 2$, dann ist $\mathbb{S}_2 = \{\mathrm{id}, (1\ 2)\}$ und damit folgt:

$$\det(A) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

d. h. $\det(A)$ ist das Produkt der Elemente der Hauptdiagonalen minus dem Produkt der Elemente der Gegendiagonalen. Z.B.

$$\det\begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix} = 5 \cdot 3 - 6 \cdot 2 = 3.$$

c. Für $n = 3$ hat $\mathbb{S}_n$ bereits sechs Elemente. Man berechnet in diesem Fall die Determinante mit der *Regel von Sarrus*:



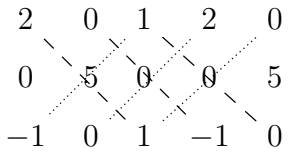
Die Produkte der Elemente längs der gestrichelten Linien tauchen bei der Berechnung der Determinante als positive Summanden auf, die Produkte der Elemente längs der gepunkteten Linien als negative Summanden. D. h., wir erhalten:

$$\begin{aligned} \det(A) \;=\; & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & -a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{aligned}$$

Wenden wir das obige Schema auf die Matrix

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 5 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

an, so erhalten wir



und damit

$$\det(A) = 2 \cdot 5 \cdot 1 - (-1) \cdot 5 \cdot 1 = 15.$$

d. Für $n = 4$ ergeben sich schon $4! = 24$ Summanden und für $n = 10$ gar $10! = 3628800$. In numerischen Anwendungen sind 1000×1000-Matrizen keine Seltenheit, so daß es sich von selbst versteht, daß dabei nicht die Definition, bei der dann für die Determinante über 10^{2567} Produkte berechnet werden müßten, zur Berechnung verwendet werden kann. In der Tat wird zur Berechnung von Determinanten über Körpern wieder der Gauß-Algorithmus eine wichtige Rolle spielen.

Haben Matrizen eine spezielle Form, so ist es unter Umständen leichter, ihre Determinante zu berechnen. Sehr einfach ist dies für obere und untere Dreiecksmatrizen.

Proposition 34.3 (Determinanten von Dreiecksmatrizen).
Ist $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ eine obere (bzw. untere) *Dreiecksmatrix*, d. h. $a_{ij} = 0$ für $i > j$ (bzw. $i < j$), dann ist

$$\det(A) = a_{11} \cdots a_{nn}$$

das Produkt der Diagonalelemente.

Beweis: Ist $\mathrm{id} \neq \sigma \in \mathbb{S}_n$, so gilt $i > \sigma(i)$ (bzw. $i < \sigma(i)$) für mindestens ein i. Wegen der Voraussetzung $a_{i\sigma(i)} = 0$ für $i > \sigma(i)$ (bzw. $i < \sigma(i)$) bleibt von den Summanden in (88) also nur der für id übrig. □

Beispiel 34.4.

$$\det \begin{pmatrix} 2 & 3 & 12 & -3 \\ 0 & -1 & 5 & -111 \\ 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 2 \cdot (-1) \cdot 3 \cdot 1 = -6.$$

Lemma 34.5 (Alternative Leibniz-Formel).
Für die Determinante von $A \in \mathrm{Mat}_n(K)$ gilt

$$\det(A) = \sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) \cdot a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}. \tag{89}$$

Beweis: Man beachte, daß für $\sigma \in \mathbb{S}_n$ auch σ^{-1} eine Permutation der Zahlen $1, \ldots, n$ ist, d. h. $\{1, \ldots, n\} = \{\sigma^{-1}(1), \ldots, \sigma^{-1}(n)\}$. Zudem wissen wir aus Satz A2.7 e., daß $\mathrm{sgn}(\sigma) = \mathrm{sgn}\left(\sigma^{-1}\right)$, und es ist gleich, ob wir über $\sigma \in \mathbb{S}_n$ summieren oder über $\sigma^{-1} \in \mathbb{S}_n$, da auf beide Weisen alle Elemente von $\mathbb{S}_n$ je einmal erreicht werden. Aus diesen Vorbetrachtungen ergibt sich:

$$\begin{aligned}
\det(A) &= \textstyle\sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) \cdot a_{\sigma^{-1}(1)\sigma(\sigma^{-1}(1))} \cdot \ldots \cdot a_{\sigma^{-1}(n)\sigma(\sigma^{-1}(n))} \\
&= \textstyle\sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) \cdot a_{\sigma^{-1}(1)1} \cdot \ldots \cdot a_{\sigma^{-1}(n)n} \\
&\overset{\text{A2.7e.}}{=} \textstyle\sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}\left(\sigma^{-1}\right) \cdot a_{\sigma^{-1}(1)1} \cdot \ldots \cdot a_{\sigma^{-1}(n)n} \\
&= \textstyle\sum_{\sigma^{-1} \in \mathbb{S}_n} \mathrm{sgn}\left(\sigma^{-1}\right) \cdot a_{\sigma^{-1}(1)1} \cdot \ldots \cdot a_{\sigma^{-1}(n)n} \\
&= \textstyle\sum_{\pi \in \mathbb{S}_n} \mathrm{sgn}(\pi) \cdot a_{\pi(1)1} \cdot \ldots \cdot a_{\pi(n)n}.
\end{aligned}$$

□

Proposition 34.6 (Die Determinante der Transponierten).
Für $A \in \mathrm{Mat}_n(K)$ gilt:

$$\det(A) = \det\left(A^t\right).$$

Beweis: Sei $A = (a_{ij})$ und $A^t = (a'_{ij})$, dann gilt $a'_{ij} = a_{ji}$. Mithin erhalten wir mit Hilfe von Lemma 34.5

$$\begin{aligned} \det(A) &= \sum_{\sigma\in\mathbb{S}_n} \mathrm{sgn}(\sigma)\cdot a_{1\sigma(1)}\cdot\ldots\cdot a_{n\sigma(n)} \\ &= \sum_{\sigma\in\mathbb{S}_n} \mathrm{sgn}(\sigma)\cdot a'_{\sigma(1)1}\cdot\ldots\cdot a'_{\sigma(n)n} \quad = \quad \det\left(A^t\right). \end{aligned}$$

□

Beispiel 34.7.
Beispiel 34.2 b. aufgreifend gilt

$$\det\begin{pmatrix} 5 & 6 \\ 2 & 3 \end{pmatrix} = \det\begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix} = 3.$$

B) Die Determinante als Volumenform

Definition 34.8 (Multilineare Abbildungen).
Es seien V und W zwei K-Vektorräume.

a. Eine Abbildung

$$f : V^n = V\times\overset{n}{\ldots}\times V \to W$$

heißt *multilinear*, falls f in jedem Argument linear ist, d. h. es gelten

$$f(x_1,\ldots,x_i+y_i,\ldots,x_n) = f(x_1,\ldots,x_i,\ldots,x_n) + f(x_1,\ldots,y_i,\ldots,x_n)$$

und

$$f(x_1,\ldots,\lambda x_i,\ldots,x_n) = \lambda\cdot f(x_1,\ldots,x_i,\ldots,x_n)$$

für jedes $i \in \{1,\ldots,n\}$ und für alle $x_1,\ldots,x_n,y_i \in V$ und $\lambda \in K$.

b. Eine multilineare Abbildung $f : V^n \to W$ heißt *alternierend*, falls für $(x_1,\ldots,x_n) \in V^n$ mit $x_i = x_j$ für ein $i \neq j$, gilt:

$$f(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_n) = 0.$$

Lemma 34.9.
Ist $f : V^n \to W$ eine alternierende multilineare Abbildung, dann gilt für $\sigma \in \mathbb{S}_n$

$$f(x_{\sigma(1)},\ldots,x_{\sigma(n)}) = \mathrm{sgn}(\sigma)\cdot f(x_1,\ldots,x_n).$$

Insbesondere gilt $f(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_n) = -f(x_1,\ldots,x_j,\ldots,x_i,\ldots,x_n)$.

Beweis: Wir betrachten zunächst den Fall, daß $\sigma = (i\ j)$ eine Transposition ist. Da f alternierend und multilinear ist, folgt die Behauptung für σ aus

$$\begin{aligned} 0 &= f(x_1, \ldots, x_i + x_j, \ldots, x_i + x_j, \ldots, x_n) \\ &= f(x_1, \ldots, x_i, \ldots, x_i, \ldots, x_n) + f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) \\ &\quad + f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n) + f(x_1, \ldots, x_j, \ldots, x_j, \ldots, x_n) \\ &= f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) + f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n). \end{aligned}$$

Ist $\sigma \in \mathbb{S}_n$ beliebig, so können wir $\sigma = \tau_1 \circ \ldots \circ \tau_k$ als Produkt von Transpositionen schreiben und die Behauptung folgt mittels Induktion nach der Anzahl k der Transpositionen. Den Induktionsanfang $k = 1$ haben wir bereits gezeigt. Ist $k \geq 2$ und setzen wir $\pi = \tau_2 \circ \ldots \tau_k$, so folgt mit der Vorüberlegung

$$\begin{aligned} f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) &= f(x_{\tau_1(\pi(1))}, \ldots, x_{\tau_1(\pi(n))}) = -f(x_{\pi(1)}, \ldots, x_{\pi(n)}) \\ &\overset{Ind.}{=} -\operatorname{sgn}(\pi) \cdot f(x_1, \ldots, x_n) \overset{Satz\ A2.7}{=} \operatorname{sgn}(\tau_1 \circ \pi) \cdot f(x_1, \ldots, x_n) \\ &= \operatorname{sgn}(\sigma) \cdot f(x_1, \ldots, x_n). \end{aligned}$$

□

Bemerkung 34.10.
Wir können $\mathrm{Mat}_n(K)$ auf recht natürliche Weise mit $K^n \times \overset{n}{\ldots} \times K^n$ identifizieren, indem wir eine Matrix $A = (a_{ij})$ mit dem n-Tupel ihrer Spaltenvektoren $(a^1, \ldots, a^n)$ gleichsetzen. Das wollen wir im folgenden tun.

Satz 34.11 (Die Determinante als Volumenform).

a. Die Determinante

$$\det : \mathrm{Mat}_n(K) \to K : A \mapsto \det(A)$$

ist eine alternierende multilineare Abbildung mit $\det(\mathbb{1}_n) = 1$.

b. Ist $f : \mathrm{Mat}_n(K) \longrightarrow K$ eine alternierende multilineare Abbildung und $A \in \mathrm{Mat}_n(K)$, so gilt

$$f(A) = f(\mathbb{1}_n) \cdot \det(A).$$

Beweis:

a. Wir werden im Beweis die Formel (89) aus Lemma 34.5 zur Berechnung der Determinante verwenden, da sie auf die Bedürfnisse der Determinante als multilineare Abbildung bezüglich der Spalten zugeschnitten ist.
Es seien $a^j = (a_{1j}, \ldots, a_{nj})^t$, $j = 1, \ldots, n$, und $b^i = (b_{1i}, \ldots, b_{ni})^t$. Wir setzen $A := (a^1 \ldots a^i \ldots a^n)$, $B := (a^1 \ldots b^i \ldots a^n)$ und $C := (a^1 \ldots \lambda a^i + \mu b^i \ldots a^n)$. Dann

gilt

$$\begin{aligned}\det(C) &= \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}(\sigma)\cdot a_{\sigma(1)1}\cdot\ldots\cdot(\lambda a_{\sigma(i)i}+\mu b_{\sigma(i)i})\cdot\ldots\cdot a_{\sigma(n)n}\\ &= \lambda\cdot\sum_{\sigma\in\mathbb{S}_n}\operatorname{sgn}(\sigma)\cdot a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(i)i}\cdot\ldots\cdot a_{\sigma(n)n}\\ &\quad +\mu\cdot\sum_{\sigma\in\mathbb{S}_n}\operatorname{sgn}(\sigma)\cdot a_{\sigma(1)1}\cdot\ldots\cdot b_{\sigma(i)i}\cdot\ldots\cdot a_{\sigma(n)n}\\ &= \lambda\cdot\det(A)+\mu\cdot\det(B),\end{aligned}$$

so daß det multilinear ist.
Sei nun $a^i = a^j$, für ein $i \neq j$. Ist $\tau = (i\ j)$, die Transposition, die i und j vertauscht, dann besitzt $\mathbb{S}_n$ nach Satz A2.7 die Zerlegung $\mathbb{S}_n = \mathbb{A}_n \cup \mathbb{A}_n\tau$. Ferner gilt für $\sigma \in \mathbb{A}_n$

$$\operatorname{sgn}(\sigma) = 1 \quad \text{und} \quad \operatorname{sgn}(\sigma\tau) = -1.$$

Wir erhalten also

$$\begin{aligned}\det(A) &= \sum_{\sigma\in\mathbb{A}_n} a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(i)i}\cdot\ldots\cdot a_{\sigma(j)j}\cdot\ldots\cdot a_{\sigma(n)n}\\ &\quad -\sum_{\sigma\in\mathbb{A}_n} a_{\sigma\tau(1)1}\cdot\ldots\cdot a_{\sigma\tau(i)i}\cdot\ldots\cdot a_{\sigma\tau(j)j}\cdot\ldots\cdot a_{\sigma\tau(n)n}\\ &= \sum_{\sigma\in\mathbb{A}_n} a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(i)i}\cdot\ldots\cdot a_{\sigma(j)j}\cdot\ldots\cdot a_{\sigma(n)n}\\ &\quad -\sum_{\sigma\in\mathbb{A}_n} a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(j)i}\cdot\ldots\cdot a_{\sigma(i)j}\cdot\ldots\cdot a_{\sigma(n)n} = 0,\end{aligned}$$

und somit ist det alternierend.
Außerdem folgt $\det(\mathbb{1}_n) = 1\cdot\ldots\cdot 1 = 1$ aus Proposition 34.3.

b. Mit den Notationen von a. gilt $a^i = \sum_{j=1}^n a_{ji}e^j$, wenn e^j der j-te Einheitsvektor ist. Aus der Multilinearität von f folgt:

$$f(A) = \sum_{j_1=1}^n a_{j_1 1} f\left(e^{j_1}\ a^2\ldots a^n\right) = \sum_{j_1=1}^n a_{j_1 1}\sum_{j_2=1}^n a_{j_2 2} f\left(e^{j_1}\ e^{j_2}\ a^3\ldots a^n\right)$$

$$= \ldots = \sum_{j_1,\ldots,j_n=1}^n a_{j_1 1}\cdot\ldots\cdot a_{j_n n} f\left(e^{j_1}\ \ldots\ e^{j_n}\right).$$

Genau dann, wenn die $j_1,\ldots,j_n$ paarweise verschieden sind, existiert eine Permutation $\sigma \in \mathbb{S}_n$ mit $\left(e^{j_1}\ldots e^{j_n}\right) = \left(e^{\sigma(1)}\ldots e^{\sigma(n)}\right)$, und wegen Lemma 34.9 gilt dann

$$f(e^{\sigma(1)}\ldots e^{\sigma(n)}) = \operatorname{sgn}(\sigma)\cdot f(e^1\ldots e^n) = \operatorname{sgn}(\sigma)\cdot f(\mathbb{1}_n).$$

Andernfalls stimmen zwei der j_i überein und $f\left(e^{j_1}\ \ldots\ e^{j_n}\right) = 0$, da f alternierend ist. Insgesamt haben wir damit gezeigt:

$$f(A) = \sum_{\sigma\in\mathbb{S}_n} a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(n)n}\cdot\operatorname{sgn}(\sigma)\cdot f(\mathbb{1}_n) = \det(A)\cdot f(\mathbb{1}_n).$$

□

Bemerkung 34.12 (Das Volumen des Parallelotops).
Eine alternierende multilineare Abbildung $f : \mathrm{Mat}_n(K) \longrightarrow K$ wird auch eine *Volumenform* genannt. Aus Satz 34.11 b. folgt, daß die Determinante die einzige Volumenform f mit $f(\mathbb{1}_n) = 1$ ist, d.h. det ist durch die Eigenschaften in Satz 34.11 a. eindeutig bestimmt.

Die Determinante hat eine wichtige geometrische Interpretation, die den Begriff *Volumenform* rechtfertigt. Seien $x_1, \ldots, x_n \in \mathbb{R}^n$ und sei

$$P(x_1, \ldots, x_n) := \{\lambda_1 x_1 + \ldots + \lambda_n x_n \in \mathbb{R}^n \mid 0 \leq \lambda_i \leq 1, i = 1, \ldots, n\}$$

das von den Vektoren $x_1, \ldots, x_n$ aufgespannte *Parallelotop* (siehe Abbildung 3).

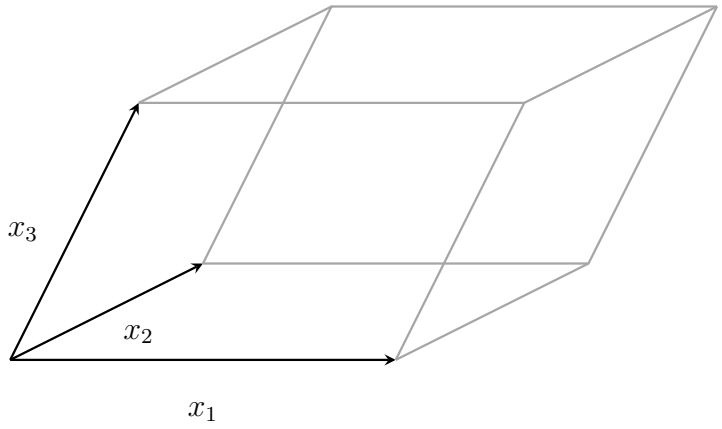


Abbildung 3: Das Parallelotop $P(x_1, x_2, x_3)$ im $\mathbb{R}^3$

Dann definiert man das n-dimensionale Volumen von $P(x_1, \ldots, x_n)$ mit Hilfe der Determinante als

$$\text{Volumen}\big(P(x_1, \ldots, x_n)\big) = |\det(x_1 \ldots x_n)|.$$

In Dimension $n = 1$ ist $|\det(x_1)| = |x_1|$ in der Tat die Länge der Strecke von 0 nach x_1, und diese ist gerade $P(x_1)$. Wir werden in Bemerkung 36.30 zeigen, daß auch in Dimension $n = 2$ und $n = 3$ das so definierte Volumen mit dem euklidischen Flächeninhalt bzw. mit dem euklidischen Volumen übereinstimmt, daß die Definition also sinnvoll ist. Sie wird im Rahmen der mehrdimensionalen Integrationstheorie und der Verallgemeinerung der Substitutionsregel eine wichtige Rolle spielen. □

C) Der Gauß-Algorithmus zur Berechnung der Determinante

Wir haben oben gesehen, daß die Leibniz-Formel ungeeignet ist zur Berechnung der Determinante einer $n \times n$-Matrix für großes n, weil die Symmetrische Gruppe dann viel zu viele Elmenete enthält. Wir wollen deshalb nun zeigen, wie die Determinante effizient mit Hilfe des Gauß-Algorithmus' berechnet werden kann.

Korollar 34.13 (Spaltenoperationen und die Determinante).
Es sei $A \in \mathrm{Mat}_n(K)$ und $\lambda \in K$.

a. Bei Vertauschung zweier Spalten von A ändert sich das Vorzeichen von $\det(A)$.

b. Bei Multiplikation einer Spalte von A mit λ multipliziert sich $\det(A)$ mit λ.

c. Bei Addition des λ-fachen einer Spalte zu einer anderen Spalte ändert sich $\det(A)$ nicht.

d. Enthält A eine Nullspalte, so ist $\det(A) = 0$.

e. Sind zwei Spalten von A gleich, so ist $\det(A) = 0$.

Beweis:

a. Das ist ein Spezialfall von Lemma 34.9, da det nach Satz 34.11 alternierend ist.

b. Dies folgt aus der Multilinearität von det, siehe Satz 34.11.

c. Für $A = \begin{pmatrix} a^1 \dots a^n \end{pmatrix}$ und $A' = \begin{pmatrix} a^1 \dots a^j + \lambda a^i \dots a^n \end{pmatrix}$ folgt aus der Multilinearität und da det alternierend ist:

$$\det(A') = \det(A) + \lambda \cdot \det(a^1 \dots a^i \dots a^i \dots a^n) = \det(A) + \lambda \cdot 0 = \det(A).$$

d. Ist eine Spalte von A Null, so folgt $\det(A) = 0$ aus b. mit $\lambda = 0$.

e. Das folgt, da det alternierend ist.

□

Da die Determinante einer Matrix gleich der Determinante der Transponierten ist, sind die Begriffe Spalte und Zeile austauschbar. Eine exaktere Formulierung bietet das folgende Korollar.

Korollar 34.14 (Zeilenoperationen und die Determinante).
Wir können $\det : \mathrm{Mat}_n(K) \to K$ auch als multilineare Abbildung auf den Zeilen einer Matrix A auffassen. Entsprechend gilt Korollar 34.13 auch für Zeilen statt Spalten.

Da sich die Determinante bei der Addition eines Vielfachen einer Zeile zu einer anderen nicht ändert, können wir den Gauß-Algorithmus zur Berechnung von Determinanten einsetzen. Dabei müssen wir uns aber merken, wie oft wir Zeilen getauscht haben und wir sollten möglichst auf das Mulitplizieren von Zeilen mit Skalaren verzichten.

Algorithmus 34.15 (Algorithmus zur Berechnung der Determinante).
INPUT: $A \in \mathrm{Mat}_n(K)$.
OUTPUT: $\det(A)$.

1. **Schritt:** Setze $d = 1$.
2. **Schritt:** Überführe A mittels Gauß-Algorithmus in nicht-reduzierte ZSF, d. h. führe im Gauß-Algorithmus 32.10 Schritt sieben nicht aus. Jedesmal, wenn dabei zwei Zeilen vertauscht werden, ersetze d durch $-d$. - Wird bei der Gaußreduktion ein Pivotelement zu Null, gib Null zurück und brich ab.
3. **Schritt:** Gib das Produkt von d mit den Diagonalelementen der ZSF zurück.

Beispiel 34.16.
Wir wollen die Determinante der Matrix

$$A = \begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 0 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$$

berechnen. Dazu überführen wir sie mittels des Gauß-Algorithmus in ZSF und merken uns die Zeilenvertauschungen.

$$\begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 0 \end{pmatrix} \xrightarrow[d=-1]{I \leftrightarrow II} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix} \xrightarrow[III \to III - 7I]{II \to II - 4I} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -21 \end{pmatrix}$$

$$\xrightarrow{III \to III - 2II} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & -9 \end{pmatrix}$$

Damit gilt dann

$$\det(A) = d \cdot 1 \cdot (-3) \cdot (-9) = -27.$$

Beispiel 34.17.
Sei $A \in \mathrm{Mat}(n+1, K)$ gegeben durch

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 & \dots & n \\ 1 & 0 & 1 & 2 & \dots & n-1 \\ 2 & 1 & 0 & 1 & \dots & n-2 \\ 3 & 2 & 1 & 0 & \dots & n-3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ n & n-1 & n-2 & n-3 & \dots & 0 \end{pmatrix}.$$

Ziehe für $i = 1, \ldots, n$ von der i-ten Zeile die $(i+1)$-te Zeile ab. Wir erhalten:

$$A' = \begin{pmatrix} -1 & 1 & 1 & 1 & \ldots & 1 & 1 \\ -1 & -1 & 1 & 1 & \ldots & 1 & 1 \\ -1 & -1 & -1 & 1 & \ldots & 1 & 1 \\ -1 & -1 & -1 & -1 & \ldots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & -1 & \ldots & -1 & 1 \\ n & n-1 & n-2 & n-3 & \ldots & 1 & 0 \end{pmatrix}.$$

Addiere nun für $i = 2, \ldots, n+1$ die erste Spalte zur i-ten Spalte. Dann erhalten wir:

$$A'' = \begin{pmatrix} -1 & 0 & 0 & 0 & \ldots & 0 & 0 \\ * & -2 & 0 & 0 & \ldots & 0 & 0 \\ * & * & -2 & 0 & \ldots & 0 & 0 \\ * & * & * & -2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ & * & * & * & \ldots & -2 & 0 \\ * & * & * & * & \ldots & * & n \end{pmatrix}.$$

Es folgt:

$$\det(A) = \det(A'') = (-1) \cdot (-2)^{n-1} \cdot n = -n \cdot (-2)^{n-1}.$$

Bemerkung 34.18.
In Beispiel 34.17 haben wir durch ganz wenige Zeilen- und Spaltenoperationen die Matrix in Dreiecksgestalt überführt. Das lag aber an der speziellen Struktur der Matrix. Im allgemeinen Fall braucht der oben beschriebene Algorithmus zur Berechnung der Determinante mit Hilfe des Gauß-Algorithmus $\sim \frac{n^3}{3}$ Multiplikationen für eine $n \times n$-Matrix. In der Definition der Determinante tauchen dagegen $n!$ Summanden von je n Produkten auf, mit $n! \sim \left(\frac{n}{e}\right)^n$, wobei e die Eulersche Zahl ist. Man sagt, daß der Gauß-Algorithmus *polynomial*, die Definition aber *exponentiell* in der Größe der Matrix ist. Grundsätzlich gelten polynomiale Algorithmen als effizient, exponentielle dagegen als unakzeptabel ineffizient. Allerdings gibt es Fälle, wo keine polynomialen Algorithmen bekannt sind.

D) Der Determinantenmultiplikationssatz

Satz 34.19 (Determinantenmultiplikationssatz).
Für Matrizen $A, B \in \mathrm{Mat}_n(K)$ gilt

$$\det(A \circ B) = \det(A) \cdot \det(B).$$

Beweis: Wähle $A \in \mathrm{Mat}_n(K)$ fest und betrachte die Abbildung

$$f : \mathrm{Mat}_n(K) \to K : B \mapsto \det(A \circ B).$$

f ist multilinear bezüglich der Spalten von B, da A auf jede Spalte von B linear wirkt. Außerdem ist f alternierend, da mit B auch $A \circ B$ zwei gleiche Spalten hat. Damit folgt aus Satz 34.11:

$$\det(A \circ B) = f(B) = f(\mathbb{1}_n) \cdot \det(B) = \det(A) \cdot \det(B).$$

□

Beispiel 34.20.
In Beispiel 34.2 b. gilt

$$A = \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}),$$

so folgt aus dem Determinantenmultiplikationssatz 34.19 und weil die beiden Matrizen auf der rechten Seite Dreiecksmatrizen sind:

$$\det \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix} = \det \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \cdot \det \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = 3 \cdot 1 = 3.$$

Das folgende Korollar ist eine Verallgemeinerung der Aussage in Aufgabe 27.14.

Korollar 34.21 (Determinante und Invertierbarkeit).
Genau dann ist $A \in \mathrm{Mat}_n(K)$ invertierbar, wenn $\det(A) \neq 0$. In diesem Fall gilt

$$\det\left(A^{-1}\right) = \frac{1}{\det(A)}.$$

Beweis: Ist A invertierbar, so gilt

$$1 = \det(\mathbb{1}_n) = \det\left(A \circ A^{-1}\right) = \det(A) \cdot \det\left(A^{-1}\right).$$

Dies zeigt, daß $\det(A)$ nicht Null sein kann, und zudem ist damit die obige Formel bewiesen.

Ist A nicht invertierbar, so sind die Spalten von A linear abhängig und durch mehrfache Addition von Vielfachen bestimmter Spalten zu einer anderen können wir eine Nullspalte erzeugen. Nach Korollar 34.13 c. ändert sich dabei der Wert der Determinante nicht, und nach Korollar 34.13 d. muß er somit 0 sein. □

Beispiel 34.22.
Die Matrix A in Beispiel 34.20 ist invertierbar und ihre Inverse hat Determinante $\frac{1}{3}$.

Dies wissen wir, ohne die Inverse auszurechnen. Diese können wir mit Hilfe von Aufgabe 27.14 berechnen. Für eine invertierbare 2×2-Matrix gilt

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_n(K)$$

gilt

$$B^{-1} = \frac{1}{ad-bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{\det(B)} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

so daß wir im Beispiel

$$A^{-1} = \frac{1}{3} \cdot \begin{pmatrix} 3 & -2 \\ -6 & 5 \end{pmatrix}$$

erhalten.

Bemerkung 34.23 (det ist ein Gruppenepimorphismus.).
In der Sprache der Vorlesung Algebraische Strukturen folgt aus Satz 34.19 und Korollar 34.21, daß

$$\det : \big(\mathrm{Gl}_n(K), \circ\big) \to \big(K^*, \cdot\big)$$

ein Gruppenepimorphismus ist. Dazu beachte man, daß det surjektiv ist wegen

$$\det \left(\begin{array}{c|ccc} \lambda & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & \mathbb{1}_{n-1} & \\ 0 & & & \end{array}\right) = \lambda.$$

E) Der Kästchensatz

Satz 34.24 (Kästchensatz).
Es sei $A \in \mathrm{Mat}_n(K)$ eine *Blockmatrix* der Form

$$A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right)$$

mit $B \in \mathrm{Mat}_k(K)$, $C \in \mathrm{Mat}(k \times l, K)$, $D \in \mathrm{Mat}_l(K)$, $0 \in \mathrm{Mat}(l \times k, K)$ und $n = k + l$. Dann gilt:

$$\det(A) = \det(B) \cdot \det(D).$$

Beweis: Man beachte, daß

$$A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right) = \left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D \end{array}\right) \circ \left(\begin{array}{c|c} B & C \\ \hline 0 & \mathbb{1}_l \end{array}\right).$$

Wegen des Determinantenmultiplikationssatzes 34.19 reicht es mithin zu zeigen:

$$\det\left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D \end{array}\right) = \det(D) \tag{90}$$

und

$$\det\left(\begin{array}{c|c} B & C \\ \hline 0 & \mathbb{1}_l \end{array}\right) = \det(B). \tag{91}$$

Die Abbildung

$$f : \mathrm{Mat}_l(K) \to K : D' \mapsto \det\left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D' \end{array}\right)$$

ist offensichtlich multilinear und alternierend, und wegen Satz 34.11 b. gilt mithin

$$\det\left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D \end{array}\right) = f(D) = f(\mathbb{1}_l) \cdot \det(D) = \det(\mathbb{1}_n) \cdot \det(D) = \det(D),$$

d. h. (90) ist erfüllt.

Analog ist die Abbildung

$$g : \mathrm{Mat}_k(K) \to K : B' \mapsto \det\left(\begin{array}{c|c} B' & C \\ \hline 0 & \mathbb{1}_l \end{array}\right)$$

alternierend und multilinear in den Spalten von B', also eine Volumenform. Wieder folgt aus Satz 34.11 mit Hilfe von Proposition 34.3, daß

$$\det\left(\begin{array}{c|c} B & C \\ \hline 0 & \mathbb{1}_l \end{array}\right) = g(B) = g(\mathbb{1}_k) \cdot \det(B) = \det\left(\begin{array}{c|c} \mathbb{1}_k & C \\ \hline 0 & \mathbb{1}_l \end{array}\right) \cdot \det(B) = \det(B),$$

womit auch (91) gezeigt ist. □

Beispiel 34.25 (Vandermonde-Determinante).
Wir wollen mit Hilfe des Gauß-Algorithmus zeigen, daß

$$det\begin{pmatrix} 1 & a_0 & a_0^2 & a_0^3 & \dots & a_0^n \\ 1 & a_1 & a_1^2 & a_1^3 & \dots & a_1^n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & a_n^3 & \dots & a_n^n \end{pmatrix} = \prod_{0 \le i < j \le n} (a_j - a_i)$$

für $a_0, \dots, a_n \in K$ und $n \ge 1$ gilt. Die Determinate dieser Matrix ist als *Vandermonde-Determinante* bekannt.

Beweis: Wir beweisen die Aussage mit Hilfe von Induktion nach n. Für $n = 1$ ist

$$\det(A) = \begin{vmatrix} 1 & a_0 \\ 1 & a_1 \end{vmatrix} = a_1 - a_0$$

und die Aussage stimmt. Sei also $n > 1$ und die Aussage sei für Matrizen dieser Gestalt der Größe n (beachte, daß A die Größe $n + 1$ hat) bereits gezeigt. Addieren wir für

$j = n+1, \ldots, 2$ zur j-ten Spalte das $-a_0$-fache der $j-1$-ten Spalte, so ändert sich die Determinante nicht und wir erhalten

$$\det(A) = \begin{vmatrix} 1 & a_0 & a_0^2 & a_0^3 & \ldots & a_0^n \\ 1 & a_1 & a_1^2 & a_1^3 & \ldots & a_1^n \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & a_n^3 & \ldots & a_n^n \end{vmatrix}$$

$$= \left|\begin{array}{c|cccc} 1 & 0 & 0 & \ldots & 0 \\ \hline 1 & a_1 - a_0 & a_1^2 - a_1 a_0 & \ldots & a_1^n - a_1^{n-1} a_0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n - a_0 & a_n^2 - a_n a_0 & \ldots & a_n^n - a_n^{n-1} a_0 \end{array}\right|$$

Aufgrund des Kästchensatzes und wegen $\det(1) = 1$ gilt dann

$$\det(A) = \begin{vmatrix} a_1 - a_0 & a_1^2 - a_1 a_0 & \ldots & a_1^n - a_1^{n-1} a_0 \\ \vdots & \vdots & & \vdots \\ a_n - a_0 & a_n^2 - a_n a_0 & \ldots & a_n^n - a_n^{n-1} a_0 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 - a_0 & (a_1 - a_0) \cdot a_1 & \ldots & (a_1 - a_0) \cdot a_1^{n-1} \\ \vdots & \vdots & & \vdots \\ a_n - a_0 & (a_n - a_0) \cdot a_n & \ldots & (a_n - a_0) \cdot a_n^{n-1} \end{vmatrix}$$

Klammern wir nun in der i-ten Zeile $a_i - a_0$ aus, so erhalten wir

$$\det(A) = \prod_{i=1}^{n} (a_i - a_0) \cdot \begin{vmatrix} 1 & a_1 & a_1^2 & a_1^3 & \ldots & a_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & a_n^3 & \ldots & a_n^{n-1} \end{vmatrix}.$$

Auf die letzte Determinante können wir Induktion anwenden und erhalten

$$\det(A) = \prod_{i=1}^{n} (a_i - a_0) \cdot \prod_{1 \le i < j \le n} (a_j - a_i) = \prod_{0 \le i < j \le n} (a_j - a_i).$$

□

F) Der Satz über die Adjunkte

Unser nächstes Ziel ist die Herleitung einer alternativen Berechnung der Determinante, die im Gegensatz zum Gaußalgorithmus ohne Division auskommt und deshalb über jedem kommutativen Ring mit Eins funktioniert (siehe Bemerkung 34.37). Zu ihrer Herleitung führen wir zunächst verschiedene Hilfmatrizen ein.

Definition 34.26.
Es sei $A = (a_{ij}) = (a^1 \dots a^n) \in \mathrm{Mat}_n(K)$, $n \geq 2$, und $b = (b_1, \dots, b_n)^t \in K^n$.
Wir definieren die *Ersetzungsmatrix*

$$A_i(b) := (a^1 \dots a^{i-1}\ b\ a^{i+1} \dots a^n),$$

in der die i-te Spalte von A durch b ersetzt wurde.
Ist $b = e_j$ der j-te Einheitsvektor, so gilt:

$$A_i(e_j) = \begin{pmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{j1} & \dots & 1 & \dots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{pmatrix}.$$

Ersetzen wir in $A_i(e_j)$ zusätzlich noch die j-te Zeile durch den i-ten Einheitsvektor, dann erhält man die Matrix

$$S_{ji}(A) = \begin{pmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{pmatrix}.$$

Streicht man in der Matrix A die j-te Zeile und die i-te Spalte, so erhält man die *Streichungsmatrix*

$$A_{ji} = \left(\begin{array}{ccc|ccc} a_{11} & \dots & a_{1\,i-1} & a_{1\,i+1} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & \\ a_{j-1\,1} & \dots & a_{j-1\,i-1} & a_{j-1\,i+1} & \dots & a_{j-1\,n} \\ \hline a_{j+1\,1} & \dots & a_{j+1\,i-1} & a_{j+1\,i+1} & \dots & a_{j+1\,n} \\ \vdots & & \vdots & & \vdots & \\ a_{n\,1} & \dots & a_{n\,i-1} & a_{n\,i+1} & \dots & a_{n\,n} \end{array}\right).$$

Lemma 34.27.
Für $A \in \mathrm{Mat}_n(K)$, $n \geq 2$, $1 \leq i, j \leq n$, gilt:

$$\det\Big(A_i(e_j)\Big) = \det\big(S_{ji}(A)\big) = (-1)^{i+j} \det(A_{ji}).$$

Beweis: $S_{ji}(A)$ entsteht aus $A_i(e_j)$ durch Subtraktion des a_{jk}-fachen der i-ten Spalte von der k-ten Spalte, $k \in \{1, \dots, n\} \setminus \{i\}$. Also gilt nach Korollar 34.13:

$$\det\Big(A_i(e_j)\Big) = \det\big(S_{ji}(A)\big).$$

Durch $i-1$ Spaltenvertauschungen und $j-1$ Zeilenvertauschungen entsteht aus $S_{ji}(A)$ die Matrix

$$\left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & A_{ji} & \\ 0 & & & \end{array}\right).$$

Also folgt aus dem Kästchensatz 34.24 unter Beachtung der Korollare 34.13 und 34.14

$$\det\big(S_{ji}(A)\big) = (-1)^{i+j}\det(A_{ji}).$$

□

In der folgenden Definition beachte man die Vertauschung der Indizes!

Definition 34.28.
Für $A \in \mathrm{Mat}_n(K)$, $n \geq 2$, $1 \leq i, j \leq n$ heißt

$$a_{ij}^{\#} := (-1)^{i+j}\det(A_{ji})$$

ein *Kofaktor* von A. Die Matrix der Kofaktoren

$$A^{\#} := \big(a_{ij}^{\#}\big) \in \mathrm{Mat}_n(K)$$

heißt die *Adjunkte* oder *Komplementärmatrix* von A.

Satz 34.29 (Satz über die Adjunkte).
Für $A \in \mathrm{Mat}_n(K)$, $n \geq 2$, gilt:

$$A^{\#} \circ A = A \circ A^{\#} = \det(A) \cdot \mathbb{1}_n.$$

Beweis: Sei $A^{\#} \circ A = (c_{ik})$. Dann gilt mit Lemma 34.27:

$$c_{ik} = \sum_{j=1}^{n} a_{ij}^{\#} \cdot a_{jk} = \sum_{j=1}^{n} a_{jk} \cdot \det\big(a^1 \dots a^{i-1}\ e_j\ a^{i+1} \dots a^n\big)$$

$$= \det\left(a^1 \dots a^{i-1} \sum_{j=1}^{n} a_{jk} e_j\ a^{i+1} \dots a^n\right)$$

$$= \det\big(a^1 \dots a^{i-1}\ a^k\ a^{i+1} \dots a^n\big) = \delta_{ik} \cdot \det(A),$$

wobei δ_{ik} das Kronecker-Symbol ist. Das dritte Gleichheitszeichen folgt aus der Multilinearität von det, das letzte, da det alternierend ist.

Der Beweis, daß $A \circ A^{\#} = \det(A) \cdot \mathbb{1}_n$ geht analog. □

Korollar 34.30.
Es sei $A \in \mathrm{Mat}_n(K)$ invertierbar, so ist

$$A^{-1} = \frac{1}{\det(A)} \cdot A^{\#}.$$

Wir wollen an dieser Stelle einmal die vielen Aussagen, die zur Invertierbarkeit einer quadratischen Matrix über einem Körper äquivalent sind, sammeln.

Korollar 34.31.
Für eine Matrix $A \in \mathrm{Mat}(n, K)$ sind gleichwertig:

a. A ist invertierbar.
b. $\mathrm{rang}(A) = n$.
c. $\det(A) \neq 0$.
d. f_A ist bijektiv.
e. f_A ist injektiv.
f. f_A ist surjektiv.
g. $\mathrm{rZSF}(A) = \mathbb{1}_n$.
h. A ist das Produkt endlich vieler Elementarmatrizen.
i. Es gibt eine Matrix $B \in \mathrm{Mat}(n, K)$ mit $A \circ B = \mathbb{1}_n$.

Beweis: Die unterschiedlichen Äquivalenzen sind in den Sätzen 29.29, 30.23, 30.24, 31.23, 32.14, 33.2 und 34.21 gezeigt worden. □

Beispiel 34.32.
Für eine 2×2-Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

gilt $\det(A) = ad - bc$ und

$$A^{\#} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Ist also $ad - bc \neq 0$, so gilt:

$$A^{-1} = \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Damit ist Aufgabe 27.14 bewiesen.

Sei nun konkret $K = \mathbb{Q}$ und

$$A = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix}.$$

Dann ist $\det(A) = 1$ und somit gilt

$$A^{-1} = \begin{pmatrix} 4 & -1 \\ -3 & 1 \end{pmatrix}.$$

G) Laplacescher Entwicklungssatz und Cramersche Regel

Der Satz über die Adjunkte führt zu einer rekursiven Berechnungsformel für die Determinante, die für theoretische Überlegungen sehr nützlich ist. Sie ist auch als rekursive Prozedur sehr einfach zu programmieren, aber nicht sehr effizient. Sie hat die gleiche Komplexität, wie die Leibnizsche Formel (88) zur Definition der Determinante.

Satz 34.33 (Laplacescher Entwicklungssatz).
Es sei $A \in \mathrm{Mat}_n(K)$.

a. Wir nennen die folgende Formel, die *Entwicklung nach der i-ten Zeile*:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}). \tag{92}$$

b. Entsprechend nennen wir die folgende Formel, die *Entwicklung nach der j-ten Spalte*:

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}). \tag{93}$$

Beweis: Nach Satz 34.29 gilt für $A \circ A^{\#} = (c_{ik})$

$$\det(A) = c_{ii} = \sum_{j=1}^{n} a_{ij} \cdot a_{ji}^{\#} = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}).$$

Damit folgt (92), und (93) zeigt man analog durch die Betrachtung von $A^{\#} \circ A$. □

Bemerkung 34.34.
Entwickelt man $A = (a_{ij})$ nach der ersten Zeile, so gilt:

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = a_{11} \cdot \begin{vmatrix} a_{22} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \dots & a_{nn} \end{vmatrix} - a_{12} \cdot \begin{vmatrix} a_{21} & a_{23} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n3} & \dots & a_{nn} \end{vmatrix}$$
$$+ \dots + (-1)^{n+1} a_{1n} \cdot \begin{vmatrix} a_{21} & \dots & a_{2\,n-1} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{n\,n-1} \end{vmatrix}.$$

Benutzt man dieses Verfahren, so entwickelt man am Besten nach Zeilen bzw. Spalten, die möglichst viele Nullen enthalten. Die Vorzeichen merkt man sich am Günstigsten mit der sogenannten *Schachbrettregel*:

$$\begin{vmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ - & + & - & + & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{vmatrix}$$

Für kleine Matrizen, insbesondere wenn die Matrix dünn besetzt ist, ist dieses Verfahren zur Berechnung der Determinante (und zur Berechnung der Inversen) durchaus anwendbar. Für größere Matrizen ist auf jeden Fall der Gaußsche Eliminationsalgorithmus vorzuziehen.

Wir berechnen nun die Determinante der Matrix

$$A = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 3 & 4 \\ 2 & 5 & 3 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$$

mit Hilfe der Entwicklung nach der ersten Zeile. Dann gilt

$$\begin{aligned} \det(A) =& 0 \cdot \det(A_{11}) - 2 \cdot \det(A_{12}) + 0 \cdot \det(A_{13}) \\ =& -2 \cdot \det \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} = (-2) \cdot (-5) = 10. \end{aligned}$$

Wir leiten nun aus dem Laplaceschen Entwicklungssatz einen Algorithmus ab.

Algorithmus 34.35 (Laplace-Entwicklung).
INPUT: $A \in \mathrm{Mat}_n(K)$.
OUTPUT: $\det(A)$.

1. **Schritt:** Initialisiere det auf Null.
2. **Schritt:** Falls $n = 1$, setze det $= a_{11}$ und gehe zu Schritt 3. Sonst tue für $i = 1, \ldots, n$:
   - Bilde eine Hilfsmatrix B durch Streichen der ersten Spalte und der i-ten Zeile von A.
   - Rufe den Algorithmus mit B auf und merke Dir das Ergebnis in einer Hilfsvariablen x.
   - Addiere zu det die Zahl $(-1)^{i+1} \cdot a_{i1} \cdot x$.
3. **Schritt:** Gib det zurück.

Der Satz über die Adjunkte liefert auch eine für theoretische Überlegungen sehr wichtige geschlossene Formel für die Lösungen eines linearen Gleichungssystems. Dies ist die berühmte *Cramersche Regel*.

Satz 34.36 (Cramersche Regel).
Es sei $A \in \mathrm{Mat}_n(K)$ invertierbar und $b \in K^n$.
Für die eindeutig bestimmte Lösung $x = (x_1, \ldots, x_n)^t \in K^n$ von $Ax = b$ gilt dann

$$\begin{aligned} x_i &= \tfrac{1}{\det(A)} \cdot \det\big(A_i(b)\big) \\ &= \tfrac{1}{\det(A)} \cdot \det \begin{pmatrix} a_{11} & \ldots & a_{1\,i-1} & b_1 & a_{1\,i+1} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n\,i-1} & b_n & a_{n\,i+1} & \ldots & a_{nn} \end{pmatrix}. \end{aligned}$$

Beweis: Wegen Korollar 34.30 ist

$$x = A^{-1}b = \frac{1}{\det(A)} \cdot A^{\#}b$$

die eindeutig bestimmte Lösung des linearen Gleichungssystems. Also folgt mit Lemma 34.27 und der Multilinearität der Determinante

$$\begin{aligned} x_i &= \tfrac{1}{\det(A)} \cdot \sum_{j=1}^{n} a_{ij}^{\#} \cdot b_j = \tfrac{1}{\det(A)} \cdot \sum_{j=1}^{n} \det\big(A_i(e_j)\big) \cdot b_j \\ &= \tfrac{1}{\det(A)} \cdot \sum_{j=1}^{n} \det\big(a^1 \ldots a^{i-1}\ e_j\ a^{i+1} \ldots a^n\big) \cdot b_j \\ &= \tfrac{1}{\det(A)} \cdot \det\big(a^1 \ldots a^{i-1}\ b\ a^{i+1} \ldots a^n\big) \\ &= \tfrac{1}{\det(A)} \cdot \det\big(A_i(b)\big). \end{aligned}$$

□

Bemerkung 34.37 (Determinanten über kommutativen Ringen mit Eins). Ist K nur ein kommutativer Ring mit Eins, so können wir die Determinante einer Matrix in $\mathrm{Mat}_n(K)$ ebenfalls durch die Leibniz-Formel definieren, und alle Aussagen dieses Abschnitts, die *ohne Division* auskommen, gelten mit dem gleichen Beweis.

Wir können den Gauß-Algorithmus 34.15 über beliebigen Ringen in der angegebenen Form *nicht* mehr anwenden, da dabei Divisionen nötig sind. Außerdem gilt Korollar 34.21 *nicht* mehr in der angegebenen Form, und ebenso gilt Korollar 34.31 in *nicht* in vollem Umfang.

Alle anderen Aussagen gelten jedoch ohne jede Änderung. Dies trifft insbesondere auf den Satz zur Adjunkten 34.29 zu, den wir später für Matrizen mit Koeffizienten in einem Polynomring anwenden wollen. Außerdem können wir den Laplaceschen Entwicklungssatz im Gegensatz zum Gaußschen Algorithmus über jedem kommutativen Ring mit Eins anwenden, um die Determinante auszurechnen. Es gibt aber auch hier geschicktere Verfahren, indem man den Gaußschen Algorithmus abwandelt zum sogenannten Bareiss Algorithmus (siehe [Coh96]).

Für die Aussage in Korollar 34.30 beachte man, daß aus dem Determinantenmultiplikationssatz 34.19 und dem Satz zur Adjunkten 34.29 unmittelbar folgt, daß eine quadratische Matrix über einem kommutativen Ring genau dann invertierbar ist, wenn $\det(A)$ invertierbar ist. In diesem Fall darf man dann auch in dem Ring durch $\det(A)$ teilen. Das trifft auf Korollar 34.30 ebenso zu wie auf die Cramersche Regel 34.36.

Betrachten wir konkret den Ring $\mathbb{Z}$ der ganzen Zahlen, dann sind nur 1 und -1 invertierbar. Mithin sind nur ganzzahlige Matrizen mit Determinante 1 oder -1 über $\mathbb{Z}$ invertierbar, d.h. nur für solche enthält die Inverse wieder nur ganze Zahlen. Ein Beispiel dafür haben wir in Beispiel 34.32 gesehen. Betrachten wir stattdessen die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$

so gilt $\det(A) = -2 \notin \{1, -1\}$ und die Einträge von

$$A^{-1} = -\frac{1}{2} \cdot \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

sind nicht mehr alle ganzzahlig, obwohl A nur ganzzahlige Einträge hatte. A ist als Matrix in $\mathrm{Mat}_2(\mathbb{Q})$ also invertierbar, als Matrix in $\mathrm{Mat}_2(\mathbb{Z})$ aber nicht.

Aufgaben

Aufgabe 34.38.
Berechne die Determinanten der folgenden Matrizen:

a. $\begin{pmatrix} 1 & -4 \\ -3 & 8 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$.

b. $\begin{pmatrix} -1 & 0 & -1 \\ 0 & 5 & 0 \\ 3 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$.

c. $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$.

Aufgabe 34.39.
Sei K ein Körper und $\lambda \in K$. Bestimme die Determinante der Matrix

$$\begin{pmatrix} 1 & \lambda & \lambda^2 & \dots & \lambda^{n-1} \\ \lambda^{n-1} & 1 & \lambda & \dots & \lambda^{n-2} \\ \lambda^{n-2} & \lambda^{n-1} & 1 & \dots & \lambda^{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda & \lambda^2 & \lambda^3 & \dots & 1 \end{pmatrix} \in \mathrm{Mat}(n \times n, K).$$

Aufgabe 34.40.
Für $n \in \mathbb{N} \setminus \{0\}$ definieren wir

$$A_n = \begin{pmatrix} 1 & 1 & 0 & \dots & \dots & \dots & 0 \\ 1 & 1 & 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 1 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 1 & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 & 1 & 1 \\ 0 & \dots & \dots & \dots & 0 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_n(\mathbb{R})$$

als die Matrix, deren Einträge auf der Diagonalen sowie auf der oberen und unteren Nebendiagonalen alle eins sind, während alle anderen Einträge null sind. Ferner setzen wir $d_n = \det(A_n)$.

a. Zeige, für $n \geq 3$ gilt die Rekursionsformel $d_n = d_{n-1} - d_{n-2}$.

b. Zeige, für $k \in \mathbb{N}$ gilt

$$d_n = \begin{cases} 1, & \text{falls } n \equiv 1 (\text{mod } 6) \text{ oder } n \equiv 0 (\text{mod } 6), \\ 0, & \text{falls } n \equiv 2 (\text{mod } 6) \text{ oder } n \equiv 5 (\text{mod } 6), \\ -1, & \text{falls } n \equiv 3 (\text{mod } 6) \text{ oder } n \equiv 4 (\text{mod } 6). \end{cases}$$

Aufgabe 34.41.

Sei V ein n-dimensionaler $\mathbb{C}$-Vektorraum und $f : V \longrightarrow V$ $\mathbb{C}$-linear. Mittels Einschränkung der Skalarmultiplikation können wir V als $\mathbb{R}$-Vektorraum und f als $\mathbb{R}$-lineare Abbildung auffassen. Des Weiteren bezeichnen wir mit $\det_{\mathbb{C}}(f)$ die Determinante von f als $\mathbb{C}$-lineare Abbildung und $\det_{\mathbb{R}}(f)$ die Determinante von f als $\mathbb{R}$-lineare Abbildung. Zeige:

$$\det_{\mathbb{R}}(f) = |\det_{\mathbb{C}}(f)|^2.$$

Hinweis: Für eine $\mathbb{C}$-Basis $(v_1, \ldots, v_n)$ von V betrachte man die zugehörige $\mathbb{R}$-Basis $(v_1, \ldots, v_n, iv_1, \ldots, iv_n)$ sowie jeweils die zugehörige Matrixdarstellung von f. Wem der allgemeine Fall zu schwer ist, der beschränke sich auf die Abbildung $f : \mathbb{C} \longrightarrow \mathbb{C} : z \longmapsto (a + ib) \cdot z$ mit $a, b \in \mathbb{R}$ fest vorgegeben. Was ist eine Basis von $\mathbb{C}$ als $\mathbb{R}$-Vektorraum?

Aufgabe 34.42.

Berechne die Determinante folgender Matrix mit Hilfe des Gauß-Algorithmus':

$$A = \begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ -2 & 0 & 3 & 1 & 3 \\ 0 & 2 & 5 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 3 & 4 & 1 \end{pmatrix} \in \text{Mat}_5(\mathbb{R}).$$

Aufgabe 34.43.

Berechne die folgende Determinante mit Hilfe des Determinantenentwicklungssatzes:

$$B = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 3 & 1 & 3 \\ 2 & 5 & 0 & -1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \in \text{Mat}_4(\mathbb{Q}).$$

Aufgabe 34.44.

Bestimme für welche $s \in \mathbb{R}$ die Abbildung

$$f_s : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (x, y, z)^t \mapsto (x + z, x + 2y + z, sx + y - z)^t$$

invertierbar ist und berechne für diese die Inverse mit Hilfe der Adjunkten.

Aufgabe 34.45.
Es sei $A \in \mathrm{Mat}_n(\mathbb{R})$ mit ungeradem n und $A^t = -A$. Zeige, A ist nicht invertierbar. Bleibt die Aussage wahr, wenn wir $\mathbb{R}$ durch einen anderen Körper ersetzen?

Aufgabe 34.46.
Zeige, ist $A \in \mathrm{Mat}_n(K)$ eine invertierbare obere Dreiecksmatrix, so ist auch A^{-1} eine obere Dreiecksmatrix.

Aufgabe 34.47.
Löse das folgende lineare Gleichungssystem mit Hilfe der Cramerschen Regel:

$$x + 2z = 3, \quad 3x + y = 5 \quad \text{und} \quad -x + y = 1.$$

§ 35 Endomorphismen und ihre Eigenwerte

In diesem Abschnitt sei V ein K-Vektorraum mit $1 \leq \dim_K(V) = n < \infty$.

A) Invarianten von Endomorphismen unter Konjugation

Bemerkung 35.1 (Endomorphismen).
Wir erinnern uns, daß K-lineare Abbildungen

$$f : V \longrightarrow V$$

auch *Endomorphismen* des K-Vektorraums V genannt werden (siehe Definition 28.19) und daß

$$\mathrm{End}_K(V) = \{f : V \longrightarrow V \mid f \text{ ist } K\text{-linear}\}$$

die *K-Algebra* der Endomorphismen von V ist (siehe Bemerkung 31.9).

Zudem wissen wir, wie sich die Matrixdarstellungen von Endomorphismen unter Basiswechsel verhalten. Sind B und D zwei Basen des Vektorraums V und ist $T = T_B^D$, so gilt (siehe Korollar 31.15)

$$M_D^D(f) = T^{-1} \circ M_B^B(f) \circ T.$$

Dabei ist es von großer Wichtigkeit, daß wir jeweils im Definitions- und Zielbereich von f *dieselbe* Basis verwenden, und das wollen wir von nun an stets tun, wenn wir Matrixdarstellungen von Endomorphismen betrachten!

Wir können deshalb Eigenschaften von Matrizen, die unter Transformationen der Form

$$A \mapsto T^{-1} \circ A \circ T$$

erhalten bleiben, auch für Endomorphismen definieren, indem wir dazu ihre Matrixdarstellungen bezüglich einer beliebigen Basis verwenden. In diesem Abschnitt wollen wir einige Beispiele hierfür kennen lernen.

Definition 35.2 (Konjugiert oder ähnlich).
Zwei quadratische Matrizen $A, B \in \mathrm{Mat}_n(K)$ heißen *konjugiert* oder *ähnlich*, wenn es eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ gibt, so daß $B = T^{-1} \circ A \circ T$ ist.

Bemerkung 35.3 (Konjugation ist eine Äquivalenzrelation).
Konjugation von Matrizen ist ein Beispiel für eine Äquivalenzrelation auf der Menge $\mathrm{Mat}_n(K)$ der quadratischen $n \times n$-Matrizen über K. D.h.

- jede Matrix ist zu sich selbst konjugiert, denn $A = \mathbb{1}_n^{-1} \cdot A \cdot \mathbb{1}_n$;

- ist A zu B konjugiert, so ist auch B zu A konjugiert, da aus $B = T^{-1} \circ A \circ T$ auch $A = (T^{-1})^{-1} \circ B \circ T^{-1}$ folgt;
- ist A zu B und B zu C konjugiert, so ist auch A zu C konjugiert, da aus $B = T^{-1} \circ A \circ T$ und $C = S^{-1} \circ B \circ S$ auch $C = (T \circ S)^{-1} \circ A \circ (T \circ S)$ folgt.

Definition 35.4 (Das Charakteristische Polynom einer Matrix).
Es sei $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ eine quadratische Matrix.

a. Wir nennen
$$\chi_A := \det(t \cdot \mathbb{1}_n - A) \in K[t]$$
das *charakteristische Polynom* von A, wobei
$$t \cdot \mathbb{1}_n - A = \begin{pmatrix} t - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & t - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & t - a_{nn} \end{pmatrix} \in \mathrm{Mat}_n\left(K[t]\right)$$
eine quadratische Matrix mit Polynomen als Einträgen ist.

b. Wir definieren die *Spur* von A als Summe der Diagonalelemente, d.h.
$$\mathrm{Spur}(A) := \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + a_{33} + \ldots + a_{nn}.$$

Beispiel 35.5.
Für die Matrix
$$A = \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix}$$
gilt
$$\chi_A = \det \begin{pmatrix} t-5 & -2 \\ -6 & t-3 \end{pmatrix} = (t-5) \cdot (t-3) - (-2) \cdot (-6) = t^2 - 8t + 3.$$
Man beachte, daß der konstante Term von χ_A gerade $\det(A) = 3$ und daß der Koeffizient von t gerade $-\mathrm{Spur}(A) = -8$ ist.

Proposition 35.6 (Charakteristisches Polynom).
Es sei $A \in \mathrm{Mat}_n(K)$ eine quadratische Matrix. Dann ist das Polynom
$$\chi_A = t^n + \alpha_{n-1} \cdot t^{n-1} + \alpha_{n-2} \cdot t^{n-2} + \ldots + \alpha_1 \cdot t + \alpha_0 \in K[t]$$
ein normiert vom Grad n mit $\alpha_{n-1} = -\mathrm{Spur}(A)$ und $\alpha_0 = (-1)^n \cdot \det(A)$.

Beweis: Ist $A = (a_{ij})$ und $t \cdot \mathbb{1}_n - A = \big(p_{ij}\big)$, dann folgt aus der Leibnitzschen Formel für die Determinante

$$\chi_A = \det(t \cdot \mathbb{1}_n - A) = (t - a_{11}) \cdots (t - a_{nn}) + \sum_{\mathrm{id} \neq \sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) \cdot p_{1\sigma(1)} \cdots p_{n\sigma(n)}.$$

Da für $\sigma \neq \mathrm{id}$ mindestens zwei Faktoren in $p_{1\sigma(1)} \cdots p_{n\sigma(n)}$ konstante Polynome sind, ergibt $\sum_{\mathrm{id} \neq \sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) \cdot p_{1\sigma(1)} \cdots p_{n\sigma(n)}$ ein Polynom vom Grad kleiner gleich $n-2$. Damit lassen sich die Koeffizienten α_n und α_{n-1} von t^n und t^{n-1} in χ_A aus $(t-a_{11}) \cdots (t-a_{nn})$ herleiten und sind wie oben angegeben $\alpha_n = 1$ und

$$\alpha_{n-1} = -a_{11} - a_{22} - \ldots - a_{nn} = -\operatorname{Spur}(A).$$

Ferner ist

$$\alpha_0 = \chi_A(0) = \det(-A) = (-1)^n \cdot \det(A)$$

der konstante Term im charakteristischen Polynom. □

Bemerkung 35.7.
Man beachte, daß es bei der Berechnung von $\chi_A(\lambda)$ für $\lambda \in K$ keinen Unterschied macht, ob wir zuerst t durch λ ersetzen und dann die Leibnitzformel zum Berechnen der Determinante anwenden oder ob wir zuerst die Determinante berechnen und dann t durch λ ersetzen. Das liegt daran, daß der Einsetzhomomorphismus mit der Multiplikation und Addition verträglich ist. Diese Tatsache haben wir im obigen Beweis bei der Berechnung des konstanten Terms des charakteristischen Polynoms verwendet.

Proposition 35.8 (Charakteristisches Polynom konjugierter Matrizen).
Sind $A, B \in \mathrm{Mat}_n(K)$ konjugiert, so gelten

$$\chi_A = \chi_B, \quad \det(A) = \det(B) \quad \text{und} \quad \operatorname{Spur}(A) = \operatorname{Spur}(B).$$

Beweis: Sei $T \in \mathrm{Gl}_n(K)$ mit $B = T^{-1} \circ A \circ T$, dann gilt auch

$$T^{-1} \circ (t \cdot \mathbb{1}_n - A) \circ T = t \cdot T^{-1} \circ \mathbb{1}_n \circ T - T^{-1} \circ A \circ T = t \cdot \mathbb{1}_n - B.$$

Um den Determinantenmultiplikationssatz 34.19 anwenden zu können, betrachten wir die Matrix $t \cdot \mathbb{1}_n - A$ als Matrix mit Einträgen im Körper $K(t)$ der rationalen Funktionen über K. Damit erhalten wir dann

$$\begin{aligned} \chi_B &= \det(t \cdot \mathbb{1}_n - B) = \det\left(T^{-1} \circ (t \cdot \mathbb{1}_n - A) \circ T\right) \\ &= \det(T^{-1}) \cdot \det(t \cdot \mathbb{1}_n - A) \cdot \det\left(T\right) \\ &= \frac{1}{\det(T)} \cdot \det(t \cdot \mathbb{1}_n - A) \cdot \det(T) = \chi_A. \end{aligned}$$

Die Aussagen zur Determinante und Spur folgen dann aus Proposition 35.8. □

Damit können wir das charakteristische Polynom, die Determinante und die Spur eines Endomorphismus definieren.

Definition 35.9 (Charakteristisches Polynom eines Endomorphismus). Sei V ein endlich-dimensionaler K-Vektorraum der Dimension mindestens Eins mit Basis B und $f \in \mathrm{End}_K(V)$. Wir definieren

- das *charakteristische Polynom* von f durch
$$\chi_f := \chi_{M_B^B(f)},$$
- die *Determinante* von f durch
$$\det(f) := \det\left(M_B^B(f)\right)$$
- und die *Spur* von f durch
$$\mathrm{Spur}(f) := \mathrm{Spur}\left(M_B^B(f)\right).$$

Da die Matrixdarstellungen eines Endomorphismus f zu verschiedenen Basen nach Korollar 31.15 konjugiert sind, sind diese Definitionen unter Berücksichtigung von Proposition 35.8 unabhängig von der Wahl der Basis B.

Beispiel 35.10.
Wir betrachten den Endomorphismus

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x, y)^t \mapsto (5x + 2y, 6x + 3y)^t.$$

Ist E die Standardbasis des $\mathbb{R}^2$, so gilt

$$M_E^E(f) = \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix}.$$

Alternativ könnte man die Basis $B = \left((1, 1)^t, (0, 1)^t\right)$ betrachten und erhält dann

$$M_B^B(f) = \begin{pmatrix} 7 & 2 \\ 2 & 1 \end{pmatrix}.$$

Das charakteristische Polynom von f erhalten wir dann als

$$\chi_f = \chi_{M_E^E(f)} = \det \begin{pmatrix} t-5 & -2 \\ -6 & t-3 \end{pmatrix} = t^2 - 8t + 3$$

oder alternativ als

$$\chi_f = \chi_{M_B^B(f)} = \det \begin{pmatrix} t-7 & -2 \\ -2 & t-1 \end{pmatrix} = (t-7) \cdot (t-1) - 4 = t^2 - 8t + 3.$$

B) Eigenwerte, Eigenvektoren und Eigenräume

Der Begriff des Eigenwertes ist von zentraler Bedeutung für die Untersuchung von Endomorphismen und quadratischen Matrizen.

Definition 35.11 (Eigenwerte und Eigenvektoren).
Sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

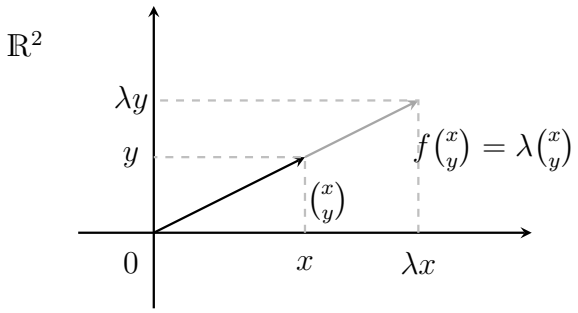
a. $\lambda \in K$ heißt *Eigenwert* von f, falls es ein $0 \neq x \in V$ mit $f(x) = \lambda x$ gib.] Der Vektor x heißt dann *Eigenvektor* zum Eigenwert λ von f, und die Menge

$$\mathrm{Eig}(f, \lambda) := \{y \in V \mid f(y) = \lambda y\}$$

heißt der *Eigenraum* von f zum Eigenwert λ.

b. $\lambda \in K$ heißt *Eigenwert* von A, falls es ein $0 \neq x \in K^n$ mit $Ax = \lambda x$ gibt. Der Vektor x heißt dann *Eigenvektor* zum Eigenwert λ von A, und die Menge $\mathrm{Eig}(A, \lambda) := \{y \in V \mid Ay = \lambda y\}$ heißt *Eigenraum* von A zum Eigenwert λ.

Bemerkung 35.12 (Geometrische Interpretation von Eigenvektoren).
Ist λ Eigenwert von f mit Eigenvektor x, so bedeutet das anschaulich, daß f in *Richtung* von x durch Multiplikation mit λ wirkt. Diese Anschauung liefert im Fall $V = \mathbb{R}^n$ und $\lambda > 0$, daß f den Vektor x um den Faktor λ streckt, falls $\lambda > 1$, und um den Faktor λ staucht, falls $0 < \lambda < 1$.



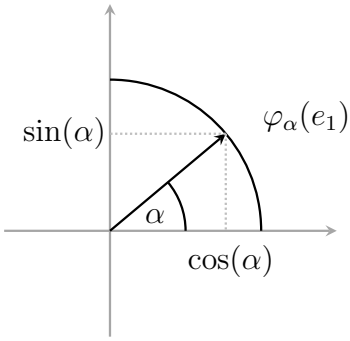
Beispiel 35.13.

a. Ist $\dim_K(V) = 1$, so ist jeder Vektor ungleich Null ein Eigenvektor von f, da f schlicht die Multiplikation mit einer Konstanten ist.

b. Ist $\dim_K(V) \geq 2$, so braucht f hingegen keine Eigenwerte und Eigenvektoren zu besitzen. Dabei hängt die Frage der Existenz wesentlich vom Grundkörper K ab. Betrachte etwa die Drehung $\varphi_\alpha = f_{A_\alpha} : \mathbb{R}^2 \to \mathbb{R}^2$ um den Winkel $\alpha \in \mathbb{R}$ aus Beispiel 28.24. Die Matrixdarstellung bezüglich der kanonischen Basis $E = (e_1, e_2)$

ist

$$A_\alpha = M_E^E(\varphi_\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$



Aus einer rein geometrischen Betrachtung folgt unmittelbar, daß φ_α bzw. A_α nur dann einen Eigenvektor besitzen können, wenn α ein ganzzahliges Vielfaches von π ist.

Bemerkung 35.14 (Eigenräume).
Es seien $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Da $f(x) = \lambda x$ für $x \in V$ und $\lambda \in K$ genau dann erfüllt ist, wenn x im Kern der linearen Abbildung $f - \lambda\,\mathrm{id}_V \in \mathrm{End}_K(V)$ liegt, gilt also

$$\mathrm{Eig}(f,\lambda) = \mathrm{Ker}(f - \lambda\,\mathrm{id}_V) = \mathrm{Ker}(\lambda\,\mathrm{id}_V - f).$$

Analog erhält man:

$$\mathrm{Eig}(A,\lambda) = \mathrm{Ker}(f_A - \lambda\,\mathrm{id}_V) = \mathrm{Lös}(A - \lambda\mathbb{1}_n, 0) = \mathrm{Lös}(\lambda\mathbb{1}_n - A, 0).$$

b. Aus der Definition folgt unmittelbar, daß $\sigma(A) = \sigma(f_A)$ und $\sigma(f) = \sigma\big(M_B^B(f)\big)$.

c. Ebenso folgt unmittelbar, daß der Eigenraum $\mathrm{Eig}(f,\lambda)$ von f zum Eigenwert λ f-invariant ist.

d. Kennt man einen Eigenwert $\lambda \in K$ von A, so kann man das lineare Gleichungssystem

$$(A - \lambda\mathbb{1}_n)x = 0 \quad \text{oder} \quad (\lambda\mathbb{1}_n - A)x = 0$$

lösen und damit eine Basis des Eigenraumes $\mathrm{Eig}(A,\lambda) = \mathrm{Lös}(A - \lambda\mathbb{1}_n, 0)$ bestimmen. D. h., bei Kenntnis des Eigenwertes λ lassen sich die Eigenvektoren von A zu λ durch Lösen eines linearen Gleichungssystems bestimmen. Aber wie kommt man zu den Eigenwerten von A?

C) Eigenwerte als Nullstellen des charakteristischen Polynoms

Satz 35.15 (Eigenwerte und das charakteristische Polynom).
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Die Eigenwerte von f sind genau die Nullstellen von χ_f in K.

b. Die Eigenwerte von A sind genau die Nullstellen von χ_A in K.

Insbesondere, f und A haben höchstens n paarweise verschiedene Eigenwerte.

Beweis: Für $\lambda \in K$ gilt unter Berücksichtigung von Korollar 30.23:

$$\begin{aligned}
\lambda \text{ ist Eigenwert von } f &\iff \mathrm{Ker}(\lambda\,\mathrm{id}_V - f) = \mathrm{Eig}(f,\lambda) \neq \{0\} \\
&\iff \lambda\,\mathrm{id}_V - f \text{ ist nicht injektiv} \\
&\overset{30.23}{\iff} \lambda\,\mathrm{id}_V - f \text{ ist nicht bijektiv} \\
&\iff \chi_f(\lambda) = \det(\lambda\,\mathrm{id}_V - f) = 0.
\end{aligned}$$

Der Beweis für die Matrizen geht analog. □

Bevor wir das charakteristische Polynom weiter untersuchen, wollen wir zunächst einige Beispiele betrachten.

Beispiel 35.16.

a. Betrachten wir zunächst die folgende Matrix:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} \in \mathrm{Mat}(3,\mathbb{Q}).$$

Mit Hilfe der Regel von Sarrus oder durch den Laplaceschen Entwicklungssatz bestimmen wir das charakteristische Polynom von A als

$$\chi_A = \det \begin{pmatrix} t & -1 & -1 \\ 1 & t-2 & -1 \\ 1 & -1 & t-2 \end{pmatrix} = t^3 - 4t^2 + 5t - 2 = (t-1)^2 \cdot (t-2).$$

Alternativ kann man allgemein die Matrix $t\mathbb{1}_n - A \in \mathrm{Mat}_n\big(\mathbb{Q}(t)\big)$ auch als Matrix über dem Körper $\mathbb{Q}(t)$ auffassen (siehe Bemerkung 26.24). Da $\mathbb{Q}(t)$ ein Körper ist, dürfen wir die Determinante mittels des Gaußschen Algorithmus' 34.15 bestimmen.

Insbesondere dürfen wir dabei durch Polynome dividieren!

$$
\begin{pmatrix} t & -1 & -1 \\ 1 & t-2 & -1 \\ 1 & -1 & t-2 \end{pmatrix} \xrightarrow[III \mapsto III - \frac{1}{t} I]{II \mapsto II - \frac{1}{t} I} \begin{pmatrix} t & -1 & -1 \\ 0 & t-2+\frac{1}{t} & \frac{1}{t}-1 \\ 0 & \frac{1}{t}-1 & t-2+\frac{1}{t} \end{pmatrix} = \tag{94}
$$

$$
\begin{pmatrix} t & -1 & -1 \\ 0 & \frac{(t-1)^2}{t} & -\frac{t-1}{t} \\ 0 & -\frac{t-1}{t} & \frac{(t-1)^2}{t} \end{pmatrix} \xrightarrow{III \mapsto III + \frac{1}{t-1} II} \begin{pmatrix} t & -1 & -1 \\ 0 & \frac{(t-1)^2}{t} & -\frac{t-1}{t} \\ 0 & 0 & t-2 \end{pmatrix}.
$$

Entsprechend erhalten wir für das charakteristische Polynom

$$\chi_A = t \cdot \tfrac{(t-1)^2}{t} \cdot (t-2) = (t-1)^2 \cdot (t-2).$$

Das charakteristische Polynom hat also die Nullstellen $\lambda = 1$ und $\lambda = 2$, wobei $\lambda = 1$ eine zweifache Nullstelle ist. Insbesondere ist also $\sigma(A) = \{1, 2\}$.
Wir können jetzt für $\lambda = 1$ und für $\lambda = 2$ jeweils den Eigenraum $\mathrm{Eig}(A, \lambda) = \mathrm{Lös}(\lambda \mathbb{1}_n - A, 0)$ mit Hilfe des Gauß-Algorithmus bestimmen.[4]
Der Algorithmus zur Bestimmung von $\mathrm{Eig}(A, 1) = \mathrm{Lös}(\mathbb{1}_n - A, 0)$ sieht vor, daß wir die Matrix zunächst auf reduzierte ZSF bringen und dann in den Nullzeilen die Diagonalelemente durch -1 ersetzen:

$$
\begin{pmatrix} -1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.
$$

Die letzten beiden Spalten, d.h. die, bei denen eine -1 auf der Diagonalen steht, bilden dann eine Basis des Eigenraumes zum Eigenwert 1:

$$\mathrm{Eig}(A, 1) = \mathrm{Lin}\left((-1, -1, 0)^t, (-1, 0, -1)^t\right).$$

$\mathrm{Eig}(A, 1)$ ist also zweidimensional.
Analog ergibt sich $\mathrm{Eig}(A, 2)$ aus

$$
\begin{pmatrix} -2 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \end{pmatrix},
$$

und damit gilt $\mathrm{Eig}(A, 2) = \mathrm{Lin}\left((-1, -1, -1)^t\right)$.

[4] Man beachte, daß es zur Berechnung der reduzierten Zeilen-Stufen-Form von $\lambda \mathbb{1}_n - A$ für $\lambda = 1$ *nicht* erlaubt ist, in (94) in der letzten Matrix t etwa durch $\lambda = 1$ zu ersetzen, um die ZSF zu erhalten, da wir bei den vorgenommenen Umformungen zur Ermittelung obiger Matrix durch das Polynom $t - 1$ dividiert haben. Dies ist über $\mathbb{Q}(t)$ eine erlaubte Operation gewesen. Ersetzen wir jedoch t durch 1, so ist die Operation nicht mehr erlaubt!

b. Wir hatten schon durch eine geometrische Argumentation gesehen, daß die Drehung um einen Winkel α im allgemeinen keinen reellen Eigenwert besitzt. Den gleichen Sachverhalt prüfen wir nun noch einmal mit algebraischen Methoden. Die Matrixdarstellung der Drehung bezüglich der kanonischen Basis von $\mathbb{R}^2$ ist

$$A_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

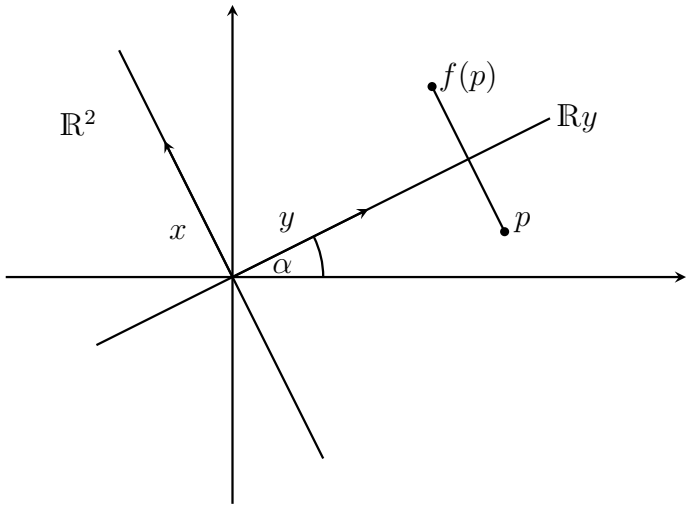
Aber

$$\chi_{A_\alpha} = \big(t - \cos(\alpha)\big)^2 + \sin^2(\alpha) = t^2 - 2\cos(\alpha)t + 1.$$

Die Nullstellen von χ_{A_α} sind $\cos(\alpha) + \sqrt{\cos^2(\alpha) - 1}$ und $\cos(\alpha) - \sqrt{\cos^2(\alpha) - 1}$. Für beide Terme gilt, sie sind genau dann reell, wenn α ein ganzzahliges Vielfaches von π ist.

Insbesondere hat die Drehung also nur dann reelle Eigenwerte, wenn α ein ganzzahliges Vielfaches von π ist, d. h. $A_\alpha = \mathbb{1}_2$ oder $A_\alpha = -\mathbb{1}_2$.

c. Es sei $f \in \mathrm{End}_{\mathbb{R}}\left(\mathbb{R}^2\right)$ die *Spiegelung* an einer Geraden $\mathrm{Lin}\,(y) = \mathbb{R} \cdot y \subset \mathbb{R}^2$ mit $0 \neq y = (y_1, y_2)^t \in \mathbb{R}^2$.



Wir setzen $x = (-y_2, y_1)^t \in \mathbb{R}^2$. Dann steht x senkrecht auf y und $B = (y, x)$ ist eine Basis von $\mathbb{R}^2$. Die Spiegelung f bildet mithin y auf sich selbst und x auf $-x$ ab, da x senkrecht auf $\mathrm{Lin}\,(y)$ steht. Damit hat f die folgende Matrixdarstellung bezüglich B

$$M_B^B(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

und das charakteristische Polynom von f ist gerade

$$\chi_f = (t-1) \cdot (t+1).$$

Die Spiegelung von f hat also Spektrum $\sigma(f) = \{-1, 1\}$.

Beschreiben wir f in den Standardkoordinaten $E = (e_1, e_2)$ von $\mathbb{R}^2$, so ist f die Spiegelung an $\mathrm{Lin}\,(e_1) = \mathbb{R} \cdot e_1$ gefolgt von der Drehung um den Winkel 2α, wenn

α der Winkel ist, den $\mathrm{Lin}\,(y)$ mit $\mathrm{Lin}\,(e_1)$ einschließt. Wir erhalten also

$$M_E^E(f) = \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix} = \begin{pmatrix} \cos(2\alpha) & -\sin(2\alpha) \\ \sin(2\alpha) & \cos(2\alpha) \end{pmatrix} \circ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Das charakteristische Polynom errechnet sich aus dieser Matrixdarstellung als

$$\big(t - \cos(2\alpha)\big) \cdot \big(t + \cos(2\alpha)\big) - \sin^2(2\alpha) = t^2 - 1 = (t-1) \cdot (t+1).$$

Korollar 35.17 (Eigenwerte einer Dreiecksmatrix).
Ist $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ eine obere oder untere Dreiecksmatrix, dann ist

$$\chi_A = (t - a_{11}) \cdot \ldots \cdot (t - a_{nn})$$

und die Einträge auf der Diagonalen sind genau die Eigenwerte von A.

Beweis: Für eine obere Dreiecksmatrix $A = (a_{ij})$ gilt

$$\chi_A = \det \begin{pmatrix} t - a_{11} & * & \ldots & \ldots & * \\ 0 & t - a_{22} & * & \ldots & * \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \ldots & \ldots & 0 & t - a_{nn} \end{pmatrix} = \prod_{i=1}^{n} (t - a_{ii}),$$

und dieses Polynom hat genau die Nullstellen $a_{11}, \ldots, a_{nn}$. Der Beweis für untere Dreiecksmatrizen geht analog. □

D) Diagonalisierbarkeit

In diesem Abschnitt wollen wir Eigenschaften eines Endomorphismus' untersuchen, die sicherstellen, daß seine Matrixdarstellung eine Diagonalmatrix und damit besonders einfach wird.

Definition 35.18 (Diagonalisierbar).
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. f heißt *diagonalisierbar* falls es eine Basis B von V gibt, so daß $M_B^B(f)$ eine Diagonalmatrix ist.

b. A heißt *diagonalisierbar* falls es eine Matrix $T \in \mathrm{Gl}_n(K)$ gibt, so daß $T^{-1} \circ A \circ T$ eine Diagonalmatrix ist.

Bei unserem Bestreben, Kriterien für die Diagonalisierbarkeit eines Endomorphismus' herzuleiten, benötigen wir die Aussage, daß Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig voneinander sind.

Lemma 35.19 (Eigenvektoren zu verschiedenen Eigenwerten).
Sind $\lambda_1, \ldots, \lambda_k$ paarweise verschiedene Eigenvektoren eines Endomorphismus $f \in \mathrm{End}_K(V)$ und sind $x_1, \ldots, x_k$ zugehörige Eigenvektoren, dann ist die Familie $(x_1, \ldots, x_k)$ linear unabhängig.

Beweis: Wir zeigen die Behauptung mit Induktion nach k. Für $k = 1$ ist die Aussage offensichtlich richtig, da ein Eigenvektor nicht der Nullvektor ist. Sei nun $k > 1$ und sei

$$\mu_1 \cdot x_1 + \ldots + \mu_k \cdot x_k = 0$$

eine Linearkombination der x_i, die null ergibt. Dann folgt

$$\begin{aligned} 0 =& f(0) = f\big(\mu_1 \cdot x_1 + \ldots + \mu_k \cdot x_k\big) \\ =& \mu_1 \cdot f(x_1) + \ldots + \mu_k \cdot f(x_k) = \mu_1 \cdot \lambda_1 \cdot x_1 + \ldots + \mu_k \cdot \lambda_k \cdot x_k \end{aligned}$$

und zudem

$$0 = \lambda_k \cdot 0 = \lambda_k \cdot (\mu_1 \cdot x_1 + \ldots + \mu_k \cdot x_k) = \mu_1 \cdot \lambda_k \cdot x_1 + \ldots + \mu_k \cdot \lambda_k \cdot x_k.$$

Ziehen wir die beiden Gleichungen voneinander ab, so erhalten wir

$$\begin{aligned} 0 =& \mu_1 \cdot \lambda_1 \cdot x_1 + \ldots + \mu_k \cdot \lambda_k \cdot x_k - \mu_1 \cdot \lambda_k \cdot x_1 + \ldots + \mu_k \cdot \lambda_k \cdot x_k \\ =& \mu_1 \cdot (\lambda_1 - \lambda_k) \cdot x_1 + \ldots + \mu_{k-1} \cdot (\lambda_1 - \lambda_{k-1}) \cdot x_{-1}. \end{aligned}$$

Nach Induktion ist die Familie $(x_1, \ldots, x_{k-1})$ aber linear unabhängig, so daß

$$\mu_i \cdot (\lambda_i - \lambda_k) = 0$$

für $i = 1, \ldots, k-1$ gelten muß. Nach Voraussetzung sind die λ_i paarweise verschieden und somit ist $\lambda_i - \lambda_k \neq 0$ für $i = 1, \ldots, k-1$, so daß notwendigerweise $\mu_i = 0$ für $i = 1, \ldots, k-1$ gelten muß. Aber dann folgt

$$0 = \mu_1 \cdot x_1 + \ldots + \mu_k \cdot x_k = \mu_k \cdot x_k$$

und somit auch $\mu_k = 0$, da x_k als Eigenvektor nicht der Nullvektor ist. Wir haben also

$$\mu_1 = \ldots = \mu_k = 0$$

gezeigt und $(x_1, \ldots, x_k)$ ist linear unabhängig. □

Wir wollen nun eine erste Charakterisierung diagonalisierbarer Endomorphismen und Matrizen geben. Das eben bewiesene Lemma wird dabei im Beweis eingehen. Für weitere Charakterisierungen der Diagonalisierbarkeit, die Begriffe wie das Minimalpolynom oder die geometrische und algebraische Vielfachheit von Eigenwerten verwenden, verweisen wir den Leser auf den Anhang (siehe Satz B1.27).

Satz 35.20 (Diagonalisierbarkeit von Endomorphismen).
Für einen Endomorphismus $f \in \mathrm{End}_K(V)$ sind die folgenden Aussagen äquivalent:

a. f ist diagonalisierbar.

b. V hat eine Basis aus Eigenvektoren von f.

c. Sind $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von f in K, dann gilt
$$V = \mathrm{Eig}(f, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(f, \lambda_r).$$

d. Sind $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von f in K, dann gilt
$$\dim_K(V) = \dim_K \mathrm{Eig}(f, \lambda_1) + \ldots + \dim_K \mathrm{Eig}(f, \lambda_r).$$

Beweis: Wir beachten zunächst, daß in der Matrixdarstellung von f bezüglich einer Basis $B = (x_1, \ldots, x_n)$ genau dann in der i-ten Spalte nur in der i-ten Zeile ein Eintrag steht, wenn $f(x_i) = \lambda \cdot x_i$ für ein $\lambda \in K$ gilt, d.h. wenn x_i ein Eigenvektor von f ist. Damit ist die Äquivalenz von a. und b. bewiesen.

b. $\Rightarrow$ c.: Nun wollen wir mit b. voraussetzen, daß wir eine Basis $B = (x_1, \ldots, x_n)$ aus Eigenvektoren von f haben. Wir wählen dann für jeden Eigenraum $\mathrm{Eig}(f, \lambda_i)$ eine Basis B_i, die die Eigenvektoren zum Eigenwert λ_i aus B bereits enthält.

Nehmen wir an, daß in einer der Basen ein Vektor x dazu gekommen ist, der noch nicht in B war. Wir können ohne Einschränkung $x \in B_r$ annehmen. Dann läßt sich x als Linearkombination
$$x = \mu_1 \cdot x_1 + \ldots + \mu_n \cdot x_n$$
in der Basis B ausdrücken. Wir setzen nun
$$y_i = \sum_{x_j \in \mathrm{Eig}(f, \lambda_i)} \mu_j \cdot x_j \in \mathrm{Eig}(f, \lambda_i)$$
für $i = 1, \ldots, r-1$
$$y_r = \sum_{x_j \in \mathrm{Eig}(f, \lambda_r)} \mu_j \cdot x_j - x \in \mathrm{Eig}(f, \lambda_r).$$
Damit gilt dann
$$y_1 + \ldots + y_r = \sum_{j=1}^{n} \mu_j \cdot x_j - x = 0. \tag{95}$$
Nach Lemma 35.19 ist die Familie $(y_i \mid y_i \neq 0, i = 1, \ldots, r)$ aber linear unabhängig, woraus wegen (95) unmittelbar
$$y_i = 0$$

für alle $i = 1, \ldots, r$ folgt. Dann ist aber

$$x = \sum_{x_j \in \mathrm{Eig}(f,\lambda_r)} \mu_j \cdot x_j$$

linear abhängig von den übrigen Vektoren in B_r, im Widerspruch dazu, daß B_r linear unabhängig ist. Es gilt also

$$B = B_1 \cup \ldots \cup B_r$$

und mit Aufgabe 29.26 ist dann

$$V = \mathrm{Eig}(f, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(f, \lambda_r)$$

die direkte Summe der von den B_i erzeugten Unterräume.

c. ⇒ d.: Aus der Dimensionsformel in Korollar 30.17 folgt mit Induktion, daß die Dimension einer direkten Summe die Summe der Dimensionen der beteiligten Unterräume ist, d.h.

$$\begin{aligned}\dim_K(V) &= \dim_K \big(\mathrm{Eig}(f, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(f, \lambda_r)\big) \\ &= \dim_K \mathrm{Eig}(f, \lambda_1) + \ldots + \dim_K \mathrm{Eig}(f, \lambda_r).\end{aligned}$$

d. ⇒ b.: Wir wählen Basen $B_i = (x_{i1}, \ldots, x_{in_i})$ für $\mathrm{Eig}(f, \lambda_i)$ für $i = 1, \ldots, r$ und zeigen, daß

$$B = B_1 \cup \ldots \cup B_r$$

linear unabhängig ist. Sei dazu

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} \mu_{ij} \cdot x_{ij} = 0$$

eine Linearkombination der Vektoren in B, die null ergibt. Wie oben muß dann wegen Lemma 35.19

$$y_i := \sum_{j=1}^{n_i} \mu_{ij} \cdot x_{ij} = 0$$

gelten, weil die Familie $(y_i \mid y_i \neq 0, i = 1, \ldots, r)$ linear unabhängig ist. Da B_i aber jeweils auch linear unabhängig ist, folgt damit

$$\mu_{ij} = 0$$

für alle $j = 1, \ldots, n_i$ und alle $i = 1, \ldots, r$. Mithin ist B linear unabhängig, und da B genau $\dim_K(V)$ viele Vektoren enthält, ist B dann auch eine Basis von V und nach Konstruktion eine Basis aus Eigenvektoren. □

Der Satz zur Diagonalisierbarkeit eines Endomorphismus kann natürlich genauso für quadratische Matrizen formuliert werden.

Korollar 35.21 (Diagonalisierbarkeit von Matrizen).
Für eine quadratische Matrix $A \in \mathrm{Mat}_n(K)$ sind die folgenden Aussagen äquivalent:

a. A ist diagonalisierbar.

b. K^n hat eine Basis aus Eigenvektoren von A.

c. Sind $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von A, dann gilt
$$K^n = \bigoplus_{i=1}^{r} \mathrm{Eig}(A, \lambda_i).$$

d. Sind $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von A, dann gilt
$$\dim_K \mathrm{Eig}(A, \lambda_1) + \ldots + \dim_K \mathrm{Eig}(A, \lambda_r) = n.$$

Insbesondere, genau dann ist $T \in \mathrm{Gl}_n(K)$ so, daß $T^{-1} \circ A \circ T$ eine Diagonalmatrix ist, wenn die Spalten von T eine Basis des K^n aus Eigenvektoren von A sind.

Beweis: Wende Satz B1.27 auf die Abbildung f_A an. □

Falls ein Endomorphismus oder eine Matrix hinreichend viele verschiedene Eigenwerte hat, so folgt aus den obigen Überlegungen unmittelbar deren Diagonalisierbarkeit.

Korollar 35.22 (Diagonalisierbarkeit).
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Hat f genau n paarweise verschiedene Eigenwerte, so ist f diagonalisierbar.

b. Hat A genau n paarweise verschiedene Eigenwerte, so ist A diagonalisierbar.

Beweis: Hat f genau n paarweise verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n \in K$, so muß
$$\chi_f = \mu_f = (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_n)$$
gelten. D.h. μ_f zerfällt in paarweise verschiedene Linearfaktoren und f ist diagonalisierbar. Der Beweis für A geht analog. □

Aus Korollar 35.21 können wir ein Verfahren ableiten, das es uns erlaubt, eine Matrix zu diagonalisieren und die Transformationsmatrix T zu berechnen. Dieses wollen wir nun als Algorithmus formulieren.

Algorithmus 35.23 (Algorithmus zur Diagonalisierung).

INPUT: $A \in \mathrm{Mat}_n(K)$.

OUTPUT: 0, falls A über K nicht diagonalisierbar ist,
$1, D, T$, falls A diagonaliserbar ist, wobei D eine zu A konjugierte Diagonalmatrix ist, und T die zugehörige Transformationsmatrix mit $T^{-1} \circ A \circ T = D$.

1. **Schritt:** Berechne das charakteristische Polynom von A.
2. **Schritt:** Faktorisiere das charakteristische Polynom über K. Ist einer der Faktoren nicht linear, ist A nicht diagonalisierbar und man gebe 0 zurück. Sind alle Faktoren linear, so liefert die Faktorisierung die Eigenwerte $\lambda_1, \ldots, \lambda_r$ sowie ihre algebraischen Vielfachheiten $n_1, \ldots, n_r$.
3. **Schritt:** Bestimme für jeden Eigenwert λ_i eine Basis des Eigenraums $\mathrm{Eig}(A, \lambda_i)$ als $\mathrm{Lös}(A - \lambda_i \mathbb{1}_n, 0)$ - vgl. Algorithmus 33.9 - sowie seine Dimension - vgl. Algorithmus 32.20.
4. **Schritt:** Ist die Summe der Dimensionen n, so schreibe man die im 3. Schritt bestimmten Basen als Spalten in eine Matrix und erhält so T. Ferner erhält man D, indem man die Eigenwerte $\lambda_1, \ldots, \lambda_r$ entsprechend ihren Vielfachheiten in der Diagonalen einer Nullmatrix einträgt.

Beispiel 35.24.

Gegeben sei die Matrix

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q}).$$

Das charakteristische Polynom von A berechnet man mit Hilfe zweifacher Laplace-Entwicklung nach der jeweils letzten Spalte als

$$\chi_A = \begin{vmatrix} t-2 & 1 & 0 & 0 \\ 0 & t-1 & 0 & 0 \\ 0 & 0 & t-2 & 0 \\ -1 & 1 & 1 & t-1 \end{vmatrix} = (t-1)\cdot(t-2)\cdot \begin{vmatrix} t-2 & 1 \\ 0 & t-1 \end{vmatrix} = (t-1)^2 \cdot (t-2)^2.$$

Damit ist also $\sigma(A) = \{1, 2\}$ mit $\mathrm{mult}(\chi_A, 1) = \mathrm{mult}(\chi_A, 2) = 2$.

Als nächstes berechnen wir den Eigenraum $\mathrm{Lös}(2\mathbb{1}_4 - A, 0)$ zum Eigenwert $\lambda = 2$:

$$2 \cdot \mathbb{1}_4 - A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 1 \end{pmatrix} \xrightarrow[\substack{IV \mapsto IV - II \\ I \mapsto I + II}]{\substack{I \leftrightarrow IV \\ I \mapsto -I}} \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{-1\text{'en einfügen}} \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Mithin ist

$$\operatorname{Eig}(A,2) = \operatorname{Lin}\left((-1,0,-1,0)^t, (-1,0,0,-1)^t\right)$$

und

$$\dim_{\mathbb{Q}} \operatorname{Eig}(A,2) = 2 = \operatorname{mult}(\chi_A, 2).$$

Dann berechnen wir den Eigenraum $\operatorname{Lös}(\mathbb{1}_4 - A, 0)$ zum Eigenwert $\lambda = 1$:

$$1 \cdot \mathbb{1}_4 - A = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 1 & 1 & 0 \end{pmatrix} \overset{\substack{IV \mapsto IV - I + III \\ I \mapsto -I \\ III \mapsto -III}}{\longmapsto} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{-1\text{'en einfügen}}{\longmapsto} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Mithin ist

$$\operatorname{Eig}(A,1) = \operatorname{Lin}\left((-1,-1,0,0)^t, (0,0,0,-1)^t\right)$$

und

$$\dim_{\mathbb{Q}} \operatorname{Eig}(A,1) = 2 = \operatorname{mult}(\chi_A, 1).$$

Also zerfällt χ_A über $\mathbb{Q}$ in Linearfaktoren und die geometrischen Vielfachheiten der Eigenwerte stimmen mit den algebraischen überein, so daß A diagonalisierbar ist. Zudem gilt für

$$T = \begin{pmatrix} -1 & -1 & -1 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 \end{pmatrix},$$

daß

$$T^{-1} \circ A \circ T = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Aufgaben

Aufgabe 35.25.
Berechne das charakteristische Polynom der Matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & 0 & 1 \\ a_0 & \dots & \dots & \dots & \dots & a_{n-2} & a_{n-1} \end{pmatrix} \in \mathrm{Mat}_n(\mathbb{R})$$

für $a_0, \dots, a_{n-1} \in K$.

Aufgabe 35.26 (Nilpotente Endomorphismen und Matrizen).
Es sei $A \in \mathrm{Mat}_n(K)$ und $f \in \mathrm{End}_K(V)$ mit $1 \leq \dim_K(V) < \infty$.

a. Zeige, gibt es ein $r \in \mathbb{N}$ mit $f^r = 0$, so gilt $\mathrm{Spur}(f) = 0$.

b. Zeige, gibt es ein $r \in \mathbb{N}$ mit $A^r = 0$, so gilt $\mathrm{Spur}(A) = 0$.

c. Finde ein Beispiel für eine Matrix wie in Teil b., bei der nicht alle Diagonalelemente Null sind.

Hinweis zum Beweis von a.: Führe Induktion über $n = \dim_K(V)$. Dazu zeige man, daß $M_B^B(f)$ für eine geeignete Wahl von B Blockgestalt mit einem Nullblock in der oberen linken Ecke hat. Aufgabe 31.34b. mit $U = \mathrm{Ker}(f)$ ist dabei hilfreich.

Aufgabe 35.27.
Es sei $1 \leq \dim_K(V) < \infty$, $f \in \mathrm{End}_K(V)$.

a. Ist $U \subseteq V$ ein f-invarianter Unterraum, dann gilt

$$\det(f) = \det(f_U) \cdot \det(f_{V/U})$$

und

$$\chi_f = \chi_{f_U} \cdot \chi_{f_{V/U}}.$$

b. Ist $V = U_1 \oplus \ldots \oplus U_k$, wobei die U_i f-invariant seien, dann gilt

$$\det(f) = \det(f_{U_1}) \cdot \ldots \cdot \det(f_{U_k})$$

und

$$\chi_f = \chi_{f_{U_1}} \cdot \ldots \cdot \chi_{f_{U_k}}.$$

Aufgabe 35.28.
Für ein Polynom $p \in K[t]$ und zwei konjugierte Matrizen $A, B \in \mathrm{Mat}_n(K)$ gilt

$$p(A) = 0 \iff p(B) = 0.$$

Aufgabe 35.29 (Zyklische Unterräume).
Zeige, $\chi_{f_U} = \mu_{f_U} = t^m$ für den Endomorphismus f_U aus Aufgabe 31.36.

Aufgabe 35.30.
Bestimme die Eigenwerte und die Eigenräume der folgenden Matrix A und entscheide, ob sie diagonalisierbar ist:

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q}).$$

Aufgabe 35.31 (Die Eigenräume bilden eine direkte Summe.).
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Sind $x_1, \ldots, x_r \in V$ Eigenvektoren von f zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r \in K$, dann ist die Familie $(x_1, \ldots, x_r)$ linear unabhängig. Insbesondere gilt

$$\mathrm{Eig}(f, \lambda_1) + \ldots + \mathrm{Eig}(f, \lambda_r) = \mathrm{Eig}(f, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(f, \lambda_r).$$

b. Sind $x_1, \ldots, x_r \in K^n$ Eigenvektoren von A zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r \in K$, dann ist die Familie $(x_1, \ldots, x_r)$ linear unabhängig. Insbesondere gilt

$$\mathrm{Eig}(A, \lambda_1) + \ldots + \mathrm{Eig}(A, \lambda_r) = \mathrm{Eig}(A, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(A, \lambda_r).$$

Aufgabe 35.32.
Sind $f, g \in \mathrm{End}_K(V)$, so gilt $\sigma(f \circ g) = \sigma(g \circ f)$.

Kapitel V

Euklidische und unitäre Räume

Im folgenden sei stets $\mathbb{K}$ einer der beiden Körper $\mathbb{R}$ oder $\mathbb{C}$.

§ 36 Euklidische und unitäre Räume

Zur Motivation beginnen wir den Abschnitt mit einigen Überlegungen zur euklidischen Geometrie in der reellen Ebene $\mathbb{R}^2$.

Wir definieren uns zunächst zwei Abbildungen

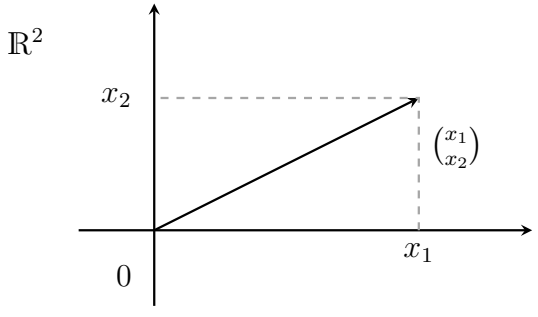
$$||\cdot|| : \mathbb{R}^2 \to \mathbb{R}_{\geq 0},$$

die einem Vektor $x = (x_1, x_2)^t \in \mathbb{R}^2$ seine Länge $||x||$ zuordnet, sowie

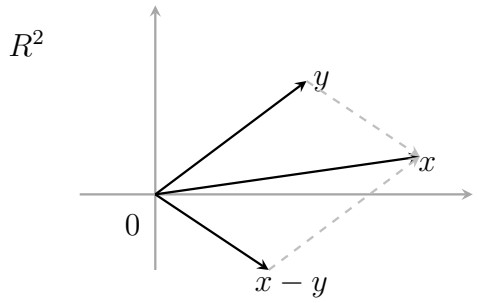
$$d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_{\geq 0},$$

die zwei Punkten $x \in \mathbb{R}^2$ und $y \in \mathbb{R}^2$ ihren Abstand $d(x, y)$ zuweist.

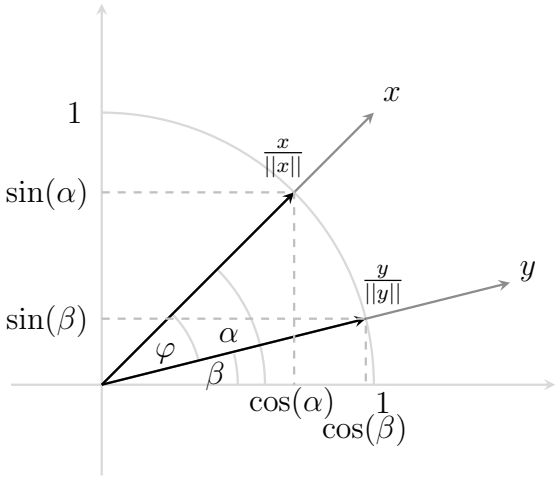
Der Satz von Pythagoras liefert dann $||x|| = \sqrt{x_1^2 + x_2^2}$.



Wir nennen $||x||$ auch die *euklidische Norm* des Vektors x. Da der Abstand der Punkte $x = (x_1, x_2)^t$ und $y = (y_1, y_2)^t$ gerade die Länge des Vektors $x - y$ ist, folgt somit $d(x, y) = ||x - y|| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$.



Mit Hilfe der Norm können wir - nach einigen geometrischen Überlegungen - auch den Winkel $\measuredangle(x, y)$, den zwei Vektoren x und y miteinander einschließen, bestimmen.



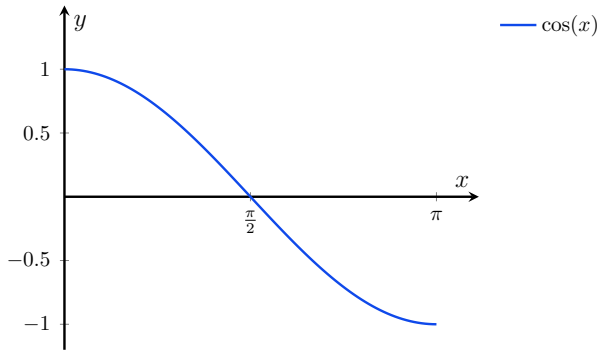
Dazu gehen wir zunächst zu den normierten Vektoren $\frac{x}{||x||}$ und $\frac{y}{||y||}$ über, die beide die Länge eins haben, wobei wir $x \neq 0$ und $y \neq 0$ voraussetzen. Mit den Bezeichnungen in der Skizze gilt dann

$$\measuredangle(x, y) = \measuredangle\left(\frac{x}{||x||}, \frac{y}{||y||}\right) = \alpha - \beta = \varphi.$$

Um φ selbst (im Bogenmaß) auszudrücken, müßte man die Länge des Kreisbogens zwischen $\frac{x}{||x||}$ und $\frac{y}{||y||}$ messen, also einer gekrümmten Linie. Dazu greifen wir auf unsere Analysiskenntnisse zurück.

Zur anschaulichen Herleitung des Winkels φ mit $0 \leq \varphi \leq \pi$, benötigen wir nur, daß die Funktion

$$\cos : [0, \pi] \to \mathbb{R} : \varphi \mapsto \cos(\varphi)$$



injektiv ist. Also reicht es, $\cos(\varphi)$ zu kennen, um den Winkel φ eindeutig beschrieben zu haben. Unter Zuhilfenahme der obigen Skizze und des Additionstheorems 12.38 für den Cosinus erhalten wir

$$\cos(\varphi) = \cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) = \frac{x_1y_1 + x_2y_2}{||x|| \cdot ||y||}.$$

Dies führt zur Definition einer weiteren Abbildung

$$\langle \cdot, \cdot \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} : (x, y) = \big((x_1, x_2)^t, (y_1, y_2)^t\big) \mapsto \langle x, y \rangle := x_1y_1 + x_2y_2,$$

welche wir *Skalarprodukt* nennen. Mit deren Hilfe erhalten wir

$$\cos(\varphi) = \frac{\langle x, y \rangle}{||x|| \cdot ||y||}$$

oder alternativ

$$\measuredangle(x, y) = \varphi = \arccos\left(\frac{\langle x, y \rangle}{||x|| \cdot ||y||}\right).$$

Wir sind also mittels recht einfacher Abbildungen in der Lage, Längen und Winkel auszudrücken. Dieses Beispiel motiviert die folgenden Begriffsbildungen.

A) Skalarprodukte

Definition 36.1 (Skalarprodukt).
Es sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{K}$$

heißt ein *Skalarprodukt* auf V, falls für alle $\lambda, \mu \in \mathbb{K}$ und $x, y, z \in V$ gilt:

(1) $\langle x, \lambda y + \mu z \rangle = \lambda \cdot \langle x, y \rangle + \mu \cdot \langle x, z \rangle$,

(2) $\langle x, y \rangle = \overline{\langle y, x \rangle}$,

(3) $\langle x, x \rangle \in \mathbb{R}$ und $\langle x, x \rangle > 0$ für $x \neq 0$.

Ist $\mathbb{K} = \mathbb{R}$, so nennen wir das Quadrupel $(V, +, \cdot, \langle \cdot, \cdot \rangle)$ einen *euklidischen Raum*; ist $\mathbb{K} = \mathbb{C}$, so nennen wir es einen *unitären Raum*.

Beispiel 36.2 (Skalarprodukte).

a. Wir nennen

$$\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} : (x,y) \mapsto x^t \circ y = \sum_{i=1}^{n} x_i \cdot y_i$$

das *kanonische Skalarprodukt* oder *Standardskalarprodukt* auf $\mathbb{R}^n$. $\big(\mathbb{R}^n, +, \cdot, \langle\cdot,\cdot\rangle\big)$ ist ein euklidischer Raum.

b. Analog definieren wir das *kanonische Skalarprodukt* auf $\mathbb{C}^n$ durch

$$\langle\cdot,\cdot\rangle : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C} : (x,y) \mapsto \overline{x}^t \circ y = x^* \circ y = \sum_{i=1}^{n} \overline{x_i} \cdot y_i,$$

und $\big(\mathbb{C}^n, +, \cdot, \langle\cdot,\cdot\rangle\big)$ ist ein unitärer Raum.

c. Sei $V = \mathcal{C}([0,1], \mathbb{R})$ der $\mathbb{R}$-Vektorraum der auf dem Intervall $[0,1]$ stetigen Funktionen. Dann wird durch

$$\langle f, g\rangle := \int_0^1 f(x)g(x)dx \in \mathbb{R}$$

für $f, g \in V$ ein Skalarprodukt definiert (siehe Aufgabe 36.31) und $\big(\mathcal{C}([0,1], \mathbb{R}), +, \cdot, \langle\cdot,\cdot\rangle\big)$ ist ein euklidischer Raum.

Bemerkung 36.3.

Wir behandeln im Folgenden oft euklidische und eines unitäre Räume parallel. Dabei machen wir uns zunutze, daß für eine reelle Zahl λ gilt $\lambda = \overline{\lambda}$. Mithin gilt in einem *reellen* Vektorraum V

$$\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle \quad \Longleftrightarrow \quad \langle \lambda x + \mu y, z\rangle = \overline{\lambda}\langle x, z\rangle + \overline{\mu}\langle y, z\rangle,$$

und für $A \in \mathrm{Mat}(n, \mathbb{R})$ gilt genau dann $A = A^t$, wenn $A = \overline{A}^t$ erfüllt ist.

B) Die euklidische Norm zu einem Skalarprodukt

Definition 36.4 (Normierter Raum).

Es sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung

$$||\cdot|| : V \to \mathbb{R}_{\geq 0},$$

heißt eine *Norm* auf V, falls für alle $x, y \in V$ und $\lambda \in \mathbb{K}$

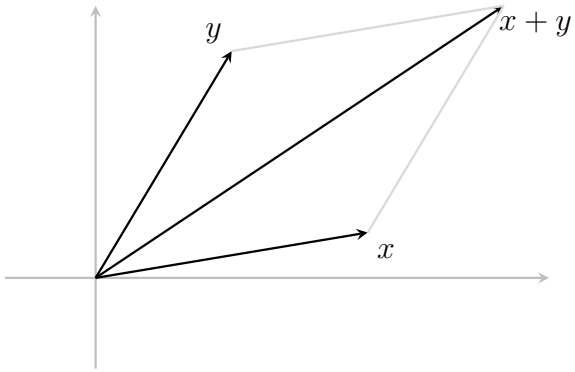
(1) $||x|| = 0 \quad \Leftrightarrow \quad x = 0,$ ("Positive Definitheit")

(2) $||\lambda x|| = |\lambda| \cdot ||x||$, und ("Homogenität")

(3) $||x + y|| \leq ||x|| + ||y||.$ ("Dreiecksungleichung")

Das Quadrupel $\big(V, +, \cdot, ||\cdot||\big)$ heißt dann ein *normierter Raum*.

Bemerkung 36.5 (Norm als Längenmaß).
Wir erinnern uns, daß eine Norm die Länge von Vektoren sowie Abstände messen soll. Bedingung (1) kann dann so interpretiert werden, daß jeder Vektor eine nicht-negative Länge hat und daß nur der Nullvektor die Länge null hat. Bedingung (2) bedeutet, daß die Streckung eines Vektors um den Faktor λ seine Länge um $|\lambda|$ strecken möge. Und Bedingung (3) kann dahingehend interpretiert werden, daß der Weg vom Ursprung über den Punkt x hin zum Punkt $x+y$ unter gar keinen Umständen kürzer ist, als der direkte Weg vom Ursprung zum Punkt $x+y$.



Diese Forderungen scheinen allesamt für eine Funktion, die die Länge von Vektoren beziehungsweise Abstände von Punkten messen soll, nicht unbillig. Und in der Tat reichen diese Forderungen auch bereits aus, um einen vernünftigen Längenbegriff zu erhalten.

Satz 36.6 (Cauchy-Schwarzsche Ungleichung).
Ist V ein euklidischer oder unitärer Raum, dann gilt für alle $x, y \in V$

$$\big|\langle x,y\rangle\big| \leq \sqrt{\langle x,x\rangle}\cdot\sqrt{\langle y,y\rangle}, \tag{96}$$

zudem gilt die Gleichheit genau dann, wenn x und y linear abhängig sind.

Beweis: Für $x=0$ oder $y=0$ ist die Aussage offensichtlich richtig. Wir können also $x,y\neq 0$ annehmen. Dann gilt für $\lambda\in\mathbb{K}$

$$0 \leq \langle x-\lambda y, x-\lambda y\rangle = \langle x,x\rangle - \overline{\lambda\langle x,y\rangle} - \lambda\langle x,y\rangle + \overline{\lambda}\lambda\langle y,y\rangle. \tag{97}$$

Wählen wir nun speziell $\lambda = \frac{\overline{\langle x,y\rangle}}{\langle y,y\rangle}\in\mathbb{K}$, dann folgt

$$\begin{aligned} 0 &\leq \langle x,x\rangle - \tfrac{\langle x,y\rangle}{\langle y,y\rangle}\cdot\overline{\langle x,y\rangle} - \tfrac{\overline{\langle x,y\rangle}}{\langle y,y\rangle}\cdot\langle x,y\rangle + \tfrac{\overline{\langle x,y\rangle}}{\langle y,y\rangle}\tfrac{\langle x,y\rangle}{\langle y,y\rangle}\cdot\langle y,y\rangle \\ &= \langle x,x\rangle - \tfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle} - \tfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle} + \tfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle}, \end{aligned}$$

also

$$\big|\langle x,y\rangle\big|^2 \leq \langle x,x\rangle\cdot\langle y,y\rangle. \tag{98}$$

Durch Ziehen der positiven Wurzel folgt die gesuchte Ungleichung (96).

Nun sind x und y genau dann linear abhängig, wenn es ein $\lambda \in \mathbb{K}$ gibt, für das $x = \lambda y$ gilt. Das wiederum ist wegen der positiven Definitheit von $\langle \cdot, \cdot \rangle$ gleichbedeutend dazu, daß es ein $\lambda \in \mathbb{K}$ gibt, für das in (97) das Gleichheitszeichen gilt. Dieses λ ist eindeutig bestimmt, und erfüllt

$$\lambda = \lambda \cdot \frac{\langle y, y \rangle}{\langle y, y \rangle} = \frac{\langle y, \lambda y \rangle}{\langle y, y \rangle} = \frac{\langle y, x \rangle}{\langle y, y \rangle} = \frac{\overline{\langle x, y \rangle}}{\langle y, y \rangle}.$$

Damit ist die Gleichheit in (97) gleichwertig zur Gleichheit in (98). □

Satz 36.7 (Die euklidische Norm zu einem Skalarprodukt).
Es sei V ein euklidischer oder unitärer Raum. Dann wird durch

$$|| \cdot || : V \to \mathbb{R}_{\geq 0} : x \mapsto \sqrt{\langle x, x \rangle}$$

eine Norm auf V definiert. Wir werden euklidische und unitäre Räume im folgenden stets mit dieser zugehörigen euklidischen Norm als normierte Räume betrachten, so daß die Cauchy-Schwarzsche Ungleichung die folgende Form hat:

$$\big|\langle x, y \rangle\big| \leq ||x|| \cdot ||y||.$$

Beweis: Seien $x \in V$ und $\lambda \in \mathbb{K}$. Aus der positiven Definitheit von $\langle \cdot, \cdot \rangle$ folgt, daß $\langle x, x \rangle \geq 0$ und somit $||x||$ definiert und stets nicht-negativ ist. Ferner folgt, daß $||x|| = 0$ genau dann gilt, wenn x der Nullvektor ist. Aus der Bilinearität bzw. Sesquilinearität von $\langle \cdot, \cdot \rangle$ leiten wir her, daß

$$\langle \lambda x, \lambda x \rangle = \overline{\lambda} \lambda \langle x, x \rangle = |\lambda|^2 \langle x, x \rangle,$$

und somit $||\lambda x|| = |\lambda| \cdot ||x||$.

Allein, die Dreiecksungleichung ist etwas schwieriger zu zeigen. Wir verwenden hierfür die Cauchy-Schwarzsche Ungleichung aus Satz 36.6. Beachten wir noch, daß für eine komplexe Zahl $c \in \mathbb{C}$ stets

$$\mathrm{Re}(c) = \tfrac{1}{2} \cdot \big(c + \overline{c}\big)$$

gilt, so erhalten wir für $x, y \in V$

$$\begin{aligned}
||x + y||^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \overline{\langle x, y \rangle} + \langle y, y \rangle \\
&= ||x||^2 + 2 \cdot \mathrm{Re}\big(\langle x, y \rangle\big) + ||y||^2 \\
&\leq ||x||^2 + 2 \cdot \big|\langle x, y \rangle\big| + ||y||^2 \\
&\leq ||x||^2 + 2 \cdot ||x|| \cdot ||y|| + ||y||^2 \\
&= \big(||x|| + ||y||\big)^2.
\end{aligned}$$

Hieraus folgt dann die Dreiecksungleichung. □

Bemerkung 36.8 (Winkel in euklidischen Räumen).
Die Cauchy-Schwarzsche Ungleichung erlaubt es uns nun, in einem beliebigen *euklidischen Raum* V Winkel zwischen zwei Vektoren zu definieren. Denn aus der Ungleichung (96) folgt für $0 \neq x, y \in V$

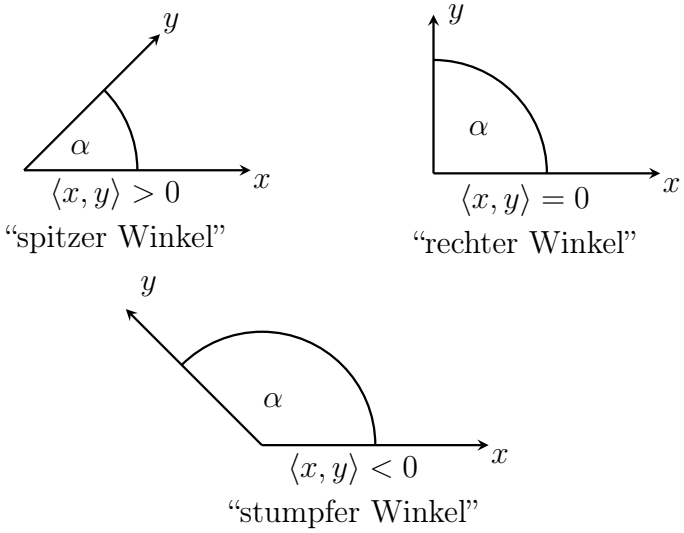
$$-1 \leq \frac{\langle x, y\rangle}{||x|| \cdot ||y||} \leq 1. \tag{99}$$

Aus Satz 16.12 wissen wir aber, daß es zu jeder reellen Zahl $-1 \leq r \leq 1$ genau einen Winkel $\alpha \in [0, \pi]$ gibt mit $r = \cos(\alpha)$, nämlich $\alpha = \arccos(r)$. Man definiert deshalb

$$\measuredangle(x, y) = \arccos\left(\frac{\langle x, y\rangle}{||x|| \cdot ||y||}\right) \in [0, \pi]$$

als den *Winkel* zwischen x und y.

Ist $\langle x, y\rangle > 0$, also $\measuredangle(x, y) \in \left[0, \frac{\pi}{2}\right[$, so spricht man von einem *spitzen Winkel*.
Ist $\langle x, y\rangle < 0$, also $\measuredangle(x, y) \in \left]\frac{\pi}{2}, \pi\right]$, so spricht man von einem *stumpfen Winkel*.
Ist $\langle x, y\rangle = 0$, also $\measuredangle(x, y) = \frac{\pi}{2}$, so spricht man von einem *rechten Winkel*.



C) Orthonormalbasen und Parsevalsche Gleichung

Definition 36.9 (Orthogonal).
Sei V ein euklidischer oder unitärer Raum, $x, y \in V$, $M, N \subseteq V$ und $U \leq V$.

a. x heißt *orthogonal* zu y, falls $\langle x, y\rangle = 0$. Wir schreiben dann $x \perp y$.

b. M heißt *orthogonal* zu N, falls $m \perp n$ für alle $m \in M$ und $n \in N$. Wir schreiben dann $M \perp N$.

c. Wir nennen $U^{\perp} := \{z \in V \mid z \perp U\}$ das *orthogonale Komplement* von U.

Lemma 36.10 (Orthogonales Komplement).
Ist V ein euklidischer oder unitärer Raum und $U \leq V$, dann ist $U^\perp \leq V$.

Beweis: Wegen $0 \in U^\perp$ ist $U^\perp \neq \emptyset$. Sind $x, y \in U^\perp$ und $\lambda, \mu \in \mathbb{K}$, so gilt für $z \in U$

$$\langle \lambda x + \mu y, z\rangle = \overline{\lambda}\langle x, z\rangle + \overline{\mu}\langle y, z\rangle = 0,$$

Also $\lambda x + \mu y \in U^\perp$. Damit ist $U^\perp$ ein Unterraum von V. □

Definition 36.11 (Orthonormalbasis).
Sei V ein euklidischer oder unitärer Raum. Eine Familie $B = (x_i \mid i \in I)$ von V heißt eine *orthonormal*, wenn

- $x_i \perp x_j$ für alle $i, j \in I$ mit $i \neq j$ und
- $||x_i|| = 1$ für alle $i \in I$ gilt,

d.h. wenn die Vektoren in B paarweise orthogonal zueinander und normiert sind. Ist B sogar eine Basis, nennt man sie eine *Orthonormalbasis*, kurz *ONB*.

Beispiel 36.12 (ONB).
Betrachten wir $\mathbb{K}^n$ mit dem kanonischen Skalarprodukt, dann ist die kanonische Basis $E = (e_1, \ldots, e_n)$ offenbar eine Orthonormalbasis von $\mathbb{K}^n$, da $\langle e_i, e_j\rangle = \delta_{ij}$ für $i, j \in \{1, \ldots, n\}$.

Lemma 36.13 (Orthogonal impliziert linear unabhängig.).
Sei V ein euklidischer oder unitärer Raum und $B = (x_i \mid i \in I)$ eine Familie paarweise orthogonaler Vektoren in $V \setminus \{0\}$. Dann ist B linear unabhängig.

Beweis: Aus $\sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i x_i = 0$ folgt für jedes $j \in I$

$$0 = \langle x_j, 0\rangle = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \langle x_j, x_i\rangle = \lambda_j \langle x_j, x_j\rangle.$$

Da $\langle x_j, x_j\rangle \neq 0$, muß also $\lambda_j = 0$ gelten. □

Proposition 36.14 (Parsevalsche Gleichung).
Sei V ein euklidischer oder unitärer Raum, $B = (x_i \mid i \in I)$ eine ONB und $x \in V$, dann gilt

$$x = \sum_{\substack{i\in I\\ \text{endlich}}} \langle x_i, x\rangle \cdot x_i. \tag{100}$$

Insbesondere sind nur endlich viele $\langle x_i, x\rangle$, $i \in I$, ungleich null.

Beweis: Da die Darstellung $x = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i x_i$ von x als endliche Linearkombination von B eindeutig ist, folgt die Behauptung aus

$$\langle x_j, x\rangle = \Big\langle x_j, \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i x_i \Big\rangle = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \langle x_j, x_i\rangle = \lambda_j \langle x_j, x_j\rangle = \lambda_j.$$

□

Bemerkung 36.15.
Ist B eine ONB von V, so erlaubt es die Gleichung (100), einen Vektor aus V als Linearkombination von B darzustellen, ohne hierzu eigens ein LGS lösen zu müssen, durch simples Einsetzen der Vektoren in das Skalarprodukt. Dieses Verfahren ist sehr effizient und von hoher praktischer Bedeutung. Die Tatsache, daß sich die Koordinaten eines Vektors bezüglich einer ONB mit Hilfe des Skalarproduktes so einfach ausdrücken lassen, spielt aber auch in vielen Beweisen eine Rolle, und ist somit ebenfalls für die Theorie von Bedeutung.

D) Das Orthonormalisierungsverfahren von Gram-Schmidt

Wir beweisen jetzt, daß jeder endlich-dimensionale euklidische bzw. unitäre Raum eine ONB besitzt. Etwas allgemeiner gilt der folgende Satz.

Satz 36.16 (Gram-Schmidt).
Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und B eine orthonormale Familie in V, dann läßt sich B zu einer ONB von V ergänzen.

Beweis: Ist B schon eine Basis von V, so sind wir fertig. Wir dürfen also annehmen, daß B keine Basis und wegen Lemma 36.13 dann auch kein Erzeugendensystem von V ist. Wir zeigen nun konstruktiv, wie wir die orthonormale Familie $B = (z_1, \ldots, z_r)$ zu einer orthonormalen Familie $(z_1, \ldots, z_{r+1})$ ergänzen können. Wenden wir dieses Verfahren dann $\dim_{\mathbb{K}}(V) - r$ mal an, so haben wir die Aussage bewiesen.

Dazu wählen wir zunächst einen Vektor x_{r+1}, der linear unabhängig von B ist. Dann setzen wir

$$y_{r+1} := x_{r+1} - \sum_{i=1}^{r} \langle z_i, x_{r+1}\rangle \cdot z_i. \tag{101}$$

Da x_{r+1} linear unabhängig von B ist, ist $y_{r+1} \neq 0$, und wir können deshalb

$$z_{r+1} := \tfrac{1}{\|y_{r+1}\|} \cdot y_{r+1} \tag{102}$$

setzen. Dann ist $||z_{r+1}|| = 1$ und außerdem gilt für $i = 1, \ldots, r$

$$\begin{aligned}\langle z_i, z_{r+1}\rangle &= \frac{1}{||y_{r+1}||} \cdot \langle z_i, y_{r+1}\rangle \\ &= \frac{1}{||y_{r+1}||} \cdot \left(\langle z_i, x_{r+1}\rangle - \sum_{j=1}^{r} \langle z_j, x_{r+1}\rangle \cdot \langle z_i, z_j\rangle\right) \\ &= \frac{1}{||y_{r+1}||} \cdot \big(\langle z_i, x_{r+1}\rangle - \langle z_i, x_{r+1}\rangle\big) = 0.\end{aligned}$$

Dann ist aber $(z_1, \ldots, z_{r+1})$ orthonormal und wir sind fertig. □

Korollar 36.17 (Existenz einer ONB).
Jeder endlich-dimensionale euklidische bzw. unitäre Raum besitzt eine ONB.

Beweis: Wende Satz 36.16 mit $B = \emptyset$ an. □

Der Beweis von Satz 36.16 ist konstruktiv und wird auch das *Gram-Schmidtsche Orthonormalisierungsverfahren* genannt. Es erlaubt, aus einem gegebenen Erzeugendensystem eine ONB zu konsturieren.

Algorithmus 36.18 (Gram-Schmidt-Orthonormalisierungsverfahren).
INPUT: $M \subseteq \mathbb{K}^n$ und ein Skalarprodukt $\langle \cdot, \cdot \rangle$ auf $\mathbb{K}^n$
OUTPUT: ONB B von $\langle M \rangle$

1. **Schritt:** Bestimme eine Basis $B = (x_1, \ldots, x_r)$ von $\mathrm{Lin}(M)$, z. B. mittels Algorithmus 32.20.
2. **Schritt:** Für $i = 1, \ldots, r$ führe man folgende Schritte aus:
   Schritt a.: berechne die Summe $y_i = x_i - \sum_{j=1}^{i-1} \langle z_j, x_i\rangle \cdot z_j$;
   Schritt b.: berechne $z_i = \frac{1}{||y_i||} \cdot y_i$;
3. **Schritt:** Gib die veränderte Basis $(z_1, \ldots, z_r)$ zurück.

Bemerkung 36.19.

a. Will man in der Praxis ein Skalarprodukt übergeben, so wird man im reellen Fall eine symmetrische Matrix A übergeben und im komplexen Fall eine hermitesche (siehe Bemerkung 37.23). Das Skalarprodukt wird dann durch

$$\langle x, y\rangle = x^t \circ A \circ \overline{y}$$

gebildet.

b. Um zu normieren, ist in Algorithmus 36.18 das Ziehen von Wurzeln notwendig. Verzichtet man jedoch auf die Normierung der Vektoren, so kommt man ohne Wurzelziehen aus. Läßt man im obigen Algorithmus Schritt 2.b. weg und ersetzt

dafür in Schritt 2.a. die rechte Seite der Gleichung durch

$$y_i = x_i - \sum_{j=1}^{i-1} \frac{\langle y_j, x_i\rangle}{\langle y_j, y_j\rangle} \cdot y_j,$$

dann liefert Algorithmus 36.18 eine Basis $(y_1, \ldots, y_r)$ paarweise orthogonaler Vektoren von $\mathrm{Lin}(M)$. Das hat den Vorteil, daß man exakt rechnen kann - etwa in SINGULAR, wenn die Eingabedaten rationale Zahlen waren.

Beispiel 36.20 (Gram-Schmidt-Orthonormalisierungsverfahren).
Es sei $B = (x_1, x_2, x_3) = \{(1,0,1)^t, (1,1,1)^t, (0,0,4)^t\} \subseteq \mathbb{R}^3$, wobei wir $\mathbb{R}^3$ mit dem kanonischen Skalarprodukt versehen betrachten. Man sieht leicht, daß B bereits eine Basis von $\mathbb{R}^3$ ist. Wir wollen hier B in eine ONB von $\mathbb{R}^3$ überführen.

Wir setzen nun $y_1 := (1,0,1)^t$, dann ist $\langle y_1, y_1\rangle = 2$ und somit ersetzen wir x_1 in B durch

$$z_1 = \frac{1}{||y_1||} \cdot y_1 = \tfrac{1}{\sqrt{2}} \cdot y_1 = \left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)^t.$$

Im nächsten Schritt setzen wir

$$y_2 = x_2 - \langle z_1, x_2\rangle \cdot z_1 = (1,1,1)^t - \tfrac{2}{\sqrt{2}} \cdot \left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)^t = (0,1,0)^t.$$

Dann ist $\langle y_2, y_2\rangle = 1$ und somit ersetzen wir x_2 in B durch $z_2 = y_2$.

Schließlich bilden wir

$$\begin{aligned} y_3 &= x_3 - \langle z_1, x_3\rangle \cdot z_1 - \langle z_2, x_3\rangle \cdot z_2 \\ &= (0,0,4)^t - \tfrac{4}{\sqrt{2}} \cdot \left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)^t - 0 \cdot (0,1,0)^t \\ &= (-2,0,2)^t, \end{aligned}$$

und erhalten $\langle y_3, y_3\rangle = 8$. Somit müssen wir x_3 durch den Vektor

$$z_3 = \tfrac{1}{||y_3||} \cdot y_3 = \tfrac{1}{\sqrt{8}} \cdot (-2,0,2)^t = \left(-\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)^t$$

ersetzen. Damit ergibt sich aus dem Gram-Schmidtschen Orthonormalisierungsverfahren die ONB

$$\left(\left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)^t, (0,1,0)^t, \left(-\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)^t\right).$$

E) Orthogonale und unitäre Matrizen

Schreibt man die Vektoren einer Orthonormalbasis als Spalten in eine Matrix, so erhält man eine invertierbare Matrix, deren Inverse sich durch einfaches Transponieren und komplexes Konjugieren berechnen läßt. Dies führt zum Begriff der orthognalen bzw. unitären Matrizen.

Definition 36.21 (Orthogonale / unitäre Matrizen).

a. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ heißt *orthogonal*, wenn $A^t \circ A = \mathbb{1}_n$ gilt. Wir nennen $\mathrm{O}(n) := \{A \in \mathrm{Mat}_n(\mathbb{R}) \mid A \text{ ist orthogonal}\}$ *orthogonale Gruppe* vom Grad n.

b. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ heißt *unitär*, wenn $A^* \circ A = \mathbb{1}_n$ gilt, und wir nennen $\mathrm{U}(n) := \{A \in \mathrm{Mat}_n(\mathbb{C}) \mid A \text{ ist unitär}\}$ die *unitäre Gruppe* vom Grad n.

Proposition 36.22 (Determinante orthogonaler / unitärer Matrizen).
Ist $A \in \mathrm{Mat}_n(\mathbb{K})$ orthogonal oder unitär, so gilt $|\det(A)| = 1$.

Beweis: Wegen $A^* \circ A = \mathbb{1}_n$ gilt

$$\begin{aligned} 1 = \det(\mathbb{1}_n) &= \det(A^* \circ A) = \det(A^*) \cdot \det(A) \\ &= \overline{\det(A^t)} \cdot \det(A) = \overline{\det(A)} \cdot \det(A) = |\det(A)|^2. \end{aligned}$$

Daraus folgt $|\det(A)| = 1$. □

Proposition 36.23 (Orthogonale / unitäre Matrizen).
Für eine quadratische Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ sind äquivalent:

a. A ist orthogonal bzw. unitär.

b. A ist invertierbar mit $A^* = A^{-1}$.

c. Die Spalten von A sind eine ONB von $\mathbb{K}^n$ mit kanonischem Skalarprodukt.

d. Die Zeilen von A sind eine ONB von $\mathbb{K}^n$ mit kanonischem Skalarprodukt.

Beweis: Die Äquivalenz von a. und b. folgt unmittelbar aus der Definition, denn $A^* \circ A = \mathbb{1}_n$ heißt, daß A^* die Inverse von A ist.

Ist a^i der i-te Spaltenvektor von A, so ist $\overline{a^i}^t$ der i-te Zeilenvektor von A^* und deshalb ist

$$\langle a^i, a^j \rangle = \overline{a^i}^t \circ a^j$$

der Eintrag an der Stelle (i, j) von $A^* \circ A$. Deshalb sind die Spalten von A genau dann eine ONB von $\mathbb{K}^n$, wenn $A^* = A^{-1}$ die Inverse von A ist. Dies zeigt die Äquivalenz von b. und c..

Ist a_i der i-te Zeilenvektor von A, so ist $\overline{a_i}^t$ der i-te Spaltenvektor von A^*. Also ist

$$\overline{\langle a_i^t, a_j^t \rangle} = a_i \circ \overline{a_j}^t$$

der Eintrag von $A \circ A^*$ an der Stelle (i,j). Dies zeigt schließlich, daß b. und d. äquivalent sind. □

Beispiel 36.24 (Orthogonale Matrix).
Die Matrix

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$$

ist orthogonal, da ihre Spalten nach Beispiel 36.20 eine ONB von $\mathbb{R}^3$ bezüglich des kanonischen Skalarproduktes sind.

Korollar 36.25 (Die orthogonale und die unitäre Gruppe).
$(\mathrm{O}(n), \circ)$ und $(\mathrm{U}(n), \circ)$ sind Gruppen.

Beweis: Es reicht, zu zeigen, daß sie Untergruppen von $\mathrm{Gl}_n(\mathbb{C})$ sind. Offenbar sind $\mathrm{O}(n)$ und $\mathrm{U}(n)$ nicht-leere Teilmengen von $\mathrm{Gl}_n(\mathbb{C})$. Sind nun A und B in $\mathrm{O}(n)$ bzw. in $\mathrm{U}(n)$, so gilt

$$(A \circ B)^* = B^* \circ A^* = B^{-1} \circ A^{-1} = (A \circ B)^{-1},$$

und

$$\left(A^{-1}\right)^* = A^{**} = A = \left(A^{-1}\right)^{-1}.$$

Mithin liegen auch $A \circ B$ und A^{-1} in $\mathrm{O}(n)$ bzw. in $\mathrm{U}(n)$. Damit ist gezeigt, daß $\mathrm{O}(n)$ und $\mathrm{U}(n)$ Untergruppen von $\mathrm{Gl}_n(\mathbb{C})$ sind. □

Bemerkung 36.26 (Die orthogonale Gruppe $\mathrm{O}(2)$).
Die Determinante

$$\det : \mathrm{O}(2) \longrightarrow \{1, -1\}$$

ist wegen des Determinantenmultiplikationssatzes ein Gruppenepimorphismus. Der Kern von det ist der Normalteiler

$$\mathrm{SO}(2) := \{A \in \mathrm{O}(2) \mid \det(A) = 1\}$$

von $\mathrm{O}(2)$ und wird die *spezielle orthogonale Gruppe* vom Grad 2 genannt. Wir werden unten zeigen, daß

$$\mathrm{SO}(2) = \{T(\alpha) \mid \alpha \in [0, 2\pi)\}$$

und

$$\mathrm{O}(2) \setminus \mathrm{SO}(2) = \{S(\alpha) \mid \alpha \in [0, 2\pi)\},$$

wobei

$$T(\alpha) := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

eine Drehung um den Winkel α ist und

$$S(\alpha) := \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}$$

eine Spiegelung an der Geraden $\mathrm{Lin}\left(\left(\cos\left(\frac{\alpha}{2}\right), \sin\left(\frac{\alpha}{2}\right)\right)^t\right)$. Insbesondere ist im Fall $n = 2$ also jede orthogonale Matrix eine Drehung oder eine Spieglung.

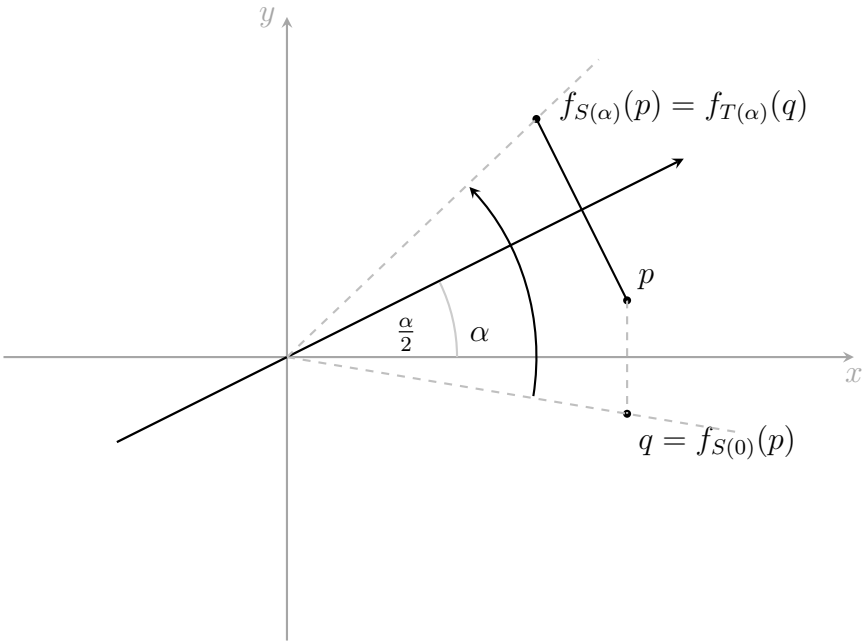


Abbildung 1: Die Spielung $S(\alpha) = T(\alpha) \circ S(0)$

Man beachte auch daß $S(\alpha) = T(\alpha) \circ S(0)$ d. h. die von $S(\alpha)$ induzierte Spiegelung ist Komposition der Spiegelung an der x-Achse gefolgt von einer Drehung um den Winkel α. Damit gilt zugleich, daß jede Drehung im $\mathbb{R}^2$ Komposition von zwei Spiegelungen ist.

Beweis: Die Matrix

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$$

ist genau dann orthogonal, wenn die beiden Spaltenvektoren $x = (a, b)^t$ und $y = (c, d)^t$ eine Orthonormalbasis von $\mathbb{R}^2$ bezüglich des Standardskalarproduktes sind. Insbesondere muß y also senkrecht auf x stehen. In der Ebene ist ein Vektor, der senkrecht steht auf x aber bis auf einen Skalarfaktor eindeutig bestimmt und $(-b, a)^t$ ist ein solcher Vektor. Es muß also

$$y = \lambda \cdot (-b, a)^t$$

gelten. Aus

$$1 = ||y|| = |\lambda| \cdot \sqrt{(-b)^2 + a^2} = |\lambda| \cdot ||x|| = |\lambda|$$

folgt dann $\lambda = 1$ oder $\lambda = -1$. Also ist die Matrix A von der Form

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{oder} \quad A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix},$$

wobei die Zahlen $a, b \in \mathbb{R}$ nur die Bedingung

$$a^2 + b^2 = ||x||^2 = 1$$

erfüllen müssen. Aus dem Satz von Pythagoras wissen wir aber, daß es dann genau einen Winkel $\alpha \in [0, 2\pi)$ gibt mit

$$a = \cos(\alpha) \quad \text{und} \quad b = \sin(\alpha),$$

und somit

$$A = T(\alpha) \quad \text{oder} \quad A = S(\alpha).$$

Beachten wir nun noch, daß

$$\det(T(\alpha)) = \cos(\alpha)^2 + \sin(\alpha)^2 = 1$$

und

$$\det(S(\alpha)) = -\cos(\alpha)^2 - \sin(\alpha)^2 = -1$$

ist, so ist

$$\mathrm{SO}(2) = \{T(\alpha) \mid \alpha \in [0, 2\pi)\}$$

und

$$\mathrm{O}(2) \setminus \mathrm{SO}(2) = \{S(\alpha) \mid \alpha \in [0, 2\pi)\},$$

gezeigt. □

F) Orthogonale Summe und orthogonale Projektion

Definition 36.27 (Orthogonale Summe).
Sei V ein euklidischer oder unitärer Raum. Wir nennen V die *orthogonale Summe* der Unterräume $U_1, \ldots, U_r$, falls $V = U_1 \oplus \ldots \oplus U_r$ und $U_i \perp U_j$ für $i \neq j$. In diesem Fall schreiben wir $V = U_1 \perp \ldots \perp U_r$.

Proposition 36.28 (Orthogonales Komplement).
Ist V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$, so gilt

$$V = U \perp U^{\perp}.$$

Insbesondere, $U^{\perp}$ ist ein Komplement von U.

Beweis: Da nach Voraussetzung $U \perp U^\perp$ gilt, bleibt $U \cap U^\perp = \{0\}$ und $V = U + U^\perp$ zu zeigen, wobei für letzteres auch $V \subseteq U + U^\perp$ reicht.

Ist $x \in U \cap U^\perp$, so gilt $\langle x, x\rangle = 0$ und damit $x = 0$, da das Skalarprodukt positiv definit ist. Also ist $U \cap U^\perp = \{0\}$.

Zudem können wir wegen Satz 36.16 eine ONB $(x_1, \ldots, x_r)$ von U wählen und diese zu einer ONB $(x_1, \ldots, x_n)$ von V ergänzen. Dann gilt aber

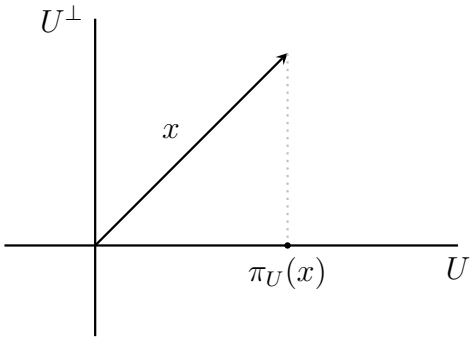
$$V = \operatorname{Lin}(x_1, \ldots, x_r) + \operatorname{Lin}(x_{r+1}, \ldots, x_n) \subseteq U + U^\perp,$$

da nach Wahl $x_{r+1}, \ldots, x_n \in U^\perp$. Hierbei beachte man, daß ein Vektor, der orthogonal zu einer Basis von U ist, automatisch orthogonal zu jedem Vektor in U ist. □

Bemerkung 36.29.
Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$. Da sich jeder Vektor $x \in V$ in eindeutiger Weise darstellen läßt als $x = u + u'$ mit $u \in U$ und $u' \in U^\perp$, können wir die *orthogonale Projektion* von V auf U

$$\pi_U : V \to V$$

definieren durch $\pi\big(u + u'\big) = u$ für $u \in U$ und $u' \in U^\perp$.



In Aufgabe 36.32 wird gezeigt, daß π_U in der Tat eine Projektion ist, d.h. π_U ist linear mit $\pi_U^2 = \pi_U$. Außerdem ist $\operatorname{Im}(\pi_U) = U$ das Bild von π_U und $\operatorname{Ker}(\pi_U) = U^\perp$ der Kern.

Bemerkung 36.30 (Determinante als Volumenform).
In Bemerkung 34.12 haben wir das Parallelotop

$$P\big(x, y, z\big) := \big\{\lambda x + \mu y + \nu z \in \mathbb{R}^3 \mid 0 \leq \lambda, \mu, \nu \leq 1\big\}$$

betrachtet, das von den Vektoren $0 \neq x, y, z \in \mathbb{R}^3$ aufgespannt wird (siehe Abbildung 2).

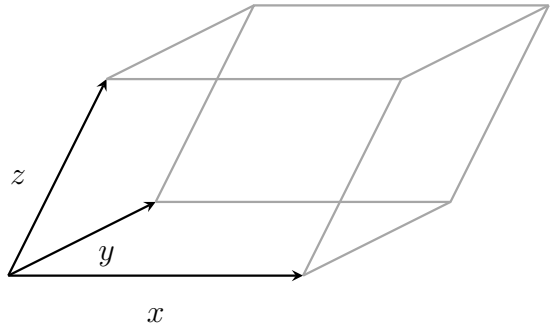


Abbildung 2: Das Parallelotop $P(x, y, z)$ im $\mathbb{R}^3$

Wir wollen zeigen, daß das Volumen

$$\text{Volumen}\big(P(x,y,z)\big) = |\det(x \;\; y \;\; z)|$$

die Determinante der Matrix ist, deren Spalten die Vektoren x, y und z sind.

Elementargeometrisch berechnet sich das Volumen von $P(x, y, z)$ als Grundfläche mal Höhe, d.h. als Fläche A des von x und y aufgespannten Parallelogramms multipliziert mit der Höhe h des Parallelogramms. Dabei berechnet sich A als Länge von x mal der Höhe h' des Parallelogramms.

Wenden wir uns zunächst letzterer Berechnung zu. Es sei $U = \text{Lin}\,(x)$ und π_U sei die orthogonale Projektion auf U. Dann ist die Höhe h' des Parallelogramms genau die Länge des Vektors $y - \pi_U(y)$ (siehe Abbildung 3).

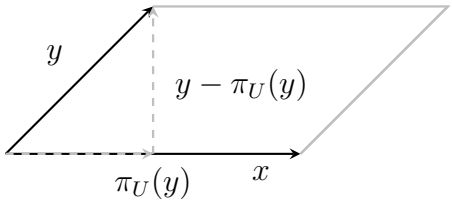


Abbildung 3: Die Höhe im Parallelogramm zu x und y.

Für die Fläche A des Parallelogramms gilt deshalb

$$A = ||x|| \cdot ||y - \pi_U(y)||.$$

Auf ähnliche Weise kann man die Höhe h des Parallelotops $P(x, y, z)$ bestimmen. Hierzu betrachten wir den Unterraum $W = \text{Lin}\,(x, y)$ und die orthogonale Projektion π_W auf W. Dann ist h die Länge des Vektors $z - \pi_W(z)$ (siehe Abbildung 4).

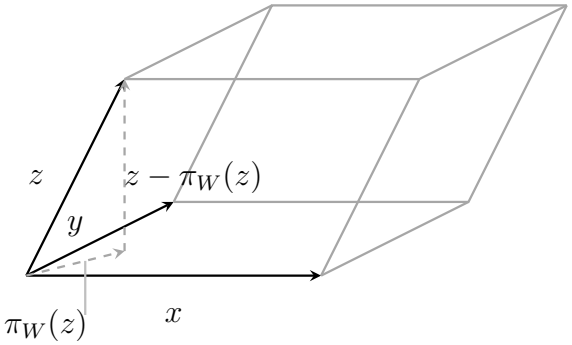


Abbildung 4: Die Höhe in $P(x,y,z)$

Für das Volumen von $P(x,y,z)$ erhalten wir deshalb

$$\text{Volumen}\big(P(x,y,z)\big) = A \cdot h = ||x|| \cdot ||y - \pi_U(y)|| \cdot ||z - \pi_W(z)||.$$

Wegen $\pi_U(y) \in U = \text{Lin}\,(x)$ und $\pi_W(z) \in W = \text{Lin}\,(x,y)$, gibt es $\lambda, \mu, \nu \in \mathbb{R}$ mit

$$\pi_U(y) = \lambda x \quad \text{und} \quad \pi_W(z) = \mu x + \nu y.$$

Dann gilt aber

$$\det(x \;\; y \;\; z) = \det(x \;\; y - \lambda x \;\; z - \mu x - \nu y) = \det(x \;\; y - \pi_U(y) \;\; z - \pi_W(z)),$$

da sich die Determinante einer Matrix nicht ändert, wenn wir Vielfache einer Spalte zu einer anderen addieren. Nun beachten wir, daß nach Konstruktion die Spalten der rechten Matrix orthogonal zueinander sind (siehe Abbildung 3 und 4). Normieren wir sie, so bilden sie eine ONB von $\mathbb{R}^3$ und die Matrix wird orthogonal. Da die Determinante einer orthogonalen Matrix Betrag 1 hat, erhalten wir also

$$\begin{aligned} |\det(x \;\; y \;\; z)| &= ||x|| \cdot ||y - \pi_U(y)|| \cdot ||z - \pi_W(z)|| \cdot \left|\det\left(\tfrac{x}{||x||} \;\; \tfrac{y-\pi_U(y)}{||y-\pi_U(y)||} \;\; \tfrac{z-\pi_W(z)}{||z-\pi_W(z)||}\right)\right| \\ &= ||x|| \cdot ||y - \pi_U(y)|| \cdot ||z - \pi_W(z)||. \end{aligned}$$

Dies beweist die Aussage und begründet den Begriff *Volumenform* im Zusammenhang mit Determinanten. Man beachte auch, daß die entsprechende Aussage für Parallelogramme analog gezeigt werden kann.

Aufgaben

Aufgabe 36.31.
Es sei $V = \mathcal{C}([0,1], \mathbb{R})$ der Vektorraum der auf $[0,1]$ stetigen Funktionen. Zeige, daß durch

$$\langle f, g \rangle := \int_0^1 f(x) \cdot g(x)\, dx$$

ein Skalarprodukt auf V definiert wird.

Aufgabe 36.32 (Orthogonale Projektion).
Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$.

a. Zeige, $\pi_U \in \mathrm{End}_{\mathbb{K}}(V)$ ist eine Projektion mit $\mathrm{Ker}(\pi_U) = U^\perp$ und $\mathrm{Im}(\pi_U) = U$.

b. Zeige, ist $\pi \in \mathrm{End}_{\mathbb{K}}(V)$ eine Projektion mit $\mathrm{Ker}(\pi) = U^\perp$ und $\mathrm{Im}(\pi) = U$, dann ist $\pi = \pi_U$.

c. Ist $(x_1, \ldots, x_r)$ eine ONB von U und $x \in V$, dann gilt

$$\pi_U(x) = \sum_{i=1}^{r} \langle x_i, x \rangle \cdot x_i.$$

Aufgabe 36.33.
Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$. Dann gilt $(U^\perp)^\perp = U$.

Aufgabe 36.34.
Zeige, durch $\langle (x_1, x_2, x_3)^t, (y_1, y_2, y_3)^t \rangle := x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2 + x_2y_3 + x_3y_2 + 2x_3y_3$ für $(x_1, x_2, x_3)^t, (y_1, y_2, y_3)^t \in \mathbb{R}^3$ wird ein Skalarprodukt auf $\mathbb{R}^3$ definiert und bestimme eine Orthonormalbasis des $\mathbb{R}^3$ bezüglich dieses Skalarproduktes.

Aufgabe 36.35.
Bestimme eine Orthonormalbasis des $\mathbb{R}$-Vektorraums $U = \{(v, w, x, y, z)^t \in \mathbb{R}^5 \mid v + w + x + y + z = 0\}$ bezüglich des Standardskalarproduktes auf $\mathbb{R}^5$.

Aufgabe 36.36 (Legendre-Polynome).
Betrachte den Vektorraum $U = \{f : [0, 1] \to \mathbb{R} : x \mapsto ax^2 + bx + c \mid a, b, c \in \mathbb{R}\}$ und bestimme eine ONB bezüglich des Skalarproduktes aus Aufgabe 36.31.

Aufgabe 36.37 (Tschebyscheff-Polynome).

a. Zeige, daß auf $V = \mathcal{C}([-1, 1], \mathbb{R})$ durch

$$V \times V \longrightarrow \mathbb{R} : (f, g) \mapsto \langle f, g \rangle := \int_{-1}^{1} f(x) \cdot g(x) \cdot \frac{1}{\sqrt{1 - x^2}}\, dx$$

ein Skalarprodukt definiert wird.

b. Berechne für den Unterraum $U\{f : [-1, 1] \to \mathbb{R} : x \mapsto ax^2 + bx + c \mid a, b, c \in \mathbb{R}\}$ von V eine Orthonormalbasis bezüglich des Skalarproduktes aus Teil a..

Hinweis, substituiere in $x = \cos(t)$, um die Konvergenz des uneigentlichen Integrals zu zeigen.

Aufgabe 36.38.
Für $V = \mathrm{Mat}_n(\mathbb{R})$ definieren wir $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R} : (A, B) \mapsto \mathrm{Spur}\left(A^t \circ B\right)$.

a. Zeige, $\langle \cdot, \cdot \rangle$ ist ein Skalarprodukt auf V.

b. Zeige, für $U = \{A \in V \mid A^t = A\}$ gilt $U^\perp = \{A \in V \mid A^t = -A\}$.

Aufgabe 36.39.
Sei V ein euklidischer oder unitärer Raum und $|| \cdot || : V \to \mathbb{R} : x \mapsto \sqrt{\langle x, x \rangle}$ die durch das Skalarprodukt definierte Norm. Zeige, für $x, y \in V$ gelten:

a. Die Parallelogrammgleichung: $||x + y||^2 + ||x - y||^2 = 2 \cdot \left(||x||^2 + ||y||^2\right)$.

b. Der Satz des Pythagoras': $x \perp y \implies ||x||^2 + ||y||^2 = ||x + y||^2$.

Aufgabe 36.40.
Zeige, für $a, b \in \mathbb{R}$ gibt es genau dann zwei normierte Vektoren $x = (u, v, a)^t \in \mathbb{R}^3$ und $y = (r, s, b)^t \in \mathbb{R}^3$ die bezüglich des Standardskalarproduktes orthogonal zueinander sind, wenn $a^2 + b^2 \leq 1$.

Aufgabe 36.41 (Der p-adische Betrag).
Sei p eine Primzahl. Für $0 \neq a \in \mathbb{Z}$ bezeichne $\nu_p(a)$ die höchste Potenz von p, die a teilt, und für $0 \neq q = \frac{b}{c} \in \mathbb{Q}$ setzen wir $\nu_p(q) = \nu_p(b) - \nu_p(c)$. Zeige, die Abbildung

$$|\cdot|_p : \mathbb{Q} \longrightarrow \mathbb{R} : q \mapsto \begin{cases} p^{-\nu_p(q)}, & \text{wenn } q \neq 0, \\ 0, & \text{wenn } q = 0. \end{cases}$$

ist positiv definit, multiplikativ und genügt der Dreiecksungleichung.

§ 37 Spektralsatz und Hauptachsentransformation

In diesem Abschnitt sei V ein euklidischer oder unitärer Raum der Dimension $1 \leq n < \infty$ mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und euklidischer Norm $|| \cdot ||$.

A) Die adjungierte Abbildung

Satz 37.1 (Die adjungierte Abbildung).
Zu jedem Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ gibt es genau ein $f^* \in \mathrm{End}_{\mathbb{K}}(V)$, so daß

$$\langle f(x), y \rangle = \langle x, f^*(y) \rangle \tag{103}$$

für alle $x, y \in V$ gilt. Ist $B = (x_1, \ldots, x_n)$ eine ONB von V, so gilt für $y \in V$

$$f^*(y) = \sum_{i=1}^{n} \langle f(x_i), y \rangle \cdot x_i. \tag{104}$$

Die Abbildung f^* heißt die *adjungierte Abbildung* zu f.

Beweis: Wir wollen zunächst zeigen, daß es einen Endomorphismus $f^* \in \mathrm{End}_{\mathbb{K}}(V)$ mit der Eigenschaft (103) gibt. Dazu wählen wir eine ONB $B = (x_1, \ldots, x_n)$ von V und definieren

$$f^* : V \longrightarrow V : y \mapsto \sum_{i=1}^{n} \langle f(x_i), y \rangle \cdot x_i,$$

d.h. wir definieren $f^*(y)$ durch die Formel in (104). Da das Skalarprodukt in der zweiten Komponente linear ist, ist f^* in der Tat eine lineare Abbildung, also ein Endomorphismus von V.

Seien nun $x, y \in V$ gegeben. Unter Anwendung der Parsevalschen Gleichung 36.14 gilt dann

$$\begin{aligned}
\langle f(x), y \rangle &\overset{36.14}{=} \left\langle f\left(\sum_{i=1}^{n} \langle x_i, x \rangle \cdot x_i\right), y \right\rangle = \left\langle \sum_{i=1}^{n} \langle x_i, x \rangle \cdot f(x_i), y \right\rangle \\
&= \sum_{i=1}^{n} \overline{\langle x_i, x \rangle} \cdot \langle f(x_i), y \rangle = \sum_{i=1}^{n} \langle x, x_i \rangle \cdot \langle f(x_i), y \rangle \\
&= \sum_{i=1}^{n} \langle x, \langle f(x_i), y \rangle \cdot x_i \rangle = \left\langle x, \sum_{i=1}^{n} \langle f(x_i), y \rangle \cdot x_i \right\rangle = \langle x, f^*(y) \rangle.
\end{aligned}$$

Damit ist gezeigt, daß der durch (104) definierte Endomorphismus die Gleichung (103) erfüllt.

Es bleibt noch die Eindeutigkeit zu zeigen. Sei dazu $h \in \mathrm{End}_{\mathbb{K}}(V)$ mit

$$\langle f(x), y\rangle = \langle x, h(y)\rangle \tag{105}$$

für alle $x, y \in V$. Wegen der Parsevalschen Gleichung 36.14 gilt für $y \in V$ dann

$$h(y) \overset{36.14}{=} \sum_{i=1}^{n} \langle x_i, h(y)\rangle \cdot x_i \overset{(105)}{=} \sum_{i=1}^{n} \langle f(x_i), y\rangle \cdot x_i \overset{(104)}{=} f^*(y).$$

Mithin stimmen f^* und h überein, so daß f^* eindeutig bestimmt ist. □

Korollar 37.2 ($f^{**} = f$).
Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$, so gilt $f^{**} = f$, d.h. $\langle f^*(x), y\rangle = \langle x, f(y)\rangle$ für $x, y \in V$.

Beweis: Für $x, y \in V$ gilt

$$\langle f^*(x), y\rangle = \overline{\langle y, f^*(x)\rangle} \overset{(103)}{=} \overline{\langle f(y), x\rangle} = \langle x, f(y)\rangle.$$

Damit erfüllt f die Bedingung, durch die die Abbildung f^{**} eindeutig festgelegt ist. Also muß $f = f^{**}$ gelten. □

Definition 37.3 (Adjungierte Matrix).
Für eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ heißt $A^* = \overline{A}^t$ die zu A *adjungierte Matrix*.
Man beachte, daß über den reellen Zahlen die adungierte Matrix gerade die Transponierte ist.

Korollar 37.4 (Matrixdarstellung der adjungierten Abbildung).
Sei B eine ONB von V und $f \in \mathrm{End}_{\mathbb{K}}(V)$, so gilt

$$M_B^B(f^*) = M_B^B(f)^* := \overline{M_B^B(f)}^t,$$

d.h. die Matrixdarstellung der adjungierten Abbildung ist die Adjungierte der Matrixdarstellung.

Beweis: Seien $M_B^B(f) = (a_{ij})_{i,j}$ und $M_B^B(f^*) = (b_{ji})_{j,i}$. Unter Berücksichtigung der Parsevalschen Gleichung 36.14 gilt

$$\sum_{i=1}^{n} a_{ij} \cdot x_i = f(x_j) \overset{36.14}{=} \sum_{i=1}^{n} \langle x_i, f(x_j)\rangle \cdot x_i \tag{106}$$

und

$$\sum_{j=1}^{n} b_{ji} \cdot x_j = f^*(x_i) \overset{36.14}{=} \sum_{j=1}^{n} \langle x_j, f^*(x_i)\rangle \cdot x_j. \tag{107}$$

Da die Darstellung als Linearkombination einer Basis eindeutig ist, erhalten wir

$$\overline{a_{ij}} \overset{(106)}{=} \overline{\langle x_i, f(x_j)\rangle} = \langle f(x_j), x_i\rangle \overset{(103)}{=} \langle x_j, f^*(x_i)\rangle \overset{(107)}{=} b_{ji}.$$

Daraus folgt

$$M_B^B(f^*) = (b_{ji})_{j,i} = \left(\overline{a_{ij}}\right)_{j,i} = \left(\overline{a_{ij}}\right)_{i,j}^t = M_B^B(f)^*.$$

□

Beispiel 37.5 (Adjungierte).
Wir betrachten $V = \mathbb{C}^2$ als unitären Raum mit dem kanonischen Skalarprodukt sowie die Abbildung

$$f : \mathbb{C}^2 \longrightarrow \mathbb{C}^2 : (x, y)^t \mapsto (2x + 4y, 2y - 4x)^t.$$

Bezüglich der kanonischen Basis E hat f die Matrixdarstellung

$$M_E^E(f) = \begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix}$$

und somit gilt

$$M_E^E(f)^* = \overline{\begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix}}^t = \begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix}^t = \begin{pmatrix} 2 & -4 \\ 4 & 2 \end{pmatrix}.$$

Da E eine ONB bezüglich des kanonischen Skalarproduktes ist, ist somit

$$f^* : \mathbb{C}^2 \longrightarrow \mathbb{C}^2 : (x, y)^t \mapsto (2x - 4y, 2y + 4x)^t$$

nach Korollar 37.4 die Adjungierte von f.

B) Spektralsatz für selbstadjungierte Endomorphismen

Definition 37.6 (Selbstadjungierter Endomorphismus).

a. $f \in \mathrm{End}_{\mathbb{K}}(V)$ heißt *selbstadjungiert* oder *hermitesch*, wenn $f = f^*$ gilt.

b. $A \in \mathrm{Mat}_n(\mathbb{C})$ heißt *hermitesch*, wenn $A = A^*$ gilt.

Proposition 37.7 (Matrixdarstellung selbstadjungierter Endomorph.).
Es sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ und B eine ONB von V. Genau dann ist f selbstadjungiert, wenn $M_B^B(f)$ symmetrisch bzw. hermitesch ist.

Beweis: Aus Korollar 37.4 wissen wir, daß $M_B^B(f^*) = M_B^B(f)^*$ gilt. Deshalb gilt

$$\begin{aligned} f \text{ selbstadjungiert} &\iff f = f^* \iff M_B^B(f) = M_B^B(f^*) = M_B^B(f)^* \\ &\iff M_B^B(f) \text{ symmetrisch bzw. hermitesch.} \end{aligned}$$

□

Lemma 37.8 (Eigenwerte selbstadjungierter Endomorphismen).
Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ selbstadjungiert, dann ist $\chi_f \in \mathbb{R}[t]$ und χ_f zerfällt über $\mathbb{R}$. Insbesondere gilt, ist $\lambda \in \sigma(f)$ ein Eigenwert von f, dann ist $\lambda \in \mathbb{R}$.

Beweis: Ist B eine ONB, dann ist $A = M_B^B(f)$ symmetrisch bzw. hermitesch und es reicht zu zeigen, daß $\chi_f = \chi_A$ in $\mathbb{R}[t]$ liegt und über $\mathbb{R}$ zerfällt.

Hierfür machen wir uns zunutze, daß wir A auf alle Fälle als eine Matrix in $\mathrm{Mat}_n(\mathbb{C})$ auffassen können und daß $A = A^*$ gilt. Über $\mathbb{C}$ zerfällt das charakteristische Polynom von A, d. h. es gibt $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ mit

$$\chi_A = (t - \lambda_1) \cdots (t - \lambda_n).$$

Es reicht, zu zeigen, daß $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ gilt. Nun gibt es zu jedem λ_i aber einen Vektor $0 \neq x_i \in \mathbb{C}^n$ mit $Ax_i = \lambda_i x_i$. Wegen $A = A^* = \overline{A}^t$ gilt für diesen Vektor

$$\begin{aligned} \lambda_i \cdot \left(\overline{x_i}^t \circ x_i\right) &= \overline{x_i}^t \circ (\lambda_i x_i) = \overline{x_i}^t \circ (Ax_i) = \overline{x_i}^t \circ A \circ x_i \\ &= \overline{x_i}^t \circ \overline{A}^t \circ x_i = \overline{Ax_i}^t \circ x_i = \overline{\lambda_i x_i}^t \circ x_i = \overline{\lambda_i} \cdot \left(\overline{x_i}^t \circ x_i\right). \end{aligned}$$

Aus $\overline{x_i}^t \circ x_i \neq 0$ folgt dann $\lambda_i = \overline{\lambda_i}$, d. h. $\lambda_i \in \mathbb{R}$. □

Satz 37.9 (Spektralsatz für selbstadjungierte Endomorphismen).
Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ ist genau dann selbstadjungiert, wenn V eine ONB aus Eigenvektoren von f besitzt und alle Eigenwerte von f reell sind.

Beweis: Sei zunächst f selbstadjungiert. Nach Lemma 37.8 ist $\chi_f \in \mathbb{R}[t]$ und zerfällt über $\mathbb{R}$.

Wir wollen nun mittels Induktion nach $n = \dim_{\mathbb{K}}(V)$ die Existenz einer ONB aus Eigenvektoren von V beweisen. Dabei ist im Fall $n = 1$ nichts zu zeigen, da jede ONB aus einem Eigenvektor besteht.

Da χ_f über $\mathbb{R}$ zerfällt, hat χ_f insbesondere eine Nullstelle $\lambda \in \mathbb{R}$, die dann ein Eigenwert von f ist und es gibt somit einen Vektor $0 \neq x \in V$ mit

$$f(x) = \lambda \cdot x.$$

Der Unterraum $U = \mathrm{Lin}(x)$ ist somit f-invariant, und wir wollen nun zeigen, daß dies auch auf sein orthogonales Komplement $U^\perp$ zutrifft. Sei dazu $y \in U^\perp$. Dann gilt

$$\langle f(y), x\rangle = \langle y, f^*(x)\rangle = \langle y, f(x)\rangle = \langle y, \lambda \cdot x\rangle = \lambda \cdot \langle y, x\rangle = 0,$$

woraus, wie gewünscht, $f(y) \in U^\perp$ folgt.

Wegen

$$\chi_f = \chi_{f_U} \cdot \chi_{f_{U^\perp}} = (t - \lambda) \cdot \chi_{f_{U^\perp}}$$

ist auch $\chi_{f_{U^\perp}}$ ein reelles Polynom, das über $\mathbb{R}$ zerfällt, und wegen

$$(f_{U^\perp})^* = (f^*)_{U^\perp} = f_{U^\perp}$$

ist $f_{U^\perp}$ zudem selbstadjungiert. Per Induktion gibt es also eine ONB $(x_2, \ldots, x_n)$ von $U^\perp$ aus Eigenvektoren $f_{U^\perp}$, die dann aber auch Eigenvektoren von f sind. Mit $x_1 = \frac{x}{||x||}$ ist schließlich

$$B = (x_1, x_2, \ldots, x_n)$$

die gesuchte ONB aus Eigenvektoren von f.

Besitzt umgekehrt V eine ONB $B = (x_1, \ldots, x_n)$ aus Eigenvektoren von V mit reellen Eigenwerten $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, so gilt

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix} = \begin{pmatrix} \overline{\lambda_1} & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \overline{\lambda_n} \end{pmatrix} = M_B^B(f)^*.$$

Mithin ist $M_B^B(f)$ symmetrisch bzw. hermitesch, und somit ist f selbstadjungiert. □

Korollar 37.10 (Spektralsatz für symmetrische / hermitesche Matrizen). Zu jeder symmetrischen bzw. hermiteschen Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ gibt es eine Matrix $T \in \mathrm{O}(n)$ bzw. $T \in \mathrm{U}(n)$ mit

$$T^{-1} \circ A \circ T = T^* \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}$$

und $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Insbesondere ist jede symmetrische bzw. hermitesche Matrix diagonalisierbar und hat nur reelle Eigenwerte.

Beweis: Die Aussage folgt aus dem Spektralsatz für selbstadjungierte Endomorphismen über $\mathbb{K}^n$ mit kanonischem Skalarprodukt angewendet auf f_A, wenn wir $T = T_E^B$ für die dortige ONB B wählen. □

Dies ist eine wichtige Ergänzung des Satzes über die Jordansche Normalform.

Beispiel 37.11.
Wir betrachten $\mathbb{R}^n$ mit dem kanonischen Skalarprodukt sowie die Matrix

$$A = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R}).$$

Da $A = A^t$ gilt, ist A symmetrisch, und wir wollen eine orthogonale Transformationsmatrix berechnen, die A auf Diagonalgestalt transformiert.

Dazu bestimmen wir zunächst das charakteristische Polynom von A als

$$\chi_A = \begin{vmatrix} t & 1 & -1 \\ 1 & t & 1 \\ -1 & 1 & t \end{vmatrix} = t^3 - 3t - 2 = (t-2)\cdot(t+1)^2.$$

Da A diagonalisierbar ist, wissen wir nun schon, daß

$$\dim_{\mathbb{R}} \operatorname{Eig}(A, 2) = \operatorname{mult}(\chi_A, 2) = 1$$

und

$$\dim_{\mathbb{R}} \operatorname{Eig}(A, -1) = \operatorname{mult}(\chi_A, -1) = 2$$

gilt. Mit Hilfe des Gauß-Algorithmus können wir dann Basen der Eigenräume von A zu den Eigenwerten 2 und -1 berechnen. Die Rechnung wollen wir hier nicht vorführen, sondern nur das Ergebnis angeben:

$$B' = \big((1,-1,1)^t\big)$$

ist eine Basis von $\operatorname{Eig}(A, 2)$ und

$$B'' = \big((1,1,0)^t, (0,1,1)^t\big)$$

ist eine Basis von $\operatorname{Eig}(A, -1)$.

Dann müssen wir B' und B'' mittels des Gram-Schmidtschen-Orthonormalisierungsverfahrens in Orthonormalbasen der jeweiligen Eigenräume überführen. Bei B' ist das sehr einfach, da wir den einzigen Vektor in B' nur normieren müssen. Wir erhalten als einzigen Vektor in der ONB von $\operatorname{Eig}(A, 2)$ deshalb

$$z_1 = \left(\tfrac{1}{\sqrt{3}}, \tfrac{-1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}\right)^t.$$

Für $B'' = (x_2, x_3)$ ist es etwas mehr Aufwand. Wir setzen zunächst

$$z_2 = \frac{1}{||x_2||} \cdot x_2 = \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0\right)^t.$$

Als nächstes setzen wir

$$y_3 = x_3 - \langle z_2, x_3\rangle \cdot z_2 = (0,1,1)^t - \tfrac{1}{\sqrt{2}} \cdot \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0\right)^t = \left(\tfrac{-1}{2}, \tfrac{1}{2}, 1\right)^t$$

und normieren diesen Vektor anschließend zu

$$z_3 = \frac{1}{||y_3||} \cdot y_3 = \left(\tfrac{-1}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}, \tfrac{2}{\sqrt{6}}\right)^t.$$

Die Vektoren z_2 und z_3 bilden eine ONB von $\operatorname{Eig}(A, -1)$, und $B = (z_1, z_2, z_3)$ ist somit eine ONB von $\mathbb{R}^3$. Schreiben wir die Vektoren als Spalten in die Matrix T, so erhalten

wir die gesuchte Transformationsmatrix

$$T = \begin{pmatrix} \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} \in \mathrm{O}(3).$$

Man rechnet folgendes leicht nach:

$$T^* \circ A \circ T = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{pmatrix} \circ \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \circ \begin{pmatrix} \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Satz 37.12 (Spektralzerlegung für selbstadjungierte Endomorphismen). Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ selbstadjungiert und seien $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ die paarweise verschiedenen Eigenwerte von f. Ferner bezeichne

$$\pi_i : V \longrightarrow V$$

die orthogonale Projektion von V auf $\mathrm{Eig}(f, \lambda_i)$.
Dann ist

$$V = \mathrm{Eig}(f, \lambda_1) \perp \ldots \perp \mathrm{Eig}(f, \lambda_r)$$

die *orthogonale Summe* der Eigenräume von f und es gilt

$$f = \lambda_1 \cdot \pi_1 + \ldots + \lambda_r \cdot \pi_r.$$

Man nennt dies die *Spektralzerlegung* von f.

Beweis: Nach dem Spektralsatz für selbstadjungierte Abbildungen besitzt f eine ONB aus Eigenvektoren, so daß die Eigenräume orthogonal aufeinander stehen und V deren direkte und damit orthogonale Summe ist (siehe auch Satz B1.27).

Ist nun $x = x_1 + \ldots + x_r \in V$ mit $x_i \in \mathrm{Eig}(f, \lambda_i)$ gegeben, so gilt

$$\pi_j(x_i) = \delta_{ij} \cdot x_i,$$

da $x_j \in \mathrm{Eig}(f, \lambda_j)$ und $x_i \perp \mathrm{Eig}(f, \lambda_j)$ für $i \neq j$, und somit

$$\pi_j(x) = \pi_j(x_1) + \ldots + \pi_j(x_r) = x_j.$$

Damit erhalten wir dann

$$\begin{aligned}(\lambda_1 \cdot \pi_1 + \ldots + \lambda_r \cdot \pi_r)(x) =& \lambda_1 \cdot \pi_1(x) + \ldots + \lambda_r \cdot \pi_r(x) \\ =& \lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r = f(x_1) + \ldots + f(x_r) = f(x).\end{aligned}$$

□

C) Hauptachsentransformationssatz für quadratische Formen

Definition 37.13 (Bilinearformen).
Es sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung

$$b : V \times V \to \mathbb{R},$$

die linear in beiden Argumenten ist, nennen wir *bilinear* oder eine *Bilinearform*, d. h. für $x, y, z \in V$ und $\lambda, \mu \in \mathbb{R}$ gilt (vgl. Definition 34.8):

$$b(\lambda x + \mu y, z) = \lambda b(x, z) + \mu b(y, z)$$

und

$$b(z, \lambda x + \mu y) = \lambda b(z, x) + \mu b(z, y).$$

Gilt zudem

$$b(x, y) = b(y, x)$$

für alle $x, y \in V$, so heißt b eine *symmetrische* Bilinearform, und wir nennen die Abbildung

$$q_b : V \longrightarrow \mathbb{R} : x \mapsto b(x, x)$$

die zu b gehörende *quadratische Form*.

Bemerkung 37.14 (Polarisierung einer Bilinearform).
Ist $b : V \times V \longrightarrow \mathbb{R}$ eine symmetrische Bilinearform, dann gilt

$$b(x, y) = \tfrac{1}{2}\big(q_b(x + y) - q_b(x) - q_b(y)\big)$$

für $x, y \in V$, wie man durch Einsetzen der Definition leicht nachrechnet. Die Bilinearform ist durch die zugehörige quadratische Form also eindeutig bestimmt. Man nennt die Gleichung die *Polarisierung* der Bilinearform.

Beispiel 37.15.
Jedes Skalarprodukt auf einem $\mathbb{R}$-Vektorraum ist eine symmetrische Bilinearform. Zudem definiert jede symmetrische Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ durch

$$b_A : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} : (x, y) \mapsto x^t \circ A \circ y$$

eine symmetrische Bilinearform. Die zu b_A gehörende quadratische Form bezeichnen wir mit

$$q_A : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto x^t \circ A \circ x.$$

Aus dem Spektralsatz für symmetrische Matrizen können wir eine Normalform für die durch eine symmetrische Bilinearform definierte quadratische Form herleiten.

Satz 37.16 (Hauptachsentransformationssatz für quadratische Formen).
Ist $A \in \mathrm{Mat}_n(\mathbb{R})$ eine symmetrische Matrix, so gibt es eine ONB $B = (x_1, \ldots, x_n)$ von $\mathbb{R}^n$ zum kanonischen Skalarprodukt mit

$$A \circ x_i = \lambda_i \cdot x_i$$

für $i = 1, \ldots, n$ und

$$q_A(x) = \sum_{i=1}^{n} \lambda_i \cdot \langle x_i, x \rangle^2.$$

Die x_i werden die *Hauptachsen* der quadratischen Form genannt.

Beweis: Aus dem Spektralsatz für selbstadjungierte Endomorphismen 37.9 wissen wir, daß es eine ONB $B = (x_1, \ldots, x_n)$ von $\mathbb{R}^n$ aus Eigenvektoren von f_A gibt. Diese wählen wir, so daß

$$A \circ x_i = f_A(x_i) = \lambda_i \cdot x_i$$

für geeignete $\lambda_i \in \mathbb{R}$ gilt. Sei $x \in V$, so folgt aus der Parsevalschen Gleichung 36.14

$$x = \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x_i.$$

Für die quadratische Form q_A ausgewertet in x ergibt sich dann

$$\begin{aligned} q_A(x) =& x^t \circ A \circ x = x^t \circ A \circ \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x_i = \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x^t \circ A \circ x_i \\ =& \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x^t \circ \lambda_i \cdot x_i = \sum_{i=1}^{n} \langle x_i, x \rangle \cdot \lambda_i \cdot \langle x, x_i \rangle = \sum_{i=1}^{n} \lambda_i \cdot \langle x, x_i \rangle^2. \end{aligned}$$

□

Bemerkung 37.17 (Geometrische Interpretation).
Ist $A \in \mathrm{Mat}_n(\mathbb{R})$ eine positiv definite symmetrische Matrix mit quadratischer Form q_A, so interessieren wir uns für die Einheitssphäre zu q_A

$$S_{q_A} = \{y \in \mathbb{R}^n \mid q_A(y) = y^t \circ A \circ y = 1\}.$$

Man beachte, daß die Bilinearform b_A in diesem Fall ein Skalarprodukt ist und daß S_{q_A} die Menge der Vektoren in $\mathbb{R}^n$ ist, die bezüglich der zu diesem Skalarprodukt gehörenden Norm die Länge 1 haben.

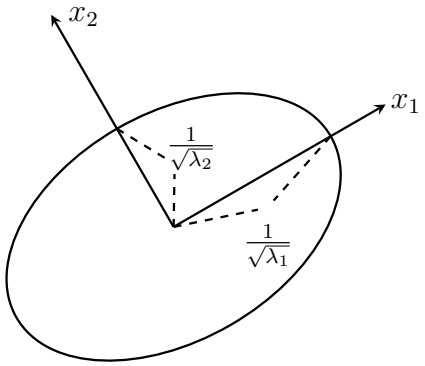


Abbildung 5: Ellipse S_{q_A} mit Hauptachsen $x_1 = Te_1$ und $x_2 = Te_2$

Der Spektralsatz für selbstadjungiert Endomorphismen liefert die Existenz einer orthogonalen Matrix $T \in \mathrm{O}(n)$, so daß

$$T^t \circ A \circ T = \lambda_1 \cdot \mathbb{1}_1 \oplus \ldots \oplus \lambda_n \cdot \mathbb{1}_1$$

eine Diagonalmatrix ist, bei der die $\lambda_i > 0$ die Eigenwerte von A sind. Die Spaltenvektoren von T sind dann ein neues orthonormales Koordinatensystem, in dem die quadratische Form die Gestalt

$$q = \lambda_1 \cdot y_1^2 + \ldots + \lambda_n \cdot y_n^2$$

hat. Die Einheitssphäre zu q ist dann ein n-dimensionales Ellipsoid

$$S_q = \{y \in \mathbb{R}^n \mid q(y) = \lambda_1 \cdot y_1^2 + \ldots + \lambda_n \cdot y_n^2 = 1\}.$$

Man kann sich S_{q_A} deshalb als eine Einheitskugel vorstellen, die in Richtung $x_i = Te_i$ um den Faktor $\frac{1}{\sqrt{\lambda_i}}$ gestreckt wurde. Die Koordinatenvektoren x_i sind dann die *Hauptachsen* des Ellipsoids (siehe Abbildung 5). Im Fall $n = 2$ besagt der Satz der Hauptachsentransformation dann, daß wir allein durch Drehen und Spiegeln die Ellipse S_{q_A} so bewegen können, daß ihre Hauptachsen mit den Koordinatenachsen zusammenfallen. Daher rührt der Begriff der *Hauptachsentransformation.*

D) Positiv definite symmetrische und hermitesche Matrizen

Skalarprodukte auf einem reellen Vektorraum sind symmetrische Bilinearformen, und auf dem $\mathbb{R}^n$ ist auch jedes Skalarprodukt als Bilinearform b_A zu einer symmetrischen Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ gegeben. Die Symmetrie von A reicht aber nicht aus, um sicherzustellen, daß b_A ein Skalarprodukt ist. Für die Definitheit der Bilinearform muß A eine weitere Bedingung erfüllen, die man positiv definit nennt. Wir wollen im vorliegenden Abschnitt Bedingungen herleiten, die sicherstellen, daß eine Matrix A positiv definit ist.

Definition 37.18 (Definite Matrizen).

a. Eine symmetrische Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ heißt:
- *positiv definit* $\Longrightarrow$ $x^t \circ A \circ x > 0$ für alle $x \in \mathbb{R}^n$;
- *negativ definit* $\Longrightarrow$ $x^t \circ A \circ x < 0$ für alle $x \in \mathbb{R}^n$;
- *indefinit* $\Longrightarrow$ $x^t \circ A \circ x > 0 > y^t \circ A \circ y$ für geeignete $x, y \in \mathbb{R}^n$.

b. Eine hermitesche Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ heißt:
- *positiv definit* $\Longrightarrow$ $\overline{x}^t \circ A \circ x > 0$ für alle $x \in \mathbb{R}^n$;
- *negativ definit* $\Longrightarrow$ $\overline{x}^t \circ A \circ x < 0$ für alle $x \in \mathbb{R}^n$;
- *indefinit* $\Longrightarrow$ $\overline{x}^t \circ A \circ x > 0 > \overline{y}^t \circ A \circ y$ für geeignete $x, y \in \mathbb{R}^n$.

c. Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ und entsteht die $k \times k$-Untermatrix $A(k)$ von A durch Streichen der letzten $n-k$ Zeilen und Spalten, so nennen wir $A(k)$ die k-te *Hauptmatrix* von A und $\det\big(A(k)\big)$ den k-ten *Hauptminor* von A.

Bemerkung 37.19 (Negativ definite Matrizen).
Man beachte, daß eine symmetrische oder hermitesche Matrix A genau dann negativ definit ist, wenn $-A$ positiv definit ist. Es reicht deshalb, ein Kriterium für positive Definitheit zu finden, um zugleich ein Kriterium für negative Definitheit zu erhalten, indem man A durch $-A$ ersetzt.

Lemma 37.20 (Positiv definite Matrizen).
Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ symmetrisch bzw. hermitesch und $T \in \mathrm{Gl}_n(\mathbb{K})$ invertierbar. Genau dann ist A positiv definit, wenn $T^* \circ A \circ T$ positiv definit ist.

Beweis: Wir beachten, daß

$$\mathbb{K}^n \setminus \{0\} = \{T \circ x \mid 0 \neq x \in \mathbb{K}^n\} \tag{108}$$

gilt, da T invertierbar ist, und daß die Matrix $T^* \circ A \circ T$ symmetrisch bzw. hermitesch ist, wegen

$$(T^* \circ A \circ T)^* = T^* \circ A^* \circ T^{**} = T^* \circ A \circ T.$$

Wir erhalten deshalb

$$\begin{aligned} A \text{ ist positiv definit} &\iff \overline{x}^t \circ A \circ x > 0 \quad \forall\, x \in \mathbb{K}^n \\ &\overset{(108)}{\iff} \overline{T \circ x}^t \circ A \circ (T \circ x) > 0 \quad \forall\, 0 \neq x \in \mathbb{K}^n \\ &\iff \overline{x}^t \circ T^* \circ A \circ T \circ x > 0 \quad \forall\, 0 \neq x \in \mathbb{K}^n \\ &\iff T^* \circ A \circ T \text{ ist positiv definit} \end{aligned}$$

□

Satz 37.21 (Hurwitz-Kriterium für positive Definitheit).
Für eine symmetrische bzw. hermitesche Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ sind die folgenden Aussagen gleichwertig:

a. A ist positiv definit.

b. Alle Eigenwerte von A sind positiv.

c. Alle Hauptminoren von A sind positiv.

Beweis:

a. $\Longleftrightarrow$ b.: Wegen des Spektralsatzes für symmetrische und hermitesche Matrizen 37.10 gibt es eine orthogonale bzw. unitäre Matrix $T \in \mathrm{Mat}_n(\mathbb{K})$, so daß

$$T^* \circ A \circ T = T^{-1} \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \tag{109}$$

gilt, wobei $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ genau die Eigenwerte von A sind.

Ist A positiv definit und ist x_i die i-te Spalte von T, so folgt aus (109)

$$\lambda_i = \overline{x_i}^t \circ A \circ x_i > 0.$$

Seien nun umgekehrt alle Eigenwerte positiv und sei $0 \neq x \in \mathbb{K}^n$ gegeben. Die Spaltenvektoren $x_1, \dots, x_n$ von T sind eine ONB $B = (x_1, \dots, x_n)$ von $\mathbb{K}^n$, da T orthogonal bzw. unitär ist. Ist $M_B(x) = (\mu_1, \dots, \mu_n)^t$ der Koordinatenvektor von x bezüglich B, so gilt

$$T \circ M_B(x) = \sum_{i=1}^{n} \mu_i \cdot x_i = x$$

und

$$\begin{aligned} \overline{x}^t \circ A \circ x &= \overline{T \circ M_B(x)}^t \circ A \circ (T \circ M_B(x)) \\ &= \overline{M_B(x)}^t \circ (T^* \circ A \circ T) \circ M_B(x) \\ &\overset{(109)}{=} \sum_{i=1}^{n} \overline{\mu_i} \cdot \mu_i \cdot \lambda_i = \sum_{i=1}^{n} |\mu_i|^2 \cdot \lambda_i > 0, \end{aligned}$$

da nicht alle μ_i null sind. Damit ist die Äquivalenz von a. und b. gezeigt.

a. $\Longrightarrow$ c.: Da wir die Äquivalenz von a. und b. bereits gezeigt haben, können wir hier beide Bedingungen voraussetzen. Wegen des Spektralsatzes für symmetrische und hermitesche Matrizen 37.10 gibt es eine Matrix $T \in \mathrm{Gl}_n(\mathbb{K})$, so daß (109) erfüllt ist, und deshalb gilt

$$\det(A) = \det(T^{-1} \circ A \circ T) = \lambda_1 \cdot \ldots \cdot \lambda_n > 0.$$

Jede positiv definite symmetrische oder hermitesche Matrix hat also eine positive Determinante.

Die k-te Hauptmatrix $A(k)$ und $x, y \in K^k$ gilt dann

$$x^t \circ A(k) \circ \overline{y} = u^t \circ A \circ \overline{v} > 0$$

für $u = (x, 0, \ldots, 0)^t, v = (y, 0, \ldots, 0)^t \in K^n$, da A positiv definit ist. Mithin ist auch die Matrix $A(k)$ positiv definit und ihre ihre Determinante, der k-te Hauptminor von A, ist somit ebenfalls positiv.

c. $\Longrightarrow$ a.: Wir führen den Beweis durch Induktion über n unter Ausnutzung der bereits gezeigten Äquivalenzen, wobei für $n = 1$ die Determinante $\det(A) > 0$ der einzige Eigenwert ist.

Sei also $n > 1$. Wegen des Spektralsatzes für symmetrische und hermitesche Matrizen 37.10 existiert für die symmetrische bzw. hermitesche Matrix $A(n-1)$ eine orthogonale bzw. unitäre Matrix $S \in \mathrm{Gl}_{n-1}(\mathbb{K})$, die $A(n-1)$ auf Diagonalgestalt transformiert:

$$S^{-1} \circ A(n-1) \circ S = S^* \circ A(n-1) \circ S = \bigoplus_{i=1}^{n-1} \lambda_i \mathbb{1}_1 =: D.$$

Da $A(n-1)$ die Induktionsvoraussetzung erfüllt, muß $A(n-1)$ dann positiv definit sein und somit sind die Eigenwerte $\lambda_1, \ldots, \lambda_{n-1}$ von $A(n-1)$ positiv.

Wir setzen nun $T = S \oplus \mathbb{1}_1 \in \mathrm{Gl}_n(\mathbb{K})$. Dann gilt

$$T^* \circ A \circ T = \left(\begin{array}{ccc|c} & & & a_1 \\ & D & & \vdots \\ & & & a_{n-1} \\ \hline \overline{a_1} & \ldots & \overline{a_{n-1}} & a_n \end{array}\right) =: B$$

für geeignete $a_1, \ldots, a_n \in \mathbb{K}$. Setzen wir ferner $c_j = -\frac{a_j}{\lambda_j}$, $j = 1, \ldots, n-1$, und

$$C = \left(\begin{array}{ccc|c} & & & c_1 \\ & \mathbb{1}_{n-1} & & \vdots \\ & & & c_{n-1} \\ \hline 0 & \ldots & 0 & 1 \end{array}\right) \in \mathrm{Gl}_n(\mathbb{K}),$$

dann folgt

$$E := (T \circ C)^* \circ A \circ (T \circ C) = C^* \circ T^* \circ A \circ T \circ C = C^* \circ B \circ C = \bigoplus_{i=1}^{n} \lambda_i \mathbb{1}_1,$$

wobei $\lambda_n \in \mathbb{R}$ geeignet ist. Man beachte, daß damit $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von E sind, und daß

$$\lambda_1 \cdot \ldots \cdot \lambda_n = \det\big((T \circ C)^* \circ A \circ (T \circ C)\big) = \det(A) \cdot |\det(C \circ T)|^2 > 0,$$

da $\det(A) > 0$ der n-te Hauptminor von A ist. Da aber $\lambda_1, \ldots, \lambda_{n-1}$ nach Voraussetzung positiv waren, ist dann auch λ_n positiv. E hat somit nur positive Eigenwerte und ist deshalb positiv definit. Aber mit Lemma 37.20 ist dann auch A positiv definit. □

Bemerkung 37.22 (Negativ definite und indefinite Matrizen).
Wie im Beweis von "a. $\Longleftrightarrow$ b." im Beweis von Satz 37.21 sieht man:

$$A \text{ ist } \textit{negativ definit} \iff A \text{ hat nur negative Eigenwerte}$$

und

$$A \text{ ist } \textit{indefinit} \iff A \text{ hat einen positiven und einen negativen Eigenwert.}$$

Bemerkung 37.23 (Skalarprodukte).
Genau dann ist

$$b_A : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} : (x, y) \mapsto x^t \circ A \circ y$$

für $A \in \mathrm{Mat}_n(\mathbb{R})$ ein Skalarprodukt auf $\mathbb{R}^n$, wenn A symmetrisch und positiv definit ist.

Genau dann ist

$$b_A^s : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C} : (x, y) \mapsto x^t \circ A \circ \overline{y}$$

für $A \in \mathrm{Mat}_n(\mathbb{C})$ ein Skalarprodukt auf $\mathbb{C}^n$, wenn A hermitesch und positiv definit ist.

Aufgaben

Aufgabe 37.24 (Lineare Funktionale).
Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum.
Dann gibt es für jedes $g \in \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$ genau ein $y \in V$, so daß für alle $x \in V$ gilt

$$g(x) = \langle y, x \rangle.$$

Aufgabe 37.25 (Die adjungierte Abbildung).
Seien V und W zwei endlich-dimensionale euklidische oder unitäre Räume mit Skalarprodukten $\langle \cdot, \cdot \rangle_V$ und $\langle \cdot, \cdot \rangle_W$. Dann gibt es zu jeder linearen Abbildung $f : V \longrightarrow W$ genau eine lineare Abbildung $f^* : W \longrightarrow V$, so daß

$$\langle f(x), y \rangle_W = \langle x, f^*(y) \rangle_V \tag{110}$$

für alle $x \in V$ und $y \in W$. Die Abbildung f^* heißt die *adjungierte Abbildung* zu f.

Aufgabe 37.26 (Matrixdarstellung der adjungierten Abbildung).
Seien V und W zwei endlich-dimensionale euklidische oder unitäre Räume mit Orthonormalbasen B bzw. D. Dann gilt für jede $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow W$

$$M_B^D(f^*) = \big(M_D^B(f)\big)^*,$$

d.h. die Matrixdarstellung der adjungierten Abbildung ist die Adjungierte der Matrixdarstellung.

Aufgabe 37.27.
Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum. Zeige, ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ mit $f^* \circ f = f \circ f^*$, so gelten

$$\mathrm{Ker}(f) = \mathrm{Ker}(f^*)$$

und

$$V = \mathrm{Ker}(f) \perp \mathrm{Im}(f).$$

Aufgabe 37.28.
Sei V ein endlich-dimensionaler unitärer Raum und $f \in \mathrm{End}_{\mathbb{C}}(V)$ mit $f^* \circ f = f \circ f^*$.
Zeige, es gibt ein Polynom $p \in \mathbb{C}[t]$ mit $f^* = p(f)$.

Aufgabe 37.29.
Es sei $V \neq 0$ ein endlich-dimensionaler unitärer Raum und $f \in \mathrm{End}_{\mathbb{C}}(V)$. Zeige, die folgenden Aussagen sind gleichwertig:

a. $f^* = -f$.

b. Für alle $x \in V$ gilt: $\langle f(x), x\rangle \in i\mathbb{R}$.

c. Es gibt eine Orthonormalbasis von V aus Eigenvektoren von f und der Realteil aller Eigenwerte ist Null.

Aufgabe 37.30.
Bestimme eine orthogonale Matrix $T \in \mathrm{O}(3)$, die die folgende symmetrische Matrix A diagonalisiert:

$$A = \begin{pmatrix} 5 & 2 & 2 \\ 2 & 2 & -4 \\ 2 & -4 & 2 \end{pmatrix}.$$

§ 38 Singulärwertzerlegung

Die essentiellen Informationen einer quadratischen Matrix sind in den Eigenwerten und Eigenvektoren der Matrix enthalten. Matrizen, für die eine Eigenwertzerlegung existiert, haben wir als *diagonalisierbar* bezeichnet. Allerdings ist nicht jede quadratische Matrix diagonalisierbar.

Aufgrund des großen praktischen Interesses, stellt sich die Frage nach einer informativen Zerlegung, die für jede (auch nicht-quadratische) Matrix berechnet werden kann. Daher beschäftigen wir uns in diesem Abschnitt mit der *Singulärwertzerlegung*. Im Gegensatz zur Eigenwertzerlegung existiert eine Singulärwertzerlegung für nicht-quadratische Matrizen. Wieder reduzieren wir die gegebene Matrix auf Diagonalgestalt, allerdings müssen wir hierfür unterschiedliche Transformationen von links und von rechts zulassen.

Die Singulärwertzerlegung hat viele Anwendungen in den Bereichen der Linearen Algebra, des Maschinellen Lernens, der Statistik, der Bildverarbeitung, der Geographie (z.B. Wettervorhersage), der Chemie, ...

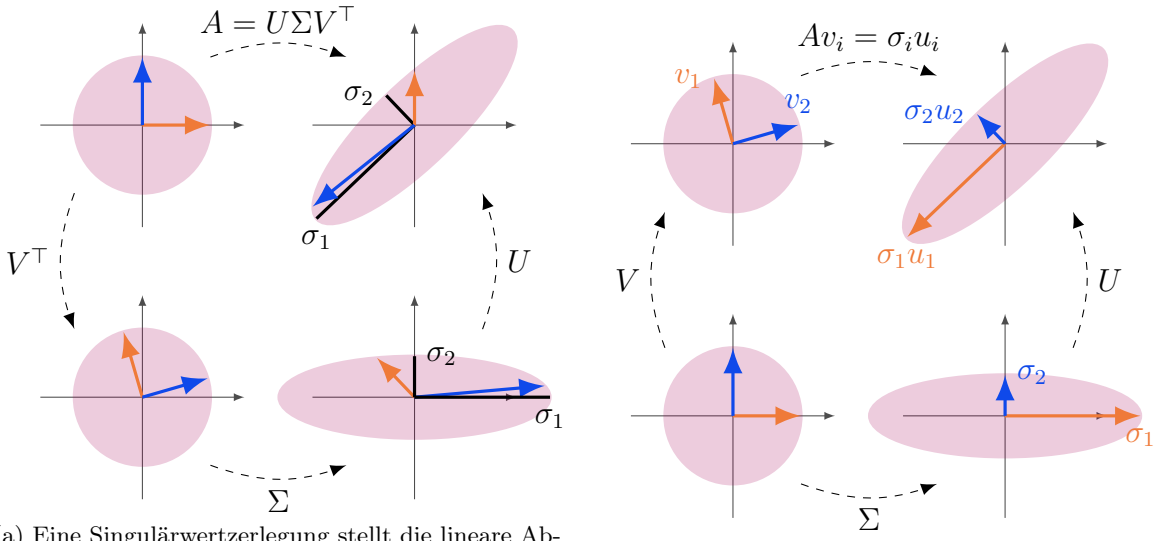
Der folgende Satz ist die zentrale Aussage in diesem Abschnitt.

Satz 38.1 (Singulärwertzerlegung).
Sei $A \in \mathbb{R}^{m\times n}$. Dann existieren orthogonale Matrizen $U \in \mathrm{O}(m)$, $V \in \mathrm{O}(n)$ und $\sigma_1, \dots, \sigma_p \in \mathbb{R}$ mit $\sigma_1 \geq \dots \geq \sigma_p \geq 0$ und $p = \min(m, n)$, so dass

$$A = U\Sigma V^\top \qquad \text{mit } \Sigma := \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_p \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \text{ gilt.}$$

Bevor wir den Satz beweisen, schauen wir uns ein paar Beispiele und die geometrische Intuition der Aussage an.

Definition 38.2.
Eine Darstellung $A = U\Sigma V^\top$ wie in Satz 38.1 heißt *Singulärwertzerlegung* von A. Die Spalten von U heißen *linke Singulärvektoren*, die Spalten von V *rechte Singulärvektoren* und die Diagonaleinträge $\sigma_1 \geq \dots \geq \sigma_{\min(m,n)} \geq 0$ von Σ heißen *Singulärwerte* von A.



(a) Eine Singulärwertzerlegung stellt die lineare Abbildung A als eine Komposition einer Drehspiegelung $V^\top$, einer Skalierungsmatrix Σ, und einer weiteren Drehspiegelung U dar.

(b) Bezüglich der Basen aus Singulärvektoren ist die lineare Abbildung A besonders einfach. Es gilt $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$.

Abbildung 6. Geometrie einer Singulärwertzerlegung.

Geometrisch stellt eine Singulärwertzerlegung die lineare Abbildung A als Komposition einer Drehspiegelung, einer Skalierung der Koordinatenachsen und einer weiteren Drehspiegelung dar. Diese Interpretation ist in Abbildung 6(a) visualisiert.

Alternativ zur Zerlegung in ein Produkt von Matrizen lässt sich eine Singulärwertzerlegung von A auch wie folgt schreiben:

$$A = U\Sigma V^\top = \sum_{i=1}^{p} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top ,$$

wobei $p = \min(m, n)$ ist. Eine weitere Darstellungsmöglichkeit ist

$$AV = U\Sigma , \quad \text{bzw.} \quad A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i \text{ für } i = 1, \dots, p .$$

Letztere zeigt, dass die lineare Abbildung A bezüglich der Basen $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ von $\mathbb{R}^n$ und $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ von $\mathbb{R}^m$ eine besonders einfache Darstellung hat, nämlich eine Skalierung der Basisvektoren. Diese Charakterisierung ist in Abbildung 6(b) visualisiert. Dabei können die Ellipsen (bzw. Ellipsoiden) auch niedrigdimensionale Einbettungen sein, z.B. eine Scheibe im $\mathbb{R}^3$.

Beispiel 38.3. a. Für eine symmetrische positiv semi-definite Matrix A liefert die Eigenwertzerlegung eine Singulärwertzerlegung.

b. Für eine symmetrische Matrix A sind die Singulärwerte gerade die Beträge der Eigenwerte von A.

c. Sei $A = U\Sigma V^\top$ eine Singulärwertzerlegung. Dann gilt:

$$AA^\top = (U\Sigma V^\top)(V\Sigma^\top U^\top) = U\Sigma\Sigma^\top U^\top \quad \text{und analog} \quad A^\top A = V\Sigma^\top \Sigma V^\top .$$

Daher entsprechen die Quadrate der Singulärwerte $\sigma_1, \dots, \sigma_p$ den Eigenwerten von $AA^\top$ und $A^\top A$ zu den Eigenvektoren $\boldsymbol{u}_1, \dots, \boldsymbol{u}_p$ bzw. $\boldsymbol{v}_1, \dots, \boldsymbol{v}_p$.

Beweis: [von Satz 38.1] Nach dem Satz über die Eigenwertzerlegung können wir die symmetrische Matrix $A^\top A$ als ein Produkt einer orthogonalen Matrix $V \in \mathrm{O}\, n$ mit Eigenvektoren $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ als Spalten und einer Diagonalmatrix $\Lambda \in \mathbb{R}^{n\times n}$ mit den Eigenwerten $\lambda_1 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0$ auf der Diagonalen schreiben:

$$A^\top A = V \Lambda V^\top .$$

Wir konstruieren die gewünschte orthogonale Matrix U mit Spalten $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ explizit. Für $i = 1, \dots, r$ setzen wir $\boldsymbol{u}_i = \frac{A\boldsymbol{v}_i}{\sigma_i}$ mit $\sigma_i = \sqrt{\lambda_i} > 0$. Diese Vektoren sind offensichtlich orthonormal:

$$\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle A\boldsymbol{v}_i, A\boldsymbol{v}_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle A^\top A \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \frac{\lambda_i}{\sigma_i \sigma_j} \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \delta_{ij} = \begin{cases} 1\,, & \text{wenn } i = j\,, \\ 0\,, & \text{sonst}\,. \end{cases}$$

Mittels des Gram–Schmidt Orthogonalisierungsverfahrens ergänzen wir $\boldsymbol{u}_1, \dots, \boldsymbol{u}_r$ zu einer Orthonormalbasis des $\mathbb{R}^m$.

Es bleibt zu zeigen, dass die Konstruktion das Richtige tut, d.h. $U\Sigma V^\top = A$. Zunächst bemerken wir, dass $\ker A = \ker(A^\top A)$ ist, da $\|A\boldsymbol{v}_i\|_2^2 = \langle A^\top A \boldsymbol{v}_i, \boldsymbol{v}_i \rangle = \lambda_i = 0$ für $i = r+1, \dots, n$ gilt. Nach dem Rangsatz folgt $r = \operatorname{rank}(A^\top A) = \operatorname{rank}(A)$. Damit verifizieren wir nun die Darstellung der Singulärwertzerlegung:

$$U\Sigma V^\top = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top = \sum_{i=1}^{n} (A\boldsymbol{v}_i) \boldsymbol{v}_i^\top = A \sum_{i=1}^{n} \boldsymbol{v}_i \boldsymbol{v}_i^\top = A\,.$$

□

Bemerkung 38.4. a. Im Englischen heißt die Singulärwertzerlegung *singular value decomposition*, kurz *SVD*.

b. Die Singulärwertzerlegung ist im Allgemeinen nicht eindeutig. Die Eigenvektoren zu gleichen Eigenwerten von $A^\top A$ können in verschiedener Reihenfolge gewählt werden. Auch die Ergänzung von $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_r)$ zu einer Basis ist bezüglich ihrer Reihenfolge nicht eindeutig bestimmt. Die Singulärwerte jedoch sind eindeutig.

Satz 38.5.
Sei $A = U\Sigma V^\top$ wie in Satz 38.1 mit orthogonalen Matrizen $U = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$, $V = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ und $\sigma_1 \geq \dots \geq \sigma_r > 0$ mit $r \leq \min\{m, n\}$. Dann gelten:

a. $\operatorname{rank}(A) = r$,

b. $\operatorname{Im}(A) = \operatorname{Lin}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_r)$,

c. $\ker(A) = \mathrm{Lin}(\boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n)$.

Beweis: Folgt direkt aus dem Beweis von Satz 38.1. □

Beispiel 38.6.
Wir betrachten eine Problemstellung der 3D-Rekonstruktion. Seien $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_q$ Punkte auf einem Objekt in $\mathbb{R}^3$. Diese Punkte werden von f verschiedenen Kameras (bzw. von einer Kamera zu verschiedenen Zeitpunkten) aufgenommen. Die "Projektion" eines Punktes $\boldsymbol{X}_i$ in die j-te Kamera bezeichnen wir mit $\boldsymbol{x}_i^{(j)} \in \mathbb{R}^2$, $j = 1, \ldots, f$ und $i = 1, \ldots, q$. Das Ziel der Anwendung ist die Rekonstruktion der 3D-Koordinaten basierend auf den in die Kameras projizierten 2D-Koordinaten.

Unter der Annahme eines affinen Kameramodells gilt für alle Punkte:

$$\boldsymbol{x}_i^{(j)} = M^{(j)} \boldsymbol{X}_i + \boldsymbol{t}^{(j)}, \quad i = 1, \ldots, q,$$

wobei $M^{(j)} \in \mathbb{R}^{2\times 3}$ und $\boldsymbol{t}^{(j)} \in \mathbb{R}^2$ ist. Da wir den Ursprung des 3D-Koordinatensystems frei wählen können, nehmen wir an, dass die Punkte $\boldsymbol{X}_i$ um den Ursprung herum zentriert sind, d.h. es gilt $\frac{1}{q}\sum_{i=1}^q \boldsymbol{X}_i = 0$. Daraus ergibt sich mit obiger Gleichung für jede Kamera j direkt, dass die projizierten Punkte $\boldsymbol{x}_i^{(j)}$ um $\boldsymbol{t}^{(j)} = \frac{1}{q}\sum_{i=1}^q \boldsymbol{x}_i^{(j)}$ zentriert sind:

$$\frac{1}{q}\sum_{i=1}^q \boldsymbol{x}_i^{(j)} = \frac{1}{q}\sum_{i=1}^q (M^{(j)}\boldsymbol{X}_i + \boldsymbol{t}^{(j)}) = M^{(j)} \underbrace{\left(\frac{1}{q}\sum_{i=1}^q \boldsymbol{X}_i\right)}_{=0} + \boldsymbol{t}^{(j)} = \boldsymbol{t}^{(j)}.$$

In der Praxis haben wir nur die 2D-Koordinaten gegeben. Wir können diese 2D-Koordinaten problemlos zentrieren. Die resultierenden zentrierten 2D-Koordinaten sind

$$\hat{\boldsymbol{x}}_i^{(j)} := \boldsymbol{x}_i^{(j)} - \frac{1}{q}\sum_{i=1}^q \boldsymbol{x}_i^{(j)}.$$

Wir erhalten das folgende Modell für den Zusammenhang von 2D- und 3D-Koordinaten:

$$\underbrace{\begin{pmatrix} \hat{\boldsymbol{x}}_1^{(1)} & \hat{\boldsymbol{x}}_2^{(1)} & \cdots & \hat{\boldsymbol{x}}_q^{(1)} \\ \hat{\boldsymbol{x}}_1^{(2)} & \hat{\boldsymbol{x}}_2^{(2)} & \cdots & \hat{\boldsymbol{x}}_q^{(2)} \\ \vdots & \vdots & & \vdots \\ \hat{\boldsymbol{x}}_1^{(f)} & \hat{\boldsymbol{x}}_2^{(f)} & \cdots & \hat{\boldsymbol{x}}_q^{(f)} \end{pmatrix}}_{=:W\in\mathbb{R}^{2f\times q}} = \underbrace{\begin{pmatrix} M^{(1)} \\ M^{(2)} \\ \vdots \\ M^{(f)} \end{pmatrix}}_{=:M\in\mathbb{R}^{2f\times 3}} \underbrace{\begin{pmatrix} \hat{\boldsymbol{X}}_1 & \hat{\boldsymbol{X}}_2 & \cdots & \hat{\boldsymbol{X}}_q \end{pmatrix}}_{=:S\in\mathbb{R}^{3\times q}},$$

wobei M die Matrix der Kamerabewegungen ist und S als Strukturmatrix bezeichnet wird. Offensichtlich ist der Rang von M höchstens 3. Daher liegen die so genannten Trajektorien[1] $(\hat{\boldsymbol{x}}_i^{(1)}, \ldots, \hat{\boldsymbol{x}}_i^{(f)})$, die Spalten von W, auf einem 3-dimensionalen Unterraum von $\mathbb{R}^{2f}$.

[1] Eine Trajektorie besteht aus allen 2D-Koordinaten eines 3D Punktes in den verschiedenen Kameras.

Ist W gegeben, so können wir M und S bis auf affine Ambiguität[2] mithilfe der Singulärwertzerlegung $W = U\Sigma V^\top$ bestimmen. Wir setzen

$$M = (\sigma_1 u_1, \sigma_2 u_2, \sigma_3 u_3) \quad \text{und} \quad S = (v_1, v_2, v_3)^\top .$$

Die Matrix S liefert also die gewünschte Rekonstruktion der 3D-Punkte.

[2] Es gilt $W = (MA)(A^{-1}S)$ für jede beliebige Matrix $A \in \mathrm{Gl}_3(\mathbb{R})$, d.h. W ist bis auf eine gemeinsame lineare Transformation von M mit A und von S mit A^{-1} eindeutig bestimmt.

Kapitel A

Einige Ergänzungen zu den Algebraischen Strukturen

§ A1 Einleitung

Strichcodes, ISBN-Nummern, Banknoten

Bezahlt man heutzutage in einem Geschäft seine Ware, so wird der Preis in aller Regel nicht per Hand eingegeben. Vielmehr wird der Strichcode, der sich an jedem Artikel befindet, eingescannt, und selbst wenn das aufgrund technischer Probleme nicht funktioniert, gibt der Kassierer nicht den Preis, sondern die zum Strichcode gehörende Ziffernfolge ein – üblicherweise aus 13 Ziffern bestehend. Der Preis wird dann aus

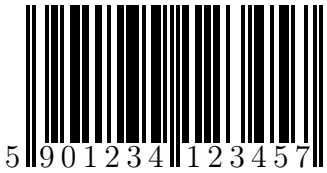


Abbildung 1. Ein EAN-13 Strichcode

einer vorhandenen Datenbank ermittelt, und in selbiger Datenbank wird vermerkt, daß das Geschäft nun einen Artikel dieses Codes weniger im Angebot hat. Wie gesagt, manchmal funktioniert das Einscannen nicht und eine solch lange Ziffernfolge abzutippen ist reichlich fehleranfällig. Im Falle eines Fehlers piepst die Kasse und der Kassierer gibt die Ziffernfolge noch mal ein. Wie kommt es, daß ein Fehler beim Eingeben immer auffällt?

Zunächst einmal bedeutet es nur, daß der falsche Code nicht in der Datenbank vorhanden ist. Das scheint auf den ersten Blick nicht zu verwundern, hat der Code doch 13 Ziffern und das Geschäft sicher nicht einmal 10000 verschiedene Artikel im Angebot. Würde man also jedem Artikel einen zufälligen Code aus 13 Ziffern geben, so könnte man ziemlich sicher sein, daß einzelne Fehler beim Abtippen auffallen würden. Nun werden

diese Strichcodes aber nicht von den Geschäften vergeben, die die Ware verkaufen, sondern vom Hersteller – oder genauer gesagt, der Hersteller läßt sie bei einer zentralen Agentur eintragen. Letzteres ist sinnvoll, denn es sollen schließlich nicht zwei Hersteller den gleichen Code verwenden. Was das für ein Geschäft bedeuten würde, das Waren von beiden bezieht, liegt auf der Hand.

Der in Europa gängigste Typ des Strichcodes ist der **EAN-13**, was soviel wie *European Article Number* der Länge 13 bedeutet. Vergeben werden sie von mehreren Organisationen, und die ersten zwei bis drei Ziffern des Codes identifizieren (im wesentlichen) das Land in dem der Code ausgegeben wurde. Einige der folgenden Ziffern identifizieren den Hersteller, der wiederum weitere Ziffern zur Identifikation seines Produktes zur Verfügung hat.[1] Daraus ergibt sich aber, daß für verschiedene Produkte eines Herstellers ein großer Teil des Strichcodes identisch ist. Außerdem wird der Hersteller auch bei dem Teil, den er selbst bestimmen kann, kaum willkürliche Ziffern vergeben, sondern den Code weiter zur systematischen Produktklassifizierung nutzen wollen. Die Idee des *zufällig* gewählten Codes ist also hinfällig, und es kann durchaus sein, daß sich der Code für eine $100g$-Tafel Schokolade sehr wenig von dem für eine $400g$-Tafel unterscheidet – im Gegensatz zum Preis.

Wir brauchen also eine neue Idee, wie man Fehler bei der Eingabe mit hoher Wahrscheinlichkeit bemerken kann, und die Idee heißt *Redundanz*! Man hängt an den Teil des Codes, den man zur Identifikation des Produktes braucht, zusätzliche (redundante) Ziffern an, deren einziger Sinn es ist, den Code fehlerresistenter zu machen. Dabei sollten möglichst wenig zusätzliche Ziffern eine möglichst hohe Sicherheit bieten. Weshalb und wie man das mit nur einer Ziffer, der sogenannten *Prüfziffer*, erreicht, hat mit Gruppen zu tun – bei **EAN-13** ganz konkret die Gruppe $\mathbb{Z}_{10}$, die wir im Kapitel 24 kennen lernen.

Ein analoges Verfahren wird bei nahezu allen Ziffern- und Buchstabencodes angewendet, die zur Identifizierung von Produkten und Personen verwendet werden: z.B. Kreditkarten, Personalausweise, **ISBN**-Nummern, Seriennummern von Banknoten, etc.. Aber nicht alle verwenden die gleiche Gruppe, z.B. haben **ISBN**-Nummern bis 2007 die Gruppe $\mathbb{Z}_{11}$ verwendet und die alten DM-Scheine die Diëdergruppe $\mathbb{D}_{10}$. Es gibt gute Gründe, verschiedene Gruppen und Verfahren zu verwenden, denn nicht alle bieten die gleiche Sicherheit!

[1] Das Bild des Strichcodes in Figur 1 stammt von der Web-Seite:

http://de.wikipedia.org/wiki/European_Article_Number

Dort findet man auch weitere Informationen zur Struktur des EAN-13 Strichcodes.

Zyklische Codes

Die oben angesprochene Prüfzifferkodierung ist ein Spezialfall des allgemeineren Problems, daß man Information über einen störanfälligen Kanal schicken möchte.

Bei der Prüfzifferkodierung war der Kanal im wesentlichen die Person, die die Ziffern eingibt bzw. der Scanner. Ziel war es, Fehler zu erkennen.

Ein Beispiel für das allgemeinere Problem ist das Abspielen und Hören von CDs. Man kann die CD als Sender, den Nutzer als Empfänger und den CD-Spieler als Kanal auffassen. Im Gegensatz zur Prüfzifferkodierung wird es uns beim Hören einer CD nicht wirklich reichen, daß ein Fehler erkannt wird. Wir wollen zweifelsohne auch, daß er korrigiert wird. Eine Möglichkeit dazu ist, den Laserstrahl die Stelle, an der ein Fehler aufgetreten ist, noch mal lesen zu lassen. Aber wenn ein Fehler z.B. durch einen Kratzer auf der CD entstanden ist, wird das nicht viel helfen. Besser wäre es, wenn kleinere Fehler nicht nur erkannt, sondern auch korrigiert werden könnten.

Wir werden in der Vorlesung die Theorie der fehlerkorrigierenden Codes nicht wirklich behandeln können. Aber die Grundidee läßt sich einfach erläutern. Die Information auf einer CD ist im wesentlichen dadurch gespeichert, daß in einem gewissen Bereich die unteren Schichten des reflektierenden Materials Löcher haben, in anderen nicht – zwei Zustände also, sagen wir 0 und 1. Gehen wir vereinfachend davon aus, daß die Länge eines Bereichs mit Loch bzw. ohne Loch, der Anzahl an Nullen oder Einsen entspricht, so besteht die Information aus einer Ziffernfolge von Nullen und Einsen. Zur sinnvollen Weiterverarbeitung wird die Information in Pakete fester Länge gebündelt, sagen wir etwa Zifferntupel der Länge n. Damit besteht die Information, die über den Kanal geht, also aus Einheiten

$$(c_0, \ldots, c_{n-1}) \in (\mathbb{Z}_2)^n,$$

d.h. aus Elementen eines Vektorraums über dem Körper $\mathbb{Z}_2$. Was ein Vektorraum ist, wird in den *Grundlagen der Mathematik* erläutert, der *Körper* $\mathbb{Z}_2$ wird in dieser Vorlesung eingeführt.

Wie bei der Prüfzifferkodierung werden nicht alle Ziffern des Tupels für die Kodierung der gesendeten Information benötigt, einige sind nur zur Sicherung der Information da. Anders ausgedrückt, nicht jedes Tupel $(c_0, \ldots, c_{n-1})$ wird eine zulässige Information sein, und es wird darauf ankommen, daß die Menge C der zulässigen Codewörter eine zusätzliche algebraische Struktur aufweist, damit es gute Methoden gibt, Fehler zu erkennen und ggf. zu beheben. Genauer gesagt, C sollte zumindest ein *Untervektorraum* sein, um Methoden der *linearen Algebra* anwenden zu können.

$$\text{Z.B.: } C = \{(0,0,0), (1,1,0), (0,1,1), (1,0,1)\} \leq \mathbb{Z}_2^3.$$

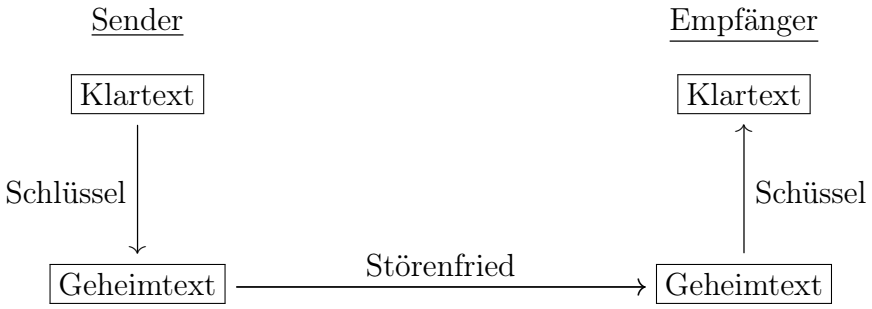
Aber besser ist es noch, einen Vektor $(c_0, \ldots, c_{n-1})$ mit dem Polynom

$$c_0 + c_1 \cdot t + c_2 \cdot t^2 + \ldots + c_{n-1} \cdot t^{n-1}$$

und den Vektorraum $\mathbb{Z}_2^n$ mit einem Faktorring des Polynomrings $\mathbb{Z}_2[t]$ zu identifizieren, vor allem wenn dann C ein *Ideal* in diesem Faktorring ist. Dann reicht nämlich im wesentlichen 1 Element, um alle Codewörter zu beschreiben, wie wir im Kapitel über den Polynomring sehen werden.

Das RSA-Verfahren und der Chinesische Restsatz

Bei den bisher angesprochenen Kodierungen ging es stets darum, Verfälschungen der Informationen zu erkennen und ggf. zu korrigieren, damit sie korrekt beim Empfänger ankommen. Diesen Zweig der Mathematik nennt man Kodierungstheorie und grenzt ihn von der Kryptographie ab. Auch letztere beschäftigt sich mit dem Schutz von Informationen die ein Sender über einen störanfälligen Kanal zu einem Empfänger schickt. Ziel ist es aber primär zu verhindern, daß ein Störenfried die Informationen mithören und *verstehen* oder *unbemerkt verändern* kann. Da wir den Kanal als unsicher annehmen, können wir das *Mithören* in aller Regel nicht verhindern. Also muß beim Verstehen und Verändern angesetzt werden. Die Grundidee ist, den Text zu verschlüsseln.



In der einfachsten Form der aus dem alten Rom überlieferten *Caesar Chiffre* vertauscht man die Buchstaben der Nachricht zyklisch, z.B.

a	b	c	d	e	$\ldots$	x	y	z
$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$		$\downarrow$	$\downarrow$	$\downarrow$
m	n	o	p	q	$\ldots$	j	k	l

Der Schlüssel besteht hierbei aus einer einzigen Zahl, nämlich um wieviel Buchstaben man das "a" nach rechts geschiftet hat; im obigen Beispiel ist dies 12. Eine solch einfache Verschlüsselung ist natürlich auch sehr einfach von einem Störenfried zu brechen. Aber sie weist ein wichtiges Merkmal auf, das auch allen der nach Caesar entwickelten Verschlüsselungsverfahren bis ins letzte Jahrhundert eigen war: der gleiche Schlüssel dient zum Verschlüsseln und zum Entschlüsseln, muß also *geheim* bleiben! Man nennt solche Verschlüsselungsverfahren deshalb *symmetrisch*, und eines ihrer wesentlichen

Sicherheitsrisiken besteht darin, daß Sender und Empfänger zunächst einmal den geheimen Schlüssel austauschen müssen, ohne dabei abgehört werden zu können.

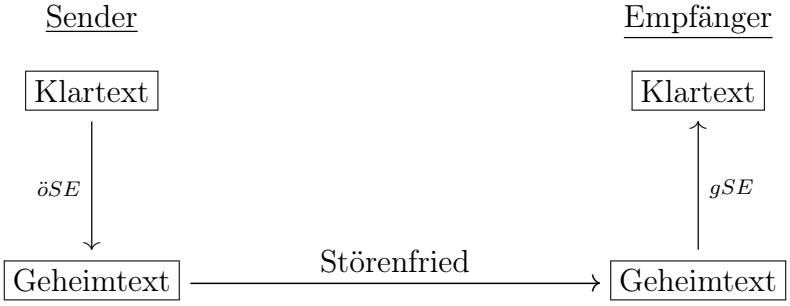
Eine Idee von Whitfield Diffie und Martin Hellman (siehe [DH76]) aus den siebziger Jahren revolutionierte die Kryptographie. Zum Ver- und Entschlüsseln sollten zwei unterschiedliche Schlüssel verwendet werden, und die Kenntnis von einem der beiden und der Nachricht sollte es nicht erlauben, auf den anderen zurückzuschließen. So könnte der Sender einen der beiden Schlüssel *geheim* halten, den anderen aber *öffentlich* bekannt geben. Damit ist es leicht, eine Nachricht so zu verschlüsseln, daß dem Empfänger jede Veränderung auffallen würde. Wir stellen dies in dem folgenden Schema dar, wobei gSS für den *geheimen Schlüssel des Senders* steht und $öSS$ für den *öffentlichen Schlüssel des Senders*:

(111)

Der Störenfried kann die Nachricht zwar abfangen, mit dem (auch ihm bekannten) öffentlichen Schlüssel entschlüsseln und kennt dann deren Inhalt. Da ihm aber der geheime Schlüssel fehlt, kann er die Nachricht nicht verfälschen, wieder verschlüsseln und gefälscht weiter schicken.

Wenn man die Nachricht geheim halten möchte, sollte der Empfänger je einen geheimen und öffentlichen Schlüssel haben. Wie dann die Verschlüsselung aussehen kann, stellen wir in folgendem Schema dar, wobei wir für den geheimen Schlüssel des Empfängers die Abkürzung gSE verwenden und für seinen öffentlichen Schlüssel die Abkürzung $öSE$:

(112)



Da der Störenfried den geheimen Schlüssel des Empfängers nicht kennt, kann er die Nachricht auch nicht entschlüsseln. Verschlüsselungsverfahren dieser Art nennt man *asymmetrisch*, oder spezieller *public key Verfahren*. Aber damit ein solches Verfahren funktionieren kann, muß es einigen wichtigen Anforderungen genügen, und um dies zu beschreiben sollten wir den Begriff der *Verschlüsselung* etwas mathematischer fassen.

Bei der Caesar Chiffre aus obigem Beispiel werden Textblöcke verschlüsselt, die aus einem einzigen Buchstaben bestehen, und man kann die Verschlüsselung als *Abbildung*

$$f_k : \mathcal{N} \longrightarrow \mathcal{N}$$

der Menge

$$\mathcal{N} = \{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z\}$$

in sich selbst auffassen, die von dem Schlüssel k abhängt (in obigem Beispiel $k = 12$) und die *Nachricht* um k Stellen verschiebt – wobei wir im Alphabet mit a weiter machen, wenn wir bei z angekommen sind. Wichtig ist dabei, daß die Funktion eine *Umkehrfunktion* besitzt (man nennt die Funktion f_k dann *bijektiv*), die es erlaubt, den Prozess rückgängig zu machen. In unserem Fall ist dies die Funktion f_{-k}, die eine Nachricht um k Stellen nach links verschiebt. Auch sie hängt von einem Schlüssel ab, und es ist im wesentlichen der gleiche Schlüssel – das Verschlüsselungsverfahren ist *symmetrisch*! Da man für jeden zulässigen Schlüssel eine Funktion f_k zum Verschlüsseln benötigt, spricht man auch von einer *Familie* von Funktionen $\{f_k \mid k \in \mathcal{S}\}$, wobei $\mathcal{S}$ die Menge der zulässigen Schlüssel sein soll. Im Fall der Caesar Chiffre könnten wir $\mathcal{S} = \{-25, -24, \ldots, 24, 25\}$ wählen.

Im Allgemeinen wird man Textblöcke größerer Länge verschlüsseln, und man wird sie in aller Regel zunächst durch einen einfachen Übersetzungsmechanismus in Ziffern überführen, um leichter die Methoden der Mathematik anwenden zu können. Bei der Caesar Chiffre könnte man z.B. die Buchstaben durch ihre Position im Alphabet ersetzen, $a = 1$, $b = 2$, etc., und man könnte $\mathcal{N}$ auf dem Weg etwa mit $\{1, 2, \ldots, 26\}$ oder gar mit $\mathbb{Z}_{26}$ gleichsetzen. Jedenfalls schadet es nichts, wenn wir vereinfachend davon ausgehen, daß die Nachricht, die wir verschlüsseln wollen aus einer Zahl besteht! Für das oben beschriebene *public key Verfahren* benötigen wir dann eine Familie von bijektiven Funktionen $\mathcal{F} = \{f_k : \mathcal{N} \to \mathcal{N} \mid k \in \mathcal{S}\}$ auf der Menge $\mathcal{N}$ der Nachrichten, so daß für jeden Schlüssel $gS \in \mathcal{S}$ ein Schlüssel $öS \in \mathcal{S}$ existiert mit

$$f_{gS} \circ f_{öS} = f_{öS} \circ f_{gS} = \mathrm{id}_{\mathcal{N}}\,. \tag{113}$$

Die Abbildung $f_{öS}$ ist dann die Inverse von f_{gS}, so daß man die Bedingung (113) auch alternativ schreiben könnte als

$$f_k \in \mathcal{S} \quad \Longrightarrow \quad f_k^{-1} \in \mathcal{S}.$$

Die beiden Eigenschaften in (113) bedeuten für die Anwendung, daß es egal ist, ob man den öffentlichen oder den geheimen Schlüssel zum *Verschlüsseln* verwendet, der jeweils andere kann zum *Entschlüsseln* verwendet werden. Das haben wir in den beiden oben beschriebenen Anwendungen (siehe (111) und (112)) bereits ausgenutzt.

Ein ungemein wichtiger Punkt dabei ist natürlich, daß man aus der Kenntnis der Familie $\mathcal{F}$ sowie eines gegebenen öffentlichen Schlüssels $öS$ *keine Chance* hat, den zugehörigen

geheimen Schlüssel gS zu bestimmen. Dabei heißt *keine Chance* nicht, daß es prinzipiell unmöglich ist, sondern daß der notwendige Rechenaufwand nicht in sinnvoller Zeit zu bewerkstelligen ist. Zugleich muß der Rechenaufwand zur Bestimmung von $f_k(n)$ bei gegebenem n und k sehr gering sein, damit man das Verfahren auch praktisch anwenden kann!

Eine solche Familie von Funktionen haben Ronald Rivest, Adi Shamir und Leonard Adleman 1977 (siehe [**RSA78**]) gefunden, und daraus ist das *RSA-Verfahren* entstanden, das aus mathematischer Sicht nicht mehr als die Primfaktorzerlegung der ganzen Zahlen und ein paar einfache Ergebnisse wie den Chinesischen Restsatz oder den Kleinen Satz von Fermat braucht – Ergebnisse, die wir im Rahmen dieser Vorlesung kennenlernen werden. Entscheidend dabei ist folgende Erkenntnis: so einfach die Zerlegung einer Zahl in Primfaktoren *im Prinzip* auch ist, so schwierig ist sie doch ganz *konkret* durchzuführen (selbst für gute Computer), wenn die Zahlen einmal mehrere hundert Ziffern besitzen!

§ A2 Die symmetrische Gruppe

Die symmetrische Gruppe $\mathrm{Sym}(M)$ der bijektiven Selbstabbildungen einer Menge M ist in gewissem Sinn die *Urmutter* aller Gruppen, da jede Gruppe isomorph zu einer Untergruppe von $\mathrm{Sym}(M)$ für ein geeignetes M ist.[2] Für eine beliebige Menge M ist $\mathrm{Sym}(M)$ allerdings wenig nützlich, da man außer der Definition kaum etwas über sie aussagen kann.

Für eine endliche Menge M ist das ganz anders. Zunächst einmal ist es egal, ob wir $\mathrm{Sym}\big(\{m_1,\dots,m_n\}\big)$, für eine beliebige n-elementige Menge $M=\{m_1,\dots,m_n\}$, betrachten oder $\mathbb{S}_n=\mathrm{Sym}\big(\{1,\dots,n\}\big)$. Die beiden Gruppen sind isomorph, und zwar so offensichtlich, daß wir keinen Unterschied machen - wir identifizieren sie. $\mathbb{S}_n$ ist für praktische Anwendungen sehr wichtig. In den Grundlagen der Mathematik wird die Gruppe $\mathbb{S}_n$ vor allem im Zusammenhang mit Determinanten benötigen.

Da die Menge $\{1,\dots,n\}$ endlich ist, können wir die Abbildungsvorschrift einer Permutation $\sigma\in\mathbb{S}_n$ leicht durch eine Art *Wertetabelle* angeben. Als solche sollte man das zweizeilige Schema in der folgenden Definition auffassen. Aus dem Schema ist unmittelbar der Definitionsbereich und der Wertebereich der Permutation ablesbar, so daß wir darauf verzichten können, diesen gesondert anzugeben. D.h. σ ist als Abbildung eindeutig durch dieses Schema bestimmt.

Definition A2.1.
Ist $\sigma\in\mathbb{S}_n$ eine *Permutation* der Menge $\{1,\dots,n\}$, so können wir σ durch das folgende zweizeilige Schema beschreiben:

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

bzw.

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \\ \sigma(a_1) & \sigma(a_2) & \dots & \sigma(a_n) \end{pmatrix},$$

falls $a_1,\dots,a_n$ irgendeine Anordnung der Zahlen $1,\dots,n$ ist.

Beispiel A2.2.
Die Gruppe $\mathbb{S}_n$ ist für $n\geq 3$ nicht abelsch, denn für die Permutationen

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}\in\mathbb{S}_3$$

gilt

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}\circ\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}\neq\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}\circ\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

[2] Dies ist die Aussage des Satzes von Cayley. Vergleiche dazu Aufgabe 22.46.

Beachte, daß es bei dem Schema nicht darauf ankommt, in welcher Reihenfolge die Zahlen von 1 bis n in der ersten Zeile stehen. Es gilt etwa:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \end{pmatrix}.$$

Es empfiehlt sich aber der Übersichtlichkeit halber für gewöhnlich, die Ziffern in aufsteigender Reihenfolge anzuordnen.

Bemerkung A2.3.
Die oben eingeführte Darstellung einer Permutation hat den angenehmen Nebeneffekt, daß man das Inverse der Permutation leicht angeben kann, indem man einfach die beiden Zeilen vertauscht. Sprich, für eine Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix} \in \mathbb{S}_n$$

ist das Inverse σ^{-1} gegeben ist durch

$$\sigma^{-1} = \begin{pmatrix} \sigma(1) & \sigma(2) & \dots & \sigma(n) \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Nun sind Mathematiker von Haus aus *faule* oder vielleicht richtiger *effiziente* Menschen, und so haben sie sich eine Schreibweise ersonnen, wie man eine Permutation darstellen kann und dabei jede der Zahlen $1, \dots, n$ höchstens *einmal* statt zweimal schreiben muß. Um dies zu bewerkstelligen, benötigen wir einen speziellen Typ von Permutation – eine, die k der Zahlen $1, \dots, n$ *zyklisch vertauscht*.

Definition A2.4.

a. Sei $\{1, \dots, n\} = \{a_1, \dots, a_k\} \cup \{b_1, \dots, b_{n-k}\}$, $k \geq 2$, und

$$\sigma = \begin{pmatrix} a_1 & a_2 & \dots & a_{k-1} & a_k & b_1 & \dots & b_{n-k} \\ a_2 & a_3 & \dots & a_k & a_1 & b_1 & \dots & b_{n-k} \end{pmatrix} \in \mathbb{S}_n,$$

so heißt σ ein **k-Zyklus**, und wir sagen, daß sie die Zahlen $a_1, \dots, a_k$ *zyklisch vertauscht*. Die Abbildungsvorschrift eines solchen k-Zyklus läßt sich deutlich kompakter durch das folgende einzeilige Schema repräsentieren:

(114) $$\sigma = (a_1 \dots a_k).$$

b. Ein $2-Zyklus$ wird auch eine **Transposition** genannt. Eine Transposition $\tau = (i\ j)$ ist mithin eine Permutation, die nur die zwei Zahlen i und j miteinander vertauscht, alle anderen aber fest läßt.

c. Das neutrale Element von $\mathbb{S}_n$, per definitionem $\mathrm{id}_{\{1,\dots,n\}}$, wollen wir der Einfachheit halber mit id bezeichnen.

Bemerkung A2.5.
Die Interpretation der Schreibweise in Gleichung (114) ist offensichtlich, das erste Element a_1 wird auf das zweite a_2 abgebildet, das zweite auf das dritte, und so weiter, bis schließlich das letzte, nämlich a_k, auf das erste, das heißt auf a_1, abgebildet wird – der *Kreis* schließt sich. Beachte hierbei, daß die Zyklen $(a_1 \dots a_k)$, $(a_k a_1 \dots a_{k-1})$, etc. stimmen überein! Um diese Mehrdeutigkeit zu vermeiden, empfiehlt es sich, einen Zyklus stets mit der kleinsten der Zahlen $a_1, \dots, a_k$ zu beginnen.

Bisher haben wir k-Zyklen nur für $k \geq 2$ definiert. Wir nun auch 1-Zyklen, etwa (1) oder (3), zulassen und definieren diese in natürlicher Weise als die Identität. □

Beispiel A2.6.
Die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix} \in \mathbb{S}_4 \quad \text{und} \quad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 2 & 5 \end{pmatrix} \in \mathbb{S}_5$$

sind jeweils 3-Zyklen, die die Zahlen $1, 4, 2$ zyklisch vertauschen. In der oben eingeführten Zyklenschreibweise gilt

$$\sigma = (1\ 4\ 2) \quad \text{und} \quad \pi = (1\ 4\ 2).$$

Damit wird der Nachteil dieser Schreibweise gegenüber dem zweizeiligen Schema deutlich – weder der Definitionsbereich noch der Wertebereich lassen sich aus der Zyklenschreibweise eindeutig ablesen. Aber diesen Preis sind wir für die gewonnene *Übersichtlichkeit* gerne bereit zu zahlen. Denn einerseits ist in Anwendungen meist zweifelsfrei bekannt, was n ist, und andererseits ist die wesentliche Information für uns letztlich, welche Zahlen durch die Permutation vertauscht werden, und nicht, welche unbewegt bleiben. □

Nun wäre die Zyklenschreibweise aber nicht sehr hilfreich, wenn wir sie nur für k-Zyklen anwenden könnten, alle anderen Permutationen aber weiterhin in dem zweireihigen Schema angegeben werden müßten. Da kommt uns die Feststellung zu Hilfe, daß jede Permutation sich als Komposition von paarweise *disjunkten* Zyklen schreiben läßt.

Satz A2.7.
Ist $\sigma \in \mathbb{S}_n$ eine Permutation, so gibt es eine disjunkte Zerlegung

$$\{1, \dots, n\} = \dot{\bigcup_{i=1}^{t}} \{a_{i1}, \dots, a_{ik_i}\},$$

so daß

$$\sigma = (a_{11} \cdots a_{1k_1}) \circ \dots \circ (a_{t1} \cdots a_{tk_t}).$$

Wir nennen diese Darstellung die *Zyklenzerlegung* von σ, und wir nennen die Zyklen *paarweise disjunkt*. Beachte auch, daß $k_1+\ldots+k_t = n$ und daß $0 \leq k_i \leq n$ für $i = 1, \ldots, t$.

Beweis: Um die Zyklen der Zyklenzerlegung zu finden, betrachten wir die Äquivalenzrelation aus Aufgabe A2.23 auf $\{1, \ldots, n\}$, die durch

$$a \sim b \quad \Longleftrightarrow \quad \exists\, m \in \mathbb{Z} \;:\; b = \sigma^m(a)$$

für $a, b \in \{1, \ldots, n\}$ gegeben ist. Für $a \in \{1, \ldots, n\}$ hat die Äquivalenzklasse von a die Form

$$\overline{a} = \big\{a, \sigma(a), \sigma^2(a), \ldots, \sigma^{k-1}(a)\big\}, \tag{115}$$

wobei

$$k = \min\{l > 0 \mid \sigma^l(a) = a\} = |\overline{a}|.$$

Gemäß Proposition 6.10 bilden die Äquivalenzklassen von $\sim$ eine disjunkte Zerlegung von $\{1, \ldots, n\}$. Wir können also Zahlen $a_{11}, \ldots, a_{t1} \in \{1, \ldots, n\}$ so wählen, daß

$$\{1, \ldots, n\} = \dot{\bigcup_{i=1}^{t}} \overline{a_{i1}}.$$

Setzen wir nun $k_i = |\overline{a_{i1}}|$ und $a_{ij} = \sigma^{j-1}(a_{i1})$, dann gilt wegen (115)

$$\{1, \ldots, n\} = \dot{\bigcup_{i=1}^{t}} \{a_{i1}, a_{i2}, \ldots, a_{ik_i}\}. \tag{116}$$

Es bleibt also noch

$$\sigma = \sigma_1 \circ \cdots \circ \sigma_t$$

zu zeigen, wobei $\sigma_i = (a_{i1} \cdots a_{ik_i})$ ein k_i-Zyklus ist. Sei dazu $b \in \{1, \ldots, n\}$, so ist $b = a_{ij} = \sigma^{j-1}(a_{i1})$ für ein $1 \leq i \leq t$ und ein $1 \leq j \leq k_i$. Wenden wir nun σ auf b an, so erhalten wir

$$\sigma(b) = \sigma(a_{ij}) = \sigma^j(a_{i1}) = \left\{ \begin{array}{ll} a_{ij+1}, & \text{falls } j < k_i, \\ a_{i1}, & \text{falls } j = k_i \end{array} \right\} = \sigma_i(b).$$

Da die Zerlegung in (116) disjunkt ist und sowohl b, als auch $\sigma_i(b)$ in $\{a_{i1}, \ldots, a_{ik_i}\}$ liegen, werden b und $\sigma_i(b)$ von allen σ_l mit $l \neq i$ fest gelassen, d.h.

$$\big(\sigma_1 \circ \cdots \circ \sigma_t\big)(b) = \sigma_i(b) = \sigma(b).$$

Damit ist die Aussage des Satzes gezeigt. □

Bemerkung A2.8.
Beachte, daß für zwei disjunkte Zyklen $\sigma = (a_1 \;\ldots\; a_k), \pi = (b_1 \;\ldots\; b_l) \in \mathbb{S}_n$ offenbar

$$\sigma \circ \pi = \pi \circ \sigma$$

gilt. Denn für $c \in \{a_1, \ldots, a_k\}$ gilt $\sigma(c) \in \{a_1, \ldots, a_k\}$ und deshalb notwendig $c, \sigma(c) \notin \{b_1, \ldots, b_l\}$, so daß

$$(\sigma \circ \pi)(c) = \sigma\big(\pi(c)\big) = \sigma(c) = \pi\big(\sigma(c)\big) = (\pi \circ \sigma)(c). \tag{117}$$

In den Fällen $c \in \{b_1, \ldots, b_l\}$ und $c \notin \{a_1, \ldots, a_k\} \cup \{b_1, \ldots, b_l\}$ zeigt man (117) analog, so daß die obige Behauptung folgt.

Zudem ist offensichtlich, daß die Zyklenzerlegung von σ bis auf die Reihenfolge der Zyklen *eindeutig* ist, da die Elemente der Zyklen von σ zyklisch vertauscht werden.

Und schließlich ist eine Permutation auch in Zyklenschreibweise leicht zu invertieren, indem man sie einfach von hinten nach vorne liest. Denn für einen k-Zyklus $\sigma = (a_1 \ \ldots \ a_k)$ ist offenbar das Inverse

$$\sigma^{-1} = (a_k \ a_{k-1} \ \ldots \ a_2 \ a_1)$$

wieder ein k-Zyklus, und somit ist für

$$\pi = (a_{11} \cdots a_{1k_1}) \circ \ldots \circ (a_{t1} \cdots a_{tk_t})$$

das Inverse gegeben durch

$$\pi^{-1} = (a_{tk_t} \cdots a_{t1}) \circ \ldots \circ (a_{1k_1} \cdots a_{11}).$$

□

Beispiel A2.9.
Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix} \in \mathbb{S}_5$$

hat die Zyklenzerlegung

$$\sigma = (1\ 2\ 5) \circ (3\ 4) = (3\ 4) \circ (1\ 2\ 5). \tag{118}$$

Ferner ist das Inverse zu σ gegeben durch

$$\sigma^{-1} = (4\ 3) \circ (5\ 2\ 1) = (1\ 5\ 2) \circ (3\ 4).$$

Eine berechtigte Frage ist, wie wir die Zyklenzerlegung in (118) gefunden haben. Wir wollen versuchen, dies so in Worte zu fassen, daß dem Leser daraus die allgemeine Vorgehensweise ersichtlich wird. Man starte mit der kleinsten Zahl, 1, und suche ihr Bild unter σ, also $\sigma(1) = 2$. Das liefert den Startteil des ersten Zyklus:

$$(1\ 2$$

Sodann betrache man das Bild von 2 unter σ, also $\sigma(2) = 5$, und erhält:

$$(1\ 2\ 5$$

Man fährt mit dem Bild von 5 unter σ, also $\sigma(5) = 1$, fort. Da dieses das erste Element des ersten Zyklus war, schließen wir den Zyklus,

$$(1\ 2\ 5),$$

und beginnen den zweiten Zyklus mit der kleinsten Zahl in $\{1, \dots, 5\}$, die noch nicht in dem ersten Zyklus vorkommt, also mit 3:

$$(1\ 2\ 5) \circ (3$$

Dann betrachten wir deren Bild unter σ, also $\sigma(3) = 4$, und setzen so unseren zweiten Zyklus fort:

$$(1\ 2\ 5) \circ (3\ 4$$

Da bereits alle fünf Elemente von $\{1, \dots, 5\}$ aufgebraucht sind, muß notwendig $\sigma(4) = 3$ gelten, was es auch tut, und wir können damit auch den zweiten Zyklus schließen:

$$\sigma = (1\ 2\ 5) \circ (3\ 4).$$

Wie gesagt, da in $\{1, \dots, 5\}$ keine Zahl mehr übrig ist, sind wir fertig und haben die Zyklenzerlegung von σ gefunden.

Das hier beschriebene Verfahren ist die praktische Umsetzung des Beweises von Satz A2.7, denn wir können die Zyklenzerlegung nun auch in folgender Form schreiben

$$\sigma = \big(1\ \ \sigma(1)\ \ \sigma^2(1)\big) \circ \big(3\ \ \sigma(3)\big),$$

wobei $\sigma^3(1) = 1$ und $\sigma^2(3) = 3$ gilt. □

Von jetzt an werden wir zwischen den beiden Darstellungsarten für Permutationen hin und her wechseln und stets die verwenden, die für unsere Zwecke am besten geeignet ist.

Bemerkung A2.10.
Für kleine Werte n ist $\mathbb{S}_n$ sehr übersichtlich, für große Werte n wird $\mathbb{S}_n$ jedoch riesig. $\mathbb{S}_1 = \{\mathrm{id}\}$ und $\mathbb{S}_2 = \{\mathrm{id}, (1\ 2)\}$. $\mathbb{S}_3 = \{\mathrm{id}, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$ hat schon sechs Elemente, $\mathbb{S}_4$ gar 24 und $\mathbb{S}_{64}$ ungefähr 10^{89}. Letztere Zahl entspricht in etwa der angenommenen Anzahl der Atome im Universum.

Proposition A2.11.
$|\mathbb{S}_n| = n! = 1 \cdot 2 \cdot 3 \cdots n$.

Bevor wir einen formalen Beweis dieser Aussage mittels Induktion geben, wollen wir ein Argument dafür geben, weshalb die Aussage richtig sein sollte. Im Prinzip handelt es sich dabei um die Idee des anschließenden formalen Beweises.

Beweisidee für Proposition A2.11: Eine Permutation $\sigma \in \mathbb{S}_n$ ist durch die Bilder $\sigma(1), \ldots, \sigma(n)$ der Zahlen $1, \ldots, n$ festgelegt, wobei jede der Zahlen $1, \ldots, n$ unter den Zahlen $\sigma(1), \ldots, \sigma(n)$ genau einmal vorkommt. Wir wollen nun zählen, wieviele Möglichkeiten es für die Definition einer solchen Permutation gibt. Zunächst müssen wir $\sigma(1)$, d.h. das Bild von 1, festlegen. Dazu haben wir noch volle n Zahlen zur Auswahl. Ist dieses festgelegt, so bleiben für das Bild $\sigma(2)$ der 2 nur noch $n-1$ Zahlen übrig. Für $\sigma(3)$ sind es schon nur noch $n-2$. Wenn man so fortfährt, hat man allgemein für $\sigma(i)$ noch $n-i+1$ Möglichkeiten, also schließlich für $\sigma(n-1)$ noch $n-(n-1)+1=2$ und für $\sigma(n)$ noch genau eine. Insgesamt gibt es deshalb

$$n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1 = n!$$

Möglichkeiten eine Permutation zu definieren, und mithin gibt es $n!$ verschiedene Permutationen. □

Unsauber an dieser Beweisidee ist der Teil "*und wenn man so fortfährt*". Ihn mathematisch sauber und korrekt zu fassen, heißt, eine Induktion zu führen.

Beweis von Proposition A2.11: Wir zeigen durch Induktion über n etwas allgemeiner:

Behauptung: Sind $M = \{m_1, \ldots, m_n\}$ und $N = \{n_1, \ldots, n_n\}$ zwei n-elementige Mengen, so hat die Menge

$$\mathrm{Iso}(M,N) := \{f : M \to N \mid f \text{ ist bijektiv}\}$$

genau $n!$ Elemente.

Induktionsanfang: Sei $n = 1$, dann gilt offensichtlich $|\,\mathrm{Iso}(M,N)| = 1 = 1!$.

Induktionsvoraussetzung: Es sei $n > 1$ beliebig, aber fest, und es gelte $|\,\mathrm{Iso}(M',N')| = (n-1)!$ für alle $n-1$-elementigen Mengen M' und N'.

Induktionsschluß: Seien nun M und N zwei n-elementige Mengen. Für $i \in \{1, \ldots, n\}$ definieren wir:

$$\mathrm{Iso}_i := \{f \in \mathrm{Iso}(M,N) \mid f(m_1) = n_i\}.$$

Offensichtlich ist die Einschränkung

$$\mathrm{Iso}_i \to \mathrm{Iso}\left(M \setminus \{m_1\}, N \setminus \{n_i\}\right) : f \mapsto f_{|M\setminus\{m_1\}}$$

bijektiv, und daher gilt nach Induktionsvoraussetzung $|\,\mathrm{Iso}_i\,| = (n-1)!$. Da nun außerdem

$$\mathrm{Iso}(M,N) = \dot{\bigcup_{i=1}^{n}} \mathrm{Iso}_i,$$

d. h. $(\mathrm{Iso}_1, \ldots, \mathrm{Iso}_n)$ ist eine disjunkte Zerlegung von $\mathrm{Iso}(M,N)$, folgt:

$$|\,\mathrm{Iso}(M,N)| = \sum_{i=1}^{n} |\,\mathrm{Iso}_i\,| = n \cdot (n-1)! = n!.$$

□

Bemerkung A2.12.
Wir wollen uns jetzt mit den Transpositionen näher beschäftigen. Zunächst ist klar, daß für eine Transposition $\tau \in \mathbb{S}_n$ gilt $\tau^{-1} = \tau$, also $\tau^2 = \mathrm{id}$.

Proposition A2.13.
Jede Permutation in $\mathbb{S}_n$, $n \geq 2$, läßt sich als Komposition von höchstens n Transpositionen darstellen.

Beweis: Ist $\sigma = (a_1 \ \ldots \ a_k)$ ein k-Zyklus mit $k \geq 2$, so gilt offenbar, daß

$$\sigma = (a_1\ a_2) \circ (a_2\ a_3) \circ \ldots \circ (a_{k-2}\ a_{k-1}) \circ (a_{k-1}\ a_k) \tag{119}$$

ein Produkt von $k-1$ Transpositionen ist. Ist nun $\mathrm{id} \neq \sigma \in \mathbb{S}_n$ beliebig, so hat σ nach Satz A2.7 eine Zyklenzerlegung der Form

$$\sigma = \sigma_1 \circ \ldots \circ \sigma_t$$

wobei $\sigma_i = (a_{i1} \cdots a_{ik_i})$ ein k_i-Zyklus ist. Da disjunkte Zyklen miteinander kommutieren, können wir ohne Einschränkung[3] annehmen, daß $k_1 \geq k_2 \geq \ldots \geq k_t$. Zudem ist σ nicht das neutrale Element id, so daß notwendig $k_1 \geq 2$ gilt und die Zahl $s = \max\{r \mid 1 \leq r \leq t, k_r \geq 2\}$ definiert ist. Damit ist aber $\sigma_i = \mathrm{id}$ für $i = s+1, \ldots, t$ und somit ist

$$\sigma = \sigma_1 \circ \ldots \circ \sigma_s$$

das Produkt von s Zyklen. Da sich σ_i als Produkt von $k_i - 1$ Transpositionen schreiben läßt, läßt sich σ als Produkt von

$$(k_1 - 1) + \ldots + (k_s - 1) = (k_1 + \ldots + k_s) - s \leq n - 1$$

Transpositionen schreiben. Die Behauptung ist also für $\sigma \neq \mathrm{id}$ gezeigt. Da aber $n \geq 2$ und zudem $\mathrm{id} = (1\ 2) \circ (1\ 2)$ das Produkt von zwei Transpositionen ist, ist die Proposition bewiesen. □

Der Beweis ist *konstruktiv*, da die Gleichung (119) angibt, wie man einen Zyklus in Transpositionen zerlegt und somit die Aufgabe, eine Permutation als Produkt von Transpositionen zu schreiben, auf die Berechnung einer Zyklenzerlegung reduziert. Allerdings haben wir zur Zerlegung (119) lediglich gesagt, daß diese *offensichtlich* gilt. Man sieht dies, indem man die beiden Abbildungen, die links und rechts des

[3] Die Aussage "*wir können ohne Einschränkung annehmen*" bedeutet, daß wir nur einen *speziellen* Fall betrachten, daß aber offensichtlich ist, wie man aus diesem Spezialfall den allgemeinen Fall herleiten würde. Letzteres tut man dann nicht explizit, da es meist mit einem hohen Notationsaufwand und vielen Indizes verbunden wäre, ohne eine tiefere Einsicht zu bringen. Man sollte allerdings nur dann etwas ohne Einschränkung annehmen, wenn man sich sicher ist, daß die übrigen Fälle in der Tat leicht aus dem Spezialfall folgen!

Gleichheitszeichens in (119) vorkommen auf die Zahlen $\{1,\dots,n\}$ anwendet – das haben wir durch *Hinschauen* getan, aber man könnte es natürlich auch formal mit allen zu betrachtenden Fällen hinschreiben. Der Beweis würde dadurch nicht verständlicher.

Der Beweis von Satz A2.7 war ziemlich technisch, und so mag es dem einen oder anderen Leser Unbehagen bereiten, daß wir diese Aussage im Beweis von Proposition A2.13 verwendet haben. Deshalb geben wir noch einen alternativen Beweis der Aussage von Proposition A2.13, der ohne Satz A2.7 auskommt, für das Finden der Zerlegung in Transpositionen aber weniger hilfreich ist.

Alternativer Beweis von Proposition A2.13: Wir führen den Beweis durch Induktion über n.

Induktionsanfang: Sei $n = 2$. Es ist $\mathbb{S}_2 = \{\mathrm{id}, (1\ 2)\}$, und $\mathrm{id} = (1\ 2) \circ (1\ 2)$, also folgt die Behauptung.

Induktionsschluß: Sei nun $n \geq 2$ gegeben, und die Behauptung gelte für n bereits. Ferner sei $\sigma \in \mathbb{S}_{n+1}$ beliebig, aber fest. Es gibt ein $i \in \{1,\dots,n+1\}$ mit $\sigma(n+1) = i$. Dann gilt mit $\tau = (n+1\ \ i)$

$$(\tau \circ \sigma)(n+1) = n+1,$$

also können wir die Einschränkung$\sigma' = (\tau \circ \sigma)_{|\{1,\dots,n\}}$ als Element von $\mathbb{S}_n$ auffassen. Mithin gilt nach Induktionsvoraussetzung, es gibt Transpositionen $\tau'_1,\dots,\tau'_k \in \mathbb{S}_n$, $k \leq n$, mit

$$\sigma' = \tau'_1 \circ \cdots \circ \tau'_k.$$

Bezeichnen wir mit τ_j die Fortsetzung von τ'_j, die definiert wird durch

$$\tau_j : \{1,\dots,n+1\} \to \{1,\dots,n+1\} : l \mapsto \begin{cases} \tau'_j(l), & \text{falls } l \leq n \\ n+1, & \text{falls } l = n+1, \end{cases}$$

so folgt unmittelbar

$$\tau \circ \sigma = \tau_1 \circ \cdots \circ \tau_k,$$

und mithin

$$\sigma = (\tau \circ \tau) \circ \sigma = \tau \circ (\tau \circ \sigma) = \tau \circ \tau_1 \circ \cdots \circ \tau_k.$$

D. h. σ ist Komposition von $k+1 \leq n+1$ Transpositionen. □

Korollar A2.14.
Jede Permutation in $\mathbb{S}_n$, $n \geq 2$, läßt sich als Produkt von Transpositionen zweier aufeinanderfolgender Zahlen schreiben.

Beweis: Wegen Proposition A2.13 reicht es, dies für eine Transposition $(i\ j)$ mit $i < j$ zu zeigen. Es gilt aber offenbar

$$\begin{aligned}(i\ j) &= (i\ i+1) \circ (i+1\ i+2) \circ \cdots \circ (j-2\ j-1) \circ (j-1\ j) \circ \\ &\quad \circ (j-2\ j-1) \circ \cdots \circ (i+1\ i+2) \circ (i\ i+1).\end{aligned}$$

□

Die Darstellung einer Permutation als Komposition von Transpositionen ist also keineswegs eindeutig. Was jedoch unabhängig ist, ist, daß eine Permutation entweder immer durch eine gerade oder immer durch eine ungerade Anzahl von Transpositionen darstellbar ist. Das wollen wir nun beweisen und definieren dazu das Vorzeichen einer Permutation.

Definition A2.15.
Es sei $\sigma \in \mathbb{S}_n$ gegeben.

a. Ein Zahlenpaar (i, j) mit $1 \leq i, j \leq n$ heißt ein *Fehlstand* von σ, falls $i < j$, aber $\sigma(i) > \sigma(j)$.

b. Wir definieren das *Signum* oder *Vorzeichen* von σ durch

$$\operatorname{sgn}(\sigma) = \begin{cases} +1, & \text{falls } \sigma \text{ eine gerade Anzahl von Fehlständen besitzt,} \\ -1, & \text{falls } \sigma \text{ eine ungerade Anzahl von Fehlständen besitzt.} \end{cases}$$

Beispiel A2.16.
Eine Transposition $\tau = (i\ j) \in \mathbb{S}_n$, mit $i < j$, hat die $2 \cdot (j - i - 1) + 1$ Fehlstände

$$(i, i+1), (i, i+2), \ldots, (i, j), (i+1, j), (i+2, j), \ldots, (j-1, j),$$

und mithin gilt $\operatorname{sgn}(\tau) = -1$.

Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

hat die Fehlstände $(1, 2)$ und $(3, 4)$. Also gilt $\operatorname{sgn}(\sigma) = 1$. □

Manchmal ist die folgende geschlossene Formel nützlich, deren Beweis als Übungsaufgabe dem Leser überlassen sei, da wir sie im folgenden nicht verwenden werden.

Bemerkung A2.17.
Für $\sigma \in \mathbb{S}_n$ gilt:

$$\operatorname{sgn}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} = \frac{\sigma(2) - \sigma(1)}{2 - 1} \cdot \frac{\sigma(3) - \sigma(1)}{3 - 1} \cdots \frac{\sigma(n) - \sigma(n-1)}{n - (n-1)}.$$

□

Weit wichtiger ist für uns die folgende Eigenschaft des Signums.

Satz A2.18.

a. Die Abbildung
$$\operatorname{sgn} : (\mathbb{S}_n, \circ) \longrightarrow \big(\{1,-1\}, \cdot\big)$$
ist ein Gruppenhomomorphismus, d.h. für $\sigma_1, \sigma_2 \in \mathbb{S}_n$ gilt
$$\operatorname{sgn}(\sigma_1 \circ \sigma_2) = \operatorname{sgn}(\sigma_1) \cdot \operatorname{sgn}(\sigma_2).$$

b. Ist $\sigma = \tau_1 \circ \cdots \circ \tau_k \in \mathbb{S}_n$ eine Komposition von k Transpositionen, dann gilt:
$$\operatorname{sgn}(\sigma) = (-1)^k.$$

c. Ist $\sigma \in \mathbb{S}_n$, so kann σ entweder nur als Produkt einer geraden Anzahl von Transpositionen geschrieben werden oder nur als Produkt einer ungeraden Anzahl von Transpositionen.

Beweis: Es sei $\sigma = \sigma' \circ \tau \in \mathbb{S}_n$ mit $\sigma' \in \mathbb{S}_n$ und $\tau = (i\ i+1)$ für ein $i \in \{1, \ldots, n-1\}$. Ist $(i, i+1)$ ein Fehlstand von σ', so hebt τ diesen auf und σ hat einen Fehlstand weniger als σ'. Ist hingegen $(i, i+1)$ kein Fehlstand von σ', so erzeugt die Komposition mit τ diesen Fehlstand neu und σ hat einen Fehlstand mehr als σ'. Damit gilt dann aber
$$\operatorname{sgn}(\sigma) = -\operatorname{sgn}(\sigma') = \operatorname{sgn}(\sigma') \cdot \operatorname{sgn}(\tau).$$
Nach Korollar A2.14 läßt sich jede Permutation als Produkt von Transpositionen aufeinanderfolgender Zahlen schreiben.

Seien nun $\sigma_1 = \tilde{\tau}_1 \circ \cdots \circ \tilde{\tau}_r$ und $\sigma_2 = \tilde{\tau}_{r+1} \circ \cdots \circ \tilde{\tau}_{r+s}$ als Produkte solcher Transpositionen aufeinanderfolgender Zahlen gegeben. Dann folgt mit Induktion über $r+s$, daß
$$\operatorname{sgn}(\sigma_1 \circ \sigma_2) = (-1)^{r+s} = (-1)^r \cdot (-1)^s = \operatorname{sgn}(\sigma_1) \cdot \operatorname{sgn}(\sigma_2).$$
Damit ist a. gezeigt und b. folgt mittels Induktion nach k.

Für c. sei schließlich $\sigma = \tau_1 \circ \cdots \circ \tau_k = \tau'_1 \circ \cdots \circ \tau'_l$ mit Transpositionen $\tau_i, \tau'_j \in \mathbb{S}_n$. Dann folgt aus b.
$$(-1)^k = \operatorname{sgn}(\sigma) = (-1)^l,$$
und deshalb sind entweder k und l beide gerade oder beide ungerade. □

Definition A2.19.
$\mathbb{A}_n := \operatorname{Ker}(\operatorname{sgn}) = \{\sigma \in \mathbb{S}_n \mid \operatorname{sgn}(\sigma) = 1\}$ heißt *alternierende Gruppe* vom Grad n.

Bemerkung A2.20.
Der Kern des Signums besteht aus allen Permutationen mit positivem Vorzeichen, man

nennt diese auch *gerade* Permutationen, und ist nach Proposition 22.22 eine Untergruppe der $\mathbb{S}_n$.

Die Menge $\{\sigma \in \mathbb{S}_n \mid \operatorname{sgn}(\sigma) = -1\}$ ist keine Untergruppe der $\mathbb{S}_n$, da sie etwa das neutrale Element id nicht enthält.

Korollar A2.21 (Die alternierende Gruppe).

a. Für $\sigma \in \mathbb{S}_n$ gilt $\operatorname{sgn}(\sigma) = \operatorname{sgn}\left(\sigma^{-1}\right)$.

b. Ist $\tau = (i\ j)$ eine Transposition, so gilt $\mathbb{S}_n = \mathbb{A}_n \dot\cup \mathbb{A}_n\tau$.

Beweis:

a. Es gilt $\operatorname{sgn}(\sigma) \cdot \operatorname{sgn}\left(\sigma^{-1}\right) = \operatorname{sgn}\left(\sigma \circ \sigma^{-1}\right) = \operatorname{sgn}(\mathrm{id}) = 1$. Daraus folgt die Behauptung, da $\operatorname{sgn}(\sigma)$ und $\operatorname{sgn}\left(\sigma^{-1}\right)$ nur die Werte 1 und -1 annehmen können.

b. Ist $\sigma \in \mathbb{S}_n$, so gilt entweder $\operatorname{sgn}(\sigma) = 1$ oder $\operatorname{sgn}(\sigma) = -1$. In ersterem Fall ist $\sigma \in \mathbb{A}_n$; in letzterem Fall ist $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau) = (-1) \cdot (-1) = 1$ und somit $\sigma \circ \tau \in \mathbb{A}_n$, dann ist aber

$$\sigma = (\sigma \circ \tau) \circ \tau \in \mathbb{A}_n\tau.$$

Dies zeigt, daß jedes Element der $\mathbb{S}_n$ in einer der beiden Mengen $\mathbb{A}_n$ oder $\mathbb{A}_n\tau$ liegt. Beachte auch, daß die Elemente der beiden Mengen verschiedenes Signum haben, so daß die Mengen in der Tat disjunkt sind.

□

Bemerkung A2.22.

Für $n \in \mathbb{Z}$ mit $n \geq 3$ können wir zwei Permutationen

$$\pi_n = (1\ 2\ \ldots\ n-1\ n)$$

und

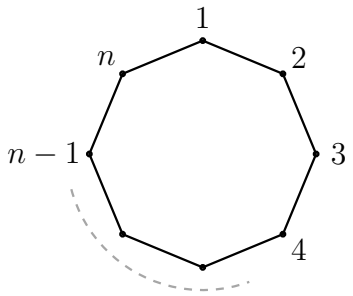
$$\sigma_n = \begin{cases} (1\ n) \circ (2\ n-1) \circ \ldots \circ \left(\frac{n}{2}\ \frac{n}{2}+1\right), & \text{falls } n \text{ gerade,} \\ (1\ n) \circ (2\ n-1) \circ \ldots \circ \left(\frac{n-1}{2}\ \frac{n+3}{2}\right), & \text{falls } n \text{ ungerade} \end{cases}$$

in $\mathbb{S}_n$ betrachten. Sie erzeugen die sogenannte *Diëdergruppe*

$$\mathbb{D}_{2n} = \langle \pi_n, \sigma_n \rangle \leq \mathbb{S}_n$$

der Ordnung $2n$.

Numeriert man die Ecken eines regulären n-Ecks im Uhrzeigersinn von 1 bis n,



so kann man π_n als Drehung des n-Ecks im Uhrzeigersinn um den Winkel $\frac{2\pi}{n}$ im Bogenmaß auffassen, die die Ecke mit Nummer 1 auf die Ecke mit Nummer 2 abbildet, die Ecke mit Nummer 2 auf die Ecke mit Nummer 3 und so weiter. Ensprechend kann man σ_n als Achsenspiegelung interpretieren. Die Diëdergruppe $\mathbb{D}_{2n}$ ist dann die volle Symmetriegruppe des regulären n-Ecks. Jedes Element entspricht entweder einer Drehung oder einer Spiegelung. (Siehe auch Beispiel 22.9.)

Die Gruppen $\mathbb{D}_8$ und $\mathbb{D}_{10}$ in den Aufgaben A2.26–24.22 sind Spezialfälle von solchen Diëdergruppen. Sie sind die Symmetriegruppen des Quadrates bzw. des regulären Fünfecks.

Aufgaben

Aufgabe A2.23.
Es sei M eine Menge und $\sigma \in \mathrm{Sym}(M)$ eine bijektive Selbstabbildung.

a. Durch
$$a \sim b \quad \Longleftrightarrow \quad \exists\, m \in \mathbb{Z} \;:\; b = \sigma^m(a)$$
für $a, b \in M$ wird eine Äquivalenzrelation auf der Menge M definiert.

b. Es sei $\overline{a}$ für $a \in M$ eine *endliche* Äquivalenzklasse bezüglich $\sim$ der Mächtigkeit $|\overline{a}| = n < \infty$. Dann gelten die folgenden Aussagen:
   (i) Das Minimum $k = \min\{l > 0 \mid \sigma^l(a) = a\}$ existiert.
   (ii) Für $q \in \mathbb{Z}$ ist $\sigma^{q\cdot k}(a) = a$.
   (iii) $\overline{a} = \big\{a, \sigma(a), \ldots, \sigma^{k-1}(a)\big\}$.
   (iv) $\overline{a}$ enthält genau k Elemente.

c. Es sei $M = \{1, \ldots, 7\}$ und $\sigma \in \mathrm{Sym}(M) = \mathbb{S}_7$ sei durch folgende Wertetabelle gegeben:

a	1	2	3	4	5	6	7
$\sigma(a)$	3	4	1	7	2	6	5

Was sind die Äquivalenzklassen bezüglich obiger Äquivalenzrelation?

Aufgabe A2.24.
Betrachte die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 6 & 3 & 7 & 5 & 1 & 4 \end{pmatrix}, \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 2 & 7 & 1 & 4 & 5 & 6 \end{pmatrix} \in \mathbb{S}_7.$$

a. Berechne $\sigma \circ \pi$, $\pi \circ \sigma$, σ^{-1}, π^{-1}.

b. Bestimme für jede der Permutationen in a. die Zyklenzerlegung.

c. Schreibe $\sigma \circ \pi$ als ein Produkt von Transpositionen.

d. Schreibe π^{-1} als ein Produkt von Transpositionen aufeinander folgender Zahlen.

e. Berechne für jede der Permutationen in a. das Signum.

Aufgabe A2.25.
Finde zwei Untergruppen von $\mathbb{S}_4$, die beide die Mächtigkeit 4 besitzen, aber nicht isomorph zueinander sind. Begründe, weshalb es Untergruppen sind und weshalb sie nicht isomorph zueinander sind.

Aufgabe A2.26.
Bestimme die Elemente der Untergruppe $\mathbb{D}_8 = \langle (1\ 2\ 3\ 4), (2\ 4) \rangle \leq \mathbb{S}_4$ von $\mathbb{S}_4$.

Aufgabe A2.27.
Bestimme die Elemente der Untergruppe $\mathbb{D}_{10} = \langle (1\ 2\ 3\ 4\ 5), (1\ 5) \circ (2\ 4) \rangle \leq \mathbb{S}_5$ von $\mathbb{S}_5$.

Aufgabe A2.28.
Zeige, daß ein k-Zyklus σ die Ordnung[4] k und das Signum $(-1)^{k-1}$ hat.

Aufgabe A2.29.
Zeige oder widerlege die Aussage, daß es in einer endlichen Gruppe G für je zwei Elemente $g, h \in G$ natürliche Zahle $a, b \in \mathbb{N}$ gibt, so daß $g \cdot h = h^a \cdot g^b$ gilt.

[4] Für den Begriff der Ordnung eines Elementes siehe Aufgabe 24.28

§ A3 Normalteiler und Faktorgruppen

Man kann auch für eine nicht-abelsche Gruppe G den Begriff der Faktorgruppe G/U bezüglich der Untergruppe U bilden. Dann muss die Untergruppe allerdings eine zusätzliche Eigenschaft erfüllen, sie muss ein Normalteiler sein. Im folgenden Abschnitt führen wir den Begriff ein.

A) Normalteiler

Im Kapitel über Äquivalenzrelationen haben wir gesehen, daß eine Äquivalenzrelation genau das richtige Mittel ist, um Ordung in eine Menge zu bringen, indem wir ihre Elemente nach vorgegebenen Gesichtspunkten zu größeren Einheiten zusammenfassen. Kurz gesagt, man verwendet Äquivalenzrelationen auf Mengen, um deren Elemente zu klassifizieren. Man erhält bei diesem Prozess wieder eine Menge, nämlich die Menge der Äquivalenzklassen. In dieser Vorlesung wollen wir uns aber stets mit Mengen beschäftigen, die eine zusätzliche Struktur tragen, und da ist es eine naheliegende Frage, ob sich die Struktur denn von der ursprünglichen Menge auf die neue Menge, die der Äquivalenzklassen, übertragen läßt. Ganz konkret, wenn G eine Gruppe ist und U eine Untergruppe von G, kann man dann auf *natürliche* Weise G/U zu einer Gruppe machen?

Dabei soll *natürlich* bedeuten, daß die Idee zur Definition der Gruppenoperation sofort ins Auge springt. Wir haben zwei Linksnebenklassen gU und hU, diese sind Teilmengen von G und das Produkt von Teilmengen von G wurde bereits in Notation 24.1 eingeführt. Natürlicher hätte man es sicher nicht definieren können. Allerdings soll dieses Produkt auch wieder eine Linksnebenklasse sein, das heißt wir müssen einen Repräsentanten davon angeben, und wenn unsere Definition des Produktes wirklich *natürlich* ist, dann sollte dieser Repräsentant sich aus den Repräsentanten der gegebenen Linksnebenklassen ergeben, d.h. wir wünschen uns, daß $gU \cdot hU = ghU$ gilt. Dies gilt leider nicht für alle Untergruppen und führt deshalb zu folgender Definition.

Definition A3.1.
Eine Untergruppe $U \leq G$ von G heißt *normal* oder *Normalteiler*, falls für alle $g \in G$ und $u \in U$ gilt

$$gug^{-1} \in U. \tag{120}$$

Wir schreiben in diesem Falle $U \trianglelefteq G$.

Bemerkung A3.2.
Um zu zeigen, daß eine *Teilmenge* $U \subseteq G$ ein Normalteiler ist, reicht es *nicht* die Eigenschaft (120) für alle $g \in G$ und $u \in U$ zu überprüfen. Zunächst muß gezeigt

werden, daß U eine *Untergruppe* von G ist! Daß dies ein wesentlicher Bestandteil der Definition des Begriffs Normalteiler ist, wird von Studienanfängern häufig übersehen.

Beispiel A3.3.
Ist G eine Gruppe, so sind die Untergruppen $\{e\}$ und G stets Normalteiler. Man nennt sie deshalb auch die *trivialen* Normalteiler.

Lemma A3.4.
Ist G eine abelsche Gruppe, so ist jede Untergruppe von G ein Normalteiler.

Beweis: Für $g \in G$ und $u \in U \leq G$ gilt $gug^{-1} = gg^{-1}u = eu = u \in U$. □

Aus diesem Lemma ergibt sich sofort folgendes Beispiel.

Beispiel A3.5.
Für jedes $n \in \mathbb{Z}$ ist die Untergruppe $n\mathbb{Z}$ von $(\mathbb{Z}, +)$ ein Normalteiler.

Proposition A3.6.
Es sei G eine Gruppe und $U \leq G$ eine Untergruppe. Die folgenden Aussagen sind äquivalent:[5]

a. $U \trianglelefteq G$ ist ein Normalteiler von G.

b. $gUg^{-1} = U$ für alle $g \in G$.

c. $gU = Ug$ für alle $g \in G$.

d. $(gU) \cdot (hU) = ghU$ für alle $g, h \in G$.

Beweis: Wir überlassen den Beweis dem Leser als Übungsaufgabe.

□

Beispiel A3.7.
Die Untergruppe $U := \{\mathrm{id}, (1\ 2)\} \subset \mathbb{S}_3$ ist kein Normalteiler der $\mathbb{S}_3$, denn für $\sigma = (2\ 3) \in \mathbb{S}_3$ gilt

$$\sigma \circ (1\ 2) \circ \sigma^{-1} = (2\ 3) \circ (1\ 2) \circ (2\ 3) = (1\ 3) \notin U.$$

Eine gute Quelle zum Auffinden von Normalteilern sind Gruppenhomomorphismen.

[5] Um die Äquivalenz von mehreren Aussagen zu zeigen, kann man einen sogenannten Ringschluß machen. Es reicht zu zeigen: "a. ⇒ b. ⇒ c. ⇒ d. ⇒ a.", denn aus "a. ⇒ b." und "b. ⇒ c." folgt z.B. "a. ⇒ c.", d.h. die scheinbar noch fehlenden Implikationen ergeben sich von selbst.

Proposition A3.8.
Ist $\alpha : G \to H$ ein Gruppenhomomorphismus, dann ist $\mathrm{Ker}(\alpha) \trianglelefteq G$.

Beweis: Wir wissen bereits aus Proposition 22.22, daß $\mathrm{Ker}(\alpha) \leq G$ eine Untergruppe von G ist. Sei nun $u \in \mathrm{Ker}(\alpha)$ und $g \in G$, dann gilt

$$\alpha\big(gug^{-1}\big) = \alpha(g) \cdot \alpha(u) \cdot \alpha\big(g^{-1}\big) \overset{u\in\mathrm{Ker}(\alpha)}{=}$$
$$\alpha(g) \cdot e_H \cdot \alpha\big(g^{-1}\big) = \alpha(g) \cdot \alpha\big(g^{-1}\big) = \alpha\big(gg^{-1}\big) = \alpha(e_G) = e_H.$$

Aber damit ist $gug^{-1} \in \mathrm{Ker}(\alpha)$ und $\mathrm{Ker}(\alpha) \trianglelefteq G$. □

Beispiel A3.9.
Betrachten wir den surjektiven Gruppenhomomorphismus (vgl. Bemerkung A2.20)

$$\mathrm{sgn} : \mathbb{S}_n \to \{-1, 1\},$$

dann gilt $\mathrm{Ker}(\mathrm{sgn}) = \mathbb{A}_n$ ist ein Normalteiler von $\mathbb{S}_n$.

Im allgemeinen ist das Produkt von zwei Untergruppen keine Untergruppe mehr. Betrachtet man etwa die Gruppe $G = \mathbb{S}_3$ sowie die Untergruppen $U = \{\mathrm{id}, (1\ 2)\}$ und $V = \{\mathrm{id}, (2\ 3)\}$, so ist $UV = \{\mathrm{id}, (1\ 2), (2\ 3), (1\ 2\ 3)\}$ und kann wegen des Satzes von Lagrange keine Untergruppe von G sein, da $|UV| = 4$ kein Teiler von $|G| = 6$ ist. Ist aber eine der beiden Untergruppen ein Normalteiler, sieht die Welt anders aus. Normalteiler sind also durchaus nützlich.

Lemma A3.10.
Es sei G eine Gruppe, $U \leq G$ und $N \trianglelefteq G$. Dann gilt:

a. $UN \leq G$.

b. $N \trianglelefteq UN$.

c. $U \cap N \trianglelefteq U$.

Beweis: Da N ein Normalteiler ist gilt nach Proposition A3.6

$$(UN) \cdot (UN) = U \cdot (NU) \cdot N = U \cdot (UN) \cdot N = (UU) \cdot (NN) = UN,$$

da $UU = U$ und $NN = N$. Damit gilt aber insbesondere, daß $g \cdot h \in UN$ für alle $g, h \in UN$.

Sei nun $g = un \in UN$ mit $u \in U$ und $n \in N$, dann ist

$$g^{-1} = n^{-1}u^{-1} \in NU \overset{\text{Prop. A3.6}}{=} UN.$$

Da ferner $e = e \cdot e \in UN$ und also UN nicht leer ist, ist UN eine Untergruppe nach dem Untergruppenkriterium.

Damit ist Teil a. gezeigt. Die beiden andere Aussagen überlassen wir dem Leser als Übungsaufgabe. □

Beispiel A3.11.
Wir betrachten die Untergruppen $U = \langle(1\ 2)\rangle$ und $V = \langle(2\ 3)\rangle$ von $\mathbb{S}_3$. Dann gilt

$$U \cdot V = \{\mathrm{id}, (1\ 2), (2\ 3), (1\ 2\ 3)\}$$

und wegen des Satzes von Lagrange kann dieses Produkt keine Untergruppe von $\mathbb{S}_3$ sein. Dies zeigt, daß die Bedinung $N \trianglelefteq G$ in Lemma A3.10 wesentlich ist.

B) Faktorgruppe

Satz 24.14 verallgemeinert sich dann unmittelbar mitsamt seinem Beweis auf den Fall einer nicht-abelschen Gruppe G mit einem Normalteiler U, da die Normalteilereigenschaft äquivalten zu der Gleichung $gU \cdot hU = ghU$ für alle $g, h \in G$ ist.

Satz A3.12.
Es sei $(G, \cdot)$ eine Gruppe und $U \trianglelefteq G$ ein Normalteiler von G. Dann gilt[6]

$$\overline{g} \cdot \overline{h} = \overline{g \cdot h}, \quad \text{für } \overline{g}, \overline{h} \in G/U. \tag{121}$$

Bezüglich der durch diese Multiplikation gegebenen zweistelligen Operation auf G/U ist G/U eine Gruppe. Das *neutrale Element* von $(G/U, \cdot)$ ist die Linksnebenklasse $\overline{e} = U$, und das zu $\overline{g} = gU \in G/U$ existierende *Inverse* Element ist die Linksnebenklasse $\overline{g^{-1}} = g^{-1}U$.
Außerdem ist die *Restklassenabbildung*

$$\pi : G \to G/U : g \mapsto \overline{g}$$

ein Epimorphismus mit $\mathrm{Ker}(\pi) = U$.
Man nennt G/U die *Faktorgruppe* von G nach U.

Beweis: Die Gleichheit in (121) folgt aus A3.6, da U ein Normalteiler ist:

$$\overline{g} \cdot \overline{h} = gU \cdot hU = ghU = \overline{gh}.$$

Zeigen wir nun, daß G/U mit dieser Operation eine Gruppe ist.

Für $\overline{g}, \overline{h}, \overline{k} \in G/U$ folgt mittels der Assoziativität der Multiplikation in G:

$$\big(\overline{g} \cdot \overline{h}\big) \cdot \overline{k} = \overline{gh} \cdot \overline{k} = \overline{(gh)k} = \overline{g(hk)} = \overline{g} \cdot \overline{hk} = \overline{g} \cdot \big(\overline{h} \cdot \overline{k}\big).$$

Außerdem ist $\overline{e} \cdot \overline{g} = \overline{eg} = \overline{g}$, so daß $\overline{e}$ das Neutrale von G/U ist, und es gilt

$$\overline{g^{-1}} \cdot \overline{g} = \overline{g^{-1} \cdot g} = \overline{e},$$

[6] Dabei ist mit $\overline{g} \cdot \overline{h} = gU \cdot hU$ einfach das Produkt von Teilmengen von G gemeint, wie es in Notation 24.1 eingeführt wurde.

und somit besitzt $\overline{g}$ ein Inverses, nämlich $\overline{g}^{-1} = \overline{g^{-1}}$. Also ist G/U eine Gruppe und die bezüglich des neutralen Elementes bzw. der Inversen getroffenen Aussagen sind ebenfalls gezeigt.

Zudem folgt aus der Definition von π

$$\pi(gh) = \overline{gh} \overset{(121)}{=} \overline{g} \cdot \overline{h} = \pi(g) \cdot \pi(h)$$

und

$$\mathrm{Ker}(\pi) = \{g \in G \mid \pi(g) = \overline{e}\} = \{g \in G \mid \overline{g} = \overline{e}\} = \overline{e} = U,$$

so daß π ein Gruppenhomomorphismus mit $\mathrm{Ker}(\pi) = U$ ist. □

Bemerkung A3.13.
Satz 24.14 zeigt, daß die Multiplikation der Linksnebenklassen auf der Menge G/U der Linksnebenklassen eine Gruppenoperation liefert, wenn U ein Normalteiler ist. Beachtet man zudem, daß das Produkt der Linksnebenklassen gU und hU auf alle Fälle das Element $gh = gehe$ enthält, so sieht man mit Proposition A3.6, daß es, wenn U kein Normalteiler ist, zwei Elemente $g, h \in G$ gibt, so daß das Produkt von gU mit hU keine Linksnebenklasse mehr ist. Das heißt, wenn U kein Normalteiler ist, kann G/U mit der obigen Multiplikation keine Gruppe sein. Es sind genau die Normalteiler, für die das Verfahren funktioniert.

Wenn für einen Normalteiler $N \trianglelefteq G$ die Menge der Linksnebenklassen G/N eine Gruppe ist, dann stellt sich die Frage, ob wir aus der Kenntnis der Untergruppen von G Rückschlüsse auf die Untergruppen von G/N ziehen können. In der Tat gibt es eine natürliche Eins-zu-Eins-Beziehung zwischen den Untergruppen einer Faktorgruppe G/N und den Untergruppen von G, die N enthalten. Entsprechendes gilt für die Normalteiler.

Proposition A3.14.
Sei $(G, \cdot)$ eine Gruppe und $N \trianglelefteq G$ ein Normalteiler, dann sind die folgenden Abbildungen *bijektiv*:

$$\left\{U \leq G \;\middle|\; N \subseteq U\right\} \longrightarrow \left\{\overline{U} \;\middle|\; \overline{U} \leq G/N\right\} : U \mapsto U/N$$

und

$$\left\{M \trianglelefteq G \;\middle|\; N \subseteq M\right\} \longrightarrow \left\{\overline{M} \;\middle|\; \overline{M} \trianglelefteq G/N\right\} : M \mapsto M/N.$$

Das heißt, die Untergruppen (bzw. Normalteiler) der Faktorgruppe G/N stehen in 1:1-Beziehung zu den Untergruppen (bzw. Normalteilern) von G, die N enthalten.

Beweis: Man beache zunächst, daß für eine Untergruppe $U \leq G$ von G, die N enthält, offenbar N ein Normalteiler von U ist. Somit ist die Faktorgruppe U/N definiert und man kann die Nebenklassen uN mit $u \in U$ als Elemente von G/N auffassen, d.h. $U/N \subseteq G/N$. Für $u, u' \in U$ gilt $uN \cdot u'N = uu'N \in U/N$ und $(uN)^{-1} = u^{-1}N \in U/N$,

so daß U/N in der Tat eine Untergruppe von G/N ist. Ist $U \trianglelefteq G$ sogar ein Normalteiler, so gilt zudem für $g \in G$ und $u \in U$ auch $gN \cdot uN \cdot (gN)^{-1} = (gng^{-1})N \in U/N$, so daß dann $U/N \trianglelefteq G/N$ sogar ein Normalteiler ist. Die beiden obigen Abbildungen sind also definiert.

Seien nun $U, V \leq G$ mit $N \subseteq U, V$ und $u \in U \setminus V$. Angenommen $uN \in V/N$, dann gäbe es ein $v \in V$ mit $uN = vN$ und insbesondere würde $u \in uN = vN \subseteq V$ gelten, im Widerspruch zur Annahme $u \notin V$. Dies zeigt, daß $U/N = V/N$ zwangsläufig $U = V$ bedingt. Die Abbildungen sind also injektiv.

Ist $\overline{U} \leq G/N$ eine Untergruppe, so setzen wir

$$U = \big\{u \in G \;\big|\; uN \in \overline{U}\big\}.$$

Sind $u, u' \in U$, so gilt $eN = N \in \overline{U}$, $uu'N = uN \cdot u'N \in \overline{U}$ und $u^{-1}N = (uN)^{-1} \in \overline{U}$, da $\overline{U}$ eine Untergruppe von G/N ist. Somit ist $e, uu', u^{-1} \in U$, und U ist eine Untergruppe von G mit $N \subseteq U$ für die nach Konstruktion $\overline{U} = U/N$ gilt. Ist $\overline{U}$ sogar ein Normalteiler von G/N, so gilt für $u \in U$ und $g \in G$ zudem $gug^{-1}N = gN \cdot uN \cdot (gN)^{-1} \in \overline{U}$. Damit ist dann $gug^{-1} \in U$ und U ist ein Normalteiler. Die beiden Abbildungen sind also auch surjektiv. □

Aus Proposition 22.16 wissen wir, daß die Untergruppen von $(\mathbb{Z}, +)$ die Form $m\mathbb{Z}$ für nicht-negative Zahlen m haben und aus Beispiel 22.9 wissen wir, daß $m\mathbb{Z}$ genau dann in $n\mathbb{Z}$ enthalten ist, wenn n ein Teiler von m ist. Wir erhalten unmittelbar folgendes Korollar.

Korollar A3.15.
Ist $n \in \mathbb{Z}_{>0}$ eine positive ganze Zahl, so gilt

$$\overline{U} \leq \mathbb{Z}_n \iff \exists\, m \in \{1, \ldots, n\} \text{ mit } m \text{ teilt } n \;:\; \overline{U} = m\mathbb{Z}/n\mathbb{Z} = \langle \overline{m}_n \rangle.$$

Insbesondere ist jede Untergruppe von $\mathbb{Z}_n$ zyklisch.

Wir wollen als nächstes die Ordnung eines Elementes $\overline{m} \in \mathbb{Z}_n$ für positive ganze Zahlen m und n bestimmen. Dazu führen wir zunächst folgende Notation ein.

Notation A3.16.
Für zwei ganze Zahlen $a, b \in \mathbb{Z}$ sei

$$\operatorname{kgv}(a,b) := \begin{cases} \min\{z > 0 \mid a \text{ und } b \text{ teilen } z\}, & \text{falls } a, b \neq 0, \\ 0, & \text{falls } a = 0 \text{ oder } b = 0. \end{cases}$$

Wir werden später (vgl. Aufgabe A6.11) sehen, daß $\operatorname{kgv}(a, b)$ ein *kleinstes gemeinsames Vielfaches* von a und b im Sinne der Definition A6.4 ist.

Korollar A3.17.
Es seien $m, n \in \mathbb{Z}_{>0}$. Dann ist

$$o(\overline{m}) = \frac{\operatorname{kgv}(m,n)}{m}$$

die Ordnung von $\overline{m} \in \mathbb{Z}_n$.

Beweis: Da $(\mathbb{Z}_n, +)$ eine endliche additive Gruppe ist, nimmt der Ausdruck zur Bestimmung der Ordnung von $\overline{m}$ in Bemerkung 24.12 folgende Form an:

$$\begin{aligned} o(\overline{m}) &= \min\left\{k > 0 \;\middle|\; k \cdot \overline{m} = \overline{0}\right\} \\ &= \min\left\{k > 0 \;\middle|\; n \text{ teilt } k \cdot m\right\} \\ &= \frac{m \cdot \min\{k > 0 \mid n \text{ teilt } k \cdot m\}}{m} \\ &= \frac{\min\left\{m \cdot k \;\middle|\; k > 0, n \text{ teilt } k \cdot m\right\}}{m} \\ &= \frac{\min\left\{l > 0 \;\middle|\; n \text{ und } m \text{ teilen } l\right\}}{m} \\ &= \frac{\operatorname{kgv}(m,n)}{m}. \end{aligned}$$

□

C) Der Homomorphiesatz

Gibt es einen Gruppenisomorphismus von einer Gruppe G in eine Gruppe H, so sind diese vom Standpunkt der Gruppentheorie nicht mehr wirklich unterscheidbar. Denn jede gruppentheoretische Eigenschaft der einen Gruppe findet sich automatisch auch in der anderen, da der Gruppenisomorphismus die Struktur erhält und bijektiv ist. Will man also eine der Gruppen studieren, so kann man genausogut auch die andere dazu betrachten, je nachdem ob einem die eine Repräsentation besser gefällt oder die andere. In diesem Sinne ist es interessant, Prinzipien kennenzulernen, die es erlauben festzustellen, wann zwei Gruppen isomorph sind und ggf. auch den Isomorphismus angeben zu können. In diesem Kapitel wollen wir die grundlegendste Methode dazu kennenlernen.

Satz A3.18 (Homomorphiesatz).
Ist $\alpha : G \to H$ ein Gruppenhomomorphismus, dann ist die durch α induzierte Abbildung

$$\widetilde{\alpha} : G/\operatorname{Ker}(\alpha) \to \operatorname{Im}(\alpha) : \overline{g} \mapsto \alpha(g)$$

wohldefiniert[7] und ein Isomorphismus. Insbesondere gilt also

$$G/\operatorname{Ker}(\alpha) \cong \operatorname{Im}(\alpha).$$

Beweis: Wir zeigen zunächst, daß $\widetilde{\alpha}$ wohldefiniert ist. Sei dazu $\overline{g} = \overline{h} \in G/\operatorname{Ker}(\alpha)$ gegeben. Dann gilt also $g^{-1}h \in \operatorname{Ker}(\alpha)$ und damit

$$e_H = \alpha\big(g^{-1}h\big) = \alpha\big(g^{-1}\big)\alpha(h) = \big(\alpha(g)\big)^{-1}\alpha(h).$$

Mithin gilt $\alpha(g) = \alpha(h)$, und $\widetilde{\alpha}$ ist somit wohldefiniert.

Für $\overline{g}, \overline{h} \in G/\operatorname{Ker}(\alpha)$ gilt ferner

$$\widetilde{\alpha}\big(\overline{g} \cdot \overline{h}\big) = \widetilde{\alpha}\big(\overline{gh}\big) = \alpha(gh) = \alpha(g)\alpha(h) = \widetilde{\alpha}\big(\overline{g}\big) \cdot \widetilde{\alpha}\big(\overline{h}\big).$$

Also ist $\widetilde{\alpha}$ ein Gruppenhomomorphismus.

$\widetilde{\alpha}$ ist offensichtlich surjektiv. Bleibt also noch zu zeigen, daß $\widetilde{\alpha}$ injektiv ist. Seien dazu $\overline{g}, \overline{h} \in G/\operatorname{Ker}(\alpha)$ mit $\alpha(g) = \widetilde{\alpha}\big(\overline{g}\big) = \widetilde{\alpha}\big(\overline{h}\big) = \alpha(h)$, so gilt

$$e_H = \big(\alpha(g)\big)^{-1}\alpha(h) = \alpha\big(g^{-1}\big)\alpha(h) = \alpha\big(g^{-1}h\big).$$

D. h. $g^{-1}h \in \operatorname{Ker}(\alpha)$, also $\overline{g} = \overline{h}$. Mithin ist $\widetilde{\alpha}$ injektiv. □

Beispiel A3.19.
Betrachte die Gruppen $(\mathbb{Z}, +)$ der ganzen Zahlen mit der Addition und $(\mathbb{C} \setminus \{0\}, \cdot)$ der komplexen Zahlen mit der Multiplikation. Aus der Vorlesung Grundlagen der Mathematik ist bekannt, daß

$$\alpha : \mathbb{Z} \longrightarrow \mathbb{C} \setminus \{0\} : z \mapsto e^{\frac{i\cdot\pi}{2}\cdot z}$$

ein Gruppenhomomorphismus ist, da das Potenzgesetz

$$e^{\frac{i\cdot\pi}{2}\cdot(z+z')} = e^{\frac{i\cdot\pi}{2}\cdot z} \cdot e^{\frac{i\cdot\pi}{2}\cdot z'}$$

[7] Der Begriff *wohldefiniert* meint im Prinzip nur, daß die getroffene Definition überhaupt eine solche ist. Doch wo ist das Problem dabei? Die Elemente von $G/\operatorname{Ker}(\alpha)$ sind nach Definition Linksnebenklassen (auch wenn wir uns bemühen wollen, das zu vergessen), und als solche besitzen sie Repräsentanten, die wir wie immer zum Rechnen und so auch zum Definieren von Abbildungen wie oben verwenden wollen. Nur besitzt für gewöhnlich jede Nebenklasse sehr viele Repräsentanten, und in einer Definition wie oben ist es a priori überhaupt nicht klar, daß die oben getroffene Zuordnung $\overline{g} \mapsto \alpha(g)$ nicht von der Wahl des Repräsentanten abhängt. D.h. wenn h ein anderer Repräsentant der gleichen Linksnebenklasse ist, also $\overline{g} = \overline{h}$, ist dann auch $\alpha(g) = \alpha(h)$? Alles andere wäre nicht gut, denn wir haben ja in unserer *Definition* nicht gesagt, welcher Repräsentant von der Linksnebenklasse verwendet werden soll! Für die *Wohldefiniertheit* der Abbildung müssen wir genau diesen Sachverhalt prüfen: $\overline{g} = \overline{h} \implies \alpha(g) = \alpha(h)$, oder bereits mit obiger Notation ausgedrückt

$$\overline{g} = \overline{h} \implies \widetilde{\alpha}(\overline{g}) = \widetilde{\alpha}(\overline{h}).$$

Dies erinnert verdächtig an die Injektivität einer Abbildung, aber Vorsicht, für diese war genau die andere Implikation zu zeigen.

gilt. Eine einfache Rechnung zeigt, daß

$$\mathrm{Im}(\alpha) = \{1, -1, i, -i\}$$

und

$$\mathrm{Ker}(\alpha) = 4 \cdot \mathbb{Z},$$

da $e^{\frac{i\cdot\pi}{2}\cdot z} = 1$ genau dann gilt, wenn $\frac{z}{2}$ ein Vielfaches von 2 ist. Aus dem Homomorphiesatz folgt mithin

$$\mathbb{Z}_4 = \mathbb{Z}/4\mathbb{Z} = \mathbb{Z}/\mathrm{Ker}(\alpha) \cong \mathrm{Im}(\alpha) = \{1, -1, i, -i\},$$

wobei die Gruppenoperation auf der linken Seite die Addition und auf der rechten Seite die Multiplikation ist.

Um die *Wohldefiniertheit* der Abbildung $\widetilde{\alpha}$ im Homomorphiesatz an einem Beispiel zu verstehen, sollte man folgendes beachten. In $\mathbb{Z}/4\mathbb{Z}$ sind die Nebenklassen $\overline{2}$ und $\overline{6}$ identisch. Mithin muß für die Abbildung

$$\widetilde{\alpha} : \mathbb{Z}/4\mathbb{Z} \longrightarrow \{1, -1, i, -i\} : \overline{z} \mapsto e^{\frac{i\cdot\pi}{2}\cdot z}$$

auch $\widetilde{\alpha}(\overline{2}) = \widetilde{\alpha}(\overline{6})$ gelten, und aufgrund der Definition von $\widetilde{\alpha}$ bedeutet das, daß notwendig $\alpha(2) = \alpha(6)$ gelten muß. Das tut's, da sich 2 und 6 um 4 unterscheiden und $e^{\frac{i\cdot\pi}{2}\cdot 4} = 1$.

Bemerkung A3.20.
Betrachten wir für $n \geq 2$ wieder den surjektiven Gruppenhomomorphismus (vgl. Bemerkung A2.20)

$$\mathrm{sgn} : \mathbb{S}_n \longrightarrow \{-1, 1\}.$$

Aus dem Homomorphiesatz A3.18 folgt insbesondere $|\mathbb{S}_n/\mathbb{A}_n| = |\{-1, 1\}| = 2$. Da nach dem Satz von Lagrange 24.10 zudem $|\mathbb{S}_n/\mathbb{A}_n| = \frac{|\mathbb{S}_n|}{|\mathbb{A}_n|}$ gilt, erhalten wir mit Proposition A2.11 die folgende Gleichung:

$$|\mathbb{A}_n| = \frac{n!}{2}.$$

Die beiden folgenden Isomorphiesätze sind leichte Anwendungen des obigen Homomorphiesatzes.

Satz A3.21 (1. Isomorphiesatz).
Ist G eine Gruppe, $U \leq G$ und $N \trianglelefteq G$. Dann ist

$$U/U \cap N \cong UN/N.$$

Beweis: Wir überlassen den Beweis dem Leser als Übungsaufgabe.

□

Satz A3.22 (2. Isomorphiesatz).
Es seien G eine Gruppe, $M \subseteq N \subseteq G$ zwei Normalteiler von G. Dann ist auch N/M ein Normalteiler von G/M und es gilt

$$(G/M)/(N/M) \cong G/N.$$

Beweis: Wir betrachten nun folgende Abbildung

$$\beta : G/M \to G/N : gM \mapsto gN,$$

und zeigen, sie ist ein Epimorphismus mit Kern N/M. Damit haben wir dann insbesondere gezeigt, daß N/M ein Normalteiler von G/M ist.[8]

Da β mittels des Repräsentanten einer Linksnebenklasse definiert ist, müssen wir zunächst wieder die Wohldefiniertheit zeigen.

***Schritt 0*:** β ist wohldefiniert.

Seien also $g, h \in G$ zwei Repräsentanten der gleichen Linksnebenklasse von M in G, d.h. $gM = hM$. Dann gilt nach Definition

$$g^{-1}h \in M \subseteq N,$$

so daß auch $gN = hN$. Die Abbildung ist also wohldefiniert.

***Schritt 1*:** β ist ein Homomorphismus.

Seien $gM, g'M \in G/M$, dann gilt

$$\beta\big(gM \cdot g'M\big) = \beta\big(gg'M\big) = gg'N = gN \cdot g'N = \beta(gM) \cdot \beta\big(g'M\big).$$

***Schritt 2*:** β ist surjektiv.

Sei $gN \in G/N$, dann ist $gN = \beta(gM) \in \mathrm{Im}(\beta)$, so daß β surjektiv ist.

***Schritt 3*:** $\mathrm{Ker}(\beta) = N/M$.

$$gM \in \mathrm{Ker}(\beta) \;\Leftrightarrow\; gN = N \;\Leftrightarrow\; g \in N \;\Leftrightarrow\; gM \in N/M.$$

Die Behauptung folgt also wieder mittels des Homomorphiesatzes:

$$(G/M)/(N/M) = (G/M)/\,\mathrm{Ker}(\beta) \cong \mathrm{Im}(\beta) = G/N.$$

□

D) Zyklische Gruppen

Wir wollen das Kapitel mit der Klassifikation zyklischer Gruppen abschließen.

[8] Beachte, daß M offenbar ein Normalteiler von N ist und somit der Quotient N/M auch tatsächlich definiert und identisch mit der Menge der Linksnebenklassen von M in G der Form nM mit $n \in N$ ist.

Satz A3.23.
Es sei $G = \langle g \rangle$ eine zyklische Gruppe.

a. Ist $|G| = \infty$, so haben wir den Gruppenisomorphismus
$$\alpha : \mathbb{Z} \xrightarrow{\cong} G : z \mapsto g^z.$$

b. Ist $|G| = n < \infty$, so haben wir den Gruppenisomorphismus
$$\overline{\alpha} : \mathbb{Z}_n \xrightarrow{\cong} G : \overline{z} \mapsto g^z.$$

Beweis: Für die Abbildung
$$\alpha : \mathbb{Z} \xrightarrow{\cong} G : z \mapsto g^z$$
und zwei ganze Zahlen $x, y \in \mathbb{Z}$ gilt
$$\alpha(x+y) = g^{x+y} = g^x \cdot g^y = \alpha(x) \cdot \alpha(y).$$
α ist also ein Gruppenhomomorphismus, und es gilt
$$\mathrm{Im}(\alpha) = \{g^z \mid z \in \mathbb{Z}\} = \langle g \rangle = G,$$
d.h. α ist surjektiv.

Ist $|G| = o(g) = \infty$, so ist
$$\{0\} = \{z \in \mathbb{Z} \mid g^z = e\} = \mathrm{Ker}(\alpha)$$
nach Bemerkung 24.12, d.h. α ist in diesem Fall auch injektiv.

Ist $|G| = o(g) = n < \infty$, so ist nach Aufgabe 22.44
$$\mathrm{Ker}(\alpha) = \{z \in \mathbb{Z} \mid g^z = e\} = n\mathbb{Z}.$$
Aus dem Homomorphiesatz folgt mithin, daß die Abbildung $\overline{\alpha}$ ein Gruppenisomorphismus ist. □

Die Klassifikation zyklischer Gruppen kann nun genutzt werden, um aus der Ordnung eines Elementes g die Ordnung all seiner Potenzen abzuleiten.

Korollar A3.24.
Ist $(G, \cdot)$ eine Gruppe, $g \in G$ mit $o(g) < \infty$ und $0 \neq m \in \mathbb{Z}$. Dann gilt
$$o\big(g^m\big) = \frac{\mathrm{kgv}\,\big(m, o(g)\big)}{|m|}.$$

Beweis: Es sei $n = o(g)$, dann ist
$$\overline{\alpha} : \mathbb{Z}_n \longrightarrow \langle g \rangle : \overline{z} \mapsto g^z$$

nach Satz A3.23 ein Gruppenisomorphismus. Aus Aufgabe 24.28 folgt dann

$$o(g^m) = o\Big(\overline{\alpha}(\overline{m})\Big) = o(\overline{m}).$$

Ist $m > 0$, so folgt die Behauptung aus Korollar A3.17. Ist $m < 0$, so ist $-m > 0$ und es folgt analog

$$o(g^{-m}) = \frac{\operatorname{kgv}(-m,n)}{-m} = \frac{\operatorname{kgv}(m,n)}{|m|}.$$

Da zudem die Ordnung eines Elementes und seines Inversen übereinstimmen, folgt die Behauptung auch im Fall $m < 0$. □

Korollar A3.25.
Ist $G = \langle g \rangle$ eine zyklische Gruppe der Ordnung $|G| = n < \infty$, dann gilt

$$U \leq G \iff \exists\, m \in \{1,\dots,n\} \text{ mit } m \text{ teilt } n \;:\; U = \langle g^m \rangle.$$

Für eine solche Untergruppe gilt zudem

$$\big|\langle g^m \rangle\big| = \frac{n}{m}.$$

Insbesondere hat G für jeden Teiler von n genau eine Untergruppe dieser Ordnung.

Beweis: Nach Satz A3.23 ist die Abbildung

$$\overline{\alpha} : \mathbb{Z}_n \longrightarrow G : \overline{z} \mapsto g^z$$

ein Gruppenisomorphismus, so daß die erste Aussage aus Korollar A3.15 folgt. Die Aussage zur Ordnung erhalten wir aus Korollar A3.24, da $\operatorname{kgv}(m,n) = n$. Schließlich beachte man noch, daß mit m auch $\frac{n}{m}$ alle Teiler von n durchläuft. □

Da die Untergruppen von $\mathbb{Z}$ zyklisch sind nach Proposition 22.16 erhalten wir mit Satz A3.23 und Korollar A3.15 folgende Aussage.

Korollar A3.26.
Jede Untergruppe einer zyklischen Gruppe ist zyklisch.

Aufgaben

Aufgabe A3.27.
Beweise Proposition A3.6.

Aufgabe A3.28.
Beweise Teil b. und c. von A3.10.

Aufgabe A3.29.
Es sei G eine endliche Gruppe und $U \leq G$ eine Untergruppe vom Index $|G : U| = 2$. Zeige, U ist ein Normalteiler von G.

Aufgabe A3.30.
Es sei G eine Gruppe und $N \leq G$ die einzige Untergruppe von G mit Ordnung $|N| = n$. Zeige, dann ist $N \trianglelefteq G$ ein Normalteiler.

Aufgabe A3.31.
Bestimme alle Normalteiler der Gruppe $\mathbb{D}_8 = \langle (1\ 2\ 3\ 4), (2\ 4) \rangle$.

Aufgabe A3.32.
Bestimme alle Normalteiler der Gruppe $\mathbb{D}_{10} = \langle (1\ 2\ 3\ 4\ ), (1\ 5) \circ (2\ 4) \rangle$.

Aufgabe A3.33.
Beweise Korollar A3.15.

Aufgabe A3.34.
Bestimme alle Untergruppen von $(\mathbb{Z}_{33}, +)$.

Aufgabe A3.35.
Betrachte für $m, n, a, b \in \mathbb{Z}_{>0}$ das Element $\left(\overline{a}_m, \overline{b}_n\right) \in \mathbb{Z}_m \times \mathbb{Z}_n$ in der Gruppe $(\mathbb{Z}_m \times \mathbb{Z}_n, +)$. Die Ordnung dieses Elementes läßt sich wie folgt berechnen

$$o\Big(\left(\overline{a}_m, \overline{b}_n\right)\Big) = \operatorname{kgv}\left(o\big(\overline{a}_m\big), o\big(\overline{b}_n\big)\right) = \operatorname{kgv}\left(\frac{\operatorname{kgv}(a,m)}{a}, \frac{\operatorname{kgv}(b,n)}{b}\right)$$

und ist ein Teiler von $\operatorname{kgv}(m, n)$. Insbesondere ist $\mathbb{Z}_m \times \mathbb{Z}_n$ nicht zyklisch, wenn m und n nicht teilerfremd sind.

Aufgabe A3.36.
Berechne die Ordnung von $\left(\overline{6}_{21}, \overline{9}_{33}\right) \in \mathbb{Z}_{21} \times \mathbb{Z}_{33}$.

Aufgabe A3.37.
Es sei σ und π zwei disjunkte Zyklen in $\mathbb{S}_n$ der Länge k bzw. l. Zeige, daß $o(\sigma \circ \pi) = \operatorname{kgv}(k, l)$.

Aufgabe A3.38.
Beweise Satz A3.21

Mit Hilfe des Homomorphiesatzes und des Satze von Lagrange kann man folgende Aufgabe lösen.

Aufgabe A3.39.
Es seien $(G,\cdot)$ und $(H,*)$ zwei endliche Gruppen teilerfremder Ordnung. Zeige, es gibt genau einen Gruppenhomomorphismus $\alpha : G \longrightarrow H$.

Aufgabe A3.40.
Es sei $(G,\cdot)$ eine Gruppe. Zeige:

a. Sind $g,h \in G$ mit $o(g) = o(h) = p$, wobei p eine Primzahl ist, dann gilt $\langle g\rangle = \langle h\rangle$ oder $\langle g\rangle \cap \langle h\rangle = \{e\}$.

b. Falls $|G| = 10$, so gibt es zwei Elemente $g,h \in G$ mit:
- $o(g) = 2$,
- $o(h) = 5$,
- $\langle h\rangle \trianglelefteq G$,
- $\langle g\rangle \cdot \langle h\rangle = G$.

Hinweis, führe in Teil b. zunächst die folgenden beiden folgenden Möglichkeiten zum Widerspruch: 1. $o(k) = 2$ für alle $e \neq k \in G$, 2. $o(k) = 5$ für alle $e \neq k \in G$.

Eine Gruppe G wie in Aufgabe A3.40 b. nennt man das semidirekte Produkt von $\langle g\rangle$ und $\langle h\rangle$. Man kann zeigen, falls $g \cdot h = h \cdot g$, so ist G isomorph zu $\mathbb{Z}_{10}$, andernfalls ist G isomorph zu $\mathbb{D}_{10}$.

Aufgabe A3.41.
Bestimme alle Gruppenhomomorphismen $\alpha : \mathbb{Z}_{10} \longrightarrow \mathbb{Z}_n$ mit $n \in \{6, 13\}$.

Aufgabe A3.42.
Ist $(G,\cdot)$ eine Gruppe und $p = |G|$ eine Primzahl, so ist G isomorph zu $(\mathbb{Z}_p, +)$.

Aufgabe A3.43.
Es sei $(G,\cdot)$ eine Gruppe, $g \in G$ und $n \in \mathbb{Z}_{>0}$. Zeige, genau dann gibt es einen Gruppenhomomorphismus $\alpha : \mathbb{Z}_n \longrightarrow G$ mit $\alpha\big(\overline{1}\big) = g$, wenn die Ordnung von g ein Teiler von n ist.

Aufgabe A3.44.
Bestimme alle Automorphismen der Gruppe $(\mathbb{Z}_{10}, +)$.

§ A4 Prüfzifferkodierung

In der Einleitung haben wir uns mit **EAN-13** Strichcodes beschäftigt, und dabei festgehalten, daß die ersten 12 Ziffern eines solchen Codes das Produkt, das ihn trägt, identifizieren. Welche Ziffern dabei welche Bedeutung haben, ist für unsere Belange nicht wichtig. Denn uns geht es letztlich nur darum, daß die 13. Ziffer des Codes nicht der Identifikation des Produktes, sondern der Absicherung des Codes gegen falsches Einscannen oder Abschreiben dient, und ob dieses sinnvoll gemacht wurde. Man nennt diese zusätzliche Ziffer auch *Prüfziffer*. Es ist ganz offensichtlich, daß eine einzige

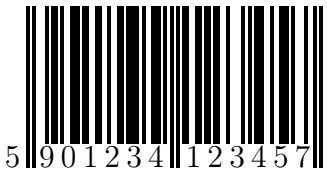


Abbildung 2. Ein EAN-13 Strichcode

zusätzliche Ziffer nicht ausreichen kann, um alle möglichen Fehler zu erkennen. Wenn man also ein *sinnvolles* Verfahren sucht, dann ist zunächst eine Analyse der häufigsten Fehler nötig, die beim Einscannen oder Abschreiben eines Codes auftreten. Das haben glücklicherweise bereits andere für uns erledigt, und eine solche Analyse zeigt, daß etwa 90% der Fehler in folgende beiden Kategorien fallen:

Typ I: "Einzelfehler" – d.h. nur eine einzelne Ziffer ist falsch erkannt. Das sind etwa 80% der auftretenden Fehler.

Typ II: "Nachbartransposition" – d.h. zwei benachbarte Ziffern sind vertauscht. Das sind etwa 10% der auftretenden Fehler.

Alle übrigen Fehlertypen, lagen bei weniger als 1%. Aufgrund dieser Analyse sollte eine gute Prüfziffer auf alle Fälle erkennen, wenn eine einzige Ziffer unter den ersten zwölfen falsch ist, und ihr sollte optimalerweise auch noch die Vertauschung zweier benachbarter Ziffern auffallen.

Die Ziffern, die in einem Strichcode verwendet werden, sind auf den ersten Blick Elemente der Menge

$$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\},$$

aber der geübte Teilnehmer der Vorlesung Algebraische Strukturen erkennt sofort, daß eine bloße *Menge* viel zu wenig Struktur hat, um damit etwas anfangen zu können, und daß man die 10 Ziffern doch viel besser als Elemente der Menge

$$\mathbb{Z}_{10} = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}, \overline{6}, \overline{7}, \overline{8}, \overline{9}\}$$

auffassen sollte. Auf diese Weise haben wir auf der Menge unserer Ziffern, die wir von jetzt an unser *Alphabet* nennen wollen, eine zusätzliche Struktur, nämlich eine Addition

bezüglich derer das Alphabet eine Gruppe ist. Diese Addition kann man ausnutzen, um die Prüfziffer mittels einer *möglichst guten* Formel aus den übrigen Ziffern zu errechnen.

Auf das konkrete Beispiel des **EAN-13** Strichcodes kommen wir später zurück. Zunächst wollen wir den Ansatz allgemeiner betrachten und für den allgemeineren Ansatz die Eigenschaften bezüglich der Fehlererkennung analysieren. Danach können wir dann schauen, ob die tatsächliche Kodierung des **EAN-13** den notwendigen Anforderungen genügt, und falls nicht, wie man ihn verbessern könnte. Die Grundidee des allgemeineren Ansatzes ist, statt $\mathbb{Z}_{10}$ eine beliebige Gruppe als Alphabet zuzulassen. Dann sollte man aber vielleicht nicht mehr von den Ziffern des Codes sprechen, sondern eher von den *Buchstaben* des Codes.

Als erste Idee für ein Verfahren zur Berechnung der Prüfziffer über $\mathbb{Z}_{10}$ könnte die *Quersumme* der ersten 12 Ziffern in $\mathbb{Z}_{10}$ dienen. Ein solches Verfahren würde aber ganz sicher keine Nachbartranspositionen erkennen, da die Gruppenoperation kommutativ ist. Diesem Mangel kann man begegnen, indem man die Elemente noch ein wenig abändert, d.h. *permutiert*, bevor man sie addiert. Dieser Gedanke führt zu folgender Definition.

Definition A4.1.
Es sei $(G,\cdot)$ eine Gruppe, $g_0 \in G$ fest gegeben, und $\pi_1,\dots,\pi_n \in \mathrm{Sym}(G)$ seien Permutationen von G.

a. Wir nennen

$$C = C_G(\pi_1,\dots,\pi_n,g_0) = \left\{(g_1,\dots,g_n) \in G^n \;\middle|\; \pi_1(g_1)\cdots\pi_n(g_n) = g_0\right\}$$

einen *Prüfziffercode* der *Länge* n auf dem *Alphabet* G.

b. Wir sagen, daß $C = C_G(\pi_1,\dots,\pi_n,g_0)$ *Fehler vom Typ I erkennt*, falls für alle $(g_1,\dots,g_n) \in C$ und $g_i' \in G$ mit $g_i' \neq g_i$ gilt $(g_1,\dots,g_{i-1},g_i',g_{i+1},\dots,g_n) \notin C$.

c. Wir sagen, daß $C = C_G(\pi_1,\dots,\pi_n,g_0)$ *Fehler vom Typ II erkennt*, falls für alle $(g_1,\dots,g_n) \in C$ und $i \in \{1,\dots,n-1\}$ mit $g_i \neq g_{i+1}$ gilt $(g_1,\dots,g_{i-1},g_{i+1},g_i,g_{i+2},\dots,g_n) \notin C$.

Wir wollen nun zunächst zeigen, daß der **EAN-13** Strichcode in dieses Schema paßt.

Beispiel A4.2 (EAN-13).
Wir überlassen es dem Leser, zu zeigen, daß die Abbildung

$$\mu_3 : \mathbb{Z}_{10} \longrightarrow \mathbb{Z}_{10} : \overline{k} \mapsto 3\cdot\overline{k} = \overline{k}+\overline{k}+\overline{k}$$

eine Permutation von $\mathbb{Z}_{10}$ ist, d.h. daß sie bijektiv ist. Man kann diese Aussage leicht verifizieren, indem man die Bilder der 10 Elemente von $\mathbb{Z}_{10}$ ausrechnet. Sie folgt allerdings auch leicht aus der Tatsache, daß die Zahlen 10 und 3 teilerfremd sind, wie wir später sehen werden.

Wir betrachten nun die Situation $(G,\cdot) = (\mathbb{Z}_{10},+)$, $g_0 = \overline{0}$, $n = 13$, $\pi_i = \mathrm{id}_{\mathbb{Z}_{10}}$ falls i ungerade ist und $\pi_i = \mu_3$ falls i gerade ist. Dann ist

$$C = C_{\mathbb{Z}_{10}}\big(\pi_1,\pi_2,\ldots,\pi_{13},\overline{0}\big)$$

der Prüfziffercode, der zu **EAN-13** gehört.

Wie läßt sich damit die 13. Ziffer aus den ersten zwölfen bestimmen? Nach Voraussetzung gilt für ein zulässiges Codewort $z_1z_2\ldots z_{12}z_{13} \in C$:

$$\overline{z_1} + 3\cdot\overline{z_2} + \overline{z_3} + \ldots + 3\cdot\overline{z_{12}} + \overline{z_{13}} = \overline{0},$$

und mithin

$$\overline{z_{13}} = -\overline{z_1} - 3\cdot\overline{z_2} - \overline{z_3}\ldots - 3\cdot\overline{z_{12}}.$$

Will man also z_{13} ermitteln, so führt man die Rechnung auf der rechten Seite des Gleichheitszeichens in $\mathbb{Z}_{10}$ durch, wählt den eindeutig bestimmten Repräsentanten $z_{13} \in \{0,1,\ldots,9\}$ der Linksnebenklasse in $\mathbb{Z}_{10}$ die man erhält.

Man kann das Verfahren auch ohne Gruppen beschreiben als: "*man berechne die abwechselnd mit* 1 *und* 3 *gewichtete Quersumme der ersten zwölf Ziffern, nenne* x *die letzte Ziffer dieser Quersumme und wähle die letzte Ziffer von* $10-x$ *als* z_{13}.". Aber die Fehlererkennungseigenschaften lassen sich mit der Gruppennotation leichter analysieren.

Betrachten wir ganz konkret das Beispiel in Figur 2, d.h. die ersten zwölf Ziffern sind 590123412345, so daß die Prüfziffer $z_{13} = 7$ sich ergibt aus

$$-\overline{5} - 3\cdot\overline{9} - \overline{0} - 3\cdot\overline{1} - \overline{2} - 3\cdot\overline{3} - \overline{4} - 3\cdot\overline{1} - \overline{2} - 3\cdot\overline{3} - \overline{4} - 3\cdot\overline{5} = \overline{-83} = \overline{7},$$

bzw. alternativ, daß sich $(5,9,0,1,2,3,4,1,2,3,4,5,7) \in C$ ergibt aus:

$$\overline{5} + 3\cdot\overline{9} + \overline{0} + 3\cdot\overline{1} + \overline{2} + 3\cdot\overline{3} + \overline{4} + 3\cdot\overline{1} + \overline{2} + 3\cdot\overline{3} + \overline{4} + 3\cdot\overline{5} + \overline{7} = \overline{0}.$$

Man beachte im Übrigen, daß $C = \mathrm{Ker}(\alpha)$ der Kern des folgenden Gruppenhomomorphismus ist,

$$\alpha : \mathbb{Z}_{10}^{13} \longrightarrow \mathbb{Z}_{10} : (z_1,\ldots,z_{13}) \mapsto z_1 + 3z_2 + z_3 + \ldots + 3z_{12} + z_{13},$$

wobei $\mathbb{Z}_{10}^{13}$ eine Gruppe bezüglich komponentenweiser Addition ist. □

Nachdem wir nun Prüfzifferkodierungen über beliebigen Gruppen eingeführt haben, sollten wir deren Fehlererkennungseigenschaften untersuchen.

Proposition A4.3 (Fehlererkennungseigenschaft).
Es sei G eine Gruppe und $C = C_G(\pi_1,\ldots,\pi_n,g_0)$ ein Prüfziffercode über dem Alphabet G.

a. C erkennt Fehler vom Typ I.

b. Für $n \geq 3$ erkennt C Fehler vom Typ II genau dann, wenn

$$g \cdot \big(\pi_{i+1} \circ \pi_i^{-1}\big)(h) \neq h \cdot \big(\pi_{i+1} \circ \pi_i^{-1}\big)(g) \tag{122}$$

für alle $i = 1, \ldots, n-1$ und für alle $g, h \in G$ mit $g \neq h$.

Beweis: a. Sei $(g_1, \ldots, g_n) \in C$ und $g_i' \in G$ mit $g_i' \neq g_i$. Angenommen $(g_1, \ldots, g_i', \ldots, g_n) \in C$, dann gilt

$$\pi_1(g_1) \cdots \pi_n(g_n) = g_0 = \pi_1(g_1) \cdots \pi_i\big(g_i'\big) \cdots \pi_n(g_n).$$

Wenden wir die Kürzungsregel auf beiden Seiten der Gleichung mehrfach an, so erhalten wir

$$\pi_i(g_i) = \pi_i\big(g_i'\big).$$

Da nach Voraussetzung π_i eine Permutation, also insbesondere injektiv ist, folgt $g_i = g_i'$ im Widerspruch zur Voraussetzung. Mithin ist $(g_1, \ldots, g_i', \ldots, g_n) \notin C$, und C erkennt Fehler vom Typ I.

b. Gehen wir zunächst davon aus, daß die Bedingung (122) erfüllt ist, und schließen daraus, daß C Fehler vom Typ II erkennt. Sei dafür $(g_1, \ldots, g_n) \in C$ gegeben mit $g_i \neq g_{i+1}$ für ein $1 \leq i \leq n-1$. Wir setzen $g = \pi_i(g_i)$ und $h = \pi_i(g_{i+1})$. Da π_i injektiv ist, gilt $g \neq h$. Aus (122) folgt damit:

$$\pi_i(g_i) \cdot \pi_{i+1}(g_{i+1}) = g \cdot \big(\pi_{i+1} \circ \pi_i^{-1}\big)(h) \neq h \cdot \big(\pi_{i+1} \circ \pi_i^{-1}\big)(g) = \pi_i(g_{i+1}) \cdot \pi_{i+1}(g_i).$$

Multiplizieren wir beide Seiten mit jeweils den gleichen Elementen aus G von links bzw. von rechts, so bleibt die Gleichheit erhalten. Auf dem Wege gilt also:

$$\pi_1(g_1) \cdots \pi_i(g_i) \cdot \pi_{i+1}(g_{i+1}) \cdots \pi_n(g_n) \neq \pi_1(g_1) \cdots \pi_i(g_{i+1}) \cdot \pi_{i+1}(g_i) \cdots \pi_n(g_n).$$

Das bedeutet aber, daß C Fehler vom Typ II erkennt.

Gehen wir nun umgekehrt davon aus, daß C Fehler vom Typ II erkennt, und zeigen, daß dann die Bedingung (122) erfüllt ist. Dazu seien $g, h \in G$ mit $g \neq h$ gegeben. Wir setzen $g_i = \pi_i^{-1}(g)$ und $g_{i+1} = \pi_i^{-1}(h)$. Da π_i bijektiv ist, folgt aus $g \neq h$, daß auch $g_i \neq g_{i+1}$. Wir wählen nun $g_j \in G$, $j \neq i, i+1$ so, daß $(g_1, \ldots, g_n) \in C$ – beachte, daß wir hier $n \geq 3$ benötigen. Nach Voraussetzung gilt dann aber

$$(g_1, \ldots, g_{i+1}, g_i, \ldots, g_n) \notin C,$$

und mithin

$$\pi_1(g_1) \cdots \pi_n(g_n) = g_0 \neq \pi_1(g_1) \cdots \pi_i(g_{i+1}) \cdot \pi_{i+1}(g_i) \cdots \pi_n(g_n).$$

Wir können nun wieder die Kürzungsregel auf beiden Seiten der Gleichung mehrfach anwenden und erhalten:

$$g \cdot \big(\pi_{i+1} \circ \pi_i^{-1}\big)(h) = \pi_i(g_i) \cdot \pi_{i+1}(g_{i+1}) \neq \pi_i(g_{i+1}) \cdot \pi_{i+1}(g_i) = h \cdot \big(\pi_{i+1} \circ \pi_i^{-1}\big)(g).$$

Damit ist die Aussage bewiesen.

□

Beispiel A4.4.
Aus Proposition A4.3 folgt, daß **EAN-13** alle Fehler vom Typ I erkennt. Wie sieht es mit Fehlern vom Typ II aus? Wir erinnern uns, daß $\pi_1 = \mathrm{id}_{\mathbb{Z}_{10}}$ und $\pi_2 = \mu_3$, die Multiplikation mit 3, ist. Mithin gilt für $g = \overline{0} \neq \overline{5} = h$

$$g + \big(\pi_2 \circ \pi_1^{-1}\big)(h) = \overline{0} + 3 \cdot \overline{5} = \overline{5} = \overline{5} + 3 \cdot \overline{0} = h + \big(\pi_2 \circ \pi_1^{-1}\big)(g).$$

Damit folgt aber aus Proposition A4.3, daß **EAN-13** nicht alle Fehler vom Typ II erkennt.

Natürlich stellt sich die Frage, welche Nachbarvertauschungen von **EAN-13** erkannt werden, und welche nicht. Eine kurze Überlegung führt zu folgender Erkenntnis: Sind $z, z' \in \{0, 1, 2, \ldots, 9\}$ zwei Ziffern, so daß $|z - z'| = 5$ gilt, dann gilt

$$2 \cdot (z - z') \equiv 0 \pmod{10} \tag{123}$$

und mithin

$$\overline{z} + 3 \cdot \overline{z'} = \overline{z'} + 3 \cdot \overline{z}. \tag{124}$$

Gilt umgekehrt (124), so gilt auch (123) und damit $|z - z'| = 5$.

Folglich erkennt **EAN-13** genau dann die Vertauschung zweier verschiedener benachbarter Ziffern nicht, wenn diese sich um 5 unterscheiden. Beim Beispiel in Figur 2 würden also auch Fehler vom Typ II erkannt. □

Proposition A4.3 hilft uns bei der Entscheidung, ob ein Prüfziffercode *alle* Fehler vom Typ II erkennt, ist dies nicht der Fall, so sagt sie aber nichts darüber aus, ob *sehr viele* solcher Fehler nicht erkannt werden, oder sehr wenige. Im Fall von **EAN-13** haben wir durch eine Zusatzüberlegung gesehen, daß vergleichsweise wenig Fehler vom Typ II übersehen werden. Könnte man die Qualität von **EAN-13** durch die Wahl anderer Gewichte, d.h. anderer Permutationen, verbessern, so daß *alle* Fehler vom Typ II erkannt werden? Dazu wollen wir zunächst einmal festhalten, daß die Gruppe $(\mathbb{Z}_{10}, +)$ abelsch ist, und daß sich für eine *abelsche* Gruppe $(G, \cdot)$ die Gleichung (122) schreiben läßt als

$$g \cdot \big(\mathrm{inv} \circ \pi_{i+1} \circ \pi_i^{-1}\big)(g) \neq h \cdot \big(\mathrm{inv} \circ \pi_{i+1} \circ \pi_i^{-1}\big)(h), \tag{125}$$

wobei

$$\mathrm{inv} : G \to G : g \mapsto g^{-1}$$

die Inversionsabbildung ist. Da inv bijektiv ist, ist $\mathrm{inv} \circ \pi_{i+1} \circ \pi_i^{-1} \in \mathrm{Sym}(G)$ eine Permutation von G. Betrachtet man nun (125), so scheint wegen Proposition A4.3 die Erkennung von Fehlern des Typs II mit Abbildungen der Form

$$g \mapsto g \cdot \pi(g)$$

zusammenzuhängen, wobei $\pi \in \mathrm{Sym}(G)$. Wir wollen diesen deshalb einen Namen geben.

Definition A4.5.
Ist G eine Gruppe und $\pi \in \mathrm{Sym}(G)$ eine Permutation von G, so nennen wir π eine *vollständige Abbildung* wenn die Abbildung

$$\pi^* : G \to G : g \mapsto g \cdot \pi(g)$$

bijektiv ist.

Bislang konnten wir nur für einen gegebenen Prüfziffercode entscheiden, ob er Fehler vom Typ II erkennt oder nicht. Die folgende Proposition gibt uns nun ein Kriterium, anhand dessen wir für *abelsche* Gruppen entscheiden können, ob es mit ihnen als Alphabet überhaupt eine Prüfzifferkodierung geben *kann*, die alle Fehler vom Typ II erkennt. Der Beweis ist *konstruktiv* in dem Sinn, daß es reicht eine *vollständige Abbildung* auf der Gruppe zu kennen, um eine solche Prüfzifferkodierung hinschreiben zu können.

Korollar A4.6.
Es sei G eine endliche *abelsche* Gruppe und $n \geq 3$. Genau dann gibt es auf G einen Prüfziffercode der Länge n, der Fehler vom Typ II erkennt, wenn es auf G eine vollständige Abbildung gibt.

Beweis: Wir setzen zunächst voraus, daß es auf G eine vollständige Abbildung $\pi \in \mathrm{Sym}(G)$ gibt. Definieren wir $g_0 = e_G$ und $\pi_i = (\mathrm{inv} \circ \pi)^i$ für $i = 1, \ldots, n$, dann wollen wir zeigen, daß $C = C_G(\pi_1, \ldots, \pi_n, g_0)$ Fehler vom Typ II erkennt.

Dazu müssen wir nur überprüfen, ob die Gleichung (125) erfüllt ist. Seien $g, h \in G$ mit $g \neq h$ gegeben, dann gilt

$$\begin{aligned} g \cdot \big(\mathrm{inv} \circ \pi_{i+1} \circ \pi_i^{-1}\big)(g) &= g \cdot \big(\mathrm{inv} \circ (\mathrm{inv} \circ \pi)^{i+1-i}\big)(g) = g \cdot \pi(g) = \pi^*(g) \\ &\neq \pi^*(h) = h \cdot \pi(h) = h \cdot \big(\mathrm{inv} \circ (\mathrm{inv} \circ \pi)^{i+1-i}\big)(h) = h \cdot \big(\mathrm{inv} \circ \pi_{i+1} \circ \pi_i^{-1}\big)(h). \end{aligned}$$

Mithin ist die Gleichung (125) erfüllt und C erkennt Fehler vom Typ II.

Setzen wir nun umgekehrt voraus, daß es auf G einen Prüfziffercode $C_G(\pi_1, \ldots, \pi_n, g_0)$ gibt, der Fehler vom Typ II erkennt. Es reicht zu zeigen, daß $\pi = \mathrm{inv} \circ \pi_2 \circ \pi_1^{-1} \in \mathrm{Sym}(G)$ eine vollständige Abbildung ist. Seien dazu $g, h \in G$ mit $g \neq h$. Aufgrund von Gleichung (125) gilt

$$\pi^*(g) = g \cdot \pi(g) = g \cdot \big(\mathrm{inv} \circ \pi_2 \circ \pi_1^{-1}\big)(g) \neq h \cdot \big(\mathrm{inv} \circ \pi_2 \circ \pi_1^{-1}\big)(h) = h \cdot \pi(h) = \pi^*(h).$$

Mithin ist π^* injektiv und damit auch bijektiv, da G endlich ist. Also ist π eine vollständige Abbildung. □

Bemerkung A4.7.

a. Wenn $|G| = 2 \cdot m$ mit m ungerade, dann gibt es *keine* vollständige Abbildung auf G.[9]Insbesondere gibt es auf $\mathbb{Z}_{10}$ also keinen Prüfziffercode, der alle Fehler vom Typ II erkennt.

b. Wenn $|G|$ ungerade ist, dann ist die Identität id_G eine vollständige Abbildung.

c. **Problem:** Es gibt keinen Prüfziffercode auf $(\mathbb{Z}_{10}, +)$, der alle Fehler vom Typ II erkennt. Wie können wir diesem Problem begegnen, wenn wir es nicht einfach hinnehmen wollen?

   Lösung 1: Man verwende eine ungerade Anzahl an Ziffern, d.h. man ersetze $\mathbb{Z}_{10}$ durch $\mathbb{Z}_m$ für ein ungerade Zahl m.
   Diese Methode wendete z.B. der ISBN-Code bis 2007 an. Er arbeitete mit $(\mathbb{Z}_{11}, +)$ als Alphabet und die *Ziffer* $\overline{10} = 10 + 11\mathbb{Z}$ wurde in einem ISBN-Code mit X bezeichnet. Dabei tauchte das X in den tatsächlich auf Büchern verwendeten ISBN-Nummern nur als Prüfziffer auf. Der ISBN-Code war der Prüfziffercode

$$C_{\mathbb{Z}_{11}}\big(\pi_1, \ldots, \pi_{10}, \overline{0}\big),$$

   wobei

$$\pi_i : \mathbb{Z}_{11} \to \mathbb{Z}_{11} : \overline{a} \mapsto (11 - i) \cdot \overline{a}.$$

   Wir überlassen es dem Leser zu überprüfen, daß der Code tatsächlich Fehler vom Typ II erkennt. Es reicht zu zeigen, daß die Gleichung (125) erfüllt ist.

   Lösung 2: Man verwende eine *nicht-abelsche* Gruppe mit zehn Elementen. Für diese gibt die Nicht-Existenz einer vollständigen Abbildung keinerlei Hinweis auf die Fehlererkennungseigenschaft von Prüfziffercodes.

Beispiel A4.8 (Seriennummern deutscher Währung).
Die Prüfziffer der Seriennummern der DM-Scheine waren kodiert mittels

$$C_{\mathbb{D}_{10}}\big(\pi^1, \ldots, \pi^{10}, \mathrm{id}_{\mathbb{D}_{10}}, (1)\big),$$

wobei

$$\mathbb{D}_{10} = \big\langle (1\ 2\ 3\ 4\ 5), (1\ 5)(2\ 4) \big\rangle \leq \mathbb{S}_5 = \mathrm{Sym}\big(\{1, \ldots, 5\}\big)$$

die *Diëdergruppe* der Ordnung 10 ist und die Permutation π weiter unten definiert wird. Wir schreiben hier das neutrale Element der $\mathbb{D}_{10}$ als Einszyklus (1) und nicht als $\mathrm{id}_{\{1,\ldots,5\}}$, um Verwechslungen mit der Permutation $\mathrm{id}_{\mathbb{D}_{10}}$ vorzubeugen.

Wie für die Gruppe $\mathbb{D}_8$ zeigt man, daß mit $\sigma = (1\ 2\ 3\ 4\ 5)$ und $\tau = (1\ 5)(2\ 4)$ die Gruppe $\mathbb{D}_{10}$ geschrieben werden kann als

$$\mathbb{D}_{10} = \big\{\sigma^0 = (1), \sigma^1, \ldots, \sigma^4, \tau \circ \sigma^0 = \tau, \tau \circ \sigma^1, \ldots, \tau \circ \sigma^4\big\}.$$

Die Gruppe ist nicht abelsch, da $\tau \circ \sigma = \sigma^{-1} \circ \tau \neq \sigma \circ \tau$.

[9]Der Beweis ist elementar, aber etwas länglich. Wir verweisen den Leser deshalb auf [Sie81].

Verhoeff hat in [Ver75] gezeigt, daß die Permutation $\pi : \mathbb{D}_{10} \to \mathbb{D}_{10}$ mit

x	σ^0	σ^1	σ^2	σ^3	σ^4	$\tau\circ\sigma^0$	$\tau\circ\sigma^1$	$\tau\circ\sigma^2$	$\tau\circ\sigma^3$	$\tau\circ\sigma^4$
$\pi(x)$	σ^1	$\tau\circ\sigma^0$	$\tau\circ\sigma^2$	$\tau\circ\sigma^1$	σ^2	$\tau\circ\sigma^3$	σ^3	σ^0	$\tau\circ\sigma^4$	σ^4

folgende Eigenschaft hat:

$$g\circ\pi(h) \neq h\circ\pi(g) \quad \text{für alle } g,h\in\mathbb{D}_{10} \text{ mit } g\neq h.$$

Aus Proposition A4.3 folgt dann, daß $C_{\mathbb{D}_{10}}\big(\pi^1,\dots,\pi^{10},\mathrm{id}_{\mathbb{D}_{10}},(1)\big)$ Fehler vom Typ II erkennt.

Natürlich hat man bei den Seriennummern der DM-Scheine keine Symbole wie σ verwendet. Stattdessen wurden die 10 Ziffern sowie zusätzlich 10 Buchstaben verwendet, die dann vermittels folgender Tabelle mit den Elementen der $\mathbb{D}_{10}$ identifiziert wurden:

σ^0	σ^1	σ^2	σ^3	σ^4	$\tau\circ\sigma^0$	$\tau\circ\sigma^1$	$\tau\circ\sigma^2$	$\tau\circ\sigma^3$	$\tau\circ\sigma^4$
0	1	2	3	4	5	6	7	8	9
A	D	G	K	L	N	S	U	Y	Z.

Wollte man überprüfen, ob eine Seriennummer zulässig ist, mußte man die Ziffern und Buchstaben durch die entsprechenden Elemente der $\mathbb{D}_{10}$ ersetzen und nachrechnen, ob das Ergebnis in der Menge $C_{\mathbb{D}_{10}}\big(\pi^1,\dots,\pi^{10},\mathrm{id}_{\mathbb{D}_{10}},(1)\big)$ liegt.

Aufgabe A4.9.
Überprüfe ob $AA6186305Z2$ eine zulässige Seriennummer für einen DM-Schein ist.

Bemerkung A4.10.
Hätte man noch eine andere Gruppe mit 10 Elementen als Alphabet verwenden können?

Nicht wirklich! Für das Alphabet kommt es nur auf den Isomorphietyp der Gruppe an, und jede Gruppe der Ordnung 10 ist entweder isomorph zu $(\mathbb{Z}_{10},+)$ oder zu $(\mathbb{D}_{10},\circ)$. Der Beweis dieser Aussage ist mit den Mitteln, die wir bislang zur Verfügung haben, etwas länglich, aber durchaus machbar.

Aufgabe A4.11.
Beweise Teil b. von Bemerkung A4.7.

§ A5 Ringe und Körper

A) Begriffsbildung und erste Beispiele

In Kapitel 22 haben wir die mathematische Struktur der *Gruppe* eingeführt, und unsere ersten Beispiele waren die additiven Gruppen $(\mathbb{Z},+)$ und $(\mathbb{Q},+)$ der ganzen bzw. der rationalen Zahlen. Auf diesen Mengen haben wir aber neben der Addition jeweils noch eine zweite zweistellige Operation, die Multiplikation, bezüglich derer in beiden Mengen wiederum interessante Rechenregeln gelten. So ist die Menge $(\mathbb{Q}\setminus\{0\},\cdot)$ wieder eine Gruppe, während $(\mathbb{Z}\setminus\{0\},\cdot)$ zu dieser Eigenschaft (nur) die multiplikativen Inversen fehlen. Wir wollen diese Beispiele nun verallgemeinern und führen dazu folgende Definition ein.

Definition A5.1.

a. Ein *Ring mit Eins* ist ein Tripel $(R,+,\cdot)$ bestehend aus einer Menge R zusammen mit zwei zweistelligen Operationen

$$+ : R\times R\to R : (a,b)\mapsto a+b, \qquad (\text{“}\textit{Addition}\text{”})$$

und

$$\cdot : R\times R\to R : (a,b)\mapsto a\cdot b, \qquad (\text{“}\textit{Multiplikation}\text{”})$$

so, daß folgende Axiome erfüllt sind:

(i) $(R,+)$ ist eine abelsche Gruppe (mit neutralem Element 0_R).

(ii) $a\cdot(b\cdot c)=(a\cdot b)\cdot c$ für alle $a,b,c\in R$. (“*Assoziativität der Multiplikation*”)

(iii) Es gibt ein Element $1_R\in R$ mit $1_R\cdot a=a\cdot 1_R=a$ für alle $a\in R$. (“*Einselement*”)

(iv) $a\cdot(b+c)=a\cdot b+a\cdot c$ und $(b+c)\cdot a=b\cdot a+c\cdot a$ für alle $a,b,c\in R$. (“*Distributivität*”)

b. Ein Ring mit Eins $(R,+,\cdot)$ heißt *kommutativ*, falls $a\cdot b=b\cdot a$ für alle $a,b\in R$.

c. Ist $(R,+,\cdot)$ ein Ring mit Eins, dann heißt $a\in R$ eine *Einheit* oder *invertierbar* in R, falls es ein $a'\in R$ gibt mit $a\cdot a'=a'\cdot a=1_R$. Wir bezeichnen mit

$$R^*=\{a\in R\mid a \text{ ist Einheit}\}$$

die Menge der Einheiten von R.

d. Ein kommutativer Ring mit Eins $(R,+,\cdot)$ heißt *Körper*, falls $1_R\in R^*=R\setminus\{0\}$.

Bemerkung A5.2.
Wir werden in Ringen für die Addition stets das Zeichen + und für die Multiplikation das Zeichen · verwenden, auch wenn wir gleichzeitig verschiedene Ringe betrachten. Wir verzichten im Folgenden deshalb darauf, die Ringoperationen jeweils anzugeben und nennen verkürzend die dem Ring $(R, +, \cdot)$ zugrunde liegende Menge R einen Ring. Zudem werden wir statt $a \cdot b$ oft auch nur ab schreiben.

Das neutrale Element von $(R, +)$ werden wir mit 0_R oder einfach mit 0 bezeichnen und das *Nullelement* von R nennen; das Einselement bezeichnen wir mit 1_R oder 1.

Ist R ein Ring und sind $a, b \in R$, so schreiben wir statt $a + (-b)$ auch kurz $a - b$.

Das Einselement in R ist eindeutig bestimmt, denn wenn $1_R, 1'_R \in R$ zwei Elemente mit der Eigenschaft des Einselementes sind, dann folgt $1_R = 1_R \cdot 1'_R = 1'_R$.

Ist $a \in R$ eine Einheit und $a', a'' \in R$ mit $a \cdot a' = a' \cdot a = 1_R$ und $a \cdot a'' = a'' \cdot a = 1_R$, so gilt

$$a' = 1_R \cdot a' = (a'' \cdot a) \cdot a' = a'' \cdot (a \cdot a') = a'' \cdot 1_R = a''.$$

Das Inverse a' zu a ist also eindeutig bestimmt und wird mit a^{-1} oder $\frac{1}{a}$ bezeichnet.

Beachte, aus der Definition folgt unmittelbar, daß $(R^*, \cdot)$ eine Gruppe ist, die sogenannte *Einheitengruppe* des Ringes R.

Unter Beachtung der Rechenregeln A5.5 kann man leicht zeigen, daß eine Tripel $(K, +, \cdot)$ genau dann ein *Körper* ist, wenn gilt:

a. $(K, +)$ ist eine abelsche Gruppe.

b. $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe.

c. Für alle $a, b, c \in K$ gilt $a \cdot (b + c) = a \cdot b + a \cdot c$.

□

Beispiel A5.3.

a. $(\mathbb{Z}, +, \cdot)$ mit der üblichen Addition und Multiplikation ist ein kommutativer Ring mit Eins, der kein Körper ist.

b. $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ mit der üblichen Addition und Multiplikation sind Körper.

c. In der Vorlesung Grundlagen der Mathematik wurden auf der Menge $\mathbb{C} = \{(x, y) \mid x, y \in \mathbb{R}\}$ die beiden zweistelligen Operationen

$$+ : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C} : \big((x, y), (x', y')\big) \mapsto (x + x', y + y')$$

und

$$\cdot : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C} : \big((x, y), (x', y')\big) \mapsto (x \cdot x' - y \cdot y', x \cdot y' + x' \cdot y)$$

eingeführt, und es wurde gezeigt, daß $(\mathbb{C}, +, \cdot)$ ein Körper ist – *Körper der komplexen Zahlen* Für die Elemente von $\mathbb{C}$ hat sich die Schreibweise $(x, y) = x + iy$ mit $i^2 = -1$ eingebürgert. Wir werden im weiteren Verlauf der Vorlesung die komplexen Zahlen und ihre Eigenschaften so, wie sie in der Vorlesung Grundlagen der Mathematik gezeigt wurden, als bekannt voraussetzen. Der Vollständigkeit halber haben wir die wichtigsten Eigenschaften im Anhang zusammengetragen.

d. Ist M eine beliebige Menge und $(R, +, \cdot)$ ein Ring mit Eins, so ist

$$R^M := \{f \mid f : M \to R \text{ ist eine Abbildung}\}$$

mit den punktweise definierten Operationen

$$+ : R^M \times R^M \to R^M : (f, g) \mapsto \big(f + g : M \to R : x \mapsto f(x) + g(x)\big),$$

und

$$\cdot : R^M \times R^M \to R^M : (f, g) \mapsto \big(f \cdot g : M \to R : x \mapsto f(x) \cdot g(x)\big),$$

ein Ring mit der Nullfunktion $0 : M \to R : x \mapsto 0_R$ als neutralem Element der Addition und der Einsfunktion $1 : M \to R : x \mapsto 1_R$ als Einselement, wie man mit etwas Fleiß nachprüft.

e. In Aufgabe 22.32 haben wir die Menge

$$\mathrm{Mat}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\middle|\; a, b, c, d \in \mathbb{R} \right\}$$

der reellen 2×2-Matrizen eingeführt und für zwei Matrizen

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$$

ihr Produkt als

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a \cdot a' + b \cdot c' & a \cdot b' + b \cdot d' \\ c \cdot a' + d \cdot c' & c \cdot b' + d \cdot d' \end{pmatrix}$$

definiert. Setzen wir zudem

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a + a' & b + b' \\ c + c' & d + d' \end{pmatrix},$$

so rechnet man mit etwas Geduld nach, daß $(\mathrm{Mat}_2(\mathbb{R}), +, \cdot)$ ein Ring mit Einselement

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

ist. Dieser Ring ist nicht-kommutativ, da

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

In der Vorlesung Grundlagen der Mathematik wird dieses Beispiel verallgemeinert auf Matrizen der Größe $n \times n$ für $n \geq 1$ über einem beliebigen Körper K, und der Nachweis, daß auf diesem Wege ein nicht-kommutativer Ring mit Eins entsteht, wird dort auf weit geschickterem Weg als durch langatmiges Nachrechnen geführt.

Beispiel A5.4.
Es seien $(R, +, \cdot)$ und $(S, +, \cdot)$ zwei kommutative Ringe mit Eins. Dann wird das kartesische Produkt $R \times S$ durch die komponentenweisen Operationen

$$(r, s) + (r', s') := (r + r', s + s')$$

und

$$(r, s) \cdot (r', s') := (r \cdot r', s \cdot s')$$

zu einem kommuativen Ring mit Eins $(1_R, 1_S)$. Das Nachrechnen der Axiome ist eine einfache Anwendung der Definition.

Für die Einheiten in $R \times S$ gilt

$$(R \times S)^* = R^* \times S^*,$$

da

$$(1_R, 1_S) = (r, s) \cdot (r', s') = (r \cdot r', s \cdot s') \iff 1_R = r \cdot r' \text{ und } 1_S = s \cdot s'.$$

Auf die gleiche Weise wird das kartesische Produkt von einer beliebigen Anzahl kommutativer Ringe mit Eins wieder ein kommutativer Ring mit Eins, und die Einheiten des kartesischen Produktes sind wieder durch das kartesische Produkt der Einheitengruppen gegeben.

Wir wollen nun einige Rechenregeln für das Rechnen in Ringen aufstellen.

Lemma A5.5 (Rechenregeln).
Es sei R ein Ring mit Eins. Für $a, b, c \in R$ gelten:

a. $-(-a) = a$.

b. $a + b = c \Leftrightarrow a = c - b$.

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

d. $0 \cdot a = a \cdot 0 = 0$.

e. $(-a) \cdot b = a \cdot (-b) = -(a \cdot b)$.

f. $(-a) \cdot (-b) = a \cdot b$.

g. $a \cdot (b - c) = a \cdot b - a \cdot c$.

h. Für $a \in R^*$ ist $a^{-1} \in R^*$ und $\left(a^{-1}\right)^{-1} = a$.

i. Ist $1_R = 0_R$, so ist $R = \{0_R\}$ der *Nullring*.

Beweis: Die Aussagen a., b. und c. folgen unmittelbar aus Lemma 22.6.

d. Für $a \in R$ gilt $0 \cdot a = (0+0) \cdot a = 0 \cdot a + 0 \cdot a$, also folgt $0 \cdot a = 0$ mittels der Kürzungsregeln in $(R, +)$. Analog sieht man $a \cdot 0 = 0$.

e. Für $a, b \in R$ gilt wegen d.:

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

also $-(a \cdot b) = (-a) \cdot b$. Die Gleichheit des Ausdrucks zu $a \cdot (-b)$ folgt analog.

f. Für $a, b \in R$ folgt unter Zuhilfenahme von a. und e.:

$$(-a) \cdot (-b) = -\big(a \cdot (-b)\big) = -\big(-(a \cdot b)\big) = a \cdot b.$$

g. Für $a, b, c \in R$ impliziert e.:

$$a \cdot (b - c) = a \cdot b + a \cdot (-c) = a \cdot b + \big(-(a \cdot c)\big) = a \cdot b - a \cdot c.$$

h. Ist $a \in R^*$ eine Einheit mit Inversem a^{-1}. Dann ist nach Definition a ein Inverses von a^{-1}, und insbesondere ist a^{-1} eine Einheit. Aus der Eindeutigkeit des Inversen (siehe Bemerkung A5.2) folgt dann, daß $a = (a^{-1})^{-1}$.

i. Ist $a \in R$, so gilt $a = 1_R \cdot a = 0_R \cdot a = 0_R$.

□

Eine wichtige Klasse kommutativer Ringe mit Eins stellen die sogenannten formalen Potenzreihenringe dar, die wir in folgenden Definition einführen wollen.

Definition A5.6.
Sei R ein kommutativer Ring mit Eins und t eine Veränderliche. Einen formalen Ausdruck der Form

$$\sum_{k=0}^{\infty} a_k \cdot t^k$$

mit $a_k \in R$ nennen wir eine *formale Potenzreihe* mit Koeffizienten in R und die Menge

$$R[[t]] := \left\{ \sum_{k=0}^{\infty} b_k \cdot t^k \;\middle|\; b_k \in R \right\}$$

den *Ring der formalen Potenzreihen* über R in der Unbestimmten t.
Für zwei formale Potenzreihen $\sum_{i=0}^{\infty} a_i \cdot t^i, \sum_{j=0}^{\infty} b_j \cdot t^j \in R[[t]]$ definieren wir ferner

$$\sum_{i=0}^{\infty} a_i \cdot t^i \;+\; \sum_{i=0}^{\infty} b_i \cdot t^i \;:=\; \sum_{i=0}^{\infty} (a_i + b_i) \cdot t^i \in R[[t]]$$

und

$$\sum_{i=0}^{\infty} a_i \cdot t^i \;\cdot\; \sum_{j=0}^{\infty} b_j \cdot t^j \;:=\; \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i \cdot b_j \right) \cdot t^k \in R[[t]].$$

Man beachte, daß aus der Definition unmittelbar folgt, daß

$$\sum_{i=0}^{\infty} a_i \cdot t^i \;=\; \sum_{i=0}^{\infty} b_i \cdot t^i \quad\Longleftrightarrow\quad a_i = b_i \;\forall\, i \in \mathbb{N}.$$

Gilt $a_i = 0$ für $i \geq n$, so schreiben wir auch abkürzend

$$\sum_{i=0}^{n} a_i \cdot t^i = \sum_{i=0}^{\infty} a_i \cdot t^i.$$

Bemerkung A5.7.
Die Definition der Multiplikation in $R[[t]]$ entspringt dem Bedürfnis, eine Art verallgemeinertes Distributivgesetz für diese *unendlichen Summen* zu haben:

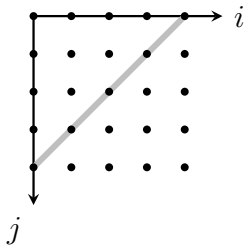
$$\text{(126)}\quad \left(\sum_{i=0}^{\infty} a_i \cdot t^i\right) \cdot \left(\sum_{j=0}^{\infty} b_j \cdot t^j\right) = \sum_{i=0}^{\infty} \left(a_i \cdot t^i \cdot \left(\sum_{j=0}^{\infty} b_j \cdot t^j\right)\right) = \sum_{i=0}^{\infty} \left(\left(\sum_{j=0}^{\infty} a_i \cdot t^i \cdot b_j \cdot t^j\right)\right) = \sum_{i=0}^{\infty} \left(\left(\sum_{j=0}^{\infty} a_i \cdot b_j \cdot t^{i+j}\right)\right).$$

Wenn die eben durchgeführten Operationen in $R[[t]]$ korrekt sind, d.h. wenn alle Gleichungen korrekt sind, dann ist unsere Notation der Elemente von $R[[t]]$ als unendliche Summen eine nützliche Notation. Der Ausdruck auf der rechten Seite in (126) ist in der angegebenen Form aber noch nicht als formale Potenzreihe, d.h. als Element von $R[[t]]$, erkennbar. Viele der t^{i+j} stimmen für verschiedene Werte von i und j überein; z.B. kann man t^2 durch $(i,j) = (2,0)$ und $(i,j) = (1,1)$ und $(i,j) = (0,2)$ erhalten. Der Koeffizient zu t^2 der Potenzreihe in (126) sollte also die Summe der $a_i \cdot b_j$ für diese (i,j) sein, d.h. $a_2 \cdot b_0 + a_1 \cdot b_1 + a_0 \cdot b_2$.

Gibt es eigentlich für jedes k nur endlich viele Paare (i,j), so daß $i + j = k$? Ja! Denn wir setzen für i und j ja voraus, daß es nicht-negative ganze Zahlen sind. Da mit der Vorgabe von k und durch die Wahl von i die Zahl j bereits als $j = k - i$ festgelegt ist und da i und j zwischen 0 und k liegen müssen, damit sowohl i als auch j nicht-negativ sind, gibt es genau die folgenden $k + 1$ Möglichkeiten:

$$(k,0),\quad (k-1,1),\quad (k-2,2),\quad \ldots\quad ,\quad (1,k-1),\quad (0,k).$$

Am deutlichsten kann man sich die Paare (i,j), deren Summe ein festes k ergibt, verdeutlichen, indem man sie als Punkte in ein Koordinatensystem mit den Achsen i und j einträgt:



Genau die Paare (i,j) auf der Diagonalen von $(k,0)$ nach $(0,k)$ haben die Eigenschaft, daß die Summe $i+j$ den Wert k ergibt. Der Koeffizient zu t^k in der rechten Seite von (126) muß deshalb

$$\sum_{i+j=k} a_i \cdot b_j = \sum_{i=0}^{k} a_i \cdot b_{k-i}$$

sein. Man spricht wegen der angegebenen graphischen Darstellung auch vom *Cantorschen Diagonalverfahren* der Multiplikation zweier Potenzreihen.

Satz A5.8.
Ist R ein kommutativer Ring mit Eins, so ist der formale Potenzreihenring $(R[[t]],+,\cdot)$ ein kommutativer Ring mit Eins $1_{R[[t]]} = t^0$.

Beweis: Nach Definition sind $+$ und $\cdot$ zwei zweistellige Operationen auf $R[[t]]$. Seien also $\sum_{i=0}^{\infty} a_i \cdot t^i, \sum_{i=0}^{\infty} b_i \cdot t^i, \sum_{i=0}^{\infty} c_i \cdot t^i \in R[[t]]$ gegeben. Dann gilt wegen der Assoziativität der Addition in R

$$\begin{aligned}\left(\sum_{i=0}^{\infty} a_i \cdot t^i + \sum_{i=0}^{\infty} b_i \cdot t^i\right) + \sum_{i=0}^{\infty} c_i \cdot t^i &= \sum_{i=0}^{\infty} \big((a_i + b_i) + c_i\big) \cdot t^i \\ &= \sum_{i=0}^{\infty} \big(a_i + (b_i + c_i)\big) \cdot t^i = \sum_{i=0}^{\infty} a_i \cdot t^i + \left(\sum_{i=0}^{\infty} b_i \cdot t^i + \sum_{i=0}^{\infty} c_i \cdot t^i\right)\end{aligned}$$

und wegen der Kommutativität der Addition in R

$$\begin{aligned}\sum_{i=0}^{\infty} a_i \cdot t^i + \sum_{i=0}^{\infty} b_i \cdot t^i = \sum_{i=0}^{\infty} (a_i + b_i) \cdot t^i \\ = \sum_{i=0}^{\infty} (b_i + a_i) \cdot t^i = \sum_{i=0}^{\infty} b_i \cdot t^i + \sum_{i=0}^{\infty} a_i \cdot t^i.\end{aligned}$$

Zudem gilt für die Nullfunktion $0_{R[[t]]} = \sum_{i=0}^{\infty} 0 \cdot t^i$

$$0_{R[[t]]} + \sum_{i=0}^{\infty} a_i \cdot t^i = \sum_{i=0}^{\infty} (0 + a_i) \cdot t^i = \sum_{i=0}^{\infty} a_i \cdot t^i,$$

und für $\sum_{i=0}^{\infty}(-a_i)\cdot t^i \in R[[t]]$ gilt

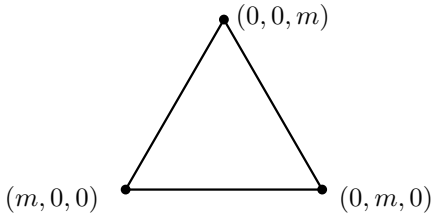
$$\sum_{i=0}^{\infty}(-a_i)\cdot t^i + \sum_{i=0}^{\infty} a_i\cdot t^i = \sum_{i=0}^{\infty}(-a_i+a_i)\cdot t^i = 0_{R[[t]]},$$

so daß $(R[[t]],+)$ eine abelsche Gruppe mit der Nullfunktion als neutralem Element ist. Man beachte nun, daß

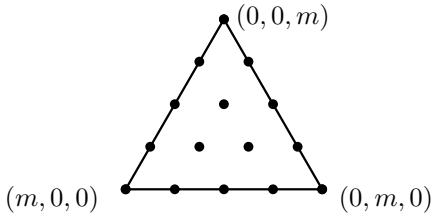
$$\sum_{k+l=m}\left(\sum_{i+j=k} a_i\cdot b_j\right)\cdot c_l = \sum_{i+j+l=m} a_i\cdot b_j\cdot c_l = \sum_{i+k=m} a_i\cdot\left(\sum_{j+l=k} b_j\cdot c_l\right), \tag{127}$$

da unter jeder der Summen jedes der Tripel (i,j,l) natürlicher Zahlen mit der Eigenschaft $i+j+l=m$ genau einmal vor kommt[10] und da in R das Assoziativgesetz der

[10]Man kann sich dies auch geometrisch veranschaulichen. Fassen wir (i,j,l) als Koordinaten des dreidimensionalen Raumes $\mathbb{R}^3$ auf, so bestimmt die Gleichung $i+j+l=m$ bei vorgegebenem m eine Ebene im Raum, nämlich die Ebene, die durch die drei Punkte $(m,0,0)$, $(0,m,0)$ und $(0,0,m)$ aufgespannt wird. Verbinden wir diese drei Punkte in dieser Ebene, so erhalten wir ein Dreieck:



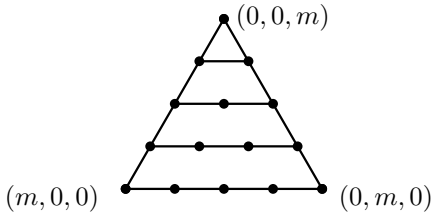
Die Punkte mit ganzzahligen Koordinaten in diesem Dreieck sind genau die Tripel nicht-negativer ganzer Zahlen, deren Summe m ist:



In der linken Seite von (127) zerlegen wir diese Menge wie folgt:

$$\bigcup_{l=0}^{m}\ \bigcup_{i+j=m-l}\{(i,j,l)\}.$$

In der inneren Summe werden also diejenigen ganzzahligen (i,j,l) in dem Dreieck zusammen gefaßt, für die die Koordinate l konstant ist und für die $i+j=m-l$ ist, d.h. die Punkte liegen auf einer Geraden parallel zur Geraden durch $(m,0,0)$ und $(0,m,0)$:



Multiplikation und das Distributivgesetz gelten. Damit gilt aber

$$\begin{aligned}\left(\sum_{i=0}^{\infty} a_i\cdot t^i \;\cdot\; \sum_{j=0}^{\infty} b_j\cdot t^j\right)\;\cdot\; \sum_{l=0}^{\infty} c_l\cdot t^l &= \sum_{k=0}^{\infty}\left(\sum_{i+j=k} a_i\cdot b_j\right)\cdot t^k \;\cdot\; \sum_{l=0}^{\infty} c_l\cdot t^l\\ &= \sum_{m=0}^{\infty}\left(\sum_{k+l=m}\left(\sum_{i+j=k} a_i\cdot b_j\right)\cdot c_l\right)\cdot t^m\\ &= \sum_{m=0}^{\infty}\left(\sum_{i+k=m} a_i\cdot \sum_{j+l=k} b_j\cdot c_l\right)\cdot t^m\\ &= \sum_{i=0}^{\infty} a_i\cdot t^i \;\cdot\; \sum_{k=0}^{\infty}\left(\sum_{j+l=k} b_j\cdot c_l\right)\cdot t^k\\ &= \sum_{i=0}^{\infty} a_i\cdot t^i \;\cdot\; \left(\sum_{j=0}^{\infty} b_j\cdot t^j \;\cdot\; \sum_{l=0}^{\infty} c_l\cdot t^l\right),\end{aligned}$$

so daß die Multiplikation auf $R[[t]]$ assoziativ ist. Ferner folgt aus der Kommutativität der Multiplikation auf R unmittelbar

$$\begin{aligned}\sum_{i=0}^{\infty} a_i\cdot t^i \;\cdot\; \sum_{j=0}^{\infty} b_j\cdot t^j &= \sum_{k=0}^{\infty}\left(\sum_{i+j=k} a_i\cdot b_j\right)\cdot t^k\\ &= \sum_{k=0}^{\infty}\left(\sum_{j+i=k} b_j\cdot a_i\right)\cdot t^k = \sum_{j=0}^{\infty} b_j\cdot t^j \;\cdot\; \sum_{i=0}^{\infty} a_i\cdot t^i.\end{aligned}$$

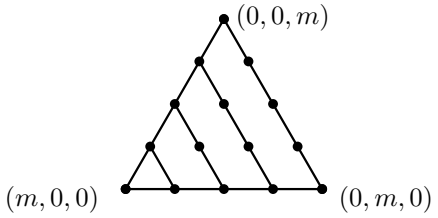
Und schließlich gilt für $1_{R[[t]]} = t^0 = \sum_{j=0}^{\infty} e_j\cdot t^j$ mit $e_0 = 1$ und $e_j = 0$ für $j \geq 1$:

$$1_{R[[t]]}\cdot \sum_{i=0}^{\infty} a_i\cdot t^i = \sum_{k=0}^{\infty}\left(\sum_{j+i=k} e_j\cdot a_i\right)\cdot t^k = \sum_{k=0}^{\infty} a_k\cdot t^k,$$

In der rechten Seite von (127) zerlegen wir diese Menge wie folgt:

$$\bigcup_{i=0}^{m}\ \bigcup_{j+l=m-i}\{(i,j,l)\}.$$

In der inneren Summe werden also diejenigen ganzzahligen (i,j,l) in dem Dreieck zusammen gefaßt, für die die Koordinate i konstant ist und für die $j+l = m-i$ ist, d.h. die Punkte liegen auf einer Geraden parallel zur Geraden durch $(0,m,0)$ und $(0,0,m)$:



In beiden Fällen wird jedes ganzzahlige Tripel (i,j,l) im Dreieck genau einmal betrachtet.

so daß t^0 unter Ausnutzung der Kommutativität der Multiplikation das Einselement von $R[[t]]$ ist.

Es bleibt nur, die Distributivität zu zeigen:

$$\begin{aligned}\sum_{i=0}^{\infty} a_i \cdot t^i \cdot \left(\sum_{j=0}^{\infty} b_j \cdot t^j + \sum_{j=0}^{\infty} c_j \cdot t^j\right) &= \sum_{i=0}^{\infty} a_i \cdot t^i \cdot \sum_{j=0}^{\infty} (b_j + c_j) \cdot t^j \\ &= \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i \cdot (b_j + c_j)\right) \cdot t^k \\ &= \sum_{k=0}^{\infty} \left(\sum_{i+j=k} (a_i \cdot b_j + a_i \cdot c_j)\right) \cdot t^k \\ &= \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i \cdot b_j\right) \cdot t^k + \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i \cdot c_j\right) \cdot t^k \\ &= \sum_{i=0}^{\infty} a_i \cdot t^i \cdot \sum_{j=0}^{\infty} b_j \cdot t^j + \sum_{i=0}^{\infty} a_i \cdot t^i \cdot \sum_{j=0}^{\infty} c_j \cdot t^j.\end{aligned}$$

Das zweite Distributivgesetz folgt mittels der Kommutativität der Multiplikation.

Damit haben wir gezeigt, daß $(R[[t]], +, \cdot)$ ein kommutativer Ring mit Eins ist. □

Aufgabe A5.9.
Es sei R ein kommutativer Ring mit Eins und $f = \sum_{k=0}^{\infty} a_k \cdot t^k \in R[[t]]$ eine formale Potenzreihe über R. Zeige, f ist genau dann eine Einheit in $R[[t]]$, wenn a_0 eine Einheit in R ist.
Hinweis, wenn a_0 eine Einheit in R ist, so ist eine Reihe $g = \sum_{k=0}^{\infty} b_k \cdot t^k$ mit $f \cdot g = t^0$ gesucht. Multipliziere die linke Seite der Gleichung aus und löse die Gleichungen, die sich für die Koeffizienten ergeben rekursiv.

B) Unterringe

Definition A5.10.
Sei R ein Ring mit Eins und $S \subseteq R$. S heißt ein *Unterring* von R, wenn

a. $1_R \in S$,

b. $a + b \in S$ für alle $a, b \in S$,

c. $-a \in S$ für alle $a \in S$, und

d. $a \cdot b \in S$ für alle $a, b \in S$.

Ist R ein Körper und S ein Unterring von R für den zusätzlich $a^{-1} \in S$ für alle $a \in S \setminus \{0\}$ gilt, so nennen wir S auch einen *Unterkörper* oder *Teilkörper* von R.

Man beachte, daß ein Unterring S von R insbesondere selbst wieder Ring ist bezüglich der Einschränkung der Addition und Multiplikation von R auf S, und daß entsprechend ein Teilkörper selbst ein Körper ist.

Beispiel A5.11.
$\mathbb{Z}$ ist ein Unterring von $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$. $\mathbb{Q}$ ist ein Teilkörper von $\mathbb{R}$ und $\mathbb{C}$. $\mathbb{R}$ ist ein Teilkörper von $\mathbb{C}$.

Das neben den ganzen Zahlen wichtigste Beispiel eines kommutativen Ringes mit Eins in dieser Vorlesung ist der Polynomring, den wir als Unterring des formalen Potenzreihenringes erhalten.

Definition A5.12.
Ist R ein kommutativer Ring, so nennen wir

$$R[t] := \left\{ \sum_{k=0}^{\infty} a_k \cdot t^k \in R[[t]] \;\middle|\; \text{nur endlich viele } a_k \text{ sind ungleich null} \right\}$$

$$= \left\{ \sum_{k=0}^{n} a_k \cdot t^k \in R[[t]] \;\middle|\; n \in \mathbb{N}, a_0, \ldots, a_n \in R \right\}$$

den *Polynomring* über R in der Unbestimmten t und die Elemente von $R[t]$ heißen *Polynome*. Für $0 \neq f = \sum_{k=0}^{\infty} a_k \cdot t^k \in R[t]$ nennen wir

$$\deg(f) = \max\{k \mid a_k \neq 0\}$$

den *Grad* des *Polynoms* f und $\operatorname{lc}(f) := a_{\deg(f)}$ seinen *Leitkoeffizienten*. Ferner setzen wir $\deg(0) = -\infty$ und $\operatorname{lc}(0) := 0$.
Beachte, daß aufgrund der in Definition A5.6 getroffenen Konvention jedes Polynom in $R[t]$ die Form

$$\sum_{k=0}^{n} a_k \cdot t^k$$

für ein $n \in \mathbb{N}$ hat.

Beispiel A5.13.
$3 \cdot t^4 - t^2 + 5 \cdot t^0 \in \mathbb{Z}[t]$ ist ein Polynom vom Grad $\deg(f) = 4$ und mit Leitkoeffizient $\operatorname{lc}(f) = 3$.

Da $R[t]$ abgeschlossen ist bezüglich Addition, Negativen und Multiplikation und da $1_{R[[t]]} = t^0 \in R[t]$ erhalten wir folgenden Satz.

Satz A5.14.
Ist R ein kommutativer Ring mit Eins, so ist $R[t]$ ein Unterring von $R[[t]]$. Insbesondere ist $R[t]$ selbst ein kommutativer Ring mit Eins.

Beweis: Seien $f = \sum_{k=0}^{m} a_k \cdot t^k, g = \sum_{k=0}^{n} b_k \cdot t^k \in R[t]$ gegeben (wobei zugelassen ist, daß alle a_k oder b_k Null sind), so ist

$$f + g = \sum_{k=0}^{\max\{m,n\}} (a_k + b_k) \cdot t^k \in R[t], \tag{128}$$

$$-f = \sum_{k=0}^{m} (-a_k) \cdot t^k \in R[t]$$

und

$$f \cdot g = \sum_{k=0}^{m+n} \left(\sum_{i=0}^{k} a_i \cdot b_{k-i} \right) \cdot t^k \in R[t], \tag{129}$$

wobei wir die Konvention verwenden, daß $a_k = 0$ für $k > m$ und $b_k = 0$ für $k > n$. Um zu sehen, daß in (129) keine Terme vom Grad größer als $n + m$ nötig sind, beachte man einfach, daß der Koeffizient von t^k für $k > n + m$ die Form

$$\sum_{i=0}^{m} a_i \cdot b_{k-i} + \sum_{i=m+1}^{k} a_i \cdot b_{k-i}$$

hat. Die zweite Summe ist Null, da alle a_i dort Null sind, während die erste Summe Null ist, da dort alle b_{k-i} Null sind. □

Die folgenden Gradformeln für Polynome ergeben sich unmittelbar aus dem obigen Beweis. Dabei verwenden wir die Konvention $m + -\infty = -\infty$ und $\max\{m, -\infty\} = m$ für alle $m \in \mathbb{N} \cup \{-\infty\}$.

Proposition A5.15 (Gradformeln).
Es sei R ein kommutativer Ring mit Eins und $f, g \in R[t]$ seien zwei Polynome. Dann gelten:

a. $\deg(f + g) \leq \max\{\deg(f), \deg(g)\}$.

b. $\deg(f \cdot g) \leq \deg(f) + \deg(g)$.

c. $\deg(f \cdot g) = \deg(f) + \deg(g)$ genau dann, wenn $\mathrm{lc}(f) \cdot \mathrm{lc}(g) \neq 0$.
In diesem Fall gilt auch $\mathrm{lc}(f \cdot g) = \mathrm{lc}(f) \cdot \mathrm{lc}(g)$.

Beweis: Ist $f = 0$ oder $g = 0$, so sind die Aussagen offenbar korrekt und wir können deshalb $f \neq 0 \neq g$ annehmen. Dann folgt a. unmittelbar aus (128) und b. aus (129).

Für c. beachte man, daß in (129) der Koeffizient von t^{m+n} gerade $a_m \cdot b_n = \mathrm{lc}(f) \cdot \mathrm{lc}(g)$ ist. □

Aufgabe A5.16. a. Es sei R ein kommutativer Ring mit Eins und $S \subset R$ ein nicht-leere Teilmenge für die gilt:

- $x + y \in S$ für alle $x, y \in S$,
- $-x \in S$ für alle $x \in S$,
- $x \cdot y \in S$ für alle $x, y \in S$ und
- $1_R \in S$.

Zeige, S ist ein kommutativer Ring mit Eins bezüglich der Einschränkung der Addition und der Multiplikation von R auf S.

b. Zeige, $\mathbb{Z}[i] := \{a + i \cdot b \in \mathbb{C} \mid a, b \in \mathbb{Z}\}$ ist ein kommutativer Ring mit Eins, wobei die Addition und die Multiplikation einfach die Addition und Multiplikation komplexer Zahlen sein sollen. Man nennt diesen Ring den *Ring der ganzen Gaußschen Zahlen*.

c. Bestimme die Einheitengruppe $\mathbb{Z}[i]^*$ des Ringes $\mathbb{Z}[i]$.

Aufgabe A5.17.

Für $\omega \in \mathbb{Z}$, $\omega \geq 2$, bezeichnen wir mit $\sqrt{-\omega}$ die komplexe Zahl $i \cdot \sqrt{\omega}$.

a. Zeige, $\mathbb{Z}\left[\sqrt{-\omega}\right] := \{a + b \cdot \sqrt{-\omega} \in \mathbb{C} \mid a, b \in \mathbb{Z}\}$ ist ein kommutativer Ring mit Eins, wobei die Addition und die Multiplikation einfach die Addition und Multiplikation komplexer Zahlen sein sollen.

b. Zeige, $\mathbb{Z}\left[\sqrt{-\omega}\right]^* = \{1, -1\}$.

C) Ringhomomorphismen

Mit einer neuen Struktur definieren wir auch gleich die strukturerhaltenden Abbildungen. Man beachte hierbei, daß zur Struktur eines Ringes mit *Eins* neben der Addition und der Multiplikation auch das Vorhandensein eines Einselementes zählt. Wir werden deshalb fordern, daß eine strukturerhaltende Abbildung verträglich ist mit der Addition und der Multiplikation und daß sie zudem das Einselement respektiert.

Definition A5.18.

Es seien R und S zwei Ringe mit Eins. Eine Abbildung $\varphi : R \to S$ heißt *Ringhomomorphismus*, falls

a. $\varphi(a + b) = \varphi(a) + \varphi(b)$ für alle $a, b \in R$,

b. $\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$ für alle $a, b \in R$ und

c. $\varphi(1_R) = 1_S$.

Ist φ ein Ringhomomorphismus, dann nennen wir φ einen

- *Monomorphismus*, falls φ injektiv ist;
- *Epimorphismus*, falls φ surjektiv ist;
- *Isomorphismus*, falls φ bijektiv ist.

Wir nennen zwei Ringe R und S *isomorph*, falls es einen Isomorphismus von R nach S gibt. Wir schreiben dann kurz $R \cong S$.

Beispiel A5.19.
Ist $S \subseteq R$ ein Unterring des Ringes R, so ist die kanonische Inklusion $i_S : S \to R$ ein Ringhomomorphismus.

Lemma A5.20.
Ist $\varphi : R \to S$ ein bijektiver Ringhomomorphismus, dann ist auch $\varphi^{-1} : S \to R$ ein Ringhomomorphismus.

Beweis: Daß φ^{-1} mit der Addition verträglich ist, folgt aus Proposition 22.22 d., da φ ein Homomorphismus von der abelschen Gruppe $(R, +)$ nach $(S, +)$ ist. Für die Verträglichkeit mit der Multiplikation kopiere man den dortigen Beweis. Schließlich gilt $\{x \in R \mid \varphi(x) = 1_S\} = \{1_R\}$ wegen der Bijektivität von φ und da $\varphi(1_R) = 1_S$, und mithin ist $\varphi^{-1}(1_S) = 1_R$. □

Lemma A5.21.
Ist $\varphi : R \longrightarrow S$ ein Ringhomomorphismus, so ist $\mathrm{Im}(\varphi)$ ein Unterring von S.

Beweis: Nach Proposition 22.22 ist $\mathrm{Im}(\varphi)$ eine Untergruppe von $(S, +)$, so daß $\mathrm{Im}(\varphi)$ abgeschlossen ist bezüglich Addition und Negativen. Zudem gilt

$$1_S = \varphi(1_R) \in \mathrm{Im}(\varphi)$$

und für $\varphi(a), \varphi(b) \in \mathrm{Im}(\varphi)$ gilt

$$\varphi(a) \cdot \varphi(b) = \varphi(a \cdot b) \in \mathrm{Im}(\varphi).$$

□

Da ein Ringhomomorphismus $\varphi : R \longrightarrow S$ nach Definition ein Gruppenhomomorphismus von $(R, +)$ nach $(S, +)$ ist, folgt aus Lemma 22.23 unmittelbar folgendes Kriterium für die Injektivität von φ, wobei $\mathrm{Ker}(\varphi) = \{a \in R \mid \varphi(a) = 0_S\}$.

Lemma A5.22.
Ein Ringhomomorphismus $\varphi : R \longrightarrow S$ ist genau dann ein Monomorphismus, wenn $\mathrm{Ker}(\varphi) = \{0_R\}$.

Bemerkung A5.23.
Ist R ein kommutativer Ring mit Eins, so ist die Abbildung

$$\iota : R \longrightarrow R[t] : a \mapsto a \cdot t^0$$

ein Ringmonomorphismus, und somit ist R isomorph zum Unterring $\mathrm{Im}(\iota) = \{a \cdot t^0 \mid a \in R\}$. Wir werden diesen Isomorphismus in Zukunft nutzen und die Elemente von R mit den konstanten Polynomen identifizieren, d.h. wir schreiben z.B. $2t^2 + 3$ anstatt $2t^2 + 3t^0$.

Aufgabe A5.24.
Es seien K ein Körper, R ein kommutativer Ring mit $1_R \neq 0_R$ und $\varphi : K \longrightarrow R$ ein Ringhomomorphismus. Zeige, φ ist ein Monomorphismus.

Aufgabe A5.25.
Es sei S ein kommutativer Ring mit Eins, $R \subseteq S$ ein Unterring und $b \in S$.

a. Wir definieren

$$f(b) = \sum_{k=0}^{n} a_k \cdot b^k \in S$$

für $f = \sum_{k=0}^{n} a_k \cdot t^k \in R[t]$. Zeige, daß die Abbildung

$$\varphi_b : R[t] \longrightarrow S : f \mapsto f(b)$$

ein Ringhomomorphismus ist. Wir nennen φ_b einen *Einsetzhomomorphismus*.

b. Ist b Nullstelle des Polynoms $g = t^n + \alpha_{n-1} \cdot t^{n-1} + \ldots + \alpha_1 \cdot t + \alpha_0 \in R[t]$, $n \geq 1$, so ist

$$\mathrm{Im}(\varphi_b) = \left\{ a_0 + a_1 \cdot b + a_2 \cdot b^2 + \ldots + a_{n-1} \cdot b^{n-1} \;\middle|\; a_0, \ldots, a_{n-1} \in R \right\}.$$

Wir bezeichnen diesen Unterring von S mit $R[b] = \mathrm{Im}(\varphi_b)$.

D) Ideale

Ist R ein kommutativer Ring mit Eins und S ein Unterring von R, so ist insbesondere $(S, +)$ ein Normalteiler der abelschen Gruppe $(R, +)$. Wir können mithin die Faktorgruppe $(R/S, +)$ bilden, wobei für zwei Nebenklassen $\overline{a}, \overline{b} \in R/S$ die Summe definiert ist als $\overline{a} + \overline{b} = \overline{a + b}$. Wir würden gerne auch die zweite Operation, die wir auf R und S

haben, auf die Faktorgruppe R/S fortsetzen durch $\overline{a} \cdot \overline{b} = \overline{a \cdot b}$ und so R/S zu einem kommutativen Ring mit Eins machen. Das geht aber schief! Denn das Nullelement von R/S muß notwendig $\overline{0}$ sein, und für ein beliebiges $\overline{a} \in R/S$ und $b \in S$ muß dann

$$S = \overline{0} = \overline{a \cdot 0} = \overline{a} \cdot \overline{b} = \overline{a \cdot b} = (a \cdot b) + S$$

gelten. Also ist $a \cdot b \in S$, d.h. S ist abgeschlossen bezüglich der Multiplikation mit beliebigen Elementen aus R. Da nach Voraussetzung $1_R \in S$, müßte für ein beliebiges $a \in R$

$$a = a \cdot 1_R \in S$$

gelten und somit $S = R$. D.h. der einzige Unterring, für den die zugehörige Faktorgruppe $(R/S, +)$ auf diesem Weg zu einem Ring mit Eins gemacht werden könnte, wäre der Ring R selbst. Dann wäre aber R/S der Nullring, und wir könnten uns die Mühe sparen.

Es bleibt uns nichts anderes übrig, als den Begriff des Unterrings durch einen anderen Begriff zu ersetzen, der es uns erlaubt, Faktorstrukturen zu bilden. Wir haben bereits gesehen, daß es wünschenswert wäre, daß es sich bei dieser neu zu definierenden Unterstruktur um eine Untergruppe von $(R, +)$ handelt, die bezüglich der Multiplikation mit beliebigen Elementen aus R abgeschlossen ist. Dies führt zu folgender Definition, bei der wir uns wie für den Rest des Kapitels auf *kommutative* Ringe mit Eins beschränken.

Definition A5.26.
Es sei R ein kommutativer Ring mit Eins und $\emptyset \neq I \subseteq R$ eine nicht-leere Teilmenge. I heißt ein *Ideal* von R, falls

(1) $a + b \in I$ für alle $a, b \in I$ und

(2) $r \cdot a \in I$ für alle $r \in R$ und $a \in I$.

Wir schreiben in diesem Fall $I \trianglelefteq R$, da Ideale für Ringe die Analoga von Normalteilern für Gruppen sind.

Bemerkung A5.27.
Es sei R ein kommutativer Ring mit Eins und $I \trianglelefteq R$. Dann ist $(I, +)$ eine Untergruppe von $(R, +)$. Dies folgt aus dem Untergruppenkriterium 22.8, da mit $-1_R \in R$ und für $a \in I$ auch $-a = -1_R \cdot a \in I$.

Beachte auch, daß aus der Definition eines Ideals per Induktion unmittelbar

$$\sum_{k=1}^{n} r_k \cdot a_k \in I$$

für alle $r_k \in R$ und $a_k \in I$ folgt.

Beispiel A5.28.

a. Ist R ein kommutativer Ring mit Eins, dann sind $\{0_R\}$ und R die *trivialen* Ideale von R.

b. $n\mathbb{Z}$ ist ein Ideal in $\mathbb{Z}$ für jedes $n \in \mathbb{Z}$, da die Menge abgeschlossen bezüglich $+$ und bezüglich Multiplikation mit ganzen Zahlen ist.

Da jedes Ideal von $(\mathbb{Z}, +, \cdot)$ insbesondere eine Untergruppe von $(\mathbb{Z}, +)$ ist folgt aus obigem Beispiel und der Klassifikation der Untergruppen von $(\mathbb{Z}, +)$ in Proposition 22.16 folgendes Korollar.

Korollar A5.29.
Für eine Teilmenge $U \subseteq \mathbb{Z}$ sind die folgenden Aussagen äquivalent:

a. U ist ein Ideal von $(\mathbb{Z}, +, \cdot)$.

b. U ist eine Untergruppe von $(\mathbb{Z}, +)$.

c. $U = n\mathbb{Z}$ für eine ganze Zahl $n \geq 0$.

Wie bei Untergruppen wollen wir auch bei Idealen wieder wissen, wie sie sich bezüglich bestimmter mengentheoretischer Operationen verhalten.

Proposition A5.30.
Ist R ein kommutativer Ring mit Eins und sind $I_\lambda \trianglelefteq R$ Ideale für $\lambda \in \Lambda$, dann ist $\bigcap_{\lambda\in\Lambda} I_\lambda \trianglelefteq R$ ein Ideal.

Beweis: Seien $a, b \in \bigcap_{\lambda\in\Lambda} I_\lambda$ und $r \in R$. Dann ist $a + b \in I_\lambda$ und $r \cdot a \in I_\lambda$, da I_λ ein Ideal ist. Mithin ist auch

$$a + b, r \cdot a \in \bigcap_{\lambda\in\Lambda} I_\lambda.$$

Zudem ist $0_R \in I_\lambda$, da $(I_\lambda, +)$ eine Untergruppe von $(R, +)$ ist, und mithin ist $0_R \in \bigcap_{\lambda\in\Lambda} I_\lambda$, so daß die Menge nicht leer ist. □

Definition A5.31.
Es sei R ein kommutativer Ring mit Eins und $M \subseteq R$ eine Teilmenge. Wir definieren das *Erzeugnis* von M als

$$\langle M\rangle_R = \bigcap_{M\subseteq I\trianglelefteq R} I,$$

den Schnitt aller Ideale von R, die M enthalten.

Proposition A5.32.
Es sei R ein kommutativer Ring mit Eins und $\emptyset \neq M \subseteq R$. Dann ist das Erzeugnis

$$\langle M\rangle_R = \left\{ \sum_{k=1}^{n} r_k \cdot a_k \;\middle|\; a_k \in M, r_k \in R, n \geq 1 \right\} \trianglelefteq R$$

von M die Menge aller endlichen *Linearkombinationen* von Elementen in M mit Koeffizienten in R.

Beweis: Wir setzen

$$J = \left\{ \sum_{k=1}^{n} r_k \cdot a_k \mid a_k \in M, r_k \in R, n \geq 1 \right\}$$

und zeigen zunächst, daß J ein Ideal von R ist.

Da M nicht die leere Menge ist, ist auch J nicht leer. Sind nun $\sum_{k=1}^{n} r_k \cdot a_k, \sum_{k=1}^{m} s_k \cdot b_k \in J$ mit $r_k, s_k \in R$ und $a_k, b_k \in M$, so setzen wir einfach $r_k = s_{k-n}$ und $a_k = b_{k-n}$ für $k = n+1, \ldots, n+m$ und erhalten

$$\sum_{k=1}^{n} r_k \cdot a_k + \sum_{k=1}^{m} s_k \cdot b_k = \sum_{k=1}^{n+m} r_k \cdot a_k \in J.$$

Da zudem für $r \in R$ auch $r \cdot r_k \in R$ gilt auch

$$r \cdot \sum_{k=1}^{n} r_k \cdot a_k = \sum_{k=1}^{n} (r \cdot r_k) \cdot a_k \in J.$$

Also ist J ein Ideal von R.

Zudem ist $M \subset J$, da $a = 1_R \cdot a \in J$ für alle $a \in M$. Also gilt nach Definition

$$\langle M\rangle_R = \bigcap_{M \subseteq I \trianglelefteq R} I \subseteq J.$$

Andererseits gilt für $\sum_{k=1}^{n} r_k \cdot a_k \in J$ mit $r_k \in R$ und $a_k \in M$, daß

$$\sum_{k=1}^{n} r_k \cdot a_k \in I$$

für jedes Ideal $I \trianglelefteq R$, das M enthält, wegen Bemerkung A5.27. Mithin gilt auch

$$J \subseteq \bigcap_{M \subseteq I \trianglelefteq R} I = \langle M\rangle_R.$$

□

Beispiel A5.33.
Ist R ein kommutativer Ring mit Eins und $a, b \in R$, dann gelten

$$\langle a\rangle_R = \{r \cdot a \mid r \in R\}$$

und

$$\langle a, b\rangle_R = \{r \cdot a + s \cdot b \mid r, s \in R\}.$$

Insbesondere gilt $n\mathbb{Z} = \langle n\rangle_{\mathbb{Z}}$.

Aufgabe A5.34.
Zeige, daß

$$I = \{f \in \mathbb{Z}[t] \mid f(5) = 0\}$$

ein Ideal von $\mathbb{Z}[t]$ ist.

Aufgabe A5.35.
Es sei R ein kommutativer Ring. R ist genau dann ein Körper, wenn R genau zwei Ideale besitzt.

Aufgabe A5.36.
Es sei R ein Ring und $I_k \trianglelefteq R$, $k \in \mathbb{N}$, seien Ideale mit der Eigenschaft

$$I_0 \subseteq I_1 \subseteq I_2 \subseteq I_3 \subseteq \ldots,$$

d.h. $I_k \subseteq I_{k+1}$ für alle $k \in \mathbb{N}$. Zeige, daß dann auch

$$\bigcup_{k\in\mathbb{N}} I_k \trianglelefteq R$$

ein Ideal in R ist.

E) Faktorringe

Satz A5.37.
Es sei R ein kommutativer Ring mit Eins und $I \trianglelefteq R$ ein Ideal. Dann wird auf der abelschen Gruppe $(R/I, +)$ durch

$$\overline{a} \cdot \overline{b} := \overline{a \cdot b}$$

für $\overline{a}, \overline{b} \in R/I$ eine zweistellige Operation definiert, und $(R/I, +, \cdot)$ ist ein kommutativer Ring mit Einselement $1_{R/I} = \overline{1_R}$. Wir nennen R/I den *Faktorring* von R nach I.

Beweis: Wir müssen zunächst zeigen, daß die Operation wohldefiniert ist, also nicht von der Wahl der Repräsentanten der Nebenklassen abhängt. Seien dazu $\overline{a} = \overline{a'}$ und $\overline{b} = \overline{b'}$, dann gilt nach Definition $a = a' + c$ und $b = b' + d$ mit $c, d \in I$. Damit folgt dann

$$a \cdot b = (a' + c) \cdot (b' + d) = a' \cdot b' + \big(a' \cdot d + c \cdot b' + c \cdot d\big)$$

mit $a' \cdot d + c \cdot b' + c \cdot d \in I$. Also gilt

$$\overline{a \cdot b} = a \cdot b + I = a' \cdot b' + I = \overline{a' \cdot b'},$$

und die Multiplikation ist wohldefiniert. Die Assoziativität und die Kommutativität der Multiplikation folgen dann aus den entsprechenden Eigenschaften der Multiplikation auf R. Zudem ist $\overline{1} \cdot \overline{a} = \overline{1 \cdot a} = \overline{a}$ für alle $\overline{a} \in R/I$, so daß $(R/I, +, \cdot)$ ein kommutativer Ring mit Eins $\overline{1}$ ist. □

Beispiel A5.38.
$(\mathbb{Z}_n, +, \cdot)$ ist ein kommutativer Ring mit Eins vermittels

$$\overline{a} + \overline{b} = \overline{a + b} \quad \text{und} \quad \overline{a} \cdot \overline{b} = \overline{a \cdot b}$$

für alle $a, b \in \mathbb{Z}$ und $n \geq 0$.

Beachte, in $\mathbb{Z}_2 = \{\overline{0}, \overline{1}\}$ ist offenbar jedes Element ungleich $\overline{0}$ eine Einheit, so daß $(\mathbb{Z}_2, +, \cdot)$ ein Körper ist. Da zudem jeder Körper K mindestens zwei Elemente, nämlich $0_K \neq 1_K$ enthalten muß, ist $\mathbb{Z}_2$ der kleinstmögliche Körper.

Aufgabe A5.39.
Welcher der folgenden Ringe ist ein Körper?

a. $\mathbb{Z}_4$.

b. $\mathbb{Z}_7$.

Beweise Deine Vermutung.

Aufgabe A5.40.
Für eine positive ganze Zahl n definieren wir die Abbildung

$$\phi_n : \mathbb{Z}[t] \longrightarrow \mathbb{Z}_n[t] : \sum_{k=0}^{m} a_k \cdot t^k \mapsto \sum_{k=0}^{m} \overline{a_k} \cdot t^k.$$

Zeige, daß ϕ_n ein Ringepimorphismus ist. Wir nennen ϕ_n *Reduktion modulo* n.

Aufgabe A5.41.

a. Bestimme $\mathbb{Z}_6^*$.

b. Bestimme $\mathbb{Z}_8^*$.

c. Bestimme $\mathbb{Z}_{15}^*$.

d. Stelle eine Vermutung auf, wann eine Element $\overline{z} \in \mathbb{Z}_n$ für $n \geq 2$ eine Einheit ist.

e. Zeige, für alle $n \geq 2$ ist $\overline{n-1} \in \mathbb{Z}_n$ eine Einheit.

F) Homomorphiesatz

Satz A5.42 (Homomorphiesatz).
Es sei $\varphi : R \longrightarrow S$ ein Ringhomomorphismus zwischen kommutativen Ringen mit Eins. Dann ist $\operatorname{Ker}(\varphi) = \{a \in R \mid \varphi(a) = 0\} \trianglelefteq R$ ein Ideal und

$$\overline{\varphi} : R/\operatorname{Ker}(\varphi) \longrightarrow \operatorname{Im}(\varphi) : \overline{a} \mapsto \varphi(a)$$

ein Isomorphismus. Insbesondere gilt $R/\operatorname{Ker}(\varphi) \cong \operatorname{Im}(\varphi)$.

Beweis: Nach Proposition 22.22 ist $(\operatorname{Ker}(\varphi), +)$ eine Untergruppe von $(R, +)$, also insbesondere nicht leer und abgeschlossen bezüglich der Addition. Sei nun $a \in \operatorname{Ker}(\varphi)$ und $r \in R$. Dann gilt

$$\varphi(r \cdot a) = \varphi(r) \cdot \varphi(a) = \varphi(r) \cdot 0 = 0.$$

Also ist $r \cdot a \in \operatorname{Ker}(\varphi)$ und $\operatorname{Ker}(\varphi)$ ist ein Ideal. Aus dem Homomorphiesatz A3.18 folgt dann, daß $\overline{\varphi}$ ein Isomorphismus abelscher Gruppen ist. Da zudem

$$\overline{\varphi}\big(\overline{a} \cdot \overline{b}\big) = \overline{\varphi}\big(\overline{a \cdot b}\big) = \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b) = \overline{\varphi}\big(\overline{a}\big) \cdot \overline{\varphi}\big(\overline{b}\big)$$

und $\overline{\varphi}\big(\overline{1}\big) = \varphi(1) = 1$ gilt, ist $\overline{\varphi}$ ein Isomorphismus von Ringen. □

Aufgabe A5.43.

a. Finde alle Ringhomomorphismen $\varphi : \mathbb{Z} \longrightarrow \mathbb{Z}_6$.

b. Finde alle Ringhomomorphismen $\varphi : \mathbb{Z}_6 \longrightarrow \mathbb{Z}$.

c. Finde alle Ringhomomorphismen $\varphi : \mathbb{Q} \longrightarrow \mathbb{R}$.

§ A6 Teilbarkeit in Ringen

Wir haben im vorigen Kapitel den Begriff des kommutativen Ringes mit Eins eingeführt. Als Modell für diesen Begriff haben uns die ganzen Zahlen gedient, und sie sollen auch das Leitbild für die in diesem Kapitel betrachteten Eigenschaften von Ringen und ihren Elementen sein.

A) Integritätsbereiche

Der zentrale Begriff wird der der Teilbarkeit sein. Eine uns vertraute Eigenschaft der natürlichen Zahlen ist, daß das Produkt zweier Zahlen nur dann Null ergeben kann, wenn eine der Zahlen bereits Null ist. Dies gilt in beliebigen Ringen nicht mehr. Betrachtet man etwa den Ring $\mathbb{Z}_4$, so ist die Restklasse $\overline{2} \neq \overline{0}$, aber $\overline{2} \cdot \overline{2} = \overline{4} = \overline{0}$. Dies hat unangenehme Folgen, denn $\overline{0}$ läßt sich somit auf mehrere Weisen als Vielfaches von $\overline{2}$ schreiben:

$$\overline{2} \cdot \overline{2} = \overline{0} = \overline{0} \cdot \overline{2}.$$

In einem solchen Ring gelten ganz offensichtlich die Kürzungsregeln für die Multiplikation nicht mehr. Diese sind aber für den Begriff der Teilbarkeit von zentraler Bedeutung, schließlich wollen wir einen *Teiler* auch *wegkürzen* können. Wir führen deshalb einen neuen Begriff für solche Ringe ein, die sich in dieser Beziehung vernünftig verhalten.

Definition A6.1.
Es sei R ein kommutativer Ring mit Eins und $a \in R$.

a. a heißt ein *Nullteiler*, falls es ein $0 \neq b \in R$ gibt mit $a \cdot b = 0$.

b. R heißt *Integritätsbereich* oder *nullteilerfrei*, falls 0 der einzige Nullteiler in R ist.

Beispiel A6.2.

a. Ist R nicht der Nullring, so ist 0 ein Nullteiler, da $0 \cdot 1 = 0$ und $1 \neq 0$.

b. Ist $a \in R^*$ eine Einheit, so ist a kein Nullteiler.
Denn da a ein Inverses $a^{-1} \in R$ besitzt, folgt aus $a \cdot b = 0$ unmittelbar

$$b = 1 \cdot b = \left(a^{-1} \cdot a\right) \cdot b = a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0 = 0.$$

c. Aus b. folgt, daß jeder Körper ein Integritätsbereich ist, da 0 das einzige Element ist, das keine Einheit ist. Insbesondere sind $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ Integritätsbereiche.

d. Jeder Unterring eines Integritätsbereichs ist ein Integritätsbereich. Insbesondere sind also $\mathbb{Z}$ und

$$\mathbb{Z}\left[\sqrt{-5}\right] = \left\{a + b \cdot \sqrt{-5} \in \mathbb{C} \;\middle|\; a, b \in \mathbb{Z}\right\}$$

Integritätsbereiche.

e. Ist R ein Integritätsbereich, so ist $R[t]$ ein Integritätsbereich und $R[t]^* = R^*$. Denn sind $f, g \in R[t] \setminus \{0\}$, so ist $\deg(f), \deg(g) \geq 0$ und $\mathrm{lc}(f) \neq 0 \neq \mathrm{lc}(g)$. Aufgrund der Gradformeln für Polynome A5.15 folgt deshalb

$$\deg(f \cdot g) = \deg(f) + \deg(g) \geq 0, \tag{130}$$

da $\mathrm{lc}(f) \cdot \mathrm{lc}(g) \neq 0$ im Integritätsbereich R, und somit ist $f \cdot g \neq 0$. Also ist $R[t]$ ein Integritätsbereich. Ist $f \in R[t]^*$ eine Einheit und g das zugehörige Inverse, so gilt $f \cdot g = t^0 = 1$ und aus (130) folgt, daß $\deg(f) = 0 = \deg(g)$. D.h. f und g sind konstante Polynome und mithin gilt $f, g \in R^*$. Ist umgekehrt $f \in R^* \subseteq R[t]$, so gibt es ein $g \in R \subseteq R[t]$ mit $f \cdot g = 1 = t^0$ und somit $f \in R[t]^*$.

f. $\mathbb{Z}_4$ ist kein Integritätsbereich, da $\overline{2}$ wegen $\overline{2} \cdot \overline{2} = \overline{0}$ ein Nullteiler ist.

Lemma A6.3 (Kürzungsregeln).
Ist R ein Integritätsbereich, so gelten die Kürzungsregeln der Multiplikation, d.h. für alle $a, b, c \in R$ mit $a \neq 0$ gilt

$$a \cdot b = a \cdot c \quad \Longrightarrow \quad b = c$$

und

$$b \cdot a = c \cdot a \quad \Longrightarrow \quad b = c.$$

Beweis: Wegen der Kommutativität der Multiplikation reicht es, eine der Kürzungsregeln zu zeigen. Seien dazu $a, b, c \in R$ mit $ab = ac$. Dann gilt

$$0 = ab - ac = a \cdot (b - c). \tag{131}$$

Da $a \neq 0$ und R ein Integritätsbereich ist, ist a kein Nullteiler. Mithin folgt aus (131), daß $b - c = 0$ und somit $b = c$ gilt. □

Nun können wir den Begriff der Teilbarkeit für Elemente eines Integritätsbereiches einführen. Dabei sollten wir beachten, daß wir in einem Integritätsbereich zunächst nur die Operationen der Addition und der Multiplikation haben, nicht aber eine Division. Wir müssen also mit ersteren auskommen um Teilbarkeit zu definieren.

Definition A6.4.
Sei R ein Integritätsbereich und $a, b \in R$.

a. Wir sagen b *teilt* a, falls es ein $c \in R$ gibt mit $a = b \cdot c$. Wir schreiben in diesem Fall $b \mid a$.

b. Wir nennen $g \in R$ einen *größten gemeinsamen Teiler* von a und b, falls die folgenden beiden Eigenschaften erfüllt sind:

(1) $g \mid a$ und $g \mid b$.

(2) Für alle $h \in R$ mit $h \mid a$ und $h \mid b$ gilt $h \mid g$.

Wir bezeichnen mit

$$\mathrm{ggT}(a,b) = \{g \in R \mid g \text{ ist größter gemeinsamer Teiler von } a \text{ und } b\}$$

die Menge der größten gemeinsamen Teiler von a und b.

c. Wir nennen $k \in R$ ein *kleinstes gemeinsames Vielfaches* von a und b, falls die folgenden beiden Eigenschaften erfüllt sind:

(1) $a \mid k$ und $b \mid k$.

(2) Für alle $l \in R$ mit $a \mid l$ und $b \mid l$ gilt $k \mid l$.

Wir bezeichnen mit

$$\mathrm{kgV}(a,b) = \{k \in R \mid k \text{ ist kleinstes gemeinsames Vielfaches von } a \text{ und } b\}$$

die Menge der kleinsten gemeinsamen Vielfachen von a und b.

d. Wir nennen a und b *teilerfremd*, wenn $1 \in \mathrm{ggT}(a,b)$.

Bemerkung A6.5.

Bei der Definition eines größten gemeinsamen Teilers g von a und b bedeutet die Bedingung (1), daß g überhaupt ein Teiler von a und von b ist. Die Bedingung (2) dient dazu den Zusatz *größter* zu rechtfertigen. Wie soll man in einem beliebigen Integritätsbereich entscheiden, wann g größer als h ist für $g, h \in R$? In $\mathbb{Z}$ kann man dazu vielleicht die bekannte Ordnungsrelation "$\leq$" heranziehen, indem man zum Beispiel für die Teiler $h = 2$ und $g = 6$ der Zahlen $a = 12$ und $b = 30$ definiert, daß wohl 6 der *größere* Teiler ist. Aber wie sollte man das etwa im Polynomring $\mathbb{Z}[t]$ machen? Ist $t + 2$ größer als t oder kleiner oder kann man sie vielleicht gar nicht vergleichen? Man macht sich deshalb zunutze, daß in den ganzen Zahlen der *größte* gemeinsame Teiler von a und b von allen gemeinsamen Teilern von a und b geteilt wird. In obigem Beispiel etwa sind 1, 2, 3 und 6 die einzigen positiven gemeinsamen Teiler von 12 und 30, und sie alle teilen $g = 6$. Man kann einen Teiler also *kleiner* als einen anderen nennen, wenn ersterer letzteren teilt. In diesem Sinne legt die Bedingung (2) in $\mathbb{Z}$ in der Tat fest, welcher der beiden Teiler g und h größer ist.

Nach Definition ist ein *größter gemeinsamer Teiler* zweier Elemente a und b also ein gemeinsamer Teiler von a und b, der von jedem anderen gemeinsamen Teiler geteilt wird. Analog ist ein *kleinstes gemeinsames Vielfaches* von a und b ein gemeinsames Vielfaches, das jedes andere gemeinsame Vielfache teilt.

Weshalb sprechen wir in der Definition von *einem* größten gemeinsamen Teiler und nicht von *dem* größten gemeinsamen Teiler? Schlicht und ergreifend, weil unsere Definition ihn nicht eindeutig bestimmt! In obigem Beispiel $a = 12, b = 30 \in \mathbb{Z}$ ist zweifellos $g = 6$

ein größter gemeinsamer Teiler von a und b. Aber auch -6 teilt a und b und wird von jedem anderen gemeinsamen Teiler von a und b geteilt. In $\mathbb{Z}$ ist nach unserer Definition ein gemeinsamer Teiler nur bis auf sein Vorzeichen (d.h. bis auf Multiplikation mit einer Einheit in $\mathbb{Z}$) eindeutig bestimmt.

In $\mathbb{Q}[t]$ wird das noch schlimmer. Betrachten wir zwei konstante Polynome $0 \neq a, g \in \mathbb{Q} \subset \mathbb{Q}[t]$, so ist

$$a = g \cdot \frac{a}{g}$$

und somit ist g ein Teiler von a. Zudem gilt für einen Teiler $c \in \mathbb{Q}[t]$ von a, daß es ein $d \in \mathbb{Q}[t]$ gibt mit $a = c \cdot d$, und aus der Gradformel $0 = \deg(a) = \deg(c) + \deg(d)$ folgt dann notwendig, daß $\deg(c) = 0$ und $c \in \mathbb{Q} \setminus \{0\}$. D.h. die Teiler von a sind genau die Elemente aus $\mathbb{Q} \setminus \{0\}$. Betrachten wir in $\mathbb{Q}[t]$ also etwa die konstanten Polynome $a = 2$ und $b = 5$, so sind die rationalen Zahlen $0 \neq q \in \mathbb{Q}$ genau die gemeinsamen Teiler von a und b und, da sie sich gegenseitig teilen, sind sie alle auch größte gemeinsame Teiler von a und b im Sinne der Definition. Da man in diesem Fall von einem größten gemeinsamen Teiler q zu einem anderen p durch Multiplikation mit der rationalen Zahl $\frac{p}{q}$ gelangt, könnte man auch sagen, daß der größte gemeinsamer Teiler nur bis auf Multiplikation mit einer rationalen Zahl ungleich Null bestimmt ist.

In den ganzen Zahlen hat man sich angewöhnt, einen der beiden größten gemeinsamen Teiler zu bevorzugen, nämlich den positiven. Für einen beliebigen Integritätsbereich gibt es dazu aber keinen einsichtigen Grund, so daß für uns jedes Element der Menge $\mathrm{ggT}(a, b)$ gleich gut ist.

Die Betrachtungen zum größten gemeinsamen Teiler lassen sich auch auf ein kleinstes gemeinsames Vielfaches k von a und b übertragen. Bedingung (1) bedeutet, daß k ein Vielfaches sowohl von a als auch von b ist, und (2) rechtfertigt den Zusatz *kleinstes*, da jedes andere Vielfache von a und b von k geteilt wird. In $\mathbb{Z}$ ist ein kleinstes gemeinsames Vielfaches wieder nur bis auf sein Vorzeichen bestimmt. □

Beispiel A6.6.

a. Für $f = t - 1 \in \mathbb{Q}[t]$ und $g = t^n - 1 \in \mathbb{Q}[t]$ mit $n \geq 1$ gilt

$$g = f \cdot \left(t^{n-1} + t^{n-2} + \ldots + t + 1\right)$$

und somit $f \mid g$.

b. Betrachten wir die komplexen Zahlen $a = 9$, $b = 2 + \sqrt{-5}$, $c = 2 - \sqrt{-5}$ und $d = 3$ in $\mathbb{Z}\left[\sqrt{-5}\right]$. Wegen

$$a = 9 = \left(2 + \sqrt{-5}\right) \cdot \left(2 - \sqrt{-5}\right) = b \cdot c$$

gilt $b \mid a$.

Wir wollen nun noch zeigen, daß d kein Teiler von b ist. Nehmen wir das Gegenteil an, d.h. $d \mid b$. Dann gibt es ein $e = x + y \cdot \sqrt{-5}$ mit $x, y \in \mathbb{Z}$, so daß

$$b = d \cdot e.$$

Mit Hilfe des Absolutbetrags komplexer Zahlen erhalten wir damit dann aber

$$9 = |b|^2 = |d|^2 \cdot |e|^2 = 9 \cdot \left(x^2 + 5 \cdot y^2\right).$$

Es folgt $x^2 + 5y^2 = 1$, und da x und y ganze Zahlen sind, muß notwendig $y = 0$ und $x \in \{1, -1\}$ gelten. Aber $b \notin \{d, -d\}$, so daß wir einen Widerspruch hergeleitet haben.

c. In $\mathbb{Z}$ gilt $\mathrm{ggT}(6, 8) = \{-2, 2\}$ und $\mathrm{kgV}(6, 8) = \{-24, 24\}$.

Viele Eigenschaften eines Elementes a in einem Integritätsbereich können ausgedrückt werden durch Eigenschaften des von a erzeugten Ideals $\langle a \rangle_R = \{r \cdot a \mid r \in R\}$. Manche Formulierung und manches Argument wird dadurch deutlich verkürzt.

Lemma A6.7.
Sei R ein Integritätsbereich und $a, b, g, k \in R$.

a. $b \mid a$ genau dann, wenn $\langle a \rangle_R \subseteq \langle b \rangle_R$.

b. Die folgenden Aussagen sind gleichwertig:
   (i) $a \mid b$ und $b \mid a$.
   (ii) $\langle a \rangle_R = \langle b \rangle_R$.
   (iii) Es gibt eine Einheit $u \in R^*$ mit $a = u \cdot b$.

c. Genau dann ist $g \in \mathrm{ggT}(a, b)$, wenn die folgenden beiden Eigenschaften erfüllt sind:
   (1) $\langle a, b \rangle_R \subseteq \langle g \rangle_R$.
   (2) Für alle $h \in R$ mit $\langle a, b \rangle_R \subseteq \langle h \rangle_R$, gilt $\langle g \rangle_R \subseteq \langle h \rangle_R$.

d. Genau dann ist $k \in \mathrm{kgV}(a, b)$, wenn die folgenden beiden Eigenschaften erfüllt sind:
   (1) $\langle k \rangle_R \subseteq \langle a \rangle_R \cap \langle b \rangle_R$.
   (2) Für alle $l \in R$ mit $\langle l \rangle_R \subseteq \langle a \rangle_R \cap \langle b \rangle_R$, gilt $\langle l \rangle_R \subseteq \langle k \rangle_R$.

Beweis: a. Falls $b \mid a$, so gibt es ein $c \in R$ mit $a = b \cdot c$. Mithin gilt für jedes $r \in R$ auch $r \cdot a = (r \cdot c) \cdot b \in \langle b \rangle_R$ und damit $\langle a \rangle_R \subseteq \langle b \rangle_R$. Gilt umgekehrt $\langle a \rangle_R \subseteq \langle b \rangle_R$, so ist $a \in \langle a \rangle_R \subseteq \langle b \rangle_R$ und mithin gibt es ein $c \in R$ mit $a = c \cdot b$. Also wird a von b geteilt.

b. Zunächst können wir ohne Einschränkung annehmen, daß $a \neq 0 \neq b$, da die drei Aussagen sonst offenbar gleichwertig sind. Setzen wir (i) voraus, so folgt (ii) aus

Teil a.. Es gelte nun (ii), so daß $a \in \langle b \rangle_R$ und $b \in \langle a \rangle_R$. Es gibt also $u, v \in R$ mit $a = u \cdot b$ und $b = v \cdot a$. Aber dann gilt

$$1 \cdot a = a = u \cdot b = (u \cdot v) \cdot a.$$

Da im Integritätsbereich R die Kürzungsregel gilt und da $a \neq 0$, folgt $1 = u \cdot v$. Da zudem R kommutativ ist, ist $u \in R^*$ eine Einheit und nach Wahl von u gilt $a = u \cdot b$, so daß (iii) erfüllt ist. Setzen wir (iii) voraus, so gilt $a = u \cdot b$ und $b = u^{-1} \cdot a$. Damit gilt aber $b \mid a$ und $a \mid b$, so daß (i) erfüllt ist.

c. Dies ist nur eine Umformulierung der Definition mit Hilfe von Teil a., wenn man beachtet, daß $\langle a, b \rangle_R \subseteq \langle g \rangle_R$ genau dann, wenn $\langle a \rangle_R \subseteq \langle g \rangle_R$ und $\langle b \rangle_R \subseteq \langle g \rangle_R$.

d. Dies ist nur eine Umformulierung der Definition mit Hilfe von Teil a..

□

Eine Verallgemeinerung der Betrachtungen zum ggT und zum kgV in $\mathbb{Z}$ und $\mathbb{Q}[t]$ ist das folgende Lemma. Zwei größte gemeinsame Teiler unterscheiden sich nur durch eine Einheit, und das gleiche gilt für zwei kleinste gemeinsame Vielfache.

Lemma A6.8.
Sei R ein Integritätsbereich, $a, b \in R$.

a. Ist $g \in \mathrm{ggT}(a, b)$, dann ist $\mathrm{ggT}(a, b) = \{u \cdot g \mid u \in R^*\}$, d.h. ein größter gemeinsamer Teiler ist bis auf Multiplikation mit Einheiten eindeutig bestimmt.

b. Ist $k \in \mathrm{kgV}(a, b)$, dann ist $\mathrm{kgV}(a, b) = \{u \cdot k \mid u \in R^*\}$, d.h. ein kleinstes gemeinsames Vielfaches ist bis auf Multiplikation mit Einheiten eindeutig bestimmt.

Beweis: Der Beweis sei dem Leser als Übungsaufgabe überlassen.

□

Aufgabe A6.9.
Beweise Lemma A6.8.

Aufgabe A6.10.

a. Bestimme die Nullteiler und die Einheiten in $\mathbb{Z}_{24}$. Ist $\mathbb{Z}_{24}$ ein Integritätsbereich?

b. Ist $3 + 4i$ ein Teiler von $7 + i$ in $\mathbb{Z}[i]$?

c. Bestimme alle größten gemeinsamen Teiler von $f = t^2 - 3t + 2$ und $g = t^3 - 2t^2 - t + 2$ in $\mathbb{Z}[t]$.

Die nächste Aufgabe zeigt, daß man in den ganzen Zahlen die Bestandteile *größter* beim größten gemeinsamen Teiler bzw. *kleinster* beim kleinsten gemeinsamen Vielfachen auch mittels der Ordnungsrelation in $\mathbb{Z}$ definieren kann.

Aufgabe A6.11.
Es seien $a, b \in \mathbb{Z}$ zwei ganze Zahlen. In Notation A3.16 haben wir die Zahl

$$\operatorname{kgv}(a,b) := \begin{cases} \min\{z > 0 \mid a \text{ und } b \text{ teilen } z\}, & \text{falls } a, b \neq 0, \\ 0, & \text{falls } a = 0 \text{ oder } b = 0, \end{cases}$$

eingeführt und ergänzen sie nun um die Zahl

$$\operatorname{ggt}(a,b) := \begin{cases} \max\{z > 0 \mid z \text{ teilt sowohl } a \text{ als auch } b\}, & \text{falls } (a,b) \neq (0,0), \\ 0, & \text{sonst.} \end{cases}$$

Zeige, $\operatorname{ggt}(a,b) \in \operatorname{ggT}(a,b)$ und $\operatorname{kgv}(a,b) \in \operatorname{kgV}(a,b)$.

Aufgabe A6.12.
Es sei R ein kommutativer Ring mit Eins, der nur endlich viele Elemente enthält. Zeige, dann ist jedes Element von R entweder eine Einheit oder ein Nullteiler.

Aufgabe A6.13.
Es sei R ein Integritätsbereich. Wir definieren eine Äquivalenzrelation auf $R \times (R \setminus \{0\})$ durch

$$(a,b) \sim (a',b') \quad :\Longleftrightarrow \quad a \cdot b' = a' \cdot b.$$

Die Äquivalenzklasse von (a,b) bezeichnen wir mit $\frac{a}{b}$, und die Menge aller Äquivalenzklassen bezeichnen wir mit $\operatorname{Quot}(R)$. Auf dieser Menge definieren wir eine Addition und eine Multiplikation durch

$$\frac{a}{b} + \frac{c}{d} := \frac{ad+bc}{bd} \quad \text{und} \quad \frac{a}{b} \cdot \frac{c}{d} := \frac{ac}{bd}.$$

Zeige:

a. $\sim$ ist eine Äquivalenzrelation.

b. Die Addition und die Multiplikation sind wohldefiniert.

c. $\big(\operatorname{Quot}(R), +, \cdot\big)$ ist ein Körper, der sogenannte *Quotientenkörper* von R.

B) Faktorielle Ringe

Wir haben bislang weder eine Aussage dazu getroffen, ob ein größter gemeinsamer Teiler zweier Elemente in einem Integritätsbereich stets existiert, noch wie man diesen ggf. ausrechnen kann. In der Tat werden wir später sehen, daß es Integritätsbereiche gibt, in denen größte gemeinsame Teiler nicht notwendig existieren. Wir wollen uns aber

zunächst dem Problem zuwenden, wie man einen größten gemeinsamen Teiler denn berechnen könnte.

Die in der Schule gängigste Methode zur Berechnung eines größten gemeinsamen Teilers zweier positiver ganzer Zahlen a und b besteht darin, diese als ein Produkt von Primzahlen zu schreiben und festzustellen, welche Primzahlen mit welchen Vielfachheiten sowohl in a, als auch in b vorkommen. Wenn wir dieses Vorgehen adaptieren wollen, müssen wir zunächst den Begriff der *Primzahl* auf beliebige Integritätsbereiche erweitern. Dazu sollten wir uns charakterisierende Eigenschaften des Begriffs Primzahl anschauen. Eine *Primzahl* ist eine *positive* ganze Zahl, die genau *zwei* positive Teiler besitzt. Dies kann man auch etwas anders ausdrücken als, $p \in \mathbb{Z}_{>1}$ ist eine Primzahl genau dann, wenn für $a, b \in \mathbb{Z}_{\geq 0}$ gilt:

$$p = a \cdot b \implies a = 1 \text{ oder } b = 1. \tag{132}$$

Denn $p = a \cdot b$ bedeutet, daß sowohl a als auch b Teiler von p sind, und für eine Primzahl können nicht beide gleich p sein.

Es gibt aber noch eine andere Eigenschaft, die Primzahlen charakterisiert, eine Eigenschaft, die man verwendet, wenn man den größten gemeinsamen Teiler auf obigem Weg ausrechnet. Wenn nämlich eine Primzahl ein Produkt teilt, so teilt sie schon einen der Faktoren. D.h. $p \in \mathbb{Z}_{>1}$ ist eine Primzahl genau dann, wenn für $a, b \in \mathbb{Z}_{\geq 0}$ gilt:

$$p \mid a \cdot b \implies p \mid a \text{ oder } p \mid b. \tag{133}$$

Den Beweis der Gleichwertigkeit der Eigenschaften (132) und (133) liefert Korollar A6.18.

Wir haben mithin zwei Möglichkeiten, den Begriff der Primzahl auf beliebige Integritätsbereiche zu verallgemeinern, und wir werden sehen, daß diese beiden Begriffe nicht notwendig übereinstimmen. Eine vernünftige Theorie der Teilbarkeit erhalten wir aber nur, wenn die beiden Begriffe übereinstimmen, denn nur dann läßt sich die Primfaktorzerlegung, wie wir sie aus den ganzen Zahlen gewohnt sind, verallgemeinern.

Ein Problem bei der Verallgemeinerung der obigen Bedingungen auf beliebige Integritätsbereiche ist das Fehlen einer Ordnungsrelation $>$, die es uns erlauben würde von *positiven* Ringelementen zu sprechen. Das erweist sich bei näherem Hinsehen jedoch als überflüssig, wenn wir in $\mathbb{Z}$ auch negative Zahlen zulassen. Die Bedingung "$= 1$" bzw. "$\neq 1$" kann man dann durch "$\in \mathbb{Z}^*$" bzw. "$\notin \mathbb{Z}^*$" ersetzen.

Definition A6.14.
Sei R ein Integritätsbereich.

a. Ein Element $0 \neq p \in R \setminus R^*$ heißt *irreduzibel*, falls aus $p = a \cdot b$ mit $a, b \in R$ folgt, daß $a \in R^*$ oder $b \in R^*$.

b. Ein Element $0 \neq p \in R \setminus R^*$ heißt *prim*, falls aus $p \mid a \cdot b$ mit $a, b \in R$ folgt, daß $p \mid a$ oder $p \mid b$.

c. R heißt ein *faktorieller* Ring[11] falls jedes $0 \neq a \in R \setminus R^*$ sich als Produkt von endlich vielen Primelementen schreiben läßt.

,

Wir werden weiter unten sehen, daß in einem faktoriellen Ring die Zerlegung in ein Produkt von Primelementen im wesentlichen eindeutig ist.

Beispiel A6.15.

a. Wir unterscheiden in $\mathbb{Z}$ zwischen *Primzahlen*, welche nach Definition positiv sind, und *Primelementen*, welche auch negativ sein können. Aufgrund der obigen Vorbetrachtungen ist eine ganze Zahl z genau dann prim, wenn sie irreduzibel ist, und das ist genau dann der Fall, wenn z oder $-z$ eine Primzahl ist. Wie angedeutet, beweisen wir diese Tatsache erst weiter unten in Korollar A6.18.

b. Ist K ein Körper und $f \in K[t]$ mit $\deg(f) = 1$, so ist f irreduzibel.
Denn $f = g \cdot h$ mit $g, h \in K[t]$ impliziert $1 = \deg(f) = \deg(g) + \deg(h)$. Mithin gilt entweder $\deg(g) = 0$ und $\deg(h) = 1$ oder es gilt $\deg(g) = 1$ und $\deg(h) = 0$. In ersterem Fall ist $g \in K \setminus \{0\} = K^* = K[t]^*$, in letzterem ist $h \in K \setminus \{0\} = K^* = K[t]^*$, wobei die Gleichheit $K \setminus \{0\} = K^*$ gilt, da K ein Körper ist.

c. Das Polynom $f = 2t + 2 \in \mathbb{Z}[t]$ ist nicht irreduzibel, da $f = 2 \cdot (t + 1)$ und weder 2 noch $t + 1$ ist eine Einheit in $\mathbb{Z}[t]$, da $\mathbb{Z}[t]^* = \mathbb{Z}^* = \{1, -1\}$.

d. Ist R ein Integritätsbereich und sind $p, q \in R$ irreduzibel mit $p \mid q$, dann ist $\langle p \rangle_R = \langle q \rangle_R$, d.h. die beiden unterscheiden sich nur um einen Einheit.
Denn $p \mid q$ bedeutet, es gibt ein $c \in R$ mit $q = p \cdot c$. Da q irreduzibel ist und p keine Einheit, muß notwendig c eine Einheit sein. Also unterscheiden sich p und q nur um eine Einheit.

Wir wollen im folgenden den Zusammenhang der Begriffe *prim* und *irreduzibel* untersuchen, und dabei unter anderem zeigen, daß diese im Ring der ganzen Zahlen übereinstimmen.

Lemma A6.16.
Ist R ein Integritätsbereich und $p \in R$ prim, so ist p irreduzibel.

[11] In der Literatur werden faktorielle Ringe auch *ZPE-Ringe* genannt, wobei ZPE für *eindeutige Primfaktorzerlegung* steht.

Beweis: Seien $a, b \in R$ gegeben mit $p = a \cdot b$, dann gilt insbesondere $p \mid a \cdot b$. Mithin gilt $p \mid a$ oder $p \mid b$. In ersterem Fall gibt es ein $c \in R$ mit $a = p \cdot c$ und somit gilt

$$p \cdot 1 = p = a \cdot b = p \cdot c \cdot b.$$

Da im Integritätsbereich R die Kürzungsregel gilt, folgt unmittelbar, daß $1 = c \cdot b$ und b eine Einheit ist. Analog folgt aus $p \mid b$, daß $a \in R^*$. Somit ist p irreduzibel. □

Beispiel A6.17.

a. Wir wollen nun ein Beispiel dafür geben, daß ein irreduzibles Element nicht notwendig prim sein muß.
Dazu betrachten wieder die komplexen Zahlen $a = 9$, $b = 2 + \sqrt{-5}$, $c = 2 - \sqrt{-5}$ und $d = 3$ in $\mathbb{Z}\left[\sqrt{-5}\right]$. Wir haben bereits in Beispiel A6.6 gesehen, daß d kein Teiler von b ist. Analog zeigt man, daß d kein Teiler von c ist. Aber $d = 3$ ist ein Teiler von $d^2 = a = b \cdot c$. Mithin ist d *nicht prim*, da es das Produkt $b \cdot c$, aber keinen der Faktoren teilt.
Sei nun $d = f \cdot g$ mit $f = x + y \cdot \sqrt{-5}$ und $g = u + v \cdot \sqrt{-5}$, $x, y, u, v \in \mathbb{Z}$. Dann gilt
$$9 = |d|^2 = |f|^2 \cdot |g|^2 = \left(x^2 + 5y^2\right) \cdot \left(u^2 + 5v^2\right)$$
mit $x^2 + 5y^2, u^2 + 5v^2 \in \mathbb{N}$. Es folgt, daß $(x^2 + 5y^2, u^2 + 5v^2) \in \{(9,1), (1,9)\}$. In ersterem Fall muß $u \in \{1, -1\}$ und $v = 0$ sein, so daß g eine Einheit in $\mathbb{Z}\left[\sqrt{-5}\right]$ ist. In letzterem Fall muß $x \in \{1, -1\}$ und $y = 0$ sein, so daß f eine Einheit in $\mathbb{Z}\left[\sqrt{-5}\right]$ ist. Es folgt, daß d *irreduzibel* ist.

b. Ist R faktoriell, so ist jedes irreduzible Element prim.
Denn falls $p \in R$ irreduzibel ist, so ist p nach Voraussetzung ein Produkt $p = q_1 \cdots q_k$ von Primelementen. Nehmen wir $k \geq 2$ an. Da q_k prim ist, ist es keine Einheit und mithin muß $q_1 \cdots q_{k-1}$ eine Einheit sein. Es gibt also ein $a \in R$ mit $1 = q_1 \cdot (q_2 \cdots q_{k-1} \cdot a)$, so daß dann q_1 eine Einheit ist, im Widerspruch zur Voraussetzung q_1 prim. Mithin ist $k = 1$ und $p = q_1$ ist prim.

c. $\mathbb{Z}\left[\sqrt{-5}\right]$ ist nicht faktoriell, da 3 irreduzibel aber nicht prim ist.

Wir haben bislang lediglich einen Körper mit nur endlich vielen Elementen kennen gelernt, $\mathbb{Z}_2$. Für Anwendungen in der Kryptographie oder der Kodierungstheorie sind endliche Körper aber von weit größerem Interesse als unendliche. Wir werden nun zeigen, daß das Beispiel $\mathbb{Z}_2$ verallgemeinert werden kann.

Korollar A6.18.
Für $0 \neq n \in \mathbb{Z}$ sind die folgenden Aussagen gleichwertig:

a. $\mathbb{Z}_n$ ist ein Körper.

b. $\mathbb{Z}_n$ ist ein Integritätsbereich.

c. n ist prim.

d. n ist irreduzibel, d.h. n ist eine Primzahl.

Beweis: **a.** $\Rightarrow$ **b.:** Ist $\mathbb{Z}_n$ ein Körper, so ist $\mathbb{Z}_n$ ein Integritätsbereich nach Beispiel A6.2.

b. $\Rightarrow$ **c.:** Gilt $n \mid a \cdot b$ mit $a, b \in \mathbb{Z}$, dann ist

$$\overline{a} \cdot \overline{b} = \overline{a \cdot b} = \overline{0} \in \mathbb{Z}_n.$$

Da nach Voraussetzung $\mathbb{Z}_n$ nullteilerfrei ist, muß $\overline{a} = \overline{0}$ oder $\overline{b} = \overline{0}$ gelten. In ersterem Fall gilt $n \mid a$, in letzterem $n \mid b$. Also ist n prim in $\mathbb{Z}$.

c. $\Rightarrow$ **d.:** Dies folgt aus Lemma A6.16.

d. $\Rightarrow$ **a.:** Ist I ein Ideal in $\mathbb{Z}_n$, so ist $(I, +)$ eine Untergruppe von $(\mathbb{Z}_n, +)$ und die Ordnung von I ist nach dem Satz von Lagrange ein Teiler der Primzahl $n = |\mathbb{Z}_n|$. Also muß $|I|$ entweder 1 oder n sein, d.h. $I = \{\overline{0}\}$ oder $I = \mathbb{Z}_n$. Da zudem $\{\overline{0}\} \neq \mathbb{Z}_n$, hat $\mathbb{Z}_n$ also genau zwei Ideale und ist nach Aufgabe A5.35 ein Körper.

□

Bemerkung A6.19.

Ist R ein faktorieller Ring und $0 \neq a \in R \setminus R^*$, dann ist die Darstellung $a = p_1 \cdots p_r$ als Produkt von Primelementen *im wesentlichen eindeutig*, d.h. sind

$$p_1 \cdot \ldots \cdot p_r = q_1 \cdot \ldots \cdot q_s \tag{134}$$

zwei solche Darstellungen, so gilt $r = s$, und nach Umordnen der q_i unterscheiden sich p_i und q_i nur noch um eine Einheit, d.h. $\langle p_i \rangle_R = \langle q_i \rangle_R$. Dies ist leicht einzusehen: da p_1 prim und ein Teiler der rechten Seite von (134) ist, muss p_1 eines der q_i teilen. Nach Umordnen der q_i können wir $p_1 \mid q_1$ annehmen. Da beide prim sind, sind sie nach Lemma A6.16 auch beide irreduzibel und unterscheiden sich nach Beispiel A6.15 nur um eine Einheit. Nun können wir p_1 aus (134) kürzen und induktiv das gleiche Verfahren auf das verbleibende Produkt anwenden. □

Bemerkung A6.20.

Es sei R ein faktorieller Ring und $a = u \cdot p_1^{m_1} \cdots p_r^{m_r}$ und $b = v \cdot p_1^{n_1} \cdots p_r^{n_r}$ seien Elemente in $R \setminus \{0\}$ mit $u, v \in R^*$ Einheiten, $p_1, \ldots, p_r$ prim, $\langle p_i \rangle_R \neq \langle p_j \rangle_R$ für $i \neq j$ und $m_1, \ldots, m_r, n_1, \ldots, n_r \in \mathbb{N}$. Dann sieht man wie für die ganzen Zahlen, daß

$$p_1^{\min\{m_1, n_1\}} \cdots p_r^{\min\{m_r, n_r\}} \in \mathrm{ggT}(a, b)$$

und

$$p_1^{\max\{m_1, n_1\}} \cdots p_r^{\max\{m_r, n_r\}} \in \mathrm{kgV}(a, b).$$

Man nennt eine Darstellung von a wie oben auch eine *Primfaktorzerlegung* von a, wenn $m_i > 0$ für alle $i = 1, \ldots, r$. Nach Bemerkung A6.19 ist sie bis auf die Reihenfolge der Faktoren und Multiplikation mit Einheiten eindeutig bestimmt. □

Aufgabe A6.21.
Es sei R ein Integritätsbereich und es gebe eine natürliche Zahl $n \geq 1$ so, daß $n \cdot 1_R = \sum_{k=1}^{n} 1_R = 0_R$, d.h. die n-fache Summe des Einselementes ergibt das Nullelement. Zeige, daß die kleinste positive ganze Zahl $p = \min\{m \in \mathbb{Z}_{>0} \mid m \cdot 1_R = 0_R\}$ mit dieser Eigenschaft irreduzibel (d.h. eine Primzahl) ist. Man nennt diese Zahl p auch die *Charakteristik* des Ringes.

Aufgabe A6.22.
Ist $p = x + y \cdot i \in \mathbb{Z}[i]$ so, daß $q := |p|^2 = x^2 + y^2$ eine Primzahl ist, dann ist p ein Primelement in $\mathbb{Z}[i]$. Finde zudem ein Beispiel für eine solche Zahl p.

Aufgabe A6.23.
Bestimme alle Polynome f in $\mathbb{Z}_2[t]$ vom Grad 4, deren Leitkoeffizient $\mathrm{lc}(f)$ und deren konstanter Koeffizient $f(0)$ beide $\overline{1}$ sind. Welche dieser Polynome sind irreduzibel?

Aufgabe A6.24.

a. Es sei $f \in \mathbb{Z}[t]$ mit $\mathrm{lc}(f) = 1$. Zeige, falls es eine Primzahl $p \in \mathbb{Z}$ gibt, so daß die Reduktion $\phi_p(f)$ von f modulo p irreduzibel in $\mathbb{Z}_p[t]$ ist (siehe Aufgabe A5.40), so ist f irreduzibel in $\mathbb{Z}[t]$.

b. Bestimme alle Polynome f in $\mathbb{Z}_2[t]$ vom Grad $0 \leq \deg(f) \leq 4$ und schreibe sie jeweils als Produkt von möglichst vielen Polynomen vom Grad größer oder gleich 1.

c. Ist $f = t^4 + 187t^3 + 5t^2 - 33t + 3001 \in \mathbb{Z}[t]$ irreduzibel?

C) Euklidische Ringe

Faktorielle Ringe verallgemeinern die ganzen Zahlen und wie in Bemerkung A6.20 gesehen, ist die *eindeutige Primfaktorzerlegung* eines Elements sehr nützlich. Allerdings wissen wir bislang von keinem anderen Ring als den ganzen Zahlen, daß er faktoriell ist. Uns fehlt ein gutes Kriterium, dies zu entscheiden, ein Kriterium, das einfacher zu Handhaben ist, als für jedes Element eine Primfaktorzerlegung anzugeben.

Stellen wir dieses Problem einmal hintan und nehmen an, wir wüßten von einem Ring bereits, daß er faktoriell ist. Unser Ausgangspunkt war, aus der Kenntnis von Primfaktorzerlegungen einen größten gemeinsamen Teiler zu bestimmen. Um dies praktisch umzusetzen, fehlt uns mithin noch ein Verfahren, das es uns erlaubt, die Primfaktorzerlegung eines Elementes tatsächlich auszurechnen. Und obwohl dies in

der Schulzeit bei den ganzen Zahlen *das* Verfahren war, um den größten gemeinsamen Teiler zweier ganzer Zahlen zu bestimmen, wage ich zu bezweifeln, daß Ihr es auf die folgenden beiden Zahlen anwenden wollt:

$$a = 12345678909876543212345678909876543 21$$

und

$$b = 27283950390827160499283950390827065.$$

Selbst mit einem Taschenrechner halte ich das Unterfangen für aussichtslos, und für Anwendungen in der Kryptographie sind diese beiden Zahlen nahezu winzig. Dort sind Zahlen mit 500 und mehr Ziffern notwendig, und die Sicherheit neuerer kryptographischer Verfahren beruht auf der Tatsache, daß es sehr schwer ist, eine Zahl in ihre Primfaktoren zu zerlegen.

Es gibt aber ein sehr einfaches und effizientes Verfahren, um den größten gemeinsamen Teiler zu bestimmen, ohne die Primfaktorzerlegung zu kennen. Dieses Verfahren wollen wir im folgenden beschreiben. Es funktioniert nicht nur für die ganzen Zahlen, sondern für jeden Integritätsbereich, in dem man eine *Division mit Rest* hat.

Definition A6.25.
Ein Integritätsbereich R heißt ein *euklidischer Ring*, wenn es eine Funktion

$$\nu : R \setminus \{0\} \longrightarrow \mathbb{N}$$

gibt, so daß es für alle $a, b \in R \setminus \{0\}$ eine *Division mit Rest* der Form

$$a = q \cdot b + r$$

mit $q, r \in R$ gibt, wobei entweder $r = 0$ oder $\nu(r) < \nu(b)$. Wir nennen ν dann eine *euklidische Funktion* von R.

Beispiel A6.26.
$\mathbb{Z}$ ist mittels $\nu : \mathbb{Z} \setminus \{0\} \longrightarrow \mathbb{N} : z \mapsto |z|$ und der wohlbekannten Division mit Rest ein euklidischer Ring.

Bevor wir zeigen, wie uns die Division mit Rest hilft, einen größten gemeinsamen Teiler zu bestimmen, wollen wir zeigen, daß es außer den ganzen Zahlen noch andere euklidische Ringe gibt. Das vielleicht wichtigste Beispiel neben $\mathbb{Z}$ sind die Polynomringe über Körpern.

Proposition A6.27 (Division mit Rest im Polynomring).
Ist R ein kommutativer Ring mit Eins und sind $0 \neq f, g \in R[t]$ mit $\mathrm{lc}(f) \in R^*$, dann gibt es Polynome $q, r \in R[t]$, so daß

$$g = q \cdot f + r \quad \text{und} \quad \deg(r) < \deg(f).$$

Dabei sind q und r eindeutig bestimmt.

Beweis: Seien $f = \sum_{i=0}^{n} a_i \cdot t^i$ und $g = \sum_{i=0}^{m} b_i \cdot t^i$ mit $m = \deg(g)$, $n = \deg(f)$ und $a_n \in R^*$ eine Einheit. Wir führen den Beweis der Existenz einer solchen Division mit Rest mittels Induktion nach m.

Falls $m = n = 0$, so sind wir fertig mit $q = \frac{b_0}{a_0}$ und $r = 0$, und falls $0 \leq m < n$, so können wir $q = 0$ und $r = g$ wählen. Diese Fälle schließen den Induktionsanfang $m = 0$ ein.

Es reicht nun, den Fall $m > 0$ und $n \leq m$ zu betrachten. Definieren wir

$$g' := g - \frac{b_m}{a_n} \cdot t^{m-n} \cdot f.$$

Dann heben sich in der Differenz die Leitterme auf, so daß $\deg(g') < \deg(g) = m$ gilt. Folglich existieren nach Induktionsannahme Polynome $q', r' \in R[t]$, so daß

$$q' \cdot f + r' = g' = g - \frac{b_m}{a_n} \cdot t^{m-n} \cdot f$$

und $\deg(r') < \deg(f)$. Also

$$g = \left(q' + \frac{b_m}{a_n} \cdot t^{m-n}\right) \cdot f + r',$$

und wir sind fertig mit $q = q' + \frac{b_m}{a_n} \cdot t^{m-n}$ und $r = r'$.

Es bleibt nur noch die Eindeutigkeit der Zerlegung zu zeigen. Nehmen wir dazu an, daß

$$g = q \cdot f + r = q' \cdot f + r'$$

mit $q, q', r, r' \in R[t]$ und $\deg(r), \deg(r') < \deg(f)$. Dann gilt

$$(q - q') \cdot f = r' - r$$

und da der Leitkoeffizient von f als Einheit kein Nullteiler ist, liefert die Gradformel

$$\deg(q - q') + \deg(f) = \deg(r' - r) \leq \max\{\deg(r), \deg(r')\} < \deg(f).$$

Das ist aber nur möglich, wenn $q - q' = 0$. Also gilt $q = q'$ und dann auch $r = r'$. □

Da in einem Körper jedes Element ungleich Null eine Einheit ist, erhalten somit unmittelbar folgendes Korollar.

Korollar A6.28.
Ist K ein Körper, so ist $K[t]$ ein euklidischer Ring mit deg als euklidischer Funktion.

Der Beweis von Proposition A6.27 ist konstruktiv, d.h. er liefert uns ein Verfahren, wie wir die Division mit Rest im Polynomring durchführen können.

Beispiel A6.29.
Seien $f = t^3 + t + 1, g = t - 1 \in \mathbb{Q}[t]$ gegeben. Wir führen Polynomdivision durch

$$\begin{array}{rl} (t^3 \quad + \quad t + 1) : (t-1) = t^2 + t + 2 + \frac{r}{t-1} \\ \underline{t^3 - t^2} \\ \quad t^2 + \quad t \\ \quad \underline{t^2 - \quad t} \\ \quad\quad 2t + \quad 1 \\ \quad\quad \underline{2t - \quad 2} \\ \quad\quad\quad 3 \quad =: r \end{array}$$

und erhalten $f = (t^2 + t + 2) \cdot g + 3$. □

Nun wollen wir den Euklidischen Algorithmus kennen lernen, der es uns erlaubt, in euklidischen Ringen einen größten gemeinsamen Teiler auszurechnen. Bevor wir den Algorithmus allgemein formulieren und beweisen, wollen wir ihn beispielhaft in den ganzen Zahlen anwenden.

Beispiel A6.30.
Wir wollen einen größten gemeinsamen Teiler der ganzen Zahlen $r_0 = 66$ und $r_1 = 15$ berechnen. Dazu führen wir Division mit Rest durch

$$r_0 = 66 = 4 \cdot 15 + 6 = q_1 \cdot r_1 + r_2$$

und erhalten den Rest $r_2 = 6$. Sodann teilen wir r_1 durch r_2 mit Rest,

$$r_1 = 15 = 2 \cdot 6 + 3 = q_2 \cdot r_2 + r_3,$$

und erhalten den Rest $r_3 = 3$. Dann teilen wir r_2 durch r_3 mit Rest,

$$r_2 = 6 = 2 \cdot 3 + 0 = q_2 \cdot r_3 + r_4,$$

und erhalten den Rest $r_4 = 0$. Das Verfahren bricht ab, da wir r_3 nicht weiter durch $r_4 = 0$ teilen können. Wir erhalten als größten gemeinsamen Teiler von r_0 und r_1

$$r_3 = 3 \in \mathrm{ggT}(66, 15) = \mathrm{ggT}(r_0, r_1),$$

d.h. den letzten Rest der sukkzessiven Division mit Rest, der nicht Null war.

Algorithmus A6.31 (Euklidischer Algorithmus).
Input: R ein euklidischer Ring mit euklidischer Funktion ν sowie $a, b \in R \setminus \{0\}$.
Output: $g \in \mathrm{ggT}(a, b)$ ein größter gemeinsamer Teiler von a und b.

1: Setze $r_0 := a$, $r_1 := b$ und $k := 2$.
2: **while** $r_{k-1} \neq 0$ **do**
3: Wähle $r_k \in R$ und $q_{k-1} \in R$ mit

$$r_{k-2} = q_{k-1} \cdot r_{k-1} + r_k \quad \text{und} \quad (r_k = 0 \text{ oder } \nu(r_k) < \nu(r_{k-1})).$$

4: Setze $k := k + 1$.
5: **end while**
6: **return** $g := r_{k-1}$

Beweis: Solange r_{k-1} nicht Null ist, können wir Division mit Rest durchführen und erhalten auf dem Wege $r_k, q_{k-1} \in R$, so daß die obigen Bedingungen erfüllt sind.

Unsere Konstruktion liefert

$$\begin{aligned} r_0 &= r_1 q_1 + r_2, & \nu(r_2) &< \nu(r_1), \\ r_1 &= r_2 q_2 + r_3, & \nu(r_3) &< \nu(r_2), \\ &\vdots \\ r_{k-2} &= r_{k-1} q_{k-1} + r_k, & \nu(r_k) &< \nu(r_{k-1}), \end{aligned}$$

und damit eine streng monoton fallende Folge natürlicher Zahlen

$$\nu(r_1) > \nu(r_2) > \nu(r_3) > \dots .$$

Da es in den natürlichen Zahlen keine unendlichen streng monoton fallenden Zahlenfolgen gibt, muß das Verfahren abbrechen, d.h. *der Algorithmus terminiert*, und das ist genau dann der Fall, wenn ein $k \geq 2$ gefunden wurde mit $r_k = 0$.

Wenden wir uns nun der *Korrektheit des Algorithmus'* zu. Dazu zeigen wir durch Induktion nach der Anzahl n der Durchläufe der While-Schleife, daß der Algorithmus einen größten gemeinsamen Teiler von $r_0 = a$ und $r_1 = b$ findet, d.h. wenn $r_n \neq 0$ und $r_{n+1} = 0$, dann ist

$$r_n \in \mathrm{ggT}(r_0, r_1).$$

Induktionsanfang: $n = 1$. Dann ist $r_2 = 0$, also $r_1 \mid r_0$ und $r_1 \in \mathrm{ggT}(r_0, r_1)$.

Induktionsschluß: Sei nun $n \geq 2$ und die Behauptung gelte für alle Paare, für die die While-Schleife einen Durchlauf weniger benötigt. Die Betrachtung der letzten $n - 1$ Durchläufe liefert mithin durch Anwendung der Induktionsvoraussetzung auf r_1 und r_2:

$$r_n \in \mathrm{ggT}(r_1, r_2).$$

Insbesondere ist r_n ein Teiler von r_1 und von r_2. Da nach Voraussetzung $r_0 = q_1 \cdot r_1 + r_2$, ist dann aber r_n auch ein Teiler von r_0.

Sei nun $r \in R$ ein weiterer Teiler von r_0 und r_1, dann gilt

$$r \mid (r_0 - q_1 \cdot r_1) = r_2,$$

und mithin ist r ein Teiler sowohl von r_1 als auch von r_2. Aber da $r_n \in \mathrm{ggT}(r_1, r_2)$ gilt dann

$$r \mid r_n,$$

und nach Definition ist deshalb $r_n \in \mathrm{ggT}(r_0, r_1)$. □

Um den Algorithmus durchführen zu können brauchen wir nur die Division mit Rest. Diese können wir auch im Polynomring über einem Körper durchführen, so daß wir größte gemeinsame Teiler auch im Polynomring ausrechnen können.

Beispiel A6.32.
Betrachte $r_0 = t^4 + t^2 \in \mathbb{Q}[t]$ und $r_1 = t^3 - 3t^2 + t - 3 \in \mathbb{Q}[t]$. Division mit Rest liefert im ersten Schritt

$$r_0 = t^4 + t^2 = (t+3) \cdot (t^3 - 3t^2 + t - 3) + (9t^2 + 9) = q_1 \cdot r_1 + r_2$$

mit Rest $r_2 = 9t^2 + 9$. Im nächsten Schritt erhalten wir

$$r_1 = t^3 - 3t^2 + t - 3 = \left(\frac{1}{9} \cdot t - \frac{1}{3}\right) \cdot (9t^2 + 9) + 0 = q_2 \cdot r_2 + r_3$$

mit Rest $r_3 = 0$. Das Verfahren bricht ab und $r_2 = 9t^2 + 9 \in \mathrm{ggT}(r_0, r_1)$ ist ein größter gemeinsamer Teiler. Man kann diesen normieren, indem man das Polynom durch seinen Leitkoeffizienten teilt und erhält dann als normierten größten gemeinsamen Teiler

$$t^2 + 1 \in \mathrm{ggT}\left(t^4 + t^2, t^3 - 3t^2 + t - 3\right).$$

Bemerkung A6.33.
Ist R ein euklidischer Ring und $a, b \in R \setminus \{0\}$, so gilt $g \in \mathrm{ggT}(a, b)$ genau dann wenn $\frac{a \cdot b}{g} \in \mathrm{kgV}(a, b)$. Man kann mit Hilfe des Euklidischen Algorithmus' in einem euklidischen Ring also auch kleinste gemeinsame Vielfache ausrechnen.

Aufgabe A6.34.
Zeige, daß $\mathbb{Z}[i] = \{x + i \cdot y \mid x, y \in \mathbb{Z}\}$ ein euklidischer Ring mit euklidischer Funktion $\nu : \mathbb{Z}[i] \longrightarrow \mathbb{N} : a \mapsto |a|^2$ ist.

Aufgabe A6.35.
Betrachte die Polynome

$$f = t^5 + 3t^4 + 2t^3 + 5t^2 + 7t + 2 \in \mathbb{Z}[t]$$

und

$$g = t^3 + t^2 + t + 1 \in \mathbb{Z}[t].$$

a. Bestimme einen größten gemeinsamen Teiler von f und g in $\mathbb{Q}[t]$ mittels des Euklidischen Algorithmus'.

b. Betrachte die Koeffizienten von f und g modulo 3 und bestimme einen größten gemeinsamen Teiler der resultierenden Polynome $\phi_3(f)$ und $\phi_3(g)$ in $\mathbb{Z}_3[t]$.

D) Der Polynomring

Wir haben im letzten Abschnitt gesehen, wie Polynomdivision funktioniert und daß es sich dabei um eine Division mit Rest handelt, die den Polynomring $K[t]$ über einem Körper K zu einem euklidischen Ring macht. In diesem Abschnitt wollen wir die Division mit Rest ausnutzen, um Nullstellen eines Polynoms als Linearfaktoren abzuspalten. Um von einer Nullstelle eines Polynoms sprechen zu können, müssen wir erlauben, in Polynomen für die Unbestimmte t Werte einzusetzen. Für allgemeine Potenzreihen hatten wir das strikt ausgeschlossen, bei Polynomen können wir es zulassen, da der resultierende Ausdruck nur endlich viele Summanden hat.

Lemma A6.36 (Einsetzhomomorphismus).
Es sei S ein kommutativer Ring mit Eins, R ein Unterring von S und $b \in S$. Die Abbildung

$$\varphi_b : R[t] \longrightarrow S : f \mapsto f(b)$$

ist ein Ringhomomorphismus, wobei

$$f(b) := \sum_{k=0}^{n} a_k \cdot b^k \in R$$

für $f = \sum_{k=0}^{n} a_k \cdot t^k \in R[t]$. Wir nennen φ_b *Einsetzhomomorphismus*, und für ein konstantes Polynom $f = a_0$ gilt $\varphi_b(f) = a_0$. Ist $S = R$, dann ist φ_b auch surjektiv.

Beweis: Seien zwei Polynome $f = \sum_{k=0}^{n} a_k \cdot t^k$ und $g = \sum_{k=0}^{m} b_k \cdot t^k$ in $R[t]$ gegeben. Wir können ohne Einschränkung annehmen, daß $m = n$ gilt. Dann gilt

$$\varphi_b(f+g) = \varphi_b\left(\sum_{k=0}^{n}(a_k+b_k)\cdot t^k\right) = \sum_{k=0}^{n}(a_k+b_k)\cdot b^k$$
$$= \sum_{k=0}^{n} a_k \cdot b^k + \sum_{k=0}^{n} b_k \cdot b^k = \varphi_b(f) + \varphi_b(g)$$

und

$$\varphi_b(f\cdot g) = \varphi_b\left(\sum_{k=0}^{2n}\sum_{i+j=k}(a_i\cdot b_j)\cdot t^k\right) = \sum_{k=0}^{2n}\sum_{i+j=k}(a_i\cdot b_j)\cdot b^k$$
$$= \sum_{k=0}^{n} a_k\cdot b^k \cdot \sum_{k=0}^{n} b_k \cdot b^k = \varphi_b(f)\cdot\varphi_b(g).$$

Außerdem gilt für ein konstantes Polynom $a_0 \cdot t^0$

$$\varphi_b(a_0 \cdot t^0) = a_0 \cdot b^0 = a_0.$$

Damit gilt aber insbesondere $\varphi_b(1) = 1$, und φ_b ist ein Ringhomomorphismus. Für die Surjektivität beachten wir nur, daß für $a \in R$ automatisch $a = \varphi_b(a \cdot t^0) \in \mathrm{Im}(\varphi_b)$. □

Bemerkung A6.37.
Der Umstand, daß φ_b ein Ringhomomorphismus ist, impliziert

$$(f+g)(b) = f(b) + g(b) \quad \text{und} \quad (f \cdot g)(b) = f(b) \cdot g(b).$$

Beachte auch, daß der Beweis von Aufgabe A5.25 a. trotz der etwas allgemeineren Voraussetzungen im wesentlichen identisch ist mit obigem Beweis.

Wir sind nun in der Lage zu definieren, was eine Nullstelle ist.

Definition A6.38.
Sei S ein kommutativer Ring mit Eins, R ein Unterring von S, $f \in R[t]$ ein Polynom und $b \in R$. Wir nennen b eine *Nullstelle* von f in S, falls $f(b) = \varphi_b(f) = 0$.

Proposition A6.39.
Sei R ein kommutativer Ring mit Eins und sei $b \in R$ eine Nullstelle des Polynoms $0 \neq g \in R[t]$ in R, so gibt es ein $q \in R[t]$ mit

$$g = q \cdot (t - b).$$

Wir nennen $t - b$ einen *Linearfaktor* des Polynoms g.

Beweis: Da der Leitkoeffizient von $f = t - b$ eine Einheit ist, liefert Division mit Rest A6.27 die Existenz zweier Polynome $q, r \in R[t]$ mit

$$g = q \cdot f + r$$

und $\deg(r) < \deg(f) = 1$. Aus der Gradbedingung folgt unmittelbar, daß $r = r_0 \cdot t^0$ ein konstantes Polynom ist. Da $\varphi_b(f) = b - b = 0$ und da b eine Nullstelle von g ist, gilt

$$r_0 = \varphi_b(r) = \varphi_b(g - q \cdot f) = \varphi_b(g) - \varphi_b(q) \cdot \varphi_b(f) = 0.$$

Also ist r das Nullpolynom und $g = q \cdot (t - b)$. □

Korollar A6.40.
Ist R ein Integritätsbereich und $0 \neq f \in R[t]$ ein Polynom vom Grad $\deg(f) \geq 2$, das eine Nullstelle in R besitzt, so ist f nicht irreduzibel.

Beweis: Ist $b \in R$ eine Nullstelle von f, so gibt es wegen Proposition A6.39 ein $q \in R[t]$ mit $f = q \cdot (t - b)$. Aus der Gradformel folgt

$$\deg(q) = \deg(f) - 1 \geq 1,$$

so daß die beiden Faktoren q und $t-b$ beides keine Einheiten in $R[t]$ sind. Also ist f nicht irreduzibel. □

Beispiel A6.41.
Wir wollen nun in zwei Beispielen sehen, wie man mit Hilfe von Polynomdivision Linearfaktoren abspalten kann.

a. Sei $f = t^3 + t^2 - 5t - 2 \in \mathbb{Q}[t]$, dann gilt offenbar $f(2) = 8 + 4 - 10 - 2 = 0$. Polynomdivision liefert:

$$\begin{array}{rl} (t^3 + & t^2 - 5t - 2) : (t-2) = t^2 + 3t + 1. \\ \underline{t^3 - 2t^2} & \\ 3t^2 - 5t & \\ \underline{3t^2 - 6t} & \\ t - & 2 \\ \underline{t - } & \underline{2} \\ - & \end{array}$$

Also gilt $f = (t^2 + 3t + 1) \cdot (t-2)$ und f ist nicht irreduzibel.

b. Ist $f = t^5 + t^4 + t^3 + t^2 + t + \overline{1} \in \mathbb{Z}_2[t]$, so gilt

$$f(\overline{1}) = \overline{1} + \overline{1} + \overline{1} + \overline{1} + \overline{1} + \overline{1} = \overline{6} = \overline{0} \in \mathbb{Z}_2.$$

Also ist $\overline{1}$ eine Nullstelle von f und f ist nicht irreduzibel. Wir können den Linearfaktor $t - \overline{1}$ mittels Polynomdivision abspalten. Dabei sollte man beachten, daß $\overline{1} = -\overline{1}$ in $\mathbb{Z}_2$, also $t - \overline{1} = t + \overline{1}$:

$$\begin{array}{rl} (t^5 + t^4 + t^3 + t^2 + t + \overline{1}) : (t + \overline{1}) = t^4 + t^2 + \overline{1} & \\ \underline{t^5 + t^4} & \\ t^3 + t^2 + t + & \overline{1} \\ \underline{t^3 + t^2} & \\ t + & \overline{1} \\ \underline{t + } & \underline{\overline{1}} \\ - & \end{array}$$

Also gilt $f = (t^4 + t^2 + \overline{1}) \cdot (t + \overline{1})$.

Satz A6.42.
Ist R ein Integritätsbereich, so hat jedes $0 \neq f \in R[t]$ höchstens $\deg(f)$ Nullstellen.

Beweis: Wir überlassen den Beweis dem Leser als Übungsaufgabe.

□

Beispiel A6.43.
Das Polynome $f = t^2 + 1$ hat in $\mathbb{R}$ keine Nullstelle, während es in $\mathbb{C}$ die Nullstellen i und $-i$ hat und sich mithin in $\mathbb{C}[t]$ als Produkt von Linearfaktoren schreiben läßt:

$$f = (t - i) \cdot (t + i).$$

Bemerkung A6.44.
Es sei K ein Körper. Aufgrund der Definition der Einsetzhomomorphismen liefert jedes Polynom $f \in K[t]$ eine Funktion

$$P_f : K \longrightarrow K : b \mapsto f(b),$$

die durch f definierte *Polynomfunktion*. Auf diesem Weg erhalten wir eine Abbildung

$$P : K[t] \longrightarrow K^K : f \mapsto P_f$$

von der Menge der Polynome über K in die Menge der Funktionen von K nach K, die einem Polynom seine Polynomfunktion zuordnet. Aus den Eigenschaften des Einsetzhomomorphismus und der Definition der Ringoperationen in K^K (siehe Beispiel A5.3) folgt

$$P_{f+g}(b) = (f + g)(b) = f(b) + g(b) = P_f(b) + P_g(b) = (P_f + P_g)(b)$$

und

$$P_{f\cdot g}(b) = (f \cdot g)(b) = f(b) \cdot g(b) = P_f(b) \cdot P_g(b) = (P_f \cdot P_g)(b).$$

Damit gilt dann aber $P_{f+g} = P_f + P_g$ und $P_{f\cdot g} = P_f \cdot P_g$. Da zudem P_1 die Einsfunktion, d.h. das Einselement von K^K, ist, ist die Abbildung P ein *Ringhomomorphismus*.

Aus der Schule sind Polynome in aller Regel nur als Polynomfunktionen bekannt, und der obige Ringhomomorphismus erlaubt es uns, unsere Polynome als Polynomfunktionen aufzufassen. Im allgemeinen ist es aber nicht richtig, daß zwei *verschiedene* Polynome auch verschiedene Polynomfunktionen definieren! Sei dazu $K = \mathbb{Z}_2$, $f = t^2 + t \in K[t]$ und $g = \overline{0} \cdot t^0 \in K[t]$. Dann gilt

$$f\left(\overline{1}\right) = \overline{1} + \overline{1} = \overline{0} \quad \text{und} \quad f\left(\overline{0}\right) = \overline{0}.$$

Da $K = \{\overline{0}, \overline{1}\}$ nur die zwei Elemente $\overline{0}$ und $\overline{1}$ enthält ist P_f die Nullfunktion, d.h. $P_f = P_g$, obwohl f nicht das Nullpolynom ist, d.h. $f \neq g$.

Ein Polynom ist im allgemeinen nicht festgelegt durch die Polynomfunktion, die es definiert. Etwas mathematischer ausgedrückt, der Ringhomomorphismus P ist im allgemeinen *nicht injektiv*. Der Ärger in obigem Beispiel rührt daher, daß K nur endlich viele Elemente besitzt und daß es somit ein Polynom ungleich 0 geben kann, daß alle diese Elemente als Nullstelle hat.

Ist hingegen K ein Körper mit unendlich vielen Elemente, so ist P injektiv.

Dazu müssen wir nur zeigen, daß der Kern von P nur das Nullpolynom enthält. Wäre $0 \neq f \in \mathrm{Ker}(P)$, so wäre P_f die Nullfunktion, d.h. jedes Element von K wäre Nullstelle

von f. Aus Satz A6.42 würde dann folgen, daß K höchstens $\deg(f)$ Elemente enthält, was im Widerspruch zur Voraussetzung steht, daß $|K| = \infty$.

Arbeitet man mit unendlichen Körper, wie z.B. $K = \mathbb{R}$ oder $K = \mathbb{C}$, dann ist es zulässig, den Polynomring mit seinem Bild unter P in K^K zu identifizieren, d.h. man kann es sich dann erlauben, nicht zwischen Polynomen und Polynomfunktionen zu unterscheiden. □

Wissen über die Existenz von Nullstellen kann hilfreich sein, um festzustellen, ob ein Polynome irreduzibel ist oder nicht.

Aufgabe A6.45. a. Ist K ein Körper und $f \in K[t]$ ein Polynom mit $deg(f) \in \{2, 3\}$. Zeige, f ist genau dann irreduzibel, wenn f keine Nullstelle hat.

b. Ist $f = t^3 + 3t + 1 \in \mathbb{Z}[t]$ irreduzibel? Falls nicht, schreibe f als Produkt von irreduziblen Polynomen.

c. Ist $f_5 = t^3 + \overline{3} \cdot t + \overline{1} \in \mathbb{Z}_5[t]$ irreduzibel? Falls nicht, schreibe f_5 als Produkt von irreduziblen Polynomen.

Aufgabe A6.46.
Beweise Satz A6.42.

Aufgabe A6.47.
Zeige, $f = t^2 + t + 1 \in \mathbb{Z}_2[t]$ ist irreduzibel und $K = \mathbb{Z}_2[t]/\langle f \rangle$ ist ein Körper mit 4 Elementen. Stelle die Additions- und Multiplikationstabelle für K auf. Was ist die Charakteristik (siehe Aufgabe A6.21) von K? Ist K isomorph zum Ring $\mathbb{Z}_4$? Ist K isomorph zum Ring $\mathbb{Z}_2 \times \mathbb{Z}_2$ mit komponentenweisen Operationen? Betrachten wir den Polynomring $K[x]$ über K in der Unbestimmten x. Ist das Polynom $g = x^2 + x + \overline{1} \in K[x]$ irreduzibel? Hat g eine Nullstelle in K?
Anmerkung, in dieser Aufgabe wollen wir die Elemente $\overline{0}$ und $\overline{1}$ in $\mathbb{Z}_2$ der Einfachheit halber mit 0 und 1 bezeichnen, wobei $1 + 1 = 0$ gilt. Das ist deshalb sinnvoll, weil auch die Elemente von $\mathbb{Z}_2[t]/\langle f \rangle$ wieder Restklassen sind, und die doppelten Restklassen (z.B. $\overline{t + \overline{1}}$) für unnötige Verwirrung sorgen.

E) Hauptidealringe

In den vorigen Abschnitten haben wir gesehen, daß größte gemeinsame Teiler sowohl in faktoriellen, als auch in euklidischen Integritätsbereichen existieren. Da stellt sich die Frage, ob es einen Zusammenhang zwischen diesen beiden Begriffen gibt. Es gibt ihn, und er führt über die sogenannten Hauptidealringe

Definition A6.48.
Ein Integritätsbereich R heißt *Hauptidealring*, wenn jedes Ideal ein Hauptideal ist, d.h. von einem Element erzeugt wird.

In einem Hauptidealring R gibt es also für jedes Ideal I ein Element $a \in I$ so, daß

$$I = \langle a \rangle_R = \{r \cdot a \mid r \in R\}.$$

Einfacher kann ein Ideal nicht mehr sein. Die Elemente in I sind alle Vielfache eines einzigen Elementes a.

Satz A6.49.
Jeder euklidische Integritätsbereich ist ein Hauptidealring.

Beweis: Sei $I \trianglelefteq R$ ein beliebiges Ideal. Wir müssen zeigen, daß $I = \langle b \rangle_R$ für ein geeignetes Element $b \in I$. Da das Nullideal von 0 erzeugt wird, können wir $I \neq \{0\}$ annehmen. Wenn $\nu : R \setminus \{0\} \to \mathbb{N}$ eine euklidische Funktion von R bezeichnet, dann wählen wir $0 \neq b \in I$ so, daß $\nu(b)$ minimal wird. Sei nun $0 \neq a \in I$ beliebig, so gibt es $q, r \in R$ so, daß $a = q \cdot b + r$ mit $r = 0$ oder $\nu(r) < \nu(b)$. Aber dann gilt

$$r = a - q \cdot b \in I,$$

da $a \in I$ und $b \in I$. Wegen der Minimalitätsbedingung, der b genügt, muß dann aber $r = 0$ gelten. Also ist $a = q \cdot b \in \langle b \rangle_R$, und somit $I = \langle b \rangle_R$. □

Da wir wissen, daß $\mathbb{Z}$ und Polynomringe über Körpern euklidisch sind, erhalten wir folgende Korollare.

Korollar A6.50.
$\mathbb{Z}$ ist ein Hauptidealring.

Korollar A6.51.
Ist K ein Körper, so ist der Polynomring $K[t]$ ein Hauptidealring.

Bemerkung A6.52.
Für die ganzen Zahlen war uns diese Aussage bereits vorher bekannt, da wir in Korollar A5.29 die Ideale in $\mathbb{Z}$ als die Untergruppen der additiven Gruppe $\mathbb{Z}$ identifiziert haben, von denen wir bereits aus Proposition 22.16 wußten, daß sie von einem Element erzeugt werden. Schaut man sich den Beweis von Proposition 22.16 an, so stellt man fest, daß er mit dem Beweis von Satz A6.49 identisch ist.

Bemerkung A6.53.
Der Ring

$$\mathbb{Z}\left[\frac{1+\sqrt{-19}}{2}\right] = \left\{a + b \cdot \frac{1+\sqrt{-19}}{2} \;\middle|\; a, b \in \mathbb{Z}\right\}$$

ist ein Hauptidealring, der *nicht* euklidisch ist. Der Beweis dieser Aussage ist mit elementaren Mitteln möglich, ist aber sehr technisch und sprengt den Rahmen dieser Vorlesung. □

Obwohl nicht jeder Hauptidealring euklidisch ist, wollen wir nun zeigen, daß auch in jedem Hauptidealring R größte gemeinsame Teiler existieren. Sind zwei Elemente a und b in R gegeben, so muß das von a und b erzeugte Ideal

$$\langle a, b\rangle_R = \{r \cdot a + s \cdot b \mid r, s \in R\}$$

nach Voraussetzung auch von einem einzigen Element erzeugt werden können. Ein solcher Erzeuger entpuppt sich als größter gemeinsamer Teiler von a und b.

Satz A6.54 (Bézout Identität).
Sei R ein Hauptidealring und $g, a, b \in R$. Die folgenden Aussagen sind gleichwertig:

a. $g \in \mathrm{ggT}(a, b)$.

b. $\langle g\rangle_R = \langle a, b\rangle_R$.

Insbesondere gibt es für $g \in \mathrm{ggT}(a, b)$ also $r, s \in R$ mit

$$g = r \cdot a + s \cdot b. \tag{135}$$

Man nennt (135) auch eine *Bézout Identität* des größten gemeinsamen Teilers g von a und b.

Beweis: Sei $g \in \mathrm{ggT}(a, b)$ und h ein Erzeuger des Ideals $\langle a, b\rangle_R$, so gilt nach Lemma A6.7

$$\langle h\rangle_R = \langle a, b\rangle_R \subseteq \langle g\rangle_R.$$

Aus dem gleichen Lemma folgt dann aber, daß h ein Teiler von a und b ist. Da g ein größter gemeinsamer Teiler ist, folgt notwendig $h \mid g$ und damit

$$\langle a, b\rangle_R = \langle h\rangle_R \supseteq \langle g\rangle_R.$$

Gilt umgekehrt die Gleichung

$$\langle a, b\rangle_R = \langle g\rangle_R,$$

so folgt aus Lemma A6.7 wieder, daß g ein Teiler von a und von b ist. Ist nun h irgend ein Teiler von a und von b, so gilt

$$\langle h\rangle_R \supseteq \langle a, b\rangle_R = \langle g\rangle_R,$$

so daß wiederum h ein Teiler von g ist. Damit ist $g \in \mathrm{ggT}(a, b)$. □

Bemerkung A6.55.
Ist R nicht nur ein Hauptidealring, sondern sogar euklidisch, so lassen sich mit Hilfe des Euklidischen Algorithmus' auch $r, s \in R$ mit $g = r \cdot a + s \cdot b$ für $g \in \mathrm{ggT}(a, b)$ berechnen. Dazu muß man sich aber die q_i und die r_i der Zwischenschritte merken und Rückeinsetzen. Wir führen dies nur am Beispiel vor.

Es seien $a = 8 \in \mathbb{Z}$ und $b = 3 \in \mathbb{Z}$. Der Euklidische Algorithmus liefert:

$$\begin{aligned} 8 &= 2 \cdot 3 + 2 \\ 3 &= 1 \cdot 2 + 1 \\ 2 &= 2 \cdot 1 + 0 \end{aligned}$$

so daß $1 \in \mathrm{ggT}(3, 8)$. Durch Rückeinsetzen erhalten wir dann:

$$1 = 3 - 1 \cdot 2 = 3 - 1 \cdot (8 - 2 \cdot 3) = 3 \cdot 3 + (-1) \cdot 8.$$

Mit Hilfe der Bézout Identität können wir die Einheitengruppe in $\mathbb{Z}_n$ bestimmen, auch wenn n keine Primzahl ist.

Proposition A6.56.
Für $0 \neq n \in \mathbb{Z}$ ist

$$\mathbb{Z}_n^* = \{\overline{a} \mid 1 \in \mathrm{ggT}(a, n)\}$$

die Einheitengruppe von $\mathbb{Z}_n$, d.h. eine Nebenklasse $\overline{a} \in \mathbb{Z}_n$ ist invertierbar genau dann, wenn 1 ein größter gemeinsamer Teiler von a und n ist.

Beweis: Sei $\overline{a} \in \mathbb{Z}_n^*$. Dann gibt es ein $\overline{b} \in \mathbb{Z}_n$ mit $\overline{1} = \overline{a} \cdot \overline{b} = \overline{a \cdot b}$. Mithin gilt

$$a \cdot b - 1 \in n\mathbb{Z}$$

ist ein Vielfaches von n. Also gibt es ein $r \in \mathbb{Z}$ mit

$$a \cdot b - 1 = r \cdot n,$$

und deshalb

$$1 = a \cdot b - r \cdot n \in \langle a, n \rangle_{\mathbb{Z}} \subseteq \mathbb{Z} = \langle 1 \rangle_{\mathbb{Z}}.$$

Aber damit ist notwendig

$$\langle 1 \rangle_{\mathbb{Z}} = \langle a, n \rangle_{\mathbb{Z}},$$

und wegen Satz A6.54 ist $1 \in \mathrm{ggT}(a, n)$.

Sei umgekehrt $1 \in \mathrm{ggT}(a, n)$, so gibt es wegen Satz A6.54 $b, r \in \mathbb{Z}$ mit

$$1 = b \cdot a + r \cdot n,$$

und damit gilt

$$\overline{1} = \overline{b \cdot a + r \cdot n} = \overline{b} \cdot \overline{a} + \overline{r} \cdot \overline{n} = \overline{b} \cdot \overline{a} \in \mathbb{Z}_n,$$

da $\overline{n} = \overline{0}$. Also ist $\overline{a} \in \mathbb{Z}_n^*$ eine Einheit. □

Bemerkung A6.57.
Der Beweis von Proposition A6.56 ist konstruktiv, d.h. er sagt uns, wie wir das Inverse von $\overline{a}$ in $\mathbb{Z}_n$ finden können, wenn a und n teilerfremd sind, nämlich mit Hilfe des Euklidischen Algorithmus. Ist $n = 8$ und $a = 3$, so haben wir in Bemerkung A6.55 mittels des Euklidischen Algorithmus' folgende Darstellung der 1 bestimmt:

$$1 = 3 \cdot 3 + (-1) \cdot 8.$$

Mithin ist $\overline{3}^{-1} = \overline{3} \in \mathbb{Z}_8$. □

Als nächstes wollen wir zeigen, daß jeder Hauptidealring faktoriell ist, so daß insbesondere auch jeder euklidische Integritätsbereich faktoriell ist. Dazu benötigen wir aber einige Vorbereitungen.

Lemma A6.58.
Sei R ein Hauptidealring, $a \in R$ irreduzibel und $b \in R \setminus \langle a \rangle_R$. Dann ist $1 \in \mathrm{ggT}(a, b)$.
Insbesondere gibt es also $r, s \in R$ so, daß $1 = r \cdot a + s \cdot b$.

Beweis: Sei $g \in \mathrm{ggT}(a, b)$. Es reicht zu zeigen, daß g eine Einheit ist. Es gilt $a \in \langle a, b \rangle_R = \langle g \rangle_R$. Folglich gilt $a = c \cdot g$ für ein geeignetes $c \in R$. Da a aber irreduzibel ist, muß entweder c eine Einheit sein oder g. Wäre c eine Einheit, so wäre $\langle a \rangle_R = \langle c \cdot g \rangle_R = \langle g \rangle_R = \langle a, b \rangle_R$ im Widerspruch zur Wahl von $b \notin \langle a \rangle_R$. Also muß g eine Einheit sein. □

Lemma A6.59.
Ist R ein Hauptidealring, so ist jedes irreduzible Element prim.

Beweis: Sei dazu $a \in R$ irreduzibel und $a \mid b \cdot c$. Angenommen $a \nmid b$ und $a \nmid c$, d.h. $b \in R \setminus \langle a \rangle_R$ und $c \in R \setminus \langle a \rangle_R$. Dann gibt es nach Lemma A6.58 Elemente $r, s, r', s' \in R$ so, daß

$$1 = r \cdot a + s \cdot b \quad \text{und} \quad 1 = r' \cdot a + s' \cdot c.$$

Mithin gilt

$$a \mid a \cdot (a \cdot r \cdot r' + r \cdot s' \cdot c + r' \cdot s \cdot b) + s \cdot s' \cdot b \cdot c = 1,$$

und a ist eine Einheit im Widerspruch zur Irreduzibilität von a. □

Satz A6.60.
Jeder Hauptidealring ist faktoriell.

Beweis: Da nach Lemma A6.59 jedes irreduzible Element prim ist, reicht es zu zeigen daß jedes $0 \neq a \in R \setminus R^*$ Produkt von endlich vielen irreduziblen Elementen ist.

Nehmen wir an es gibt ein Element $0 \neq a_0 \in R \setminus R^*$, welches sich nicht als Produkt von endlich vielen irreduziblen Elementen schreiben läßt. Dann ist a_0 insbesondere nicht selbst irreduzibel. Mithin gibt es Elemente $0 \neq a_1, b_1 \in R \setminus R^*$ so, daß $a_0 = a_1 \cdot b_1$. Da a_0 nicht Produkt von endlich vielen irreduziblen Elementen ist, muß dies ebenfalls für mindestens einen der Faktoren a_1 und b_1 gelten. Wir können ohne Einschränkung annehmen, daß es für a_1 gilt, und erhalten dann

$$\langle a_0 \rangle_R \subsetneqq \langle a_1 \rangle_R,$$

da b_1 keine Einheit ist. Wir können nun mit a_1 auf die gleiche Weise verfahren wie mit a_0, und auf diesem Wege konstruieren wir induktiv eine aufsteigende Kette von Idealen

$$\langle a_0 \rangle_R \subsetneqq \langle a_1 \rangle_R \subsetneqq \langle a_2 \rangle_R \subsetneqq \langle a_3 \rangle_R \subsetneqq \dots . \tag{136}$$

Betrachten wir nun die Vereinigung

$$I = \bigcup_{i=0}^{\infty} \langle a_i \rangle_R$$

all dieser Ideale, so erhalten wir wieder ein Ideal. Denn sind $b, c \in I$, so gibt es $i, j \in \mathbb{N}$ so, daß $b \in \langle a_i \rangle_R$ und $c \in \langle a_j \rangle_R$. Ohne Einschränkung gilt $i \leq j$ und damit $\langle a_i \rangle_R \subseteq \langle a_j \rangle_R$. Aber dann sind b und c beide in $\langle a_j \rangle_R$ und da dieses ein Ideal ist gilt auch

$$b + c \in \langle a_j \rangle_R \subseteq I.$$

Mithin ist I abgeschlossen bezüglich der Addition. Außerdem gilt

$$r \cdot b \in \langle a_i \rangle_R \subseteq I$$

für $r \in R$. Dies zeigt, daß I in der Tat ein Ideal ist.

Da R ein Hauptidealring ist, ist I ein Hauptideal. Es gibt also ein $s \in R$ so, daß $I = \langle s \rangle_R$. Aber dann gibt es ein $i \in \mathbb{N}$, so daß $s \in \langle a_i \rangle_R$ und folglich

$$\langle a_{i+1} \rangle_R \subseteq I = \langle s \rangle_R \subseteq \langle a_i \rangle_R,$$

im Widerspruch zu (136). □

Bemerkung A6.61.
Der Widerspruch im Beweis des obigen Satzes leitet sich aus dem Umstand her, daß es in einem Hauptidealring keine echt aufsteigende Kette von Idealen

$$\langle a_0 \rangle_R \subsetneqq \langle a_1 \rangle_R \subsetneqq \langle a_2 \rangle_R \subsetneqq \langle a_3 \rangle_R \subsetneqq \dots$$

geben kann. Ringe, in denen jede aufsteigende Kette von Idealen nach endlich vielen Schritten abbrechen muß, nennt man *noethersch*. Hauptidealringe sind also Beispiele für noethersche Ringe. In der kommutativen Algebra werden noethersche Ringe genauer untersucht. □

Korollar A6.62.
Ist K ein Körper, so ist $K[t]$ faktoriell, d.h. jedes Polynom in $K[t]$ besitzt eine im wesentlichen eindeutige Primfaktorzerlegung.

Beispiel A6.63.
Das Polynom $f = t^4 + \overline{3} \cdot t^3 + \overline{2} \in \mathbb{Z}_5[t]$ hat die Primfaktorzerlegung

$$f = \left(t + \overline{1}\right)^2 \cdot \left(t^2 + t + \overline{2}\right).$$

Man beachte dabei, daß $t^2 + t + \overline{2}$ nach Aufgabe A6.45 irreduzibel ist, da das Polynom keine Nullstelle in $\mathbb{Z}_5$ besitzt.

Wir haben in Korollar A6.51 gesehen, daß der Polynomring über einem Körper ein Hauptidealring ist. Die Aussage folgender Aufgabe, zeigt, daß die Bedingung an K nicht nur hinreichend, sondern auch notwendig ist.

Aufgabe A6.64.
Für einen Integritätsbereich R sind die folgenden Aussagen gleichwertig:

a. R ist ein Körper.

b. $R[t]$ ist ein euklidischer Ring.

c. $R[t]$ ist ein Hauptidealring.

Aufgabe A6.65.
Es sei K ein Körper und $I \lhd K[[t]]$ ein Ideal mit $I \neq \{0\}$ und $I \neq K[[t]]$. Zeige, es gibt ein $n \geq 1$ mit $I = \langle t^n \rangle_{K[[t]]}$. Ist $K[[t]]$ faktoriell?

F) Der Fundamentalsatz der elementaren Zahlentheorie

Aus Satz A6.60 und Korollar A6.51 erhalten wir unmittelbar folgende Resultate.

Korollar A6.66.
$\mathbb{Z}$ ist faktoriell.

Eine etwas ausführlichere Fassung dieser Aussage ist der *Fundamentalsatz der elementaren Zahlentheorie*, der unter Berücksichtigung von Bemerkung A6.19 und Bemerkung A6.20 folgt, da $\mathbb{Z}^* = \{1, -1\}$.

Korollar A6.67 (Fundamentalsatz der elementaren Zahlentheorie).
Für jedes $0 \neq z \in \mathbb{Z}$ gibt es eindeutig bestimmte, paarweise verschiedene Primzahlen $p_1, \ldots, p_k$ und eindeutig bestimmte positive ganze Zahlen $n_1, \ldots, n_k \in \mathbb{Z}_{>0}$, so daß

$$z = \operatorname{sgn}(z) \cdot p_1^{n_1} \cdots p_k^{n_k},$$

wobei

$$\operatorname{sgn}(z) := \begin{cases} 1, & z > 0, \\ -1, & z < 0. \end{cases}$$

Bezeichnen wir mit $\mathbb{P}$ die Menge der Primzahlen und führen wir für ein Primzahl $p \in \mathbb{P}$ die Notation

$$n_p(z) = \max\left\{n \in \mathbb{N} \;\middle|\; p^n \mid z\right\}$$

ein, so gilt

$$n_p(z) = \begin{cases} n_i, & p = p_i, \\ 0, & \text{sonst} \end{cases}$$

und

$$z = \operatorname{sgn}(z) \cdot \prod_{p \in \mathbb{P}} p^{n_p(z)}.$$

Man beachte bei der Formulierung im Fundamentalsatz, daß das Produkt $\prod_{p\in\mathbb{P}} p^{n_p(z)}$ zwar unendlich viele Faktoren hat, daß aber nur endlich viele davon ungleich Eins sind. Insofern ist das Produkt definiert, indem man nur die Faktoren ungleich Eins berücksichtigt.

Aus Bemerkung A6.20 erhalten wir folgendes Korollar zur Bestimmung von größten gemeinsamen Teilern und kleinsten gemeinsamen Vielfachen mittels Primfaktorzerlegung.

Korollar A6.68.
Es seien $a, b \in \mathbb{Z} \setminus \{0\}$, so gilt

$$\operatorname{ggt}(a,b) = \prod_{p \in \mathbb{P}} p^{\min\{n_p(a), n_p(b)\}}$$

und

$$\operatorname{kgv}(a,b) = \prod_{p \in \mathbb{P}} p^{\max\{n_p(a), n_p(b)\}}.$$

Damit gilt insbesondere

$$|a \cdot b| = \operatorname{kgv}(a,b) \cdot \operatorname{ggt}(a,b).$$

Wir wollen den Teilabschnitt mit der alten Erkenntnis abschließen, daß Primzahlen keine Seltenheit sind.

Satz A6.69 (Euklid).
Es gibt unendlich viele Primzahlen in $\mathbb{Z}$.

Beweis: Da 2 eine Primzahl ist, gibt es eine Primzahl. Nehmen wir nun an, daß es nur endlich viele Primzahlen $p_1, \ldots, p_r \in \mathbb{Z}$ gibt, und betrachten wir die Zahl

$$z = p_1 \cdots p_r + 1 > 1.$$

Aufgrund des Fundamentalsatzes der elementaren Zahlentheorie besitzt z eine Primfaktorzerlegung und es muß mithin mindestens eine Primzahl geben, die z teilt. D.h. es gibt ein i so, daß $p_i \mid z$. Aber dann gilt auch

$$p_i \mid z - p_1 \cdots p_r = 1,$$

was nur möglich ist, wenn p_i eine Einheit ist. Letzteres steht im Widerspruch zur Voraussetzung, daß p_i eine Primzahl ist.

Dies zeigt, daß es unendlich viele Primzahlen geben muß. □

G) Der chinesische Restsatz

Wir wollen in diesem Abschnitt folgende Frage beantworten. Gibt es Polynome $f, g \in \mathbb{Z}[t] \setminus \mathbb{Z}^*$, so daß

$$h := t^4 + 6t^3 + 17t^2 + 24t + 27 = f \cdot g,$$

d.h. ist h nicht irreduzibel in $\mathbb{Z}[t]$? Wir beachten zunächst, daß für die Leitkoeffizienten von f und g notwendig

$$\mathrm{lc}(f) \cdot \mathrm{lc}(g) = \mathrm{lc}(f \cdot g) = 1$$

gilt, so daß wir ohne Einschränkung $\mathrm{lc}(f) = 1 = \mathrm{lc}(g)$ annehmen können.

Wir gehen das Problem nun durch *Reduktion des Polynoms h modulo einer Primzahl p* an, d.h. wir betrachten das Bild von h unter der Abbildung

$$\phi_p : \mathbb{Z}[t] \longrightarrow \mathbb{Z}_p[t] : \sum_{k=0}^{n} a_k \cdot t^k \mapsto \sum_{k=0}^{n} \overline{a_k} \cdot t^k.$$

Nach Aufgabe A5.40 diese Abbildung ist ein Ringhomomorphismus, so daß die Gleichung $h = f \cdot g$ notwendig zu

$$\phi_p(h) = \phi_p(f) \cdot \phi_p(g)$$

führt.

In obigem Beispiel betrachten wir h modulo der Primzahlen 2 und 7, und erhalten

$$\begin{aligned}\phi_2(h) &= t^4 + \overline{6} \cdot t^3 + \overline{17} \cdot t^2 + \overline{24} \cdot t + \overline{27} \\ &= t^4 + t^2 + \overline{1} = \left(t^2 + t + \overline{1}\right)^2 \in \mathbb{Z}_2[t]\end{aligned}$$

und

$$\begin{aligned}\phi_7(h) &= t^4 + \overline{6}\cdot t^3 + \overline{17}\cdot t^2 + \overline{24}\cdot t + \overline{27}\\ &= t^4 + \overline{6}\cdot t^3 + \overline{3}\cdot t^2 + \overline{3}\cdot t + \overline{6}\\ &= \left(t^2 + \overline{5}\cdot t + \overline{2}\right)\cdot\left(t^2 + t + \overline{3}\right) \in \mathbb{Z}_7[t].\end{aligned}$$

Die Faktorisierung von $\phi_2(h)$ in $\mathbb{Z}_2[t]$ und von $\phi_7(h)$ in $\mathbb{Z}_7[t]$ erhält man, indem man die Produkte aller Paare zweier Polynome vom Grad höchstens drei durchprobiert, deren Grade sich zu vier addieren. Da es nur endlich viele sind, ist das kein wirkliches Problem, obwohl es durchaus etwas Zeit in Anspruch nimmt, wenn man das von Hand tun will. Die gefundenen Faktoren von $\phi_2(h)$ und von $\phi_7(h)$ sind irreduzibel nach Aufgabe A6.45, da sie jeweils Grad zwei haben, ohne eine Nullstelle zu besitzen. Letzteres ist wieder ein einfacher Test in $\mathbb{Z}_2$ bzw. in $\mathbb{Z}_7$.

Wenn es also Polynome f und g wie oben gibt, so müssen sie notwendig beide Grad zwei haben, d.h.

$$f = t^2 + b_1\cdot t + b_0 \quad \text{und} \quad g = t^2 + c_1\cdot t + c_0,$$

und ferner muß für die Reduktion modulo 2 bzw. 7 gelten

$$\phi_2(f) = \phi_2(g) = t^2 + t + \overline{1}$$

sowie ohne Einschränkung

$$\phi_7(f) = t^2 + \overline{5}\cdot t + \overline{2} \quad \text{und} \quad \phi_7(g) = t^2 + t + \overline{3}.$$

Wir suchen also Zahlen $b_0, b_1, c_0, c_1 \in \mathbb{Z}$ die folgende Kongruenzgleichungssysteme erfüllen:

$$\begin{aligned} b_0 &\equiv 1 \pmod 2\\ b_0 &\equiv 2 \pmod 7\end{aligned} \tag{137}$$

$$\begin{aligned} b_1 &\equiv 1 \pmod 2\\ b_1 &\equiv 5 \pmod 7\end{aligned} \tag{138}$$

$$\begin{aligned} c_0 &\equiv 1 \pmod 2\\ c_0 &\equiv 3 \pmod 7\end{aligned} \tag{139}$$

$$\begin{aligned} c_1 &\equiv 1 \pmod 2\\ c_1 &\equiv 1 \pmod 7\end{aligned} \tag{140}$$

Sind wir in der Lage, ein Kongruenzgleichungssystem wie (137) zu lösen? Die Antwort darauf gibt der chinesische Restsatz, ein algorithmisches Verfahren zur Lösung solcher Kongruenzgleichungssysteme, das in China bereits im 3. Jahrhundert bekannt war.

Die folgenden Lemmata sind wichtige Bausteine für den Beweis des chinesischen Restsatzes.

Lemma A6.70.
Seien $n_1, \ldots, n_r \in \mathbb{Z} \setminus \{0\}$ paarweise teilerfremd und sei $N_i = \frac{n_1 \cdots n_r}{n_i}$.
Dann sind n_i und N_i teilerfremd und $\overline{N_i} \in \mathbb{Z}_{n_i}^*$ für $i \in \{1, \ldots, n\}$.

Beweis: Sei $i \in \{1, \ldots, r\}$ gegeben. Für $j \neq i$ sind n_i und n_j teilerfremd. Dies bedeutet $1 \in \mathrm{ggT}(n_i, n_j)$, und wegen der Bézout Identität existieren mithin $s_j, r_j \in \mathbb{Z}$ so, daß

$$1 = n_i \cdot r_j + n_j \cdot s_j.$$

Wenn wir j alle Indizes von 1 bis r außer i durchlaufen lassen, so können wir die Zahl 1 in folgender Weise als Produkt von $r-1$ Faktoren schreiben:

$$1 = \prod_{j \neq i} 1 = \prod_{j \neq i} (n_i \cdot r_j + n_j \cdot s_j). \tag{141}$$

Multiplizieren wir das Produkt auf der rechten Seite aus, so erhalten wir eine Summe, in der bis auf einen einzigen Term jeder Term n_i als Faktor enthält. Der eine Term, der n_i nicht als Faktor hat, ist

$$\prod_{j \neq i} (n_j \cdot s_j) = N_i \cdot \prod_{j \neq i} s_j.$$

Spalten wir n_i von den verbleibenden Termen ab, so erhalten wir eine Zahl $z \in \mathbb{Z}$, so daß (141) folgende Form annimmt:

$$1 = n_i \cdot z + N_i \cdot \prod_{j \neq i} s_j \in \langle n_i, N_i \rangle_{\mathbb{Z}}.$$

Aber damit gilt $\langle n_i, N_i \rangle_{\mathbb{Z}} = \langle 1 \rangle_{\mathbb{Z}}$ und wegen Satz A6.54 ist $1 \in \mathrm{ggT}(n_i, N_i)$, d.h. n_i und N_i sind teilerfremd. Aus Proposition A6.56 folgt schließlich, daß N_i eine Einheit in $\mathbb{Z}_{n_i}$ ist. □

Lemma A6.71.
Sind $n_1, \ldots, n_r \in \mathbb{Z} \setminus \{0\}$ paarweise teilerfremd und $a \in \mathbb{Z} \setminus \{0\}$ mit $n_i \mid a$ für $i = 1, \ldots, r$, so gilt:

$$n_1 \cdots n_r \mid a.$$

Beweis: Wir führen den Beweis durch Induktion über r, wobei die Aussage für $r = 1$ trivialerweise erfüllt ist. Wir können also $r \geq 2$ annehmen.

Mit der Notation von Lemma A6.70 gilt dann nach Induktionsvoraussetzung

$$N_r = n_1 \cdots n_{r-1} \mid a.$$

Mithin gibt es ganze Zahlen $b, c \in \mathbb{Z}$ mit $a = n_r \cdot b$ und $a = N_r \cdot c$. Da nach Lemma A6.70 zudem n_r und N_r teilerfremd sind, liefert die Bézout Identität ganze Zahlen $x, y \in \mathbb{Z}$ mit

$$x \cdot n_r + y \cdot N_r = 1.$$

Kombinieren wir die drei Gleichungen, so erhalten wir:

$$a = a \cdot (x \cdot n_r + y \cdot N_r) = a \cdot x \cdot n_r + a \cdot y \cdot N_r$$
$$= N_r \cdot c \cdot x \cdot n_r + n_r \cdot b \cdot y \cdot N_r = n_r \cdot N_r \cdot (c \cdot x + b \cdot y).$$

Mithin wird a von $N_r \cdot n_r = n_1 \cdots n_r$ geteilt. □

Satz A6.72 (Chinesischer Restsatz).
Seien $n_1, \ldots, n_r \in \mathbb{Z} \setminus \{0\}$ paarweise teilerfremd, $N = n_1 \cdots n_r$ und $N_i = \frac{N}{n_i}$.

a. Zu beliebig vorgegebenen ganzen Zahlen $a_1, \ldots, a_r \in \mathbb{Z}$ existiert eine Lösung $x \in \mathbb{Z}$ des *Kongruenzgleichungssystems*

$$\begin{aligned} x &\equiv a_1 \pmod{n_1}, \\ x &\equiv a_2 \pmod{n_2}, \\ &\vdots \\ x &\equiv a_r \pmod{n_r}. \end{aligned} \tag{142}$$

b. Ist $\overline{x_i} = \overline{N_i}^{-1} \in \mathbb{Z}_{n_i}$ für $i = 1, \ldots, r$, so ist

$$x' = \sum_{i=1}^{r} N_i \cdot x_i \cdot a_i \in \mathbb{Z} \tag{143}$$

eine Lösung von (142).

c. Genau dann ist $x'' \in \mathbb{Z}$ eine Lösung von (142), wenn x'' sich von x' nur um ein Vielfaches von N unterscheidet. Insbesondere ist die Lösung von (142) modulo N eindeutig bestimmt.

Beweis: Wir zeigen zunächst, daß x' eine Lösung von (142) ist und beweisen damit a. und b.. Nach Lemma A6.70 existiert für $i = 1, \ldots, r$ ein $x_i \in \mathbb{Z}$ mit $\overline{x_i} = \overline{N_i}^{-1} \in \mathbb{Z}_{n_i}$. Wir können deshalb

$$x' := \sum_{j=1}^{r} N_j \cdot x_j \cdot a_j$$

betrachten. Wegen $n_i \mid N_j$ für $j \neq i$ gilt aber in $\mathbb{Z}_{n_i}$ die Gleichung

$$\overline{x'} = \sum_{j=1}^{r} \overline{N_j} \cdot \overline{x_j} \cdot \overline{a_j} = \overline{N_i} \cdot \overline{x_i} \cdot \overline{a_i} = \overline{a_i} \in \mathbb{Z}_{n_i},$$

d.h.

$$x' \equiv a_i \pmod{n_i}.$$

Es bleibt also zu zeigen, daß $x' + N\mathbb{Z}$ die Menge der Lösungen von (142) ist. Sei $x'' \in \mathbb{Z}$ eine beliebige Lösung von (142). Dann gilt für $i = 1, \ldots, r$

$$x' - a_i, x'' - a_i \in n_i\mathbb{Z}.$$

Damit gilt aber $x' - x'' \in n_i\mathbb{Z}$, d. h. $n_i \mid (x' - x'')$, für alle $i = 1, \ldots, r$. Aus Lemma A6.71 folgt dann $N \mid (x' - x'')$, d. h. $x' - x'' \in N\mathbb{Z}$, und damit

$$x' \equiv x'' \pmod{N}.$$

Ist umgekehrt $x'' = x' + N \cdot z$ für ein $z \in \mathbb{Z}$, so gilt $N \mid x' - x''$ und damit $n_i \mid x' - x''$ für alle $i = 1, \ldots, r$. Da wir bereits wissen, daß $x' \equiv a_i \pmod{n_i}$ gilt, d.h. $n_i \mid x' - a_i$, so folgt

$$n_i \mid \big((x' - a_i) - (x' - x'')\big) = (x'' - a_i),$$

d.h. $x'' \equiv a_i \pmod{n_i}$ für alle $i = 1, \ldots, r$. Also ist x'' dann auch eine Lösung von (142). □

Bemerkung A6.73.
Da wir das Inverse von $\overline{N_i}$ in $\mathbb{Z}_{n_i}$ mit Hilfe des Euklidischen Algorithmus berechnen können (siehe Bemerkung A6.57), sind wir auch in der Lage ein Kongruenzgleichungssystem der Form (142) mit Hilfe der Formel (143) zu lösen.

In Anwendungen werden die n_i meist paarweise verschiedene Primzahlen sein, wie in dem Eingangsbeispiel des Abschnitts.

Man kann die Aussage des chinesischen Restsatzes auch etwas algebraischer formulieren, wenn man das karthesische Produkt

$$\mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \ldots \times \mathbb{Z}_{n_r}$$

mit komponentenweiser Addition und Multiplikation als kommutativen Ring mit Eins betrachtet. Diese algebraische Formulierung wird in der Vorlesung Elementare Zahlentheorie eine wichtige Rolle spielen.

In folgendem Korollar werden wir die Restklasse von x in $\mathbb{Z}_m$ ausnahmsweise mit $\overline{x}_m$ statt mit $\overline{x}$ bezeichnen, um deutlich zu machen, in welchem Ring sie lebt.

Korollar A6.74 (Chinesischer Restsatz).
Es seien $n_1, \ldots, n_r \in \mathbb{Z}_{>0}$ paarweise teilerfremde positive Zahlen, dann ist die Abbildung

$$\overline{\alpha} : \mathbb{Z}_{n_1 \cdots n_r} \longrightarrow \mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_r} : \overline{x}_{n_1 \cdots n_r} \mapsto \big(\overline{x}_{n_1}, \ldots, \overline{x}_{n_r}\big)$$

ein Isomorphismus kommutativer Ringe mit Eins.
Ferner ist die induzierte Abbildung

$$\mathbb{Z}^*_{n_1 \cdots n_r} \longrightarrow \mathbb{Z}^*_{n_1} \times \ldots \times \mathbb{Z}^*_{n_r} : \overline{x}_{n_1 \cdots n_r} \mapsto \big(\overline{x}_{n_1}, \ldots, \overline{x}_{n_r}\big)$$

ein Isomorphismus der Einheitengruppen der Ringe.

Beweis: Die Abbildung

$$\alpha : \mathbb{Z} \longrightarrow \mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_r} : x \mapsto \left(\overline{x}_{n_1}, \ldots, \overline{x}_{n_r}\right)$$

ist offensichtlich ein Ringhomomorphismus. Der Chinesische Restsatz A6.72 besagt nun, daß α surjektiv ist mit

$$\mathrm{Ker}(\alpha) = \langle n_1 \cdots n_r \rangle_{\mathbb{Z}}.$$

Die erste Behauptung folgt dann mit dem Homomorphiesatz A3.18.

Da ein Isomorphismus von Ringen Einheiten auf Einheiten abbildet, gilt

$$\overline{\alpha}\left(\mathbb{Z}^*_{n_1 \cdots n_r}\right) = \left(\mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_r}\right)^* = \mathbb{Z}^*_{n_1} \times \ldots \times \mathbb{Z}^*_{n_r},$$

wobei die letzte Gleichheit aus Beispiel A5.4 folgt. Da die Abbildung $\overline{\alpha}$ die Multiplikation respektiert, induziert sie somit einen Gruppenisomorphismus der multiplikativen Gruppen. □

Beispiel A6.75.
Wir wollen nun die Kongruenzgleichungssysteme (137), (138), (139) und (140) lösen. Ersteres hat die Form:

$$\begin{aligned} b_0 &\equiv 1 \pmod 2 \\ b_0 &\equiv 2 \pmod 7 \end{aligned}$$

Dabei ist in der Notation des chinesischen Restsatzes $n_1 = N_2 = 2$, $n_2 = N_1 = 7$, $a_1 = 1$ und $a_2 = 2$. Auch ohne den Euklidischen Algorithmus anzuwenden sehen wir

$$1 = 4 \cdot 2 + (-1) \cdot 7.$$

Mithin gilt $\overline{x_1} = \overline{1} = \overline{-1} = \overline{N_1}^{-1} \in \mathbb{Z}_2$ und $\overline{x_2} = \overline{4} = \overline{N_2}^{-1} \in \mathbb{Z}_7$. Die gesuchte Lösung b_0 läßt sich bis auf ein Vielfaches von $N = 2 \cdot 7 = 14$ mithin beschreiben als

$$b_0 \equiv x_1 \cdot N_1 \cdot a_1 + x_2 \cdot N_2 \cdot a_2 = 1 \cdot 7 \cdot 1 + 4 \cdot 2 \cdot 2 = 23 \equiv 9 \pmod{14}.$$

Für die verbleibenden drei Kongruenzgleichungssysteme bleiben die $\overline{n_i}$, $\overline{N_i}$ und $\overline{x_i}$ unverändert, und nur die a_i werden ausgetauscht, so daß wie die Lösungen modulo $N = 14$ unmittelbar angeben können:

$$b_1 \equiv x_1 \cdot N_1 \cdot 1 + x_2 \cdot N_2 \cdot 5 = 1 \cdot 7 \cdot 1 + 4 \cdot 2 \cdot 5 = 47 \equiv 5 \pmod{14},$$

$$c_0 \equiv x_1 \cdot N_1 \cdot 1 + x_2 \cdot N_2 \cdot 3 = 1 \cdot 7 \cdot 1 + 4 \cdot 2 \cdot 3 = 31 \equiv 3 \pmod{14}$$

und

$$c_1 \equiv x_1 \cdot N_1 \cdot 1 + x_2 \cdot N_2 \cdot 1 = 1 \cdot 7 \cdot 1 + 4 \cdot 2 \cdot 1 = 15 \equiv 1 \pmod{14}.$$

Wüßten wir aus irgendwelchen Zusatzüberlegungen bereits, daß die Koeffizienten zwischen 0 und 13 zu liegen, so könnten wir die Polynome f und g mit Gewissheit angeben, nämlich

$$f = t^2 + b_1 \cdot t + b_0 = t^2 + 5t + 9$$

und

$$g = t^2 + c_1 \cdot t + c_0 = t^2 + t + 3.$$

Da dies nicht der Fall ist, bleibt uns nur, unser Ergebnis zu testen, und in der Tat gilt

$$f \cdot g = (t^2 + 5t + 9) \cdot (t^2 + t + 3) = t^4 + 6t^3 + 17t^2 + 24t + 27 = h.$$

□

Bemerkung A6.76.

Das in der Einleitung zu diesem Abschnitt angegebene und in Beispiel A6.75 fortgeführte Beispiel, wie man die Zerlegung eines Polynoms in $\mathbb{Z}[t]$ in irreduzible Faktoren erreichen kann, funktioniert nicht nur zufällig. Es gibt Sätze, die es erlauben, aus den Koeffizienten des Polynoms h die Größe der Koeffizienten potenzieller Teiler von h abzuschätzen. Wählt man nun hinreichend viele paarweise verschiedene Primzahlen, so daß deren Produkt größer als diese Schranke ist, so kann man im wesentlichen in der angegebenen Weise die Zerlegung von f in irreduzible Polynome in $\mathbb{Z}[t]$ bestimmen. Wenn man dann noch ein weiteres Resultat verwendet, welches sagt, daß ein in $\mathbb{Z}[t]$ irreduzibles Polynom auch in $\mathbb{Q}[t]$ irreduzibel ist, so kann man auf diesem Weg Polynome in $\mathbb{Q}[t]$ in Primfaktoren zerlegen, indem man zunächst den Hauptnenner der Koeffizienten ausklammert.

Man beachte, daß es für die Polynomringe $\mathbb{R}[t]$ und $\mathbb{C}[t]$ kein derartiges Verfahren gibt, was einer der wesentlichen Gründe für die Notwendigkeit numerischer Verfahren ist. □

Wir wollen das Kapitel mit einem etwas längeren Beispiel zum chinesischen Restsatz abschließen.

Beispiel A6.77.

Gegeben sei das folgende Kongruenzgleichungssystem:

$$\begin{aligned} x &\equiv a_1 = 1 \pmod 2, \\ x &\equiv a_2 = 2 \pmod 3, \\ x &\equiv a_3 = 4 \pmod 7. \end{aligned}$$

Es sind $n_1 = 2, n_2 = 3, n_3 = 7$ paarweise teilerfremd, und $N = 2 \cdot 3 \cdot 7 = 42$, $N_1 = 21$, $N_2 = 14$ und $N_3 = 6$.

Die Berechnung der Inversen von $\overline{N_i}$ in $\mathbb{Z}_{n_i}$ geschieht mit Hilfe des Euklidischen Algorithmus'. Da n_i und N_i teilerfremd sind, gilt wegen der Bézout Identität

$$x_i N_i + y_i n_i = 1$$

für geeignete $x_i \in \mathbb{Z}$ (und $y_i \in \mathbb{Z}$, die hier nicht interessieren):

$$\overline{x_1} = \overline{21}^{-1} = \overline{1}^{-1} = \overline{1} \in \mathbb{Z}_2,$$

$$\overline{x_2} = \overline{14}^{-1} = \overline{2}^{-1} = \overline{2} \in \mathbb{Z}_3,$$

und

$$\overline{x_3} = \overline{6}^{-1} = \overline{6} \in \mathbb{Z}_7.$$

Es folgt:

$$\begin{aligned} x &\equiv N_1 \cdot x_1 \cdot a_1 + N_2 \cdot x_2 \cdot a_2 + N_3 \cdot x_3 \cdot a_3 \\ &= 21 \cdot 1 \cdot 1 + 14 \cdot 2 \cdot 2 + 6 \cdot 4 \cdot 6 = 221 \equiv 11 \pmod{42}. \end{aligned}$$

Also ist $x = 11$ die modulo 42 eindeutig bestimmte Lösung, und die Menge aller Lösungen ist

$$\{11 + z \cdot 42 \mid z \in \mathbb{Z}\}.$$

□

Bemerkung A6.78.

Die Voraussetzung des chinesischen Restsatzes, daß die n_i paarweise teilerfremd sein sollen, ist nicht nur für unseren Beweis notwendig. Ohne diese Voraussetzung ist die Aussage im allgemeinen falsch, wie folgendes Beispiel zeigt: $n_1 = 2$, $n_2 = 4$, $a_1 = 0$, $a_2 = 1$, dann impliziert $x \equiv a_1 \pmod 2$, daß x eine gerade Zahl ist, während $x \equiv a_2 \pmod 4$ nur für eine ungerade Zahl möglich ist. Es kann also keine ganze Zahl x geben, die beide Kongruenzgleichungen zugleich erfüllt, was daran liegt, daß $n_1 = 2$ und $n_2 = 4$ nicht teilerfremd sind.

Kapitel B

Einige Ergänzungen zur linearen Algebra

§ B1 Diagonalisierbarkeit und Trigonalisierbarkeit

In diesem Abschnitt sei V ein K-Vektorraum mit $1 \leq \dim_K(V) = n < \infty$.

Bemerkung B1.1 (Normalformen bezüglich Konjugation als Ziel).
Die nächsten beiden Abschnitte Abschnitte sind folgender Aufgabe gewidmet:

Finde für $f \in \mathrm{End}_K(V)$ eine Basis B so, daß $M_B^B(f)$ eine besonders einfache Gestalt hat und wichtige Eigenschaften von f direkt aus $M_B^B(f)$ ersichtlich sind!

Alternativ kann man die Frage auch für quadratische Matrizen formulieren:

Finde für $A \in \mathrm{Mat}_n(K)$ ein invertierbares $T \in \mathrm{Gl}_n(K)$ so, daß $T^{-1} \circ A \circ T$ eine besonders einfache Gestalt hat und wichtige Eigenschaften von A sofort sichtbar sind!

Solche *einfachen Repräsentanten* der Äquivalenzklassen bezüglich Konjugation nennt man dann *Normalformen bezüglich Konjugation*. Eine Diagonalmatrix als Normalform wie in Satz B1.27 wäre optimal einfach, ist aber sehr speziell, wie selbiger Satz zeigt. Wir wollen nun untersuchen, was unter welchen Voraussetzung erreichbar ist, und suchen dabei nach Normalformen mit möglichst vielen Nullen.

Ich möchte an dieser Stelle daran erinnern, daß wir uns schon mal eine ähnliche Aufgabe gestellt haben. Wir wollten Basen B und D finden, so daß die Matrixdarstellung $M_D^B(f)$ möglichst einfache Gestalt hat, oder alternativ invertierbare Matrizen S und T, so daß $S \circ A \circ T$ möglichst einfach ist. Die Aufgabe haben wir in Satz 31.27 und Korollar 31.28

gelöst und festgestellt, daß wir stets eine Matrix der Form

$$\begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}$$

erhalten können, wobei r der Rang von f bzw. von A ist. Aus dieser Form kann man über die Abbildung bzw. die Matrix außer dem Rang keine interessante Information mehr ablesen. Das ist der Grund, weshalb es wichtig ist, daß wir uns von nun an auf die Situation $B = D$ bei Matrixdarstellungen bzw. $S = T^{-1}$ bei Matrizen beschränken! Und wir haben oben schon gesehen, daß bei solchen Transformationen interessante Eigenschaften wie die Determinante, die Spur und das charakteristische Polynom erhalten bleiben.

Definition B1.2 (Diagonalisierbar und trigonalisierbar).
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. f heißt *trigonalisierbar*, falls es eine Basis B von V gibt, so daß $M_B^B(f)$ eine obere Dreiecksmatrix ist.

b. A heißt *trigonalisierbar*, falls es eine Matrix $T \in \mathrm{Gl}_n(K)$ gibt, so daß $T^{-1} \circ A \circ T$ eine obere Dreiecksmatrix ist.

A) Trigonalisierbarkeit

Satz B1.3 (Trigonalisierbarkeit).
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Genau dann ist f trigonalisierbar, wenn χ_f über K in Linearfaktoren zerfällt.

b. Genau dann ist A trigonalisierbar, wenn χ_A über K in Linearfaktoren zerfällt.

Beweis: Ist f trigonalisierbar, so gibt es eine Basis B mit

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & * & \dots & \dots & * \\ 0 & \lambda_2 & * & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_{n-1} & * \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix}.$$

Damit folgt, daß das charakteristische Polynom

$$\chi_f = (t - \lambda_1) \cdots (t - \lambda_n)$$

von f über K in Linearfaktoren zerfällt.

Zerfalle nun umgekehrt das charakteristische Polynom von f in Linearfaktoren $\chi_f = (t-\lambda_1)\cdots(t-\lambda_n)$. Wir beweisen mit Induktion über $n = \dim_K(V)$, daß dann f trigonalisierbar ist. Im Fall $n = 1$ ist f nach Beispiel 35.13 sogar diagonalisierbar.

Sei also $n > 1$ und sei $0 \neq x_1 \in V$ ein Eigenvektor von f zum Eigenwert λ_1. Wir setzen $U := \mathrm{Lin}(x_1) \leq V$. Wegen $f(x_1) = \lambda_1 x_1 \in U$ ist U ein f-invarianter Unterraum von V und $\chi_{f_U} = t - \lambda_1$. Mithin folgt aus Aufgabe 35.27

$$\chi_{f_{V/U}} = (t-\lambda_2)\cdots(t-\lambda_n),$$

d. h. das charakteristische Polynom von $f_{V/U}$ zerfällt über K in Linearfaktoren. Da $\dim_K(V/U) = n-1 < n$, existiert per Induktion eine Basis $B'' = (\overline{x_2}, \ldots, \overline{x_n})$ von V/U, so daß $M_{B''}^{B''}(f_{V/U})$ eine obere Dreiecksmatrix ist. Dann ist aber $B = (x_1, \ldots, x_n)$ eine Basis von V und mit $B' = (x_1)$ gilt wegen Aufgabe 31.34

$$M_B^B(f) = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right) = \left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right).$$

Damit ist $M_B^B(f)$ eine obere Dreiecksmatrix und f ist trigonalisierbar.

Die Aussage für A erhalten wir aus der entsprechenden Aussage für f_A. □

Bemerkung B1.4.
Ist K ein algebraisch abgeschlossener Körper, etwa $K = \mathbb{C}$, so sind somit jede Matrix A und jeder Endomorphismus f trigonalisierbar. Eine vergleichbare Aussage für die Diagonalisierbarkeit gilt nicht.

Beispiel B1.5.

a. Die Drehmatrix

$$A_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

hat das charakteristische Polynom $\chi_{A_\alpha} = t^2 - 2\cos(\alpha)t + 1 = (t-\lambda)\cdot(t-\overline{\lambda})$ mit $\lambda = \cos(\alpha) + i\sin(\alpha) \in \mathbb{C}$, $\alpha \in \mathbb{R}$. Damit hat χ_{A_α} also keine reellen Nullstellen, wenn α kein ganzzahliges Vielfaches von π ist, und somit ist A_α über $\mathbb{R}$ nicht trigonalisierbar.
Hingegen zerfällt χ_{A_α} über $\mathbb{C}$ in Linearfaktoren, so daß A_α über $\mathbb{C}$ trigonalisierbar sein muß. In der Tat ist A_α sogar diagonalisierbar mit Diagonalgestalt

$$\begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix}.$$

Ist α kein ganzzahliges Vielfaches von π, so besitzt A_α zwei verschiedene Eigenwerte, so daß zugehörige Eigenvektoren nach Aufgabe 35.31 eine Basis von $\mathbb{C}^2$ bilden müssen und diese transformieren A_α in obige Diagonalmatrix. Ist α hingegen ein

ganzzahliges Vielfaches von π, so ist $A_\alpha = \mathbb{1}_2$ oder $A_\alpha = -\mathbb{1}_2$ und hat bereits Diagonalgestalt.

b. Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C})$$

ist hingegen auch über $\mathbb{C}$ nicht diagonalisierbar. Denn, gäbe es eine Matrix $T \in \mathrm{Gl}_2(\mathbb{C})$ mit

$$T^{-1} \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C}),$$

dann wäre

$$\begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^2 = T^{-1} \circ A^2 \circ T = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

also wären $\lambda_1 = \lambda_2 = 0$. Aber damit würde gelten:

$$0 = \mathrm{rang} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \mathrm{rang}\left(T^{-1} \circ A \circ T\right) = \mathrm{rang}(A) = 1,$$

da $T \in \mathrm{Gl}_2(\mathbb{C})$. Dies ist jedoch ein Widerspruch.

B) Geometrische und algebraische Vielfachheiten von Eigenwerten

Sucht man bei einem trigonalisierbaren Endomorphismus eine Basis, für die die obere Dreiecksgestalt der Matrixdarstellung noch besser strukturiert ist, dann braucht die in diesem Abschnitt eingeführten Begriffe der geometrischen und algebraischen Vielfachheit eines Eigenwertes, die auch im weiteren Verlauf des Abschnitts eine wichtige Rolle spielen.

Definition B1.6 (Vielfachheit von Eigenwerten).
Es sei $f \in \mathrm{End}_K(V)$, $A \in \mathrm{Mat}_n(K)$ und $\lambda \in K$.

a. $\mathrm{mult}(\chi_f, \lambda)$ heißt *algebraische Vielfachheit* von λ als Eigenwert von f.
$\dim_K \mathrm{Eig}(f, \lambda)$ heißt *geometrische Vielfachheit* von λ als Eigenwert von f.

b. $\mathrm{mult}(\chi_A, \lambda)$ heißt *algebraische Vielfachheit* von λ als Eigenwert von A.
$\dim_K \mathrm{Eig}(A, \lambda)$ heißt *geometrische Vielfachheit* von λ als Eigenwert von A.

Die algebraischen Vielfachheiten nennt man auch *arithmetische Vielfachheiten*.

Beispiel B1.7.
Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C})$$

aus Beispiel B1.5 hat nur den Eigenwert 0, da $\chi_A = t^2$. Die algebraische Vielfachheit von 0 als Eigenwert von A ist

$$\mathrm{mult}(\chi_A, 0) = \mathrm{mult}(t^2, 0) = 2,$$

während die geometrische Vielfachheit

$$\dim_{\mathbb{C}} \mathrm{Eig}(A, 0) = \dim_{\mathbb{C}} \mathrm{Lös}(A, 0) = \dim_{\mathbb{C}} \mathrm{Lin}\left((1, 0)^t\right) = 1$$

ist.

Lemma B1.8 (Geometrische und algebraische Vielfachheit).
Es sei $f \in \mathrm{End}_K(V)$, $A \in \mathrm{Mat}_n(K)$ und $\lambda \in K$. Dann gilt stets

$$\dim_K \mathrm{Eig}(f, \lambda) \leq \mathrm{mult}(\chi_f, \lambda) \quad \text{und} \quad \dim_K \mathrm{Eig}(A, \lambda) \leq \mathrm{mult}(\chi_A, \lambda),$$

d.h. die geometrische Vielfachheit eines Eigenwertes ist stets nach oben durch die algebraische Vielfachheit beschränkt.

Beweis: Man beachte, daß $U := \mathrm{Eig}(f, \lambda)$ ein f-invarianter Unterraum ist und daß $f_U = \lambda \cdot \mathrm{id}_U$ gilt. Mithin ist

$$\chi_{f_U} = \chi_{\lambda \cdot \mathrm{id}_U} = \det(t \cdot \mathrm{id}_U - \lambda \cdot \mathrm{id}_U) = \det\left((t - \lambda) \cdot \mathrm{id}_U\right) = (t - \lambda)^s$$

wobei $s = \dim_K(U) = \dim_K \mathrm{Eig}(f, \lambda)$. Außerdem gilt nach Aufgabe 35.27

$$\chi_f = \chi_{f_U} \cdot \chi_{f_{V/U}} = (t - \lambda)^s \cdot \chi_{f_{V/U}}.$$

Daraus folgt unmittelbar

$$\mathrm{mult}(f, \lambda) \geq s = \dim_K \mathrm{Eig}(f, \lambda).$$

Die analoge Aussage für A folgt hieraus mit $f = f_A$. □

Lemma B1.9 (Eigenwerte bei konjugierten Matrizen).
Für $A, B \in \mathrm{Mat}_n(K)$ und $T \in \mathrm{Gl}_n(K)$ mit $B = T^{-1} \circ A \circ T$ sowie $\lambda \in K$ gelten:

a. $\sigma(A) = \sigma(B)$.

b. $\mathrm{mult}(\chi_A, \lambda) = \mathrm{mult}(\chi_B, \lambda)$.

c. $\dim_K \mathrm{Eig}(A, \lambda) = \dim_K \mathrm{Eig}(B, \lambda)$.

d. $x \in \mathrm{Eig}(A, \lambda) \iff T^{-1}x \in \mathrm{Eig}(B, \lambda)$.

D.h. konjugierte Matrizen haben die gleichen Eigenwerte und für jeden Eigenwert stimmen ihre geometrischen Vielfachheiten ebenso überein wie ihre algebraischen Vielfachheiten.

Beweis: Nach Proposition 35.8 haben A und B die gleichen charakteristischen Polynome. Mithin stimmen wegen Satz 35.15 die Eigenwerte von A und B sowie deren algebraische Vielfachheiten überein. Damit sind a. und b. gezeigt. Ferner gilt

$$\begin{aligned} x \in \mathrm{Eig}(A,\lambda) &\iff \lambda x = Ax = ATT^{-1}x \\ &\iff \lambda T^{-1}x = T^{-1}ATT^{-1}x = BT^{-1}x \iff T^{-1}x \in \mathrm{Eig}(B,\lambda). \end{aligned}$$

Damit ist d. gezeigt und außerdem folgt, daß der Isomorphismus $f_{T^{-1}}$ den Eigenraum $\mathrm{Eig}(A,\lambda)$ isomorph auf den Eigenraum $\mathrm{Eig}(B,\lambda)$ abbildet. Die beiden müssen also die gleiche Dimension haben, womit auch c. gezeigt ist. □

C) Diagonalblockmatrizen

Definition B1.10 (Diagonalblockmatrizen).
Wir werden im Folgenden sehr häufig mit Blockmatrizen der folgenden Form arbeiten:

$$A = \left(\begin{array}{c|ccc} A_1 & 0 & \cdots & 0 \\ \hline 0 & \boxed{A_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{A_r} \end{array}\right) \in \mathrm{Mat}_n(K),$$

wobei $A_i \in \mathrm{Mat}_{n_i}(K)$, $i = 1,\ldots,r$ mit $n = n_1 + \ldots + n_r$. Es empfiehlt sich deshalb, eine Kurzschreibweise für solche *Diagonalblockmatrizen* einzuführen. Wir schreiben kurz:

$$A = A_1 \oplus \ldots \oplus A_r = \bigoplus_{i=1}^{r} A_i.$$

Bemerkung B1.11 (Diagonalblockmatrizen).

a. Man beachte, daß es bei der obigen Schreibweise für Diagonalblockmatrizen auf die Reihenfolge der Summation ankommt, daß aber Matrizen, die durch Änderung der Summationsreihenfolge entstehen, zueinander konjugiert sind!

b. Mit Hilfe dieser Notation gilt beispielsweise, daß eine Matrix A genau dann diagonalisierbar ist, wenn es Körperelemente $\lambda_1,\ldots,\lambda_r \in K$ und positive natürliche

Zahlen $n_1, \ldots, n_r \in \mathbb{N}$ gibt sowie eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ mit

$$T^{-1} \circ A \circ T = \bigoplus_{i=1}^{r} \lambda_i \mathbb{1}_{n_i}.$$

c. Ist $A = \bigoplus_{i=1}^{r} A_i$ eine Diagonalblockmatrix, so verifiziert man leicht, daß für $k \in \mathbb{N}$ gilt $A^k = \bigoplus_{i=1}^{r} A_i^k$, und damit, daß für ein Polynom $p \in K[t]$ gilt

$$p(A) = \bigoplus_{i=1}^{r} p(A_i).$$

Insbesondere gilt also für eine Diagonalmatrix $D = \bigoplus_{i=1}^{n} \lambda_i \mathbb{1}_1$, daß

$$p(D) = \bigoplus_{i=1}^{n} p(\lambda_i)\mathbb{1}_1 = \begin{pmatrix} p(\lambda_1) & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & p(\lambda_n) \end{pmatrix}$$

In der Tat kann man sogar zeigen, daß für eine Blockmatrix der Form

$$A = \left(\begin{array}{c|ccc} A_1 & * & \cdots & * \\ \hline 0 & \boxed{A_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{A_r} \end{array}\right) \in \mathrm{Mat}_n(K),$$

gilt, daß

$$p(A) = \left(\begin{array}{c|ccc} p(A_1) & * & \cdots & * \\ \hline 0 & \boxed{p(A_2)} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{p(A_r)} \end{array}\right) \in \mathrm{Mat}_n(K),$$

wobei sich die Sterne oberhalb der Blöcke verändert haben.
Damit gilt insbesondere, daß $p(A)$ eine obere Dreiecksmatrix ist, falls A eine solche war.

D) Der Satz von Cayley-Hamilton

Da $\dim_K \big(\mathrm{Mat}_n(K)\big) = n^2$ gilt, sind die $n^2 + 1$ Matrizen

$$\mathbb{1}_n = A^0, A^1, A^2, \ldots, A^{n^2}$$

in $\mathrm{Mat}_n(K)$ linear abhängig. D. h. es existieren $\lambda_0, \ldots, \lambda_{n^2} \in K$, nicht alle null, mit

$$\lambda_0 A^0 + \lambda_1 A^1 + \ldots + \lambda_{n^2} A^{n^2} = 0 \in \mathrm{Mat}_n(K).$$

Ein einfaches Dimensionsargument zeigt also, es gibt ein Polynom $0 \neq p = \lambda_{n^2} t^{n^2} + \ldots + \lambda_0 \in K[t]$ vom Grad kleiner gleich n^2 mit $p(A) = 0$. Der folgende wichtige Satz

von Cayley-Hamilton besagt nun, daß es sogar ein Polynom vom Grad n gibt, das A annulliert.

Satz B1.12 (Cayley-Hamilton).
Für $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$ gilt $\chi_f(f) = 0$ und $\chi_A(A) = 0$.

Beweis: Da für eine Basis D von V die Abbildung $M_D^D : \mathrm{End}_K(V) \longrightarrow \mathrm{Mat}_n(K)$ ein K-Algebrenhomomorphismus ist, gilt

$$M_D^D\big(\chi_f(f)\big) = \chi_f\big(M_D^D(f)\big).$$

Dann gilt aber $\chi_f(f) = 0$ genau dann, wenn

$$0 = M_D^D\big(\chi_f(f)\big) = \chi_f\big(M_D^D(f)\big) = \chi_{M_D^D(f)}\big(M_D^D(f)\big).$$

Es reicht deshalb, die Aussage für Matrizen zu beweisen.

Betrachte dazu die Matrix

$$B_t := t \cdot \mathbb{1}_n - A \in \mathrm{Mat}_n(K[t])$$

sowie die Adjunkte $B_t^\# = (p_{ij}) \in \mathrm{Mat}_n(K[t])$ von B_t, die auch *Busadjunkte* von A genannt wird. Nach dem Satz über die Adjunkte 34.29 in $\mathrm{Mat}_n(K[t])$ gilt die Adjunktengleichung

$$B_t \circ B_t^\# = (t\mathbb{1}_n - A) \circ (t\mathbb{1}_n - A)^\# = \det(t\mathbb{1}_n - A) \cdot \mathbb{1}_n = \chi_A \cdot \mathbb{1}_n. \tag{144}$$

Man beachte nun noch, daß die Einträge von $B_t^\#$ Determinanten von gewissen $(n-1) \times (n-1)$-Matrizen von B_t sind, also Polynome vom Grad höchstens $n-1$. Wir können nun $B_t^\#$ auch als Polynom schreiben, dessen Koeffizienten Matrizen sind, und dieses Polynom hat dann höchstens den Grad $n-1$, d. h. es gibt Matrizen $B_0, \ldots, B_{n-1} \in \mathrm{Mat}_n(K)$ mit

$$B_t^\# = B_{n-1}t^{n-1} + \ldots + B_1 t + B_0.$$

Ist $\chi_A = t^n + \alpha_{n-1}t^{n-1} + \ldots + \alpha_0$, so folgt aus der Adjunktengleichung (144)

$$(\mathbb{1}_n t - A) \circ \big(B_{n-1}t^{n-1} + \ldots + B_1 t + B_0\big) \stackrel{!}{=} \mathbb{1}_n t^n + \alpha_{n-1}\mathbb{1}_n t^{n-1} + \ldots + \alpha_0 \mathbb{1}_n \tag{145}$$

durch Koeffizientenvergleich für die t^i, $i = 0, \ldots, n$:

$$\begin{aligned} B_{n-1} &= \mathbb{1}_n \\ -AB_{n-1} + B_{n-2} &= \alpha_{n-1}\mathbb{1}_n \\ -AB_{n-2} + B_{n-3} &= \alpha_{n-2}\mathbb{1}_n \\ &\vdots \\ -AB_1 + B_0 &= \alpha_1 \mathbb{1}_n \\ -AB_0 &= \alpha_0 \mathbb{1}_n \end{aligned} \tag{146}$$

Multipliziert man die i-te Zeile in (146) mit A^{n-i+1} und summiert die beiden Seiten auf, so erhält man die Behauptung:

$$\begin{aligned} A^n B_{n-1} &= A^n \\ -A^n B_{n-1} + A^{n-1} B_{n-2} &= \alpha_{n-1} A^{n-1} \\ -A^{n-1} B_{n-2} + A^{n-2} B_{n-3} &= \alpha_{n-2} A^{n-2} \\ &\vdots \\ -A^2 B_1 + A B_0 &= \alpha_1 A \\ -A B_0 &= \alpha_0 \mathbb{1}_n \\ \hline 0 &= \chi_A(A). \end{aligned}$$

□

Bemerkung B1.13.

a. Man beachte, daß der folgende *offensichtliche* Beweis für $\chi_A(A) = 0$, nämlich

$$\chi_A(A) = \det(A * \mathbb{1}_n - A) = \det(0) = 0$$

falsch ist, da "$*$" beim Einsetzen von A in $\det(t\mathbb{1}_n - A) \in K[t]$ eben *nicht* die Matrixmultiplikation ist! Man beachte ferner, daß die Gleichung auch schon deshalb keinen Sinn ergeben kann, da $\chi_A(A)$ die Nullmatrix ist, während $\det(0)$ die Null in K ist.

b. Kennt man das charakteristische Polynom $\chi_A = t^n + \alpha_{n-1} t^{n-1} + \ldots + \alpha_0$, so läßt sich daraus mittels (145) und der Rekursionsformel (146) die Busadjunkte

$$(t\mathbb{1}_n - A)^{\#} = B_{n-1} t^{n-1} + \ldots + B_1 t + B_0$$

von A bestimmen. Für die B_{n-k}, $k = 1, \ldots, n$, gilt dabei explizit:

$$B_{n-k} = A^{k-1} + \alpha_{n-1} A^{k-2} + \ldots + \alpha_{n-k+1} A^0,$$

und wegen $\alpha_0 = (-1)^n \cdot \det(A)$ erhalten wir die Adjunkte von A als

$$A^{\#} = (-1)^{n+1} \cdot B_0 = (-1)^{n+1} \cdot \left(A^{n-1} + \alpha_{n-1} A^{n-2} + \ldots + \alpha_1 A^0\right).$$

Diese Formel zur Berechnung der Adjunkten von A ist weit effizienter, als die Determinanten sämtlicher Streichungsmatrizen zu berechnen.

E) Das Minimalpolynom

Das Minimalpolynom eines Endomorphismus bzw. einer quadratischen Matrix sind Spezialfälle der folgenden allgemeineren Situation.

Proposition B1.14 (Das Minimalpolynom).
Es sei L eine K-Algebra und $b \in L$.

a. Der *Einsetzhomomorphismus*

$$\phi_b : K[t] \longrightarrow L : f \mapsto f(b)$$

ist ein K-Algebrenhomomorphismus, d.h. $(f+g)(b) = f(b)+g(b)$, $(f\cdot g)(b) = f(b)\cdot g(b)$, $(\lambda\cdot f)(b) = \lambda\cdot f(b)$ und $1(b) = 1$.

b. Es gibt ein eindeutig bestimmtes normiertes Polynom $\mu_b \in K[t]$, so daß

$$\mu_b K[t] := \{\mu_b \cdot g \mid g \in K[t]\} \stackrel{!}{=} \{h \in K[t] \mid h(b) = 0\} =: \operatorname{Ker}(\phi_b).$$

μ_b heißt das *Minimalpolynom* von b.

c. Gibt es ein $0 \neq h \in K[t]$ mit $h(b) = 0$, so ist μ_b das normierte Nicht-Null-Polynom kleinsten Grades, das b als Nullstelle hat.

Beweis: Daß der Einsetzhomomorphismus ein K-Algebrenhomomorphismus ist, folgt unmittelbar aus den Definitionen. (siehe auch Lemma A6.36. Mithin ist der Kern

$$\operatorname{Ker}(\phi_b) = \{h \in K[t] \mid h(b) = 0\}$$

von ϕ_b ein Ideal im Ring $K[t]$, siehe Satz A5.42. Aus Korollar A6.51 wissen wir, daß $K[t]$ ein Hauptidealring ist. Es gibt also ein normiertes Polynom $\mu_b \in K[t]$ mit

$$\operatorname{Ker}(\phi_b) = \mu_b K[t] = \{\mu_b \cdot g \mid g \in K[t]\}.$$

Die Eindeutigkeit von μ_b folgt dann leicht aus den Kürzungsregeln in $K[t]$. Zudem folgt aus der Gradformel unmittelbar, daß jedes Nicht-Null-Polynom in $\mu_b K[t]$ mindestens den Grad von μ_b hat. □

Daraus folgt dann der folgende Satz unmittelbar.

Satz B1.15 (Das Minimalpolynom).
Es sei $f \in \operatorname{End}_K(V)$ und $A \in \operatorname{Mat}_n(K)$.

a. Es gibt ein eindeutiges normiertes Polynom $0 \neq \mu_f \in K[t]$ mit $\mu_f(f) = 0$, so daß

$$\mu_f \cdot K[t] = \{p \in K[t] \mid p(f) = 0\}.$$

Insbesondere ist μ_f das Nicht-Null-Polynom kleinsten Grades mit $\mu_f(f) = 0$. Wir nennen μ_f das *Minimalpolynom* von f.

b. Es gibt ein eindeutiges normiertes Polynom $0 \neq \mu_A \in K[t]$ mit $\mu_A(A) = 0$, so daß

$$\mu_A \cdot K[t] = \{p \in K[t] \mid p(A) = 0\}.$$

Insbesondere ist μ_A das Nicht-Null-Polynom kleinsten Grades mit $\mu_A(A) = 0$. Wir nennen μ_A das *Minimalpolynom* von A.

Beweis: Die Aussage folgt unmittelbar aus Proposition B1.14, da $\mathrm{End}_K(V)$ sowie $\mathrm{Mat}_n(K)$ beides K-Algebren sind. Man beachte dabei, daß das Minimalpolynom nicht Null ist, da nach dem Satz von Cayley-Hamilton der Kern des Einsetzhomomorphismus nicht Null ist. □

Bemerkung B1.16 (Minimalpolynome konjugierter Matrizen).

a. Sei $f \in \mathrm{End}_K(V)$, B eine Basis von V und $p \in K[t]$, dann gilt

$$M_B^B\big(p(f)\big) = p\big(M_B^B(f)\big),$$

da M_B^B ein K-Algebrenhomomorphismus ist.
Insbesondere gilt daher $p(f) = 0$ genau dann, wenn $p\big(M_B^B(f)\big) = 0$, und deshalb

$$\mu_f = \mu_{M_B^B(f)}.$$

Entsprechend gilt dann auch $\mu_{f_A} = \mu_{M_E^E(f_A)} = \mu_A$, wobei E die kanonische Basis von K^n bezeichnet.

b. Konjugierte Matrizen haben dasselbe Minimalpolynom, denn wegen Aufgabe 35.28 und Satz B1.15 gilt für konjugierte Matrizen $A, B \in \mathrm{Mat}_n(K)$

$$\mu_A \cdot K[t] = \{p \in K[t] \mid p(A) = 0\} = \{p \in K[t] \mid p(B) = 0\} = \mu_B \cdot K[t].$$

Korollar B1.17 (Primfaktorzerlegung von χ und μ).
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Hat das charakteristische Polynom von f die Primfaktorzerlegung

$$\chi_f = p_1^{n_1} \cdot \ldots \cdot p_r^{n_r},$$

so hat das Minimalpolynom von f eine Primfaktorzerlegung der Form

$$\mu_f = p_1^{m_1} \cdot \ldots \cdot p_r^{m_r}$$

mit $1 \le m_i \le n_i$ für $i = 1, \ldots, r$.

b. Hat das charakteristische Polynom von A die Primfaktorzerlegung

$$\chi_A = p_1^{n_1} \cdot \ldots \cdot p_r^{n_r},$$

so hat das Minimalpolynom von A eine Primfaktorzerlegung der Form

$$\mu_A = p_1^{m_1} \cdot \ldots \cdot p_r^{m_r}$$

mit $1 \le m_i \le n_i$ für $i = 1, \ldots, r$.

Beweis: Der Satz von Cayley-Hamilton besagt, daß $\chi_A \in \mathrm{Ker}(\phi_A) = \mu_A K[t]$ gilt. Also gibt es ein $h \in K[t]$ mit $p_1^{n_1} \cdot \ldots \cdot p_r^{n_r} = \chi_A = \mu_A \cdot h$.

Daraus folgt $\mu_A = p_1^{m_1} \cdots p_r^{m_r}$ für $0 \leq m_i \leq n_i$, $i = 1, \ldots, r$, geeignet. Wir müssen zeigen, daß jedes p_i in μ_A auch vorkommt, d. h., daß $m_i \geq 1$ für alle $i = 1, \ldots, r$.

Nehmen wir an, daß es ein i mit $m_i = 0$ gibt. Dann sind μ_A und p_i teilerfremde Polynome, also gibt es wegen der Bézout-Identität A6.54 Polynome $p, q \in K[t]$ mit

$$1 = p \cdot \mu_A + q \cdot p_i.$$

Wir führen die Annahme zum Widerspruch, indem wir zum algebraischen Abschluß $\overline{K}$ von K übergehen. Da p_i vom Grad $\deg(p_i) \geq 1$ ist, besitzt es eine Nullstelle $\lambda \in \overline{K}$. Setzen wir λ in der obigen Gleichung für t ein, so erhalten wir

$$1 = p(\lambda) \cdot \mu_A(\lambda) + q(\lambda) \cdot p_i(\lambda) = p(\lambda) \cdot \mu_A(\lambda).$$

Also muß $\mu_A(\lambda) \neq 0$ gelten.

Wegen $p_i(\lambda) = 0$, ist dann aber auch $\chi_A(\lambda) = 0$ und somit ist λ ein Eigenwert von $A \in \mathrm{Mat}_n\left(\overline{K}\right)$. Sei nun $0 \neq x \in \overline{K}^n$ ein Eigenvektor von A zum Eigenwert λ. Dann gilt für das Polynom $\mu_A = \sum_{i=0}^m a_i t^i \in K[t] \subseteq \overline{K}[t]$

$$\mu_A(A)x = \sum_{i=0}^{m} a_i\left(A^i x\right) = \sum_{i=0}^{m} a_i\left(\lambda^i x\right) = \mu_A(\lambda) \cdot x \neq 0,$$

im Widerspruch zu $\mu_A(A) = 0$. Also muß $m_i \geq 1$ gelten.

Die analoge Aussage für f folgt aus der entsprechenden Aussage für eine Matrixdarstellung $M_B^B(f)$. □

Korollar B1.18 (Die Eigenwerte sind die Nullstellen des Minimalpolynoms.).
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Die Eigenwerte von f sind genau die Nullstellen des Minimalpolynoms μ_f, und χ_f zerfällt genau dann über K in Linearfaktoren, wenn μ_f zerfällt.

b. Die Eigenwerte von A sind genau die Nullstellen des Minimalpolynoms μ_A, und χ_A zerfällt genau dann über K in Linearfaktoren, wenn μ_A zerfällt.

Beweis: Genau dann ist $\lambda \in K$ ein Eigenwert von f, wenn $t - \lambda$ ein Primfaktor von χ_f ist, und dies ist wegen Korollar B1.17 genau dann der Fall, wenn $t - \lambda$ ein Primfaktor von μ_f ist, d.h. wenn λ eine Nullstelle von μ_f ist. Wegen Korollar B1.17 zerfällt χ_f zudem genau dann über K in Linearfaktoren, wenn μ_f in Linearfaktoren zerfällt. Die entsprechenden Aussagen für eine Matrix A folgen aus a. mit $f = f_A$.

□

Beispiel B1.19.

a. Ist $A = \lambda \mathbb{1}_n \in \mathrm{Mat}_n(K)$ eine Diagonalmatrix mit gleichen Diagonalelementen, so gilt wegen $\lambda \mathbb{1}_n - A = 0$ offenbar $\chi_A = (t-\lambda)^n$ und $\mu_A = t - \lambda$.

b. Sei $\lambda \in K$ und

$$J := J_n(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix} \in \mathrm{Mat}_n(K),$$

d. h. $J_n(\lambda)$ hat auf der Hauptdiagonalen den Wert λ und auf der oberen Nebendiagonalen Einsen stehen, ansonsten nur Nullen. Wir nennen $J_n(\lambda)$ einen *Jordanblock* (oder eine *Jordanzelle* oder ein *Jordankästchen*) der Größe n zum Eigenwert λ . Offenbar gilt wieder

$$\chi_J = (t-\lambda)^n.$$

Nach Korollar B1.17 ist mithin $\mu_J = (t-\lambda)^m$ für ein $1 \le m \le n$. Dabei ist m die kleinste natürliche Zahl mit $(J - \lambda \mathbb{1}_n)^m = 0$. Nun ist aber

$$J - \lambda \mathbb{1}_n = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix} =: N$$

und man sieht mittels einer einfachen Induktion, daß $N^k \neq 0$ für $k < n$, aber $N^n = 0$ (vgl. Aufgabe 27.16). Also gilt

$$\mu_J = (t-\lambda)^n.$$

c. Ist $A = A_1 \oplus \ldots \oplus A_r \in \mathrm{Mat}_n(K)$ eine Diagonalblockmatrix mit $A_i \in \mathrm{Mat}_{n_i}(K)$, so folgt aus der Definition des charakteristischen Polynoms unmittelbar (vgl. Aufgabe 35.27)

$$\chi_A = \prod_{i=1}^{r} \chi_{A_i}.$$

Die entsprechende Formel für das Minimalpolynom gilt nicht. Sei etwa $A_1 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{Mat}_2(K)$ und $A_2 = (1) \in \mathrm{Mat}_1(K)$, dann gilt für $A = A_1 \oplus A_2$

$$\mu_A = (t-1)^2 \neq (t-1)^3 = \mu_{A_1} \cdot \mu_{A_2}.$$

Man kann zeigen, daß für eine Diagonalblockmatrix wie oben μ_A ein kleinstes gemeinsames Vielfaches von $\mu_{A_1}, \ldots, \mu_{A_r}$ im Sinne der Vorlesung Algebraische Strukturen ist. Darauf wollen wir hier nicht näher eingehen.

Bemerkung B1.20 (Berechnung des Minimalpolynoms).
Zur praktischen Berechnung des Minimalpolynoms von $A \in \mathrm{Mat}_n(K)$ kann man wie folgt vorgehen. Aufgrund des Satzes von Cayley-Hamilton wissen wir, daß die Matrizen $A^0, \ldots, A^n$ linear abhängig sind. Fassen wir die Matrix A^i als einen *langen* Spaltenvektor in K^{n^2} auf und bezeichnen wir diesen mit x_i, dann suchen wir das minimale m, so daß $x_0, \ldots, x_m$ linear abhängig sind, und wir suchen ferner geeignete $\beta_0, \ldots, \beta_{m-1}$ mit

$$x_m + \beta_{m-1}x_{m-1} + \ldots + \beta_0 x_0 = 0.$$

Dies ist dann gleichbedeutend damit, daß

$$t^m + \beta_{m-1}t^{m-1} + \ldots + \beta_0 \in K[t]$$

das gesuchte Minimalpolynom von A ist.

Bezeichne $X = (x_0 \ldots x_n) \in \mathrm{Mat}\left(n^2 \times (n+1), K\right)$ die Matrix, deren Spalten $x_0, \ldots, x_n$ sind, dann suchen wir eine Lösung des linearen Gleichungssystems

$$X \cdot \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix} = 0 \in K^{n^2} \tag{147}$$

mit $\beta_{m+1} = \ldots = \beta_n = 0$ und $\beta_m = 1$ und so, daß m minimal mit dieser Eigenschaft ist. Da $(x_0, \ldots, x_{m-1})$ nach Definition von m linear unabhängig, $(x_0, \ldots, x_m)$ aber linear abhängig ist, bedeutet dies, daß in einer ZSF von X die Zahlen $1, \ldots, m$ Pivotindizes sind, während $m+1$ kein Pivotindex mehr ist.

Berechnet man eine Basis des Lösungsraums von (147) mittels des Algorithmus 33.9, so erhalten wir den gesuchten Koeffizientenvektor β als das Negative des ersten Basisvektors, d.h. des ersten Vektors mit einer -1 auf der Diagonalen.

Dies führt zu folgendem Algorithmus zur Berechnung des Minimalpolynoms einer Matrix $A \in \mathrm{Mat}_n(K)$.

Algorithmus B1.21 (Algorithmus zur Berechnung des Minimalpolynoms).
INPUT: $A \in \mathrm{Mat}_n(K)$
OUTPUT: μ_A

1. **Schritt:** Falls A nicht quadratisch ist, gib 0 zurück.
2. **Schritt:** Bilde die Potenzen $A^0, \ldots, A^n$ und schreibe die Matrizen in Form von Spaltenvektoren der Länge n^2 in eine Matrix $X \in \mathrm{Mat}\left(n^2 \times (n+1), K\right)$.
3. **Schritt:** Berechne eine Basis von $\mathrm{Lös}(X, 0)$.
4. **Schritt:** Verwende die Negativen der Koeffizienten des ersten Basisvektors als Koeffizienten eines Polynoms und gib dieses zurück.

F) Die Hauptraumzerlegung

Für unsere weiteren Betrachtungen brauchen wir einen neuen Begriff, der auch im folgenden Abschnitt für die Jordansche Normalform von Bedeutung sein wird. Für $\lambda \in K$ haben wir aufsteigende Ketten von Unterräumen von V (vgl. Aufgabe 30.27)

$$\operatorname{Ker}\left(f - \lambda \operatorname{id}_V\right) \subseteq \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^2\right) \subseteq \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^3\right) \subseteq \ldots \subseteq V$$

und

$$\operatorname{Lös}\left(A - \lambda \mathbb{1}_n, 0\right) \subseteq \operatorname{Lös}\left((A - \lambda \mathbb{1}_n)^2, 0\right) \subseteq \operatorname{Lös}\left((A - \lambda \mathbb{1}_n)^3, 0\right) \subseteq \ldots \subseteq V$$

Die Vereinigung all dieser Unterräume ist offenbar wieder ein Unterraum und führt zu folgender Definition.

Definition B1.22 (Hauptraum).
Es sei $f \in \operatorname{End}_K(V)$, $A \in \operatorname{Mat}_n(K)$ und $\lambda \in K$. Dann heißen

$$\operatorname{Hau}(f, \lambda) = \bigcup_{k \in \mathbb{N}} \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^k\right) \quad \text{und} \quad \operatorname{Hau}(A, \lambda) = \bigcup_{k \in \mathbb{N}} \operatorname{Lös}\left((A - \lambda \mathbb{1}_n)^k, 0\right)$$

der *Hauptraum* oder *verallgemeinerte Eigenraum* von f bzw. A zu λ.

Lemma B1.23 (Nilpotenzindex und Hauptraum).
Es sei $\lambda \in K$ gegeben.

a. Es gibt eine natürliche Zahl $0 \leq m \leq n$ mit

$$\operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^0\right) \subsetneqq \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^1\right) \subsetneqq \ldots \subsetneqq \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^m\right)$$

und für $k > m$

$$\operatorname{Hau}(f, \lambda) = \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^m\right) = \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^k\right).$$

Die Zahl m heißt *Nilpotenzindex* von $f - \lambda \operatorname{id}_V$ und erfüllt $m \leq \operatorname{mult}(\mu_f, \lambda)$.

b. Für jedes $k \in \mathbb{N}$ ist $\operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^k\right)$ ein f-invarianter Unterraum von V. Insbesondere sind also Eigenräume und Haupträume von f auch f-invariant.

Die entsprechenden Aussagen für eine Matrix $A \in \operatorname{Mat}_n(K)$ gelten analog.

Beweis: Durch Betrachtung von f_A ergibt sich die Aussage für eine Matrix A unmittelbar aus der entsprechenden Aussage für Endomorphismen.

a. Aus Aufgabe 30.27 wissen wir, daß es eine natürliche Zahl $0 \leq m \leq n$ gibt mit

$$\operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^0\right) \subsetneqq \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^1\right) \subsetneqq \ldots \subsetneqq \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^m\right)$$

und

$$\operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^m\right) = \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^k\right)$$

für $k > m$. Aus der Definition des Hauptraumes folgt dann unmittelbar

$$\mathrm{Hau}(f,\lambda) = \bigcup_{k\in\mathbb{N}} \mathrm{Ker}\left((f-\lambda\,\mathrm{id}_V)^k\right) = \mathrm{Ker}\left((f-\lambda\,\mathrm{id}_V)^m\right).$$

Es bleibt zu zeigen, daß

$$m \le \mathrm{mult}(\mu_f,\lambda).$$

Angenommen, $l := \mathrm{mult}(\mu_f,\lambda) < m$. Dann gibt es ein $x \in \mathrm{Ker}\left((f-\lambda\,\mathrm{id}_V)^m\right)$ mit $y := (f-\lambda\,\mathrm{id}_V)^l(x) \neq 0$. Da λ eine l-fache Nullstelle von μ_f ist, gibt es ein $h \in K[t]$ mit

$$\mu_f = h\cdot(t-\lambda)^l,$$

wobei λ keine Nullstelle von h ist. Deshalb sind h und $(t-\lambda)^{m-l}$ teilerfremd und nach der Bézout-Identität A6.54 gibt es Polynome $p,q \in K[t]$ mit

$$p\cdot(t-\lambda)^{m-l} + q\cdot h = 1.$$

Es folgt

$$\big(p(f)\circ(f-\lambda\,\mathrm{id}_V)^{m-l} + q(f)\circ h(f)\big)(y) = \mathrm{id}_V(y) = y. \tag{148}$$

Andererseits gilt aber $(f-\lambda\,\mathrm{id}_V)^{m-l}(y) = (f-\lambda\,\mathrm{id}_V)^m(x) = 0$ sowie

$$h(f)(y) = \big(h(f)\circ(f-\lambda\,\mathrm{id}_V)^l\big)(x) = \mu_f(f)(x) = 0.$$

Aus (148) folgt damit $y = 0$, im Widerspruch zur Voraussetzung. Damit haben wir $l \ge m$ gezeigt.

b. Da f mit Potenzen von f und mit der Identität vertauschbar ist, gilt für $k \in \mathbb{N}$ und $x \in \mathrm{Ker}\left((f-\lambda\,\mathrm{id}_V)^k\right)$

$$(f-\lambda\,\mathrm{id}_V)^k\big(f(x)\big) = f\big((f-\lambda\,\mathrm{id}_V)^k(x)\big) = f(0) = 0,$$

woraus die Behauptung folgt.

□

Satz B1.24 (Hauptraumzerlegung).
Es sei $f \in \mathrm{End}_K(V)$ so, daß χ_f über K in Linearfaktoren zerfällt, d. h. es gibt paarweise verschiedene $\lambda_1,\dots,\lambda_r \in K$ und $0 < m_i \le n_i$, so daß

$$\chi_f = (t-\lambda_1)^{n_1}\cdots(t-\lambda_r)^{n_r} \quad \text{und} \quad \mu_f = (t-\lambda_1)^{m_1}\cdots(t-\lambda_r)^{m_r}.$$

Dann gelten:

a. $V = \mathrm{Hau}(f,\lambda_1)\oplus\ldots\oplus\mathrm{Hau}(f,\lambda_r)$,

b. $n_i = \mathrm{mult}(\chi_f,\lambda_i) = \dim_K\big(\mathrm{Hau}(f,\lambda_i)\big)$ und

c. $m_i = \mathrm{mult}(\mu_f,\lambda_i)$ ist der Nilpotenzindex von $f-\lambda_i\,\mathrm{id}_V$.

Die analogen Aussagen für eine Matrix $A \in \mathrm{Mat}_n(K)$, deren charakteristisches Polynom zerfällt, gelten analog.

Beweis: Wir setzen $V_i := \mathrm{Hau}(f, \lambda_i)$, wählen eine Basis B_i von V_i und setzen

$$q_i := \frac{\mu_f}{(t-\lambda_i)^{m_i}} = \prod_{j \neq i} (t-\lambda_j)^{m_j}.$$

Zeige: $V = V_1 + \ldots + V_r$ und $B = B_1 \cup \ldots \cup B_r$ erzeugt V.

Offenbar gibt es keinen Primfaktor, den alle $q_1, \ldots, q_r$ gemeinsam haben, so daß aus der Bézout-Identität A6.54 die Existenz von Polynomen $p_1, \ldots, p_r \in K[t]$ folgt mit

$$p_1 q_1 + \ldots + p_r q_r = 1.$$

Setzt man $Q_i := q_i p_i$, dann folgt

$$Q_1(f) + \ldots + Q_r(f) = \mathrm{id}_V .$$

Sei $x \in V$ beliebig. Wegen $(f - \lambda_i \mathrm{id}_V)^{m_i} \circ Q_i(f) = p_i(f) \circ \mu_f(f) = 0$ gilt

$$Q_i(f)(x) \in \mathrm{Ker}(f - \lambda_i \mathrm{id}_V)^{m_i} = V_i,$$

und somit

$$x = Q_1(f)(x) + \ldots + Q_r(f)(x) \in V_1 + \ldots + V_r.$$

Damit ist $V = V_1 + \ldots + V_r$ gezeigt, und insbesondere ist $B = B_1 \cup \ldots \cup B_r$ ein Erzeugendensystem von V, da die B_i die V_i erzeugen.

Zeige: $\dim_K(V_i) = |B_i| \leq n_i$.

Wir wissen aus Lemma B1.23, daß V_i ein f-invarianter Unterraum ist, und aus der Definition von $V_i = \mathrm{Ker}\left((f - \lambda \mathrm{id}_V)^{m_i}\right)$ folgt unmittelbar

$$(f_{V_i} - \lambda_i \mathrm{id}_{V_i})^{m_i} = 0.$$

Also muß nach Satz B1.15 $(t-\lambda_i)^{m_i}$ ein Vielfaches des Minimalpolynoms von f_{V_i} sein, d.h. $\mu_{f_{V_i}} = (t-\lambda_i)^m$ für ein $0 \leq m \leq m_i$. Wegen Korollar B1.17 gilt dann aber auch

$$\chi_{f_{V_i}} = (t-\lambda_i)^k$$

und dabei muß dann $k = \dim_K(V_i) = |B_i|$ gelten. Aus Aufgabe 35.27 wissen wir, daß χ_f ein Vielfaches von $\chi_{f_{V_i}}$ ist, und mithin muß $k \leq n_i$ gelten.

Zeige: $V = V_1 \oplus \ldots \oplus V_r$ und $n_i = \dim_K(V_i)$ (d.h. Teil a. und b.).

Da B ein Erzeugendensystem von V ist und da $|B_i| \leq n_i$ gilt, erhalten wir

$$n \leq |B| \leq |B_1| + \ldots + |B_r| \leq n_1 + \ldots + n_r = \deg(\chi_f) = n.$$

Dies zeigt, daß alle beteiligten Ungleichheitszeichen in der Tat Gleichheitszeichen waren, d.h. $\dim_K(V_i) = |B_i| = n_i$ für $i = 1, \ldots, n$. Aber als Erzeugendensystem mit genau $n = \dim_K(V)$ Elementen ist B dann eine Basis von V und aus Aufgabe 29.26 folgt dann, daß V die direkte Summe der V_i ist.

Zeige: m_i ist der Nilpotenzindex von $f - \lambda_i \mathrm{id}_V$ (d.h. Teil c.).

Für den Nilpotenzindex m von $f - \lambda_i \mathrm{id}_V$ gilt nach Lemma B1.23 $m_i \geq m$.

Angenommen, $m_i > m$. Sei nun $x = x_1 + \ldots + x_r \in V$ beliebig mit $x_j \in V_j$ für $j = 1, \ldots, r$. Dann gilt $(f - \lambda_i \mathrm{id}_V)^m(x_i) = 0$ und für $j \neq i$ gilt $q_i(f)(x_j) = 0$, also folgt für $p := (t - \lambda_i)^m q_i \in K[t]$

$$p(f)(x) = 0.$$

Also ist $p(f) = 0$ die Nullabbildung, aber wegen $0 \leq \deg(p) < \deg(\mu_f)$ ist dies ein Widerspruch zur Definition des Minimalpolynoms in Satz B1.15.

Die entsprechende Aussage für eine Matrix A läßt sich unmittelbar auf die Aussage für f_A zurückführen. □

Bemerkung B1.25 (Hauptraumzerlegung).

Man kann im Beweis von Satz B1.24 auch ohne Rückgriff auf Aufgabe 29.26 zeigen, daß die Summe $V_1 + \ldots + V_r$ eine direkte Summe ist:

Mit der Notation aus dem Beweis gilt für $x \in V_j$ und $i \neq j$

$$Q_i(f)(x) = p_i(f) \circ \prod_{k \neq i,j} (f - \lambda_k \mathrm{id}_V)^{m_k} \circ (f - \lambda_j \mathrm{id}_V)^{m_j}(x) = 0$$

und für $x \in V_i$ gilt deshalb

$$Q_i(f)(x) = \sum_{j=1}^{r} Q_j(f)(x) = \mathrm{id}_V(x) = x.$$

Sind nun also $x = x_1 + \ldots + x_r = y_1 + \ldots + y_r$ zwei Darstellung des Vektors x mit $x_i, y_i \in V_i$, $i = 1, \ldots, r$, dann gilt

$$0 = Q_i(f)(x - x) = Q_i(f)\big((x_1 - y_1) + \ldots + (x_r - y_r)\big) = Q_i(f)(x_i - y_i) = x_i - y_i,$$

d.h. $x_i = y_i$ für alle $i = 1, \ldots, r$. Die Summe ist also direkt.

Man beachte, daß wir hier $Q_i(f)^2 = Q_i(f)$ gezeigt haben, und daß damit $Q_i(f)$ die Projektion auf V_i mit Kern $\mathrm{Ker}\big(Q_i(f)\big) = \bigoplus_{j \neq i} V_j$ ist - vgl. Aufgabe 28.45.

Aus Satz B1.24 Teil b. und c. folgt, da die Haupträume von f f-invariant sind, unmittelbar das folgende Korollar.

Korollar B1.26.

Sei f wie in Satz B1.24, dann gilt

$$\chi_{f_{\mathrm{Hau}(f,\lambda_i)}} = (t - \lambda_i)^{n_i}$$

und

$$\mu_{f_{\mathrm{Hau}(f,\lambda_i)}} = (t - \lambda_i)^{m_i}.$$

G) Diagonalisierbarkeit

Satz B1.27 (Diagonalisierbarkeit von Endomorphismen).
Für einen Endomorphismus $f \in \mathrm{End}_K(V)$ sind die folgenden Aussagen äquivalent:

a. f ist diagonalisierbar.

b. V hat eine Basis aus Eigenvektoren von f.

c. Sind $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von f in K, dann gilt
$$V = \bigoplus_{i=1}^{r} \mathrm{Eig}(f, \lambda_i).$$

d. Das charakteristische Polynom von f zerfällt über K in Linearfaktoren und für jeden Eigenwert λ stimmen algebraische und geometrische Vielfachheit überein.

e. Das Minimalpolynom von f zerfällt über K in paarweise verschiedene Linearfaktoren.

Beweis:
a. ⇒ e.: Ist f diagonalisierbar, dann gibt es eine Basis B von V, so daß
$$M_B^B(f) = \bigoplus_{i=1}^{r} \lambda_i \mathbb{1}_{n_i},$$
mit $\lambda_i \neq \lambda_j$ für $i \neq j$. Setzen wir $p = (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_r) \in K[t]$, so ist
$$p\big(M_B^B(f)\big) = p(\lambda_1 \mathbb{1}_{n_1}) \oplus \ldots \oplus p(\lambda_r \mathbb{1}_{n_r})$$
wegen Bemerkung B1.11 eine Diagonalblockmatrix, und für die Blöcke gilt
$$p(\lambda_i \mathbb{1}_{n_i}) = p(\lambda_i) \cdot \mathbb{1}_{n_i} = 0.$$
Also ist schon $p(f) = 0$ erfüllt und p ist ein Vielfaches von μ_f. Dann zerfällt μ_f aber in paarweise verschiedene Linearfaktoren.

e. ⇒ d.: Zerfällt μ_f über K in paarweise verschiedene Linearfaktoren
$$\mu_f = (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_r),$$
so zerfällt wegen Korollar B1.18 auch χ_f in Linearfaktoren
$$\chi_f = (t - \lambda_1)^{n_1} \cdot \ldots \cdot (t - \lambda_r)^{n_r}.$$
Aus dem Satz zur Hauptraumzerlegung B1.24 folgt zudem $\mathrm{Hau}(f, \lambda_i) = \mathrm{Eig}(f, \lambda_i)$, da der Nilpotenzindex von $f - \lambda_i \,\mathrm{id}_V$ eins ist, und
$$\mathrm{mult}(\chi_f, \lambda_i) = n_i = \dim_K \mathrm{Hau}(f, \lambda_i) = \dim_K \mathrm{Eig}(f, \lambda_i),$$

d.h. die geometrische und die algebraische Vielfachheit jedes Eigenwertes stimmen überein.

d. ⇒ c.: Das charakteristische Polynom habe die Primfaktorzerlegung

$$\chi_f = (t-\lambda_1)^{n_1} \cdot \ldots \cdot (t-\lambda_r)^{n_r}.$$

Aus dem Satz über die Hauptraumzerlegung und der Voraussetzung folgt dann

$$\dim_K \operatorname{Hau}(f,\lambda_i) = n_i = \operatorname{mult}(\chi_f,\lambda_i) = \dim_K \operatorname{Eig}(f,\lambda_i),$$

und da stets $\operatorname{Eig}(f,\lambda_i) \subseteq \operatorname{Hau}(f,\lambda_i)$ gilt, folgt

$$\operatorname{Eig}(f,\lambda_i) = \operatorname{Hau}(f,\lambda_i).$$

Aus dem Satz über die Hauptraumzerlegung folgt dann aber wiederum

$$V = \operatorname{Hau}(f,\lambda_1) \oplus \ldots \oplus \operatorname{Hau}(f,\lambda_r) = \operatorname{Eig}(f,\lambda_1) \oplus \ldots \oplus \operatorname{Eig}(f,\lambda_r).$$

c. ⇒ b.: Es sei B_i eine Basis von $\operatorname{Eig}(f,\lambda_i)$, dann ist nach Aufgabe 29.26 $B = B_1 \cup \ldots \cup B_r$ eine Basis von V, die aus Eigenvektoren besteht.

b. ⇒ a.: Ist $B = (x_1,\ldots,x_n)$ eine Basis von V aus Eigenvektoren, so ist $f(x_i) = \lambda_i \cdot x_i$ für ein geeignetes $\lambda_i \in K$. Damit ist dann aber $M_B^B(f)$ eine Diagonalmatrix mit den Diagonaleinträgen $\lambda_1,\ldots,\lambda_n$. □

Korollar B1.28 (Diagonalisierbarkeit von Matrizen).
Für eine quadratische Matrix $A \in \operatorname{Mat}_n(K)$ sind die folgenden Aussagen äquivalent:

a. A ist diagonalisierbar.

b. K^n hat eine Basis aus Eigenvektoren von A.

c. Sind $\lambda_1,\ldots,\lambda_r$ die paarweise verschiedenen Eigenwerte von A, dann gilt
$$K^n = \bigoplus_{i=1}^{r} \operatorname{Eig}(A,\lambda_i).$$

d. Das charakteristische Polynom von A zerfällt über K in Linearfaktoren und für jeden Eigenwert λ stimmen algebraische und geometrische Vielfachheit überein.

e. Das Minimalpolynom von A zerfällt über K in paarweise verschiedene Linearfaktoren.

Insbesondere, genau dann ist $T \in \operatorname{Gl}_n(K)$ so, daß $T^{-1} \circ A \circ T$ eine Diagonalmatrix ist, wenn die Spalten von T eine Basis des K^n aus Eigenvektoren von A sind.

Beweis: Wende Satz B1.27 auf die Abbildung f_A an. □

Falls ein Endomorphismus oder eine Matrix hinreichend viele verschiedene Eigenwerte hat, so folgt aus den obigen Überlegungen unmittelbar deren Diagonalisierbarkeit.

Korollar B1.29 (Diagonalisierbarkeit).
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Hat f genau n paarweise verschiedene Eigenwerte, so ist f diagonalisierbar.

b. Hat A genau n paarweise verschiedene Eigenwerte, so ist A diagonalisierbar.

Beweis: Hat f genau n paarweise verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n \in K$, so muß

$$\chi_f = \mu_f = (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_n)$$

gelten. D.h. μ_f zerfällt in paarweise verschiedene Linearfaktoren und f ist diagonalisierbar. Der Beweis für A geht analog. □

Aus Korollar B1.28 können wir ein Verfahren ableiten, das es uns erlaubt, eine Matrix zu diagonalisieren und die Transformationsmatrix T zu berechnen.

Algorithmus B1.30 (Algorithmus zur Diagonalisierung).
INPUT: $A \in \mathrm{Mat}_n(K)$.
OUTPUT: 0, falls A über K nicht diagonalisierbar ist,
$1, D, T$, falls A diagonaliserbar ist, wobei D eine zu A konjugierte Diagonalmatrix ist, und T die zugehörige Transformationsmatrix mit $T^{-1} \circ A \circ T = D$.

1. Schritt: Berechne das charakteristische Polynom von A.

2. Schritt: Faktorisiere das charakteristische Polynom über K. Ist einer der Faktoren nicht linear, ist A nicht diagonalisierbar (nicht einmal trigonalisierbar) und man gebe 0 zurück. Sind alle Faktoren linear, so liefert die Faktorisierung die Eigenwerte $\lambda_1, \ldots, \lambda_r$ sowie ihre algebraischen Vielfachheiten $n_1, \ldots, n_r$.

3. Schritt: Bestimme für jeden Eigenwert λ_i eine Basis des Eigenraums $\mathrm{Eig}(A, \lambda_i)$ als $\mathrm{Lös}(A - \lambda_i \mathbb{1}_n, 0)$ - vgl. Algorithmus 33.9 - sowie seine Dimension - vgl. Algorithmus 32.20 -, d. h. die geometrische Vielfachheit von λ_i.

4. Schritt: Stimmt für jeden Eigenwert die algebraische Vielfachheit mit der geometrischen überein, so schreibe man die im 3. Schritt bestimmten Basen als Spalten in eine Matrix und erhält so T. Ferner erhält man D, indem man die Eigenwerte $\lambda_1, \ldots, \lambda_r$ entsprechend ihren algebraischen Vielfachheiten in der Diagonalen einer Nullmatrix einträgt.

Beispiel B1.31.
Gegeben sei die Matrix

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q}).$$

Das charakteristische Polynom von A berechnet man mit Hilfe zweifacher Laplace-Entwicklung nach der jeweils letzten Spalte als

$$\chi_A = \begin{vmatrix} t-2 & 1 & 0 & 0 \\ 0 & t-1 & 0 & 0 \\ 0 & 0 & t-2 & 0 \\ -1 & 1 & 1 & t-1 \end{vmatrix} = (t-1)\cdot(t-2)\cdot \begin{vmatrix} t-2 & 1 \\ 0 & t-1 \end{vmatrix} = (t-1)^2\cdot(t-2)^2.$$

Damit ist also $\sigma(A) = \{1, 2\}$ mit $\mathrm{mult}(\chi_A, 1) = \mathrm{mult}(\chi_A, 2) = 2$.

Als nächstes berechnen wir den Eigenraum $\mathrm{Lös}(2\mathbb{1}_4 - A, 0)$ zum Eigenwert $\lambda = 2$:

$$2\cdot\mathbb{1}_4 - A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 1 \end{pmatrix} \overset{\substack{I\leftrightarrow IV \\ I\mapsto -I \\ IV\mapsto IV-II \\ I\mapsto I+II}}{\longmapsto} \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{-1\text{'en einfügen}}{\longmapsto} \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Mithin ist

$$\mathrm{Eig}(A, 2) = \mathrm{Lin}\left((-1, 0, -1, 0)^t, (-1, 0, 0, -1)^t\right)$$

und

$$\dim_{\mathbb{Q}} \mathrm{Eig}(A, 2) = 2 = \mathrm{mult}(\chi_A, 2).$$

Dann berechnen wir den Eigenraum $\mathrm{Lös}(\mathbb{1}_4 - A, 0)$ zum Eigenwert $\lambda = 1$:

$$1\cdot\mathbb{1}_4 - A = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 1 & 1 & 0 \end{pmatrix} \overset{\substack{IV\mapsto IV-I+III \\ I\mapsto -I \\ III\mapsto -III}}{\longmapsto} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{-1\text{'en einfügen}}{\longmapsto} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Mithin ist

$$\mathrm{Eig}(A, 1) = \mathrm{Lin}\left((-1, -1, 0, 0)^t, (0, 0, 0, -1)^t\right)$$

und

$$\dim_{\mathbb{Q}} \mathrm{Eig}(A, 1) = 2 = \mathrm{mult}(\chi_A, 1).$$

Also zerfällt χ_A über $\mathbb{Q}$ in Linearfaktoren und die geometrischen Vielfachheiten der Eigenwerte stimmen mit den algebraischen überein, so daß A diagonalisierbar ist. Zudem

gilt für

$$T = \begin{pmatrix} -1 & -1 & -1 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 \end{pmatrix},$$

daß

$$T^{-1} \circ A \circ T = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Aufgaben

Aufgabe B1.32.
Es sei $E = (E_{11}, E_{12}, E_{21}, E_{22})$ die kanonische Basis von $V = \mathrm{Mat}_2(K)$ und $T = E_{11} + E_{12} + E_{22} \in \mathrm{Gl}_2(K)$. Zeige, daß der Endomorphismus $f : V \to V : A \mapsto T \circ A \circ T^{-1}$ trigonalisierbar, aber nicht diagonalisierbar ist, und bestimme eine Basis B von V, so daß $M_B^B(f)$ eine obere Dreiecksmatrix ist.

Aufgabe B1.33.
Zeige, ist $A \in \mathrm{Gl}_n(K)$, so gibt es ein Polynom $p \in K[t]$ mit $A^{-1} = p(A)$.

Aufgabe B1.34.
Zeige, ist $1 \leq \dim_K(V) = n < \infty$, so sind für $\mathcal{A} \subseteq \mathrm{End}_K(V)$ die folgenden beiden Aussagen gleichwertig:

a. $\mathcal{A}$ ist simultan diagonalisierbar, d. h. es gibt eine Basis B von V, so daß für alle $f \in \mathcal{A}$ gilt $M_B^B(f)$ ist eine Diagonalmatrix.

b. Für alle $f \in \mathcal{A}$ gilt, f ist diagonalisierbar, und für alle $f, g \in \mathcal{A}$ gilt, $f \circ g = g \circ f$.

Hinweis: Führe für "b. $\Rightarrow$ a." Induktion über n und zerlege dazu V in zwei invariante Unterräume kleinerer Dimension.

§ B2 Die Jordansche Normalform

Eine Matrix $A \in \mathrm{Mat}_n(K)$, deren charakteristisches Polynom in Linearfaktoren zerfällt, was etwa für einen algebraisch abgeschlossenen Körper wie $\mathbb{C}$ stets der Fall ist, ist zu einer Matrix konjugiert, die besonders einfach gebaut ist, der sog. Jordanschen Normalform von A. Aus der Jordanschen Normalform lassen sich Invarianten von A einfach ablesen und diese Invarianten bestimmen die Matrix A bis auf Konjugation eindeutig.

Satz B2.1 (Jordansche Normalform eines Endomorphismus).
Es sei $f \in \mathrm{End}_K(V)$ ein Endomorphismus, dessen charakteristisches Polynom über K zerfällt, $\chi_f = (t-\lambda_1)^{n_1} \cdot \ldots \cdot (t-\lambda_r)^{n_r}$, und es sei $\mu_f = (t-\lambda_1)^{m_1} \cdot \ldots \cdot (t-\lambda_r)^{m_r}$. Dann gibt es für jedes $i = 1, \ldots, r$ und $1 \le j \le m_i$, je eine natürliche Zahl t_{ij} und es gibt eine Basis B so, daß

(1) $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K \mathrm{Hau}(f, \lambda_i)$,

(2) $\sum_{j=1}^{m_i} t_{ij} = \dim_K \mathrm{Eig}(f, \lambda_i)$,

(3) $t_{im_i} \ge 1$ und

$$J_f := M_B^B(f) = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i).$$

J_f heißt *Jordansche Normalform* von f, und die t_{ij} werden *Elementarteiler* von f zum Eigenwert λ_i genannt.

Korollar B2.2 (Jordansche Normalform einer quadratischen Matrix).
Es sei $A \in \mathrm{Mat}_n(K)$ eine quadratische Matrix, deren charakteristisches Polynom über K zerfällt, $\chi_A = (t-\lambda_1)^{n_1} \cdot \ldots \cdot (t-\lambda_r)^{n_r}$, und es sei $\mu_A = (t-\lambda_1)^{m_1} \cdot \ldots \cdot (t-\lambda_r)^{m_r}$.
Dann gibt es für jedes $i = 1, \ldots, r$ und $1 \le j \le m_i$, je eine natürliche Zahl t_{ij} und es gibt ein invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ so, daß

(1) $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K \mathrm{Hau}(A, \lambda_i)$,

(2) $\sum_{j=1}^{m_i} t_{ij} = \dim_K \mathrm{Eig}(A, \lambda_i)$,

(3) $t_{im_i} \ge 1$ und

$$J_A := T^{-1} \circ A \circ T = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i).$$

J_A heißt *Jordansche Normalform* von A, und die t_{ij} werden *Elementarteiler* von A zum Eigenwert λ_i genannt.

Beweis: Der Beweis folgt aus Satz B2.1 mit $f = f_A$. □

Es scheint angebracht, den Satz zunächst etwas zu erläutern, um ihn verständlicher zu machen.

Bemerkung B2.3 (Jordansche Normalform).

a. Ziel des Abschnittes ist es, zu zeigen, daß eine Matrix A, deren charakteristisches Polynom zerfällt, konjugiert zu einer Matrix von besonders einfacher Gestalt ist. Der obige Satz sagt nun, daß in der Tat A konjugiert ist zu einer Diagonalblockmatrix, deren Diagonalblöcke, die $J_j(\lambda_i)$, alle Jordanblöcke sind, also obere Dreiecksmatrizen, die auf der Diagonalen stets den gleichen Wert λ_i stehen haben, auf der oberen Nebendiagonalen nur Einsen und ansonsten nur Nullen (vgl. Beispiel B1.19).
Dabei gelten:
   - Die natürlichen Zahlen t_{ij} geben an, wieviele Jordanblöcke der Größe $j \times j$ zum Eigenwert λ_i denn vorkommen.
   - $j \leq m_i$ bedeutet, die maximale Größe eines Jordanblockes ist $m_i \times m_i$.
   - $t_{im_i} \geq 1$ besagt, daß auch mindestens ein Block der maximalen Größe $m_i \times m_i$ vorkommt. D. h. die Vielfachheit von λ_i als Nullstelle von μ_A gibt die maximale Größe eines vorkommenden Jordanblockes in J_A zum Eigenwert λ_i an.
   - Die Summe $\sum_{j=1}^{m_i} j \cdot t_{ij}$ gibt gerade an, wie oft der Eigenwert λ_i auf der Diagonalen der Diagonalblockmatrix vorkommt, und da diese das gleiche charakteristische Polynom wie A besitzt, muß die Summe mithin n_i, also die algebraische Vielfachheit von λ_i als Eigenwert von A, sein.
   - Und $\sum_{j=1}^{m_i} t_{ij} = \dim_K \operatorname{Eig}(A, \lambda_i)$ bedeutet schließlich, daß die Anzahl der Jordanblöcke zum Eigenwert λ_i, die in J_A vorkommen, der Dimension des Eigenraumes von A zum Eigenwert λ_i entspricht, d.h. seiner geometrischen Vielfachheit.

b. Schon die direkte Summenschreibweise der Jordanschen Normalform bringt zum Ausdruck, daß die Jordansche Normalform nur bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt sein kann, und in der Tat ist sie es auch, d. h.:

Zwei Jordansche Normalformen sind genau dann konjugiert, wenn ihre Eigenwerte und die zugehörigen Elementarteiler übereinstimmen.

Es ist leicht einsichtig, daß eine Vertauschung der Blöcke durch Konjugation mit einer Reihe von Permutationsmatrizen erreicht werden kann, daß mithin zwei Jordansche Normalformen, deren Eigenwerte mit zugehörigen Elementarteilern übereinstimmen, zueinander konjugiert sind.
Seien umgekehrt zwei Jordansche Normalformen zueinander konjugiert, dann stimmen zunächst die charakteristischen Polynome und damit die Eigenwerte überein. Ferner folgt aus Aufgabe B2.21, daß die Elementarteiler übereinstimmen, da für eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ und ein $k \in \mathbb{N}$ gilt

$$\begin{aligned}\operatorname{rang}\left(\left(T^{-1} \circ A \circ T - \lambda \mathbb{1}_n\right)^k\right) &= \operatorname{rang}\left(T^{-1} \circ (A - \lambda \mathbb{1}_n)^k \circ T\right) \\ &= \operatorname{rang}\left((A - \lambda \mathbb{1}_n)^k\right).\end{aligned}$$

Damit ist natürlich auch die Jordansche Normalform eines Endomorphismus bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt.

c. Wir wollen folgende Notation einführen, die die Jordanblöcke von A (bzw. f) zu einem Eigenwert λ_i zusammenfaßt:

$$J_A(\lambda_i) := \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i) \quad \text{bzw.} \quad J_f(\lambda_i) := \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i)$$

Dann gilt

$$J_A = \bigoplus_{i=1}^{r} J_A(\lambda_i) \quad \text{bzw.} \quad J_f = \bigoplus_{i=1}^{r} J_f(\lambda_i).$$

Beispiel B2.4 (Jordansche Normalform).
Wir wollen nun in einigen einfachen Fällen die Jordansche Normalform bestimmen.

a. Die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q})$$

ist eine obere Dreiecksmatrix und ihr charakteristisches Polynom

$$\chi_A = (t-1) \cdot (t-5) \cdot (t-8) \cdot (t-2)$$

zerfällt in paarweise verschiedene Linearfaktoren. Da zu jedem der Eigenwerte ein Jordanblock gehören muß und da die Matrix J_A nicht mehr als vier Jordanblöcke aufnehmen kann, gilt also

$$J_A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Über die Transformationsmatrix T ist damit noch nichts gesagt. Da die Matrix J_A aber eine Diagonalmatrix ist, wissen wir aus Korollar B1.28 bereits, daß die Spalten von T Eigenvektoren zu den vier Eigenwerten sein müssen. Wir könnten T also leicht berechnen.

b. Die Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q})$$

hat offenbar den Rang eins. Deshalb gilt für die geometrische Vielfachheit von 0 als Eigenwert von A

$$\dim_K \mathrm{Eig}(A,0) = \dim_K \mathrm{Lös}(A,0) = 4 - \mathrm{rang}(A) = 3. \tag{149}$$

Da die algebraische Vielfachheit von 0 als Eigenwert von A mindestens so groß sein muß wie die geometrische, besitzt im charakteristischen Polynom χ_A von A der Linearfaktor t also mindestens Vielfachheit 3. Deshalb gibt es ein $\lambda \in \mathbb{Q}$ mit

$$\chi_A = t^3 \cdot (t - \lambda) = t^4 - \lambda \cdot t^3.$$

Aus Lemma 35.6 wissen wir aber, daß der zweithöchste Koeffizient des charakteristischen Polynoms das Negative der Spur der Matrix ist, d.h. $\lambda = \mathrm{Spur}(A) = 4$. Wir haben also

$$\chi_A = t^3 \cdot (t - 4).$$

Aus (149) folgt, daß es drei Jordanblöcke zum Eigenwert 0 geben muß, und außerdem muß es einen Jordanblock zum Eigenwert 4 geben. Da aber wieder höchstens vier Jordanblöcke in J_A passen, gilt

$$J_A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Die Transformationsmatrix T enthält als Spalten also auch wieder Eigenvektoren und läßt sich so leicht berechnen.

c. Wir betrachten wieder die Matrix A aus dem vorherigen Teil, fassen sie nun aber als Matrix über dem Körper $\mathbb{F}_2$ auf, d.h.

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{F}_2).$$

Wie oben sieht man, daß die Matrix Rang eins hat und somit 0 die geometrische Vielfachheit 3 besitzt. Und mit den gleichen Argumenten erhalten wir

$$\chi_A = t^3 \cdot \big(t - \mathrm{Spur}(A)\big).$$

Allerdings ist die Spur diesmal

$$\mathrm{Spur}(A) = 1 + 1 + 1 + 1 = 0 \in \mathbb{F}_2,$$

so daß wir

$$\chi_A = t^4$$

erhalten. 0 hat die geometrische Vielfachheit 3 und hat somit exakt drei Jordanblöcke zum Eigenwert 0, und da A keine anderen Eigenwerte besitzt, muß einer dieser Jordanblöcke diesmal die Größe zwei haben! Wir erhalten also

$$J_A = \left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Die Matrix A ist diesmal also *nicht* diagonalisierbar und wir wissen deshalb auch noch nicht, wie wir die Transformationsmatrix T bestimmen sollten!

d. Es sei $A \in \mathrm{Mat}_3(\mathbb{Q})$ mit der Eigenschaft $A^3 - A^2 = 0$. Was können wir über die Jordansche Normalform von A sagen?
A ist eine Nullstelle des Polynoms

$$p = t^3 - t^2 = t^2 \cdot (t - 1).$$

Das Minimalpolynom von A muß nach Satz B1.15 ein Teiler von p sein, so daß für μ_A nur folgende Möglichkeiten in Betracht kommen:

$$\mu_A \in \{t, t-1, t\cdot(t-1), t^2, t^2\cdot(t-1)\}.$$

Daraus ergeben sich für die Jordansche Normalform bis auf die Reihenfolge der Jordanblöcke folgenden Möglichkeiten:

μ_A	t	$t-1$	$t\cdot(t-1)$	t^2	$t^2\cdot(t-1)$
J_A	$\left(\begin{array}{cc\|c} 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right)$	$\left(\begin{array}{cc\|c} 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)$	$\left(\begin{array}{cc\|c} 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)$ $\left(\begin{array}{cc\|c} 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 1 \end{array}\right)$	$\left(\begin{array}{cc\|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$	$\left(\begin{array}{cc\|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 1 \end{array}\right)$

Dabei ist die Situation für $\mu_A = t$ oder $\mu_A = t - 1$ klar, da dann schon A selbst die angegebene Jordansche Normalform sein muß, wie man durch einsetzen von A in die Gleichung sieht.

Ist $\mu_A = t \cdot (t-1)$, so zerfällt das Minimalpolynom in paarweise verschiedene Linearfaktoren und A ist nach Korollar B1.28 diagonalisierbar. Zudem muß für jeden Eigenwert mindestens ein Jordanblock vorkommen, so daß genau die beiden angegebenen Matrizen in Frage kommen.
Wenn $\mu_A = t^2$ ist, so muß ein Jordanblock der Größe zwei zum Eigenwert 0 vorkommen und da nur Blöcke zum Eigenwert 0 vorkommen können, sind wir dann auch schon fertig. $\mu_A = t^2 \cdot (t-1)$ geht analog.

Wir werden Satz B2.1 zunächst für nilpotente Endomorphismen zeigen, d. h. für Endomorphismen, die nur einen Eigenwert, nämlich $\lambda = 0$, besitzen, und den allgemeinen Fall dann auf diesen zurückführen.

Definition B2.5 (Nilpotent).
Wir nennen einen Endomorphismus $f \in \mathrm{End}_K(V)$ bzw. eine Matrix $A \in \mathrm{Mat}_n(K)$ *nilpotent*, wenn es ein $r \in \mathbb{N}$ gibt, so daß $f^r = 0$ bzw. $A^r = 0$. Offenbar gilt dann $\mu_f = t^m$ bzw. $\mu_A = t^m$ für ein $1 \le m \le r$.

Beispiel B2.6 (Ein nilpotentes Jordankästchen).
Sei $f \in \mathrm{End}_K(V)$ und sei $B = (x_1, \ldots, x_n)$ eine Basis von V, so daß

$$M_B^B(f) = J_n(0) := \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & 0 \end{pmatrix}$$

ein Jordankästchen der Größe n zum Eigenwert 0 ist, dann folgt aus der Matrixdarstellung zunächst

$$f(x_{i+1}) = x_i$$

und damit

$$x_i = f^{n-i}(x_n)$$

für $i = 1, \ldots, n-1$. Das heißt, V ist ein zyklischer Unterraum mit seiner kanonischen Basis

$$B = \big(f^{n-1}(x_n), f^{n-2}(x_n), \ldots, f(x_n), x_n\big)$$

wie wir sie in Aufgabe 31.36 betrachtet haben. Man beachte auch, daß die Matrix $M_B^B(f)$ und damit der Endomorphismus f nilpotent mit Nilpotenzindex n ist (siehe Aufgabe 27.16).

Wir wollen im folgenden zeigen, daß die Jordansche Normalform eines nilpotenten Endomorphismus eine Blockdiagonalmatrix ist, deren Diagonalblöcke Jordankästchen

der obigen Form sind. Das Beispiel sagt uns also, welche Gestalt der Anteil der Basis haben muß, der zu einem solchen Kästchen gehört!

Definition B2.7 (Partitionen).
Eine *Partition* der positiven natürlichen Zahl n ist ein Tupel $P = (k_1, \ldots, k_m)$ natürlicher Zahlen, so daß $k_1 \geq k_2 \geq \ldots \geq k_m \geq 1$ und $k_1 + k_2 + \ldots + k_m = n$.

Lemma B2.8 (Die duale Partition).
Ist $P = (k_1, \ldots, k_m)$ eine Partition von n, ist auch $P^* = (l_1, \ldots, l_s)$ mit $s = k_1$ und

$$l_i = |\{j \mid 1 \leq j \leq m, k_j \geq i\}|$$

eine Partition von n, die sogenannte *duale Partition* zu P.

Beweis: Man kann die Partition P durch ihr *Young-Diagramm* veranschaulichen. Dieses besteht aus n Kästen, die in r Reihen übereinander angeordnet sind, wobei in der i-ten Reihe genau k_i Kästchen sind (siehe Abbildung 1).

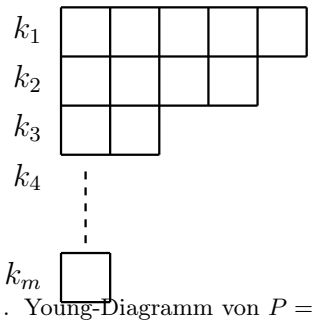


Abbildung 1. Young-Diagramm von $P = (k_1, \ldots, k_m)$

Die Anzahl an Kästchen in der i-ten Spalte ist dann genau l_i. Damit ist die Summe der l_i gerade n und $l_1 \geq l_2 \geq \ldots \geq l_s$. □

Beispiel B2.9.
$P = (5, 4, 2, 2, 2)$ ist eine Partition von $n = 15$ mit dem folgenden Young-Diagramm (siehe Abbildung 2).

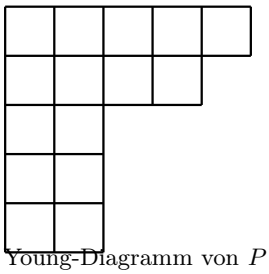


Abbildung 2. Young-Diagramm von $P = (5, 4, 2, 2, 2)$

Das Young-Diagramm der dualen Partition $P^* = (5, 5, 2, 2, 1)$ entsteht durch Spiegelung an der Winkelhalbierenden (siehe Abbildung 3).

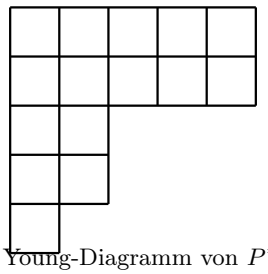


ABBILDUNG 3. Young-Diagramm von $P^* = (5, 5, 2, 2, 1)$

Bemerkung B2.10 (Anzahl der Partitionen von n).
Die Funktion

$$\pi : \mathbb{Z}_{>0} \longrightarrow \mathbb{Z}_{>0},$$

die einer natürlichen Zahl n die Anzahl der Partitionen von n zuordnet, ist eine interessante zahlentheoretische Funktion. Wir wollen einige Werte von π zur Veranschaulichung ihrer Komplexität angeben:

n	1	2	3	4	5	6	7	8	9	10	100
$\pi(n)$	1	2	3	5	7	11	15	22	30	42	190569292

Für große n gilt

$$\pi(n) \approx \tfrac{1}{4\sqrt{3n}} e^{\pi\sqrt{\frac{2n}{3}}}.$$

Das folgende Lemma zusammen mit Bemerkung B2.3 besagt, daß $\pi(n)$ zugleich die Anzahl der Konjugationsklassen nilpotenter Matrizen der Größe $n \times n$ angibt.

Lemma B2.11 (Jordansche Normalform nilpotenter Endomorphismen).
Sei $f \in \mathrm{End}_K(V)$ ein nilpotenter Endomorphismus mit $m_f = t^m$.

a. Setzen wir $U_i = \mathrm{Ker}(f^i)$, $i = 0, \ldots, m$, dann induziert f für $i = 2, \ldots, m$ eine injektive lineare Abbildung

$$f_i : U_i/U_{i-1} \longrightarrow U_{i-1}/U_{i-2} : \overline{x} \mapsto \overline{f(x)}.$$

Zudem ist $P = (k_1, \ldots, k_m)$ eine Partition von n mit

$$k_i = \dim_K(U_i/U_{i-1}) = \mathrm{rang}(f^{i-1}) - \mathrm{rang}(f^i),$$

die wir die *Jordan-Partition* des nilpotenten Endomorphismus nennen wollen.

b. Ist $P^* = (l_1, \ldots, l_s)$ die zu P duale Partition, dann gibt es eine Basis B von V, so daß

$$M_B^B(f) = J_{l_1}(0) \oplus J_{l_2}(0) \oplus \ldots \oplus J_{l_s}(0).$$

Die analogen Aussagen für Matrizen $A \in \mathrm{Mat}(n \times n, K)$ gelten ebenfalls.

Beweis: Wir beweisen zunächst Teil a. und beachten dazu, daß wir aus Lemma B1.23

$$U_0 \subsetneqq U_1 \subsetneqq \ldots \subsetneqq U_m$$

wissen.

Wir müssen zunächst zeigen, daß f_i wohldefiniert ist. Für $\overline{x} = \overline{y} \in U_i/U_{i-1}$ ist

$$x - y \in U_{i-1} = \mathrm{Ker}(f^{i-1}),$$

so daß

$$f(x) - f(y) = f(x - y) \in \mathrm{Ker}(f^{i-2}) = U_{i-2}$$

folgt, d.h. $\overline{f(x)} = \overline{f(y)} \in U_{i-1}/U_{i-2}$. Da mit $x \in U_i = \mathrm{Ker}(f^i)$ zudem $f(x) \in \mathrm{Ker}(f^{i-1}) = U_{i-1}$ gilt, ist f_i wohldefiniert.

Mit f ist dann aber auch f_i eine lineare Abbildung und für die Injektivität reicht es, zu zeigen

$$\mathrm{Ker}(f_i) = \{\overline{0}\}.$$

Nun ist aber $\overline{x} \in \mathrm{Ker}(f_i)$ gleichwertig zu $f(x) \in U_{i-2} = \mathrm{Ker}(f^{i-2})$, was wiederum nur für $x \in \mathrm{Ker}(f^{i-1}) = U_{i-1}$, d.h. für $\overline{x} = \overline{0} \in U_i/U_{i-1}$, zutrifft.

Wir müssen noch zeigen, daß $P = (k_1, \ldots, k_m)$ eine Partition von n ist. Aus

$$0 \neq U_m/U_{m-1} \hookrightarrow U_{m-1}/U_{m-2} \hookrightarrow \ldots \hookrightarrow U_1/U_0 = U_1$$

folgt auch

$$1 \leq \dim_K(U_m/U_{m-1}) \leq \dim_K(U_{m-1}/U_{m-2}) \leq \ldots \leq \dim_K(U_1/U_0),$$

d.h.

$$1 \leq k_m \leq k_{m-1} \leq \ldots \leq k_1.$$

Man beachte dabei, daß $U_m/U_{m-1} \neq 0$ gilt, weil m der Nilpotenzindex von f ist!

Außerdem folgt aus der Dimensionsformel für Vektorräume

$$n = \dim_K(V) = \dim_K(V/U_{m-1}) + \dim_K(U_{m-1}) = \dim_K(U_m/U_{m-1}) + \dim_K(U_{m-1}).$$

Mit Induktion nach m folgt dann

$$\begin{aligned} n =& \dim_K(U_m/U_{m-1}) + \dim_K(U_{m-1}) \\ =& \dim_K(U_m/U_{m-1}) + \dim_K(U_{m-1}/U_{m-2}) + \ldots + \dim_K(U_1/U_0) + \dim_K(U_0) \\ =& k_m + k_{m-1} + \ldots + k_1 + 0. \end{aligned}$$

Damit ist P eine Partition von n und Teil a. ist bewiesen.

Wenden wir uns nun Teil b. zu und konstruieren die Basis B.

Dazu wählen wir zunächst Vektoren

$$x_1^m, \ldots, x_{k_m}^m, \tag{150}$$

deren Restklassen eine Basis von U_m/U_{m-1} bilden. Dann wenden wir f auf diese an und erhalten Vektoren

$$x_1^{m-1} := f(x_1^m), \ldots, x_{k_m}^{m-1} := f(x_{k_m}^m) \in U_{m-1},$$

deren Restklassen in U_{m-1}/U_{m-2} linear unabhängig sind, weil die Abbildung

$$f_m : U_m/U_{m-1} \hookrightarrow U_{m-1}/U_{m-2}$$

eine injektive lineare Abbildung ist. Nun ergänzen wir die Restklassen von diesen durch die Restklassen von Vektoren

$$x^{m-1}_{k_m+1}, \ldots, x^{m-1}_{k_{m-1}},$$

zu einer Basis von U_{m-1}/U_{m-2}. Mit den so gewonnenen Vektoren

$$x^{m-1}_1, \ldots, x^{m-1}_{k_{m-1}}$$

verfahren wir analog und konstruieren so rekursiv Vektoren

$$x^i_1, \ldots, x^i_{k_i}, \tag{151}$$

deren Restklassen jeweils eine Basis für U_i/U_{i-1} sind, für $i = 1, \ldots, m$. Diese $n = k_1 + \ldots + k_m$ Vektoren ordnen wir der Übersichtlichkeit halber in dem Young-Diagramm von P an (siehe Abbildung 4). Wir können die Vektoren im Young-Diagramm auch als

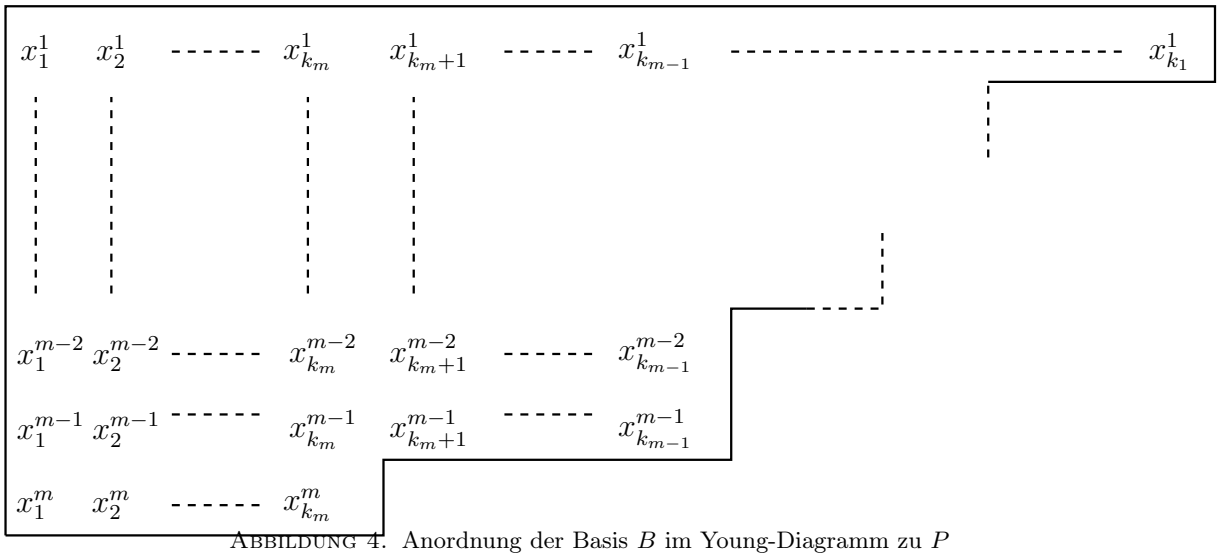


Abbildung 4. Anordnung der Basis B im Young-Diagramm zu P

Bilder unter der Abbildung f schreiben und erhalten Abbildung 5. Schließlich benennen wir die Vektoren um, wie in Abbildung 6 angegeben, d.h. wir lesen das Diagramm aus, indem wir, in der linken oberen Ecke beginnend, die Spalten sukzessive von oben nach unten durchlaufen. Wir müssen nun nur noch zeigen, daß

$$B = (x_1, \ldots, x_n)$$

linear unabhängig ist, dann ist B eine Basis des n-dimensionalen Vektorraums V und die Matrix-Darstellung hat offenbar die gewünschte Gestalt, wie wir aus Abbildung 5 sehen. Dazu beachten wir, daß die Spalten des Diagramms jeweils die kanonische Basis eines zyklischen Unterraums sind und somit einen Jordanblock liefern. Das zeigt insbesondere, daß die zu P duale Partition die Größen der Jordanblöcke liefert.

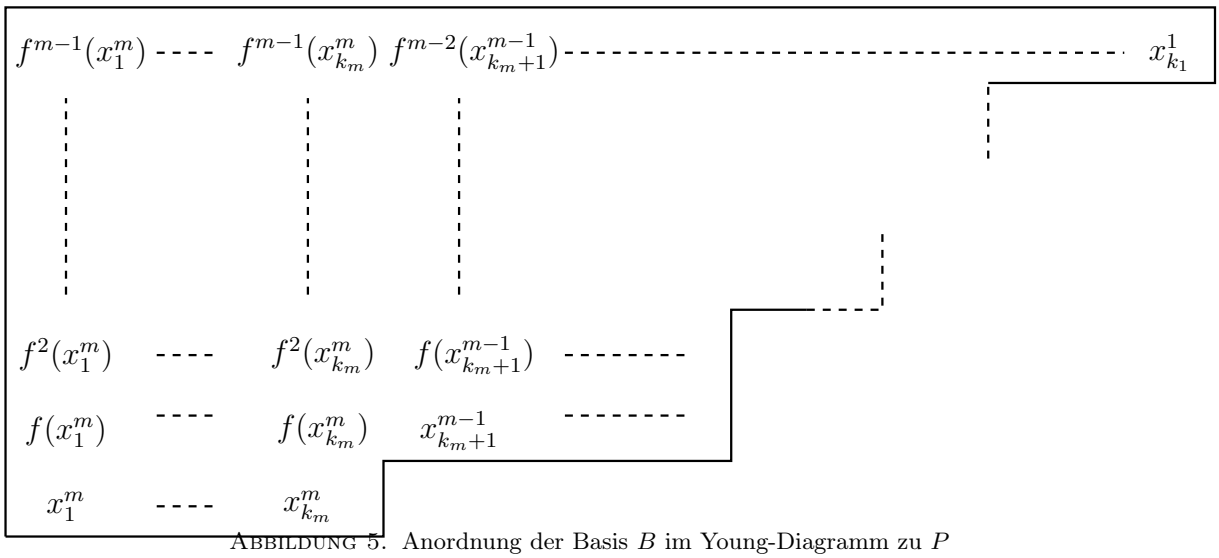


ABBILDUNG 5. Anordnung der Basis B im Young-Diagramm zu P

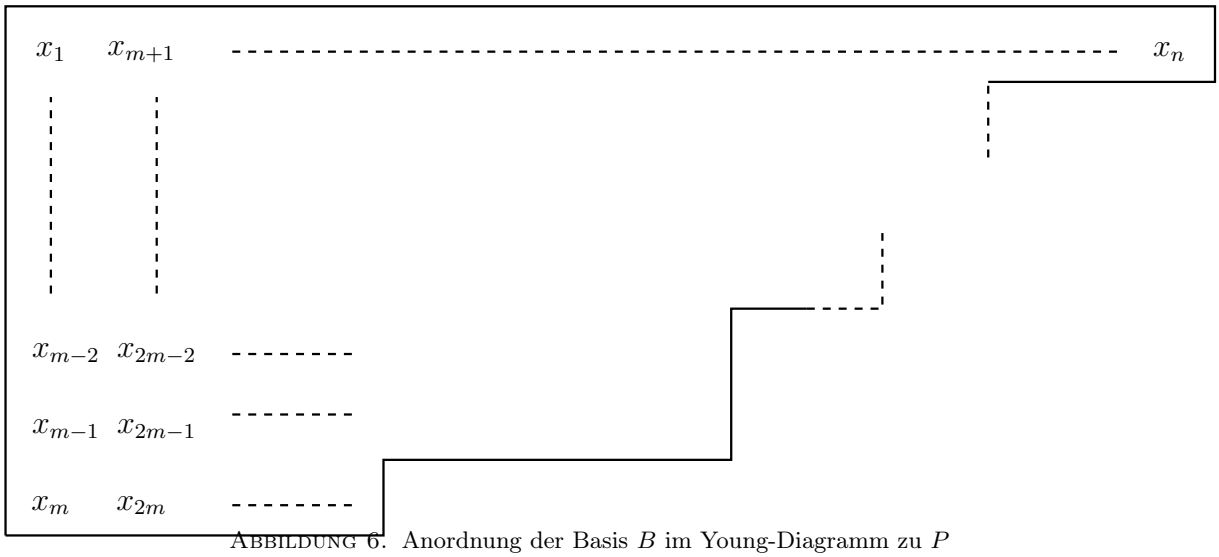


ABBILDUNG 6. Anordnung der Basis B im Young-Diagramm zu P

Um zu zeigen, daß B eine Basis ist, betrachten wir eine Linearkombination

$$\sum_{i=1}^{m}\sum_{j=1}^{k_j}\lambda_{i,j}\cdot x_j^i = 0 \tag{152}$$

der Vektoren in B, die den Nullvektor ergibt. Beachten wir, daß $x_j^i \in U_{m-1}$ für $i < m$ gilt, und betrachten wir die Gleichung in U_m/U_{m-1}, so reduziert sie sich auf

$$\sum_{j=1}^{k_m}\lambda_{m,j}\cdot\overline{x_j^m} = \overline{0} \in U_m/U_{m-1}.$$

Da die Vektoren $\overline{x_1^m},\ldots,\overline{x_{k_m}^m}$ linear unabhängig sind (siehe (150)), gilt also

$$\lambda_{m,1} = \ldots = \lambda_{m,k_m} = 0$$

und (152) reduziert sich zu

$$\sum_{i=1}^{m-1}\sum_{j=1}^{k_j}\lambda_{i,j}\cdot x_j^i = 0.$$

Diese Gleichung können wir mit demselben Argument in U_{m-1}/U_{m-2} betrachten und erhalten

$$\sum_{j=1}^{k_{m-1}}\lambda_{m-1,i}\cdot\overline{x_j^{m-1}} = \overline{0}\in U_{m-1}/U_{m-2}.$$

Die Restklassen der beteiligten Vektoren sind nach Konstruktion (siehe (151)) linear unabhängig in U_{m-1}/U_{m-2} und somit gilt

$$\lambda_{m-1,1} = \ldots = \lambda_{m-1,k_{m-1}} = 0.$$

Fahren wir so fort erhalten wir insgesamt, daß alle $\lambda_{i,j}$ Null sein müssen und B ist linear unabhängig. □

Wir können damit nun auch Satz B2.1 für nilpotente Endomorphismen beweisen.

Lemma B2.12 (Jordansche Normalform nilpotenter Endomorphismen). Es sei $f\in\mathrm{End}_K(V)$ ein nilpotenter Endomorphismus mit $\mu_f = t^m$. Dann gibt es für jedes $1\le j\le m$ je eine natürliche Zahl t_j und es gibt eine Basis B so, daß

(1) $\sum_{j=1}^{m} j\cdot t_j = n = \dim_K \mathrm{Hau}(f,0) = \dim_K(V)$,

(2) $\sum_{j=1}^{m} t_j = \dim_K \mathrm{Eig}(f,0)$,

(3) $t_m\ge 1$ und

$$J_f := M_B^B(f) = \bigoplus_{j=1}^{m}\bigoplus_{k=1}^{t_j} J_j(0).$$

Beweis: Sei die Partition $P=(k_1,\ldots,k_m)$ wie in Lemma B2.11 gegeben. Setzen wir

$$t_j := k_j - k_{j+1}$$

für $j=1,\ldots,m$ mit der Konvention $k_{m+1}=0$, dann ist t_j gerade die Anzahl der Jordanblöcke der Größe $j\times j$ in der Matrixdarstellung

$$M_B^B(f) = J_{l_1}(0)\oplus\ldots\oplus J_{l_s}(0) \tag{153}$$

in Lemma B2.11 (siehe Abbildung 4). Mithin gilt

$$\sum_{j=1}^{m} t_j = k_1 - k_{m+1} = k_1 = \dim_K(U_1) = \dim_K \mathrm{Ker}(f) = \dim_K \mathrm{Eig}(f,0)$$

und

$$\sum_{j=1}^{m} j \cdot t_j = n = \dim_K(V) = \dim_K \operatorname{Hau}(f, 0),$$

weil die Summe der Größen der Kästchen mit ihren Vielfachheiten die Größe der Gesamtmatrix ist. Außerdem ist

$$t_m = k_m - k_{m+1} = k_m - 0 = k_m \geq 1$$

und die Matrixdarstellung in (153) kann dann auch geschrieben werden als

$$M_B^B(f) = \bigoplus_{j=1}^{m} \bigoplus_{k=1}^{t_j} J_j(0).$$

□

Bemerkung B2.13 (Jordanbasis einer nilpotenten Matrix).
Ist A eine nilpotente Matrix mit $\mu_A = t^m$ und bestimmt man wie im Beweis von Lemma B2.11 (siehe auch Abbildung 5) linear unabhängige Familien

$$B_{j,l} = \left(A^{j-1}x_l^j, A^{j-2}x_l^j, \ldots, Ax_l^j, x_l^j\right) \subset K^n$$

in $\text{Lös}(A^j, 0)$ für $j = 1, \ldots, m$ und $l = k_{j+1} + 1, \ldots, k_j$, dann ist die Matrix $T \in \mathrm{Gl}_n(K)$, deren Spalten gerade all diese Vektoren sind, eine Transformationsmatrix, die A in Jordansche Normalform überführt.

Beispiel B2.14 (Jordansche Normalform einer nilpotenten Matrix).
Wir wollen nun für die folgende nilpotente Matrix

$$A = \begin{pmatrix} -2 & 0 & 0 & 1 & 0 \\ -2 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ -4 & 0 & -1 & 2 & 0 \\ -2 & -1 & 0 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_5(\mathbb{Q})$$

die Jordansche Normalform sowie die Transformationsmatrix $T \in \mathrm{Gl}_5(\mathbb{Q})$ bestimmen.

Dazu berechnen wir zunächst den Nilpotenzindex von A und merken uns die Potenzen A^k von A, da wir sie anschließend benötigen:

$$A^2 = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

und

$$A^3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

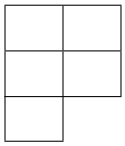
Der Nilpotenzindex von A ist also $m = 3$, $\mu_A = t^3$ und

$$\mathrm{Hau}(A, 0) = \mathbb{Q}^5.$$

Damit muß in der Jordanschen Normalform von A also ein Jordanblock $J_3(0)$ der Größe $m = 3$ vorkommen, und aufgrund der geringen Größe der Matrix A bleiben damit nur die beiden folgenden Möglichkeiten für die Jordansche Normalform übrig:

$$J_A = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad \text{oder} \quad J_A = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Bestimmen wir die Ränge der Potenzen $A^0, \ldots, A^3$ von A (siehe weiter unten), so können wir die Jordan-Partition



von A dann als

$$\begin{aligned} P =& \big(\operatorname{rang}(A^0) - \operatorname{rang}(A^1), \operatorname{rang}(A^1) - \operatorname{rang}(A^2), \operatorname{rang}(A^2) - \operatorname{rang}(A^3)\big) \\ =& (5 - 3, 3 - 1, 1 - 0) = (2, 2, 1) \end{aligned}$$

berechnen, und wir erhalten als duale Partition

$$P^* = (3, 2),$$

woraus sich unmittelbar die Jordansche Normalform

$$J_A = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

ergibt sowie die Elementarteiler

$$t_1 = 0, \quad t_2 = 1, \quad t_3 = 1.$$

Um die Jordanbasis B von f_A oder alternativ die Transformationsmatrix T von A zu bestimmen, reicht es gemäß Lemma B2.11 im wesentlichen, geeignete Basen der

Vektorräume U_j/U_{j-1} für $j = 3, 2, 1$ zu bestimmen, wobei $U_j = \text{Lös}(A^j, 0) = \text{Ker}(f_A^j)$ ist.

Wir beginnen mit $j = 3$ und $U_3/U_2 = \mathbb{Q}^5/U_2$, wobei wir $U_3 = \mathbb{Q}^5$ beachten.

Um eine Basis von U_3 zu finden, muß man einerseits eine Basis von U_2 berechnen und diese dann zu einer Basis von U_3 ergänzen, indem man sie mit Steinitz in eine Basis von U_3 hineintauscht. Bestimmen wir also zunächst eine Basis B_2' von $U_2 = \text{Lös}(A^2, 0)$. Aufgrund der einfachen Form von A^2 mit Rang 1 geschieht dies durch einfaches Draufschauen — vier der Einheitsvektoren tun es offenbar:

$$B_2' = \big((1,0,0,0,0)^t, (0,1,0,0,0)^t, (0,0,0,1,0)^t, (0,0,0,0,1)^t\big).$$

Für $U_3 = \text{Lös}(A^3, 0) = \mathbb{Q}^5$ ist es noch einfacher, eine Basis zu bestimmen, die kanonische Basis tut's:

$$B_3' = \big((1,0,0,0,0)^t, (0,1,0,0,0)^t, (0,0,1,0,0)^t, (0,0,0,1,0)^t, (0,0,0,0,1)^t\big).$$

Damit ist es in dem vorliegenden Beispiel auch denkbar einfach, die Basis B_2' in die Basis B_3' hineinzutauschen, es fehlt nämlich einfach der Vektor e_3, und wir setzen deshalb

$$x_1^3 = e_3 = (0,0,1,0,0)^t.$$

Damit erhalten wir $k_3 = \dim_{\mathbb{Q}}(U_3/U_2) = 1$ und die erste Teilbasis der Jordanbasis:

$$B_{3,1} = \big(A^2x_1^3, Ax_1^3, x_1^3\big) = \big((-1,0,0,-2,0)^t, (0,-1,0,-1,0)^t, (0,0,1,0,0)^t\big).$$

Diese können wir in das Young-Diagramm der Jordan-Partition eintragen:

$A^2x_1^3$	
Ax_1^3	
x_1^3	

Als nächstes betrachten wir $j = 2$ und U_2/U_1.

Um eine geeignete Basis von U_2/U_1 zu bestimmen, müssen wir eine Basis von U_1 berechnen und diese zusammen mit der zweiten Ebene des bereits befüllten Young-Diagramms, d.h. mit der $(Ax_1^3) = (Ae_3)$, zu einer Basis von U_2 ergänzen. Dazu berechnen wir zunächst eine Basis B_1' von $U_1 = \text{Lös}(A, 0)$:

$$A = \begin{pmatrix} -2 & 0 & 0 & 1 & 0 \\ -2 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ -4 & 0 & -1 & 2 & 0 \\ -2 & -1 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{\text{Gauß}} \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{einfügen}]{-1\text{'en}} \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

Die letzten beiden Spaltenvektoren bilden also eine Basis von $\text{Lös}(A, 0)$. Wir dürfen die Vektoren aber auch mit einem Skalar multiplizieren, um schönere Vektoren zu erhalten,

und tun dies. Unsere Basis von $U_1 = \mathrm{Eig}(A,0) = \mathrm{Lös}(A,0)$ ist dann

$$B_1' = \big((1,0,0,2,0)^t, (0,1,0,0,1)^t\big).$$

Wir müssen also die Familie

$$B_2'' = B_1' \cup (Ae_3) = \big((1,0,0,2,0)^t, (0,1,0,0,1)^t, (0,-1,0,-1,0)^t\big),$$

zu einer Basis von U_2 ergänzen. Dazu können wir sie mit Hilfe des Austauschsatzes von Steinitz in die Basis B_2' von U_2 hineintauschen, oder alternativ kann man auch einfach genau hinschauen. Man sieht nämlich leicht, daß der Vektor

$$x_2^2 = (0,0,0,0,1)^t$$

von den drei Vektoren in B_2'' linear unabhängig ist, und somit ergänzt er B_2'' zu einer Basis von U_2. Von der linearen Unabhängigkeit der vier Vektoren kann man sich auch überzeugen, indem man die Vektoren in eine Matrix schreibt und den Rang bestimmt, was schneller ist als dreimal Steinitz und trotzdem ausreicht. Wir überlassen die Rechnung dem Leser. Nachdem wir nun x_2^2 bestimmt haben, erhalten wir die zweite Teilbasis

$$B_{2,2} = \big(Ax_2^2, x_2^2\big) = \big((0,1,0,0,1)^t, (0,0,0,0,1)^t\big)$$

der Jordanbasis und das fertig ausgefüllte Young-Diagramm der Jordan-Partition:

$$\begin{array}{|c|c|}\hline A^2x_1^3 & Ax_2^2 \\ \hline Ax_1^3 & x_2^2 \\ \hline x_1^3 \\ \cline{1-1}\end{array} \quad = \quad \begin{array}{|c|c|}\hline x_1 & x_4 \\ \hline x_2 & x_5 \\ \hline x_3 \\ \cline{1-1}\end{array}$$

Im Prinzip bliebe noch der Fall $\underline{j=1}$ zu untersuchen, aber da die zu P duale Partition nur zwei Spalten hat und damit $t_1 = 0$ gilt, sind wir fertig.

Wir haben also die Jordanbasis $B = B_{3,1} \cup B_{2,2}$ bestimmt und damit auch die Transformationsmatrix T, deren Spalten die Vektoren in B sind. Wir haben

$$T = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \in \mathrm{Gl}_5(\mathbb{Q})$$

mit

$$T^{-1} = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & -1 & 0 \\ -2 & -1 & 0 & 1 & 1 \end{pmatrix},$$

und es gilt

$$J_A = T^{-1} \circ A \circ T = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Kommen wir nun zum Beweis von Satz B2.1.

Beweis von Satz B2.1: Nach Satz B1.24 zerfällt V in die direkte Summe der Haupträume $V_i := \operatorname{Hau}(f, \lambda_i)$, $i = 1, \ldots, r$, und diese sind nach Lemma B1.23 invariant unter f und $f - \lambda_i \operatorname{id}_V$. Betrachten wir nun die Abbildungen

$$(f - \lambda_i \operatorname{id}_V)_{V_i} : V_i \to V_i$$

für $i = 1, \ldots, r$, so sind diese nilpotent mit $\chi_{(f-\lambda_i \operatorname{id}_V)_{V_i}} = t^{n_i}$ und $\mu_{(f-\lambda_i \operatorname{id}_V)_{V_i}} = t^{m_i}$ (vgl. Korollar B1.26). Nach Lemma B2.12 gibt es dann aber für jedes $i = 1, \ldots, r$ Basen B_i von V_i und natürliche Zahlen t_{ij}, $j = 1, \ldots, m_i$, so daß gilt

(1) $\displaystyle\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K \operatorname{Hau}(f, \lambda_i)$,

(2) $\displaystyle\sum_{j=1}^{m_i} t_{ij} = \dim_K \operatorname{Eig}\big((f - \lambda_i \operatorname{id}_V)_{V_i}, 0\big) = \dim_K \operatorname{Eig}(f, \lambda_i)$,

(3) $t_{im_i} \geq 1$ und

$$\begin{aligned} M_{B_i}^{B_i}(f_{V_i}) &= \lambda_i \mathbb{1}_{n_i} + M_{B_i}^{B_i}\big((f - \lambda_i \operatorname{id}_V)_{V_i}\big) \\ &= \lambda_i \mathbb{1}_{n_i} + \left(\bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(0)\right) = \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i). \end{aligned}$$

Damit folgt die Behauptung, da für $B = B_1 \cup \ldots \cup B_r$ gilt

$$M_B^B(f) = \bigoplus_{i=1}^{r} M_{B_i}^{B_i}(f_{V_i}).$$

□

Wie wir schon gesehen haben, ist der Beweis zur Berechnung der Jordanschen Normalform algorithmisch. Wir wollen nun den Algorithmus beschreiben, mit Hilfe dessen man die Jordansche Normalform einer Matrix A inklusive der zugehörigen Transformationsmatrix bestimmen kann.

Algorithmus B2.15 (Jordansche Normalform - I).
INPUT: $A \in \operatorname{Mat}_n(\mathbb{Q})$ mit μ_A zerfällt in Linearfaktoren.
OUTPUT: J_A und eine Transformationsmatrix $T \in \operatorname{Gl}_n(K)$ mit $T^{-1} \circ A \circ T = J_A$.

1. Schritt: Bestimme das Minimalpolynom μ_A von A und faktorisiere es.

2. Schritt: Wenn μ_A nicht in Linearfaktoren zerfällt, gebe man eine Fehlermeldung zurück, andernfalls gilt $\mu_A = \prod_{i=1}^{r}(t-\lambda_i)^{m_i}$.

3. Schritt: Für $i = 1, \ldots, r$ bilde man die Matrix $A_i = A - \lambda_i \mathbb{1}_n$ und führe folgende Schritte aus:

Schritt a.: Berechne die Partition $P = (k_1, \ldots, k_{m_i})$ von $n - \operatorname{rang}(A_i^{m_i})$ mit $k_j = \operatorname{rang}(A_i^{j-1}) - \operatorname{rang}(A_i^j)$ gemäß Lemma B2.11 sowie das zugehörige Young-Diagramm.

Schritt b.: Bestimme eine Basis B_{m_i} von Lös $\left(A_i^{m_i}, 0\right)$ sowie eine Basis B_{m_i-1} von Lös $\left(A_i^{m_i-1}, 0\right)$.

Schritt c.: Tausche B_{m_i-1} mittels des Satzes von Steinitz in B_{m_i} hinein und bestimme die in B_{m_i} verbliebenen Vektoren $x_1^{m_i}, \ldots, x_{k_{m_i}}^{m_i}$.

Schritt d.: Dann fülle man die ersten k_{m_i} Spalten des Young-Diagramms von P durch die Vektoren $A_i^{m_i-1}x_l^{m_i}, \ldots, A_i^0 x_l^{m_i}$ auf, $l = 1, \ldots, k_{m_i}$, wie in Abbildung 7.

Schritt e.: Für $j = m_i - 1, \ldots, 1$ führe man folgendes aus:

- bestimme eine Basis B_{j-1} von Lös $\left(A_i^{j-1}, 0\right)$;
- tausche B_{j-1} sowie die auf der j-ten Ebene des Young-Diagramms bereits eingetragenen Vektoren mittels des Satzes von Steinitz in B_j hinein;
- bestimme die in B_j verbliebenen Vektoren $x_{k_{j+1}+1}^j, \ldots, x_{k_j}^j$;
- für $l = k_{j+1} + 1, \ldots, k_j$ fülle die Spalten des Young-Diagramms von P mit den Vektoren $A_i^{j-1}x_l^j, \ldots, A_i^0 x_l^j$.

Schritt f.: Füge die Vektoren aus dem Young-Diagramm als Spalten in die Matrix T ein, beginnend in der linken oberen Ecke und die Spalten des Young-Diagramms von oben nach unten nacheinander durchlaufend.

4. Schritt: Gib $T^{-1} \circ A \circ T$ und T zurück.

Beispiel B2.16 (Jordansche Normalform).
Wir wollen nun die Jordansche Normalform und die Transformationsmatrix von

$$A = \begin{pmatrix} 2 & 1 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \operatorname{Mat}_4(\mathbb{Q})$$

bestimmen.

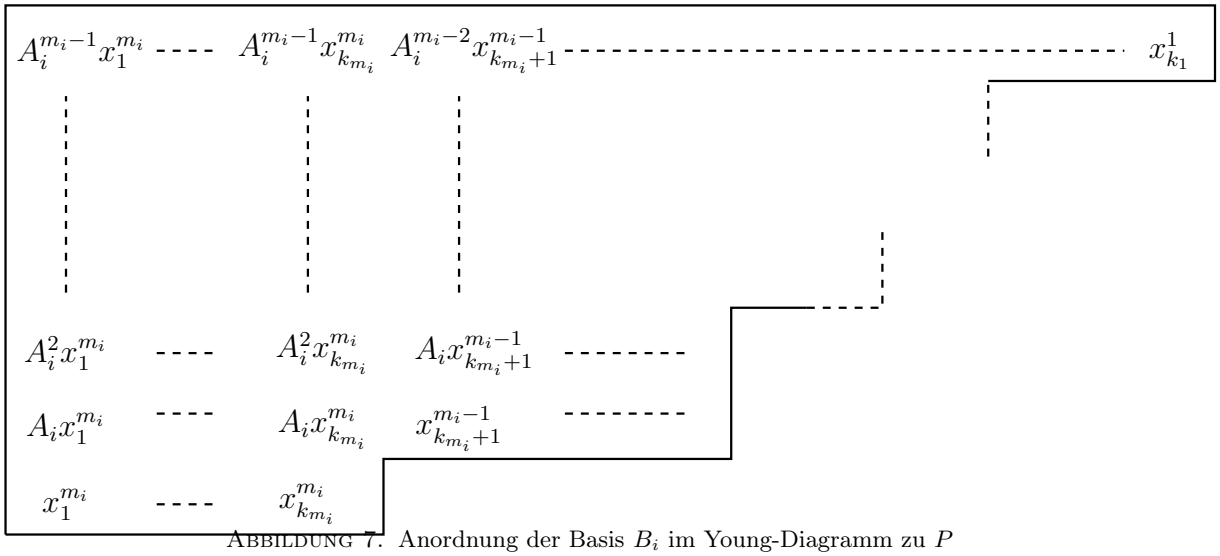


ABBILDUNG 7. Anordnung der Basis B_i im Young-Diagramm zu P

Das charakteristische Polynom berechnet man mit Hilfe des Kästchensatzes als

$$\chi_A = \begin{vmatrix} t-2 & -1 & -1 & -2 \\ 0 & t-1 & 0 & 0 \\ 0 & -1 & t-1 & -1 \\ 0 & 0 & 0 & t-2 \end{vmatrix} = (t-2)\cdot \begin{vmatrix} t-1 & 0 \\ -1 & t-1 \end{vmatrix} \cdot (t-2) = (t-2)^2\cdot(t-1)^2.$$

Dann berechnen wir eine Basis von $\mathrm{Eig}(A,1) = \mathrm{Lös}(A-\mathbb{1}_4, 0)$:

$$A-\mathbb{1}_4 = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \overset{\text{Gauß}}{\longmapsto} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{\substack{-1\text{'en} \\ \text{einfügen}}}{\longmapsto} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Also ist $\mathrm{Eig}(A,1) = \mathrm{Lin}\left((1,0,-1,0)^t\right)$, und damit stimmen die geometrische und die algebraische Vielfachheit von 1 als Eigenwert von A nicht überein. Wir müssen auch noch $\mathrm{Hau}(A,1) = \mathrm{Lös}\left((A-\mathbb{1}_4)^2, 0\right)$ bestimmen:

$$(A-\mathbb{1}_4)^2 = \begin{pmatrix} 1 & 2 & 1 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \overset{\text{Gauß}}{\longmapsto} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{\substack{-1\text{'en} \\ \text{einfügen}}}{\longmapsto} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Als Basis für $\mathrm{Hau}(A,1)$ erhalten wir also

$$B_1 = \left((2,-1,0,0)^t, (1,0,-1,0)^t\right).$$

Der erste der beiden Vektoren ist nicht in $\mathrm{Eig}(A,1)$, so daß wir ihn als x_2 wählen können. Damit erhalten wir

$$x_1 = (A-\mathbb{1}_4)x_2 = (1,0,-1,0)^t, \quad x_2 = (2,-1,0,0)^t$$

als die ersten beiden Spalten von T.

Nun wenden wir uns der Berechnung von $\mathrm{Eig}(A,2)$ zu:

$$A-2\cdot\mathbb{1}_4=\begin{pmatrix}0&1&1&2\\0&-1&0&0\\0&1&-1&1\\0&0&0&0\end{pmatrix}\overset{\text{Gauß}}{\longmapsto}\begin{pmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{pmatrix}\overset{\substack{-1\text{'en}\\\text{einfügen}}}{\longmapsto}\begin{pmatrix}-1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}.$$

Also ist $\mathrm{Eig}(A,2)=\mathrm{Lin}\left((-1,0,0,0)^t\right)$ und somit stimmen wieder die geometrische und die algebraische Vielfachheit des Eigenwertes nicht überein. Wir müssen also wieder $\mathrm{Hau}(A,2)=\mathrm{Lös}\left((A-\mathbb{1}_4)^2,0\right)$ berechnen:

$$(A-2\cdot\mathbb{1}_4)^2=\begin{pmatrix}0&0&-1&1\\0&1&0&0\\0&-2&1&-1\\0&0&0&0\end{pmatrix}\overset{\text{Gauß}}{\longmapsto}\begin{pmatrix}0&1&0&0\\0&0&1&-1\\0&0&0&0\\0&0&0&0\end{pmatrix}\overset{\substack{-1\text{'en}\\\text{einfügen}}}{\longmapsto}\begin{pmatrix}-1&0&0&0\\0&1&0&0\\0&0&1&-1\\0&0&0&-1\end{pmatrix}.$$

Wir erhalten als Basis für $\mathrm{Hau}(A,2)$ also

$$B_2\big((-1,0,0,0)^t,(0,0,-1,-1)^t\big),$$

und somit ist $x_4=(0,0,1,1)^t$ im Hauptraum, aber nicht im Eigenraum von 2. Wir erhalten deshalb

$$x_3=(A-2\cdot\mathbb{1}_4)x_4=(3,0,0,0)^t,\quad x_4=(0,0,1,1)^t$$

als die Spalten 3 und 4 der Matrix T.

Insgesamt haben wir also

$$T=\begin{pmatrix}1&2&3&0\\0&-1&0&0\\-1&0&0&1\\0&0&0&1\end{pmatrix}\in\mathrm{Gl}_4(\mathbb{Q})$$

mit

$$T^{-1}=\begin{pmatrix}0&0&-1&1\\0&-1&0&0\\\frac{1}{3}&\frac{2}{3}&\frac{1}{3}&-\frac{1}{3}\\0&0&0&1\end{pmatrix},$$

und für die Jordansche Normalform erhalten wir

$$T^{-1}\circ A\circ T=\left(\begin{array}{cc|cc}1&1&0&0\\0&1&0&0\\\hline0&0&2&1\\0&0&0&2\end{array}\right).$$

Will man nur die Normalform von A, aber nicht die Transformationsmatrix wissen, dann reicht es, die Elementarteiler zu bestimmen, was mit Hilfe von Aufgabe B2.21 sehr viel einfacher zu bewerkstelligen ist. Dies führt auf folgenden Algorithmus zur Bestimmung der Jordanschen Normalform einer Matrix A, deren charakteristisches Polynom zerfällt.

Algorithmus B2.17 (Jordansche Normalform - II).

INPUT: $A \in \mathrm{Mat}_n(\mathbb{Q})$ mit μ_A zerfällt in Linearfaktoren

OUTPUT: Liste mit den Eigenwerten von A und den Elementarteilern

1. **Schritt:** Bestimme das Minimalpolynom μ_A von A und faktorisiere es.
2. **Schritt:** Wenn μ_A nicht in Linearfaktoren zerfällt, gib eine Fehlermeldung zurück.
3. **Schritt:** Für jeden Eigenwert λ_i mit $\mathrm{mult}(\mu_A, \lambda_i) = m_i$ bestimme man für $j = 0, \ldots, m_i + 1$ die Zahlen $\mathrm{rang}\left((A - \lambda_i \mathbb{1}_n)^j\right)$ und berechne daraus den Vektor der Elementarteiler $(t_{i1}, \ldots, t_{im_i})$. Den Eigenwert und den Vektor der Elementarteiler speichere man als i-ten Eintrag in einer Liste `nf`.
4. **Schritt:** Man gebe die Liste `nf` zurück.

Bemerkung B2.18 (Jordanzerlegung einer Matrix).

Es sei $J = (a_{ij})$ eine Matrix in Jordanscher Normalform. $S = (s_{ij})$ bezeichne die Diagonalmatrix, die entsteht, wenn man in J alle Nicht-Diagonalelemente zu Null setzt, d. h. $s_{ii} = a_{ii}$ und $s_{ij} = 0$ für $i \neq j$. Ferner setzen wir $N = J - S$, d. h. N ist eine Matrix, die nur auf der oberen Nebendiagonalen Elemente ungleich Null besitzen kann.

Dann ist N *nilpotent*, und es gelten

$$J = S + N \quad \text{mit} \quad N \circ S = S \circ N.$$

Man nennt dies auch die *Jordan-Zerlegung* von J.

Um die Aussage einzusehen, beachte man, daß für $i = 1, \ldots, r$ und $1 \leq j \leq m_i$ gilt

$$J_j(\lambda_i) = \lambda_i \mathbb{1}_j + J_j(0).$$

Damit gilt

$$S = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} \lambda_i \mathbb{1}_j$$

und

$$N = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(0).$$

Aber damit folgt unmittelbar

$$N \circ S = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} \lambda_i J_j(0) = S \circ N.$$

Allgemeiner nennt man die Darstellung einer Matrix $A \in \mathrm{Mat}_n(K)$ als $A = S + N$ mit N nilpotent und S diagonalisierbar (auch *halbeinfach* genannt, engl. semi-simple, daher das S) und $S \circ N = N \circ S$ eine Jordan-Zerlegung von A. Solche Zerlegungen von Objekten in einen halbeinfachen und einen nilpotenten Anteil spielen auch in anderen Bereichen der Mathematik eine Rolle - siehe etwa Lie-Algebren oder Jordan-Algebren.

Bemerkung B2.19 (Anwendungsmöglichkeit der Jordanschen Normalform). Anwendung findet die Jordansche Normalform zum Beispiel in der Theorie der linearen Differentialgleichungen, wo ein Fundamentalsystem mit Hilfe der Exponentialabbildung einer Matrix beschrieben wird. Diese kann mit Hilfe der Jordanschen Normalform von A berechnet werden.

Wir haben oben einen *konstruktiven* Beweis des Satzes der Jordanschen Normalform für nilpotente Endomorphismen B2.12 gegeben. Aus dem Beweis ließ sich unmittelbar ableiten, wie man die Jordansche Normalform eines Endomorphismus sowie die zugehörige Jordanbasis berechnen kann. Bezahlt haben wir dies mit einer Vielzahl an Indizes, die den Beweis nicht unbedingt übersichtlich machen. Wir wollen das Kapitel damit abschließen, einen alternativen *nicht-konstruktiven* Beweis des Satzes B2.12 zu geben. Dazu erinnern wir uns zunächst der Notationen

$$I_{f,x} = \{p \in K[t] \mid p(f)(x) = 0\}$$

und

$$U_{f,x} = \{p(f)(x) \mid p \in K[t]\}$$

aus Aufgabe B2.27 und beweisen dann folgende Hilfsaussage.

Lemma B2.20 (Zyklische Unterräume haben invariante Komplemente.). Ist $f \in \mathrm{End}_K(V)$ mit $\mu_f = t^m$, so gibt es einen Vektor $0 \neq x \in V$, so daß der f-invariante Vektorraum $U_{f,x}$ Dimension m hat und ein f-invariantes direktes Komplement besitzt.

Beweis: Da m nach Satz B1.24 der Nilpotenzindex von f ist, gibt es nach Lemma B1.23 ein

$$x \in \ker(f^m) \setminus \ker(f^{m-1})$$

und nach Aufgabe B2.27 gilt dann

$$I_{f,x} = \{t^m \cdot p \mid p \in K[t]\} \tag{154}$$

und

$$U_{f,x} = \mathrm{Lin}\left(f^{m-1}(x), f^{m-2}(x), \ldots, f(x), x\right)$$

ist ein f-invarianter zyklischer Unterraum von V, der nach Aufgabe 31.36 die Dimension m hat.

Wir betrachten nun die nicht-leere Menge

$$M = \{U \leq V \mid f(U) \subseteq U, U \cap U_{f,x} = 0\}$$

und wählen in dieser ein $U \in M$ von maximaler Dimension. Wir wollen zeigen, daß der f-invariante Unterraum U ein Komplement von $U_{f,x}$ ist. Dazu setzen wir

$$W = U \oplus U_{f,x}$$

und zeigen zunächst die folgende Behauptung:

Behauptung.
Wenn $y \in V$ ist mit $f(y) \in W$, dann gilt schon $y \in W$.

Nach Voraussetzung gibt es einen Vektor $u \in U$ und ein Polynom $p \in K[t]$ mit

$$f(y) = u + p(f)(x). \tag{155}$$

Da das Minimalpolynom von f die Form $\mu_f = t^m$ hat, gilt

$$\begin{aligned} 0 = \mu_f(f)(y) = f^m(y) = f^{m-1}\big(f(y)\big) &= f^{m-1}\big(u + p(f)(x)\big) \\ &= f^{m-1}(u) + \big(f^{m-1} \circ p(f)\big)(x) = f^{m-1}(u) + q(f)(x) \in U + U_{f,x}, \end{aligned}$$

wobei $q = t^{m-1} \cdot p$ ist. Damit ist aber

$$q(f)(x) = -f^{m-1}(u) \in U \cap U_{f,x} = \{0\},$$

und wir erhalten

$$q(f)(x) = 0.$$

Letzteres impliziert

$$t^{m-1} \cdot p = q \in I_{f,x} = \{t^m \cdot g \mid g \in K[t]\}.$$

Mithin muß p durch t teilbar sein und es gibt ein Polynom $g \in K[t]$ mit

$$p = t \cdot g. \tag{156}$$

Wir setzen nun

$$z = y - g(f)(x). \tag{157}$$

Dieser Vektor z erfüllt wegen (155), (156) und (157) die Gleichung:

$$f(z) \overset{(157)}{=} f\big(y - g(f)(x)\big) = f(y) - \big(f \circ g(f)\big)(x) \overset{(156)}{=} f(y) - p(f)(x) \overset{(155)}{=} u. \tag{158}$$

Gelingt es uns, zu zeigen, daß z in W liegt, so ist auch

$$y = z + g(f)(x) \in W + U_{f,x} = W$$

und die Behauptung ist gezeigt. Wir nehmen also

$$z \notin W \tag{159}$$

an, und wollen dies zum Widerspruch führen.

Der Unterraum

$$U' = U + U_{f,z}$$

ist als Summe zweier f-invarianter Unterräume f-invariant und seine Dimension ist größer als die von U, da U' zusätzlich zu U noch den Vektor z enthält, denn aus $z \in U$ würde auch $z \in W$ folgen. Die Maximalitätsbedinung der Wahl von U impliziert also, daß

$$U' \cap U_{f,x} \neq 0$$

gilt, und wir finden mithin zwei Polynome $h, k \in K[t]$ sowie einen Vektor $v \in U$, so daß

$$0 \neq v + h(f)(z) = k(f)(x) \in U' \cap U_{f,x}. \tag{160}$$

Daraus folgt unmittelbar

$$h(f)(z) = k(f)(x) - v \in U_{f,x} + U = W. \tag{161}$$

Wir betrachten nun zunächst den Fall, daß t kein Teiler von h ist, so daß die Bézout-Identität A6.54 uns zwei Polynome $r, s \in K[t]$ schenkt mit

$$1 = r \cdot t + s \cdot h.$$

Daraus ergibt sich dann die Gleichung

$$\begin{aligned} z &= \mathrm{id}(z) = (r \cdot t + s \cdot h)(f)(z) = r(f)\big(f(z)\big) + s(f)\big(h(f)(z)\big) \\ &\overset{(158)}{=} r(f)(u) + s(f)\big(h(f)(z)\big) \overset{(161)}{\in} r(f)(U) + s(f)(W) = U + W = W, \end{aligned}$$

da U und W f-invariant sind. Dies steht aber im Widerspruch zur Annahme (159).

Als nächstes betrachten wir den Fall, daß t ein Teiler von h ist und erhalten ein Polynom $r \in K[t]$ mit $h = r \cdot t$. Damit gilt dann aber die Gleichung

$$U_{f,x} \ni k(f)(x) = v + h(f)(z) = v + r(f)\big(f(z)\big) \overset{(158)}{=} v + r(f)(u) \in U,$$

und wegen (160) hätten wir somit einen Vektor ungleich 0 in $U \cap U_{f,x} = \{0\}$ gefunden, was ein offensichtlicher Widerspruch ist.

Damit haben wir die obige Behauptung gezeigt, und wir wollen nun daraus

$$W = V$$

herleiten, was gleichbedeutend dazu ist, daß U ein f-invariantes direktes Komplement von $U_{f,x}$ in V ist. Sei dazu $y \in V$ beliebig gegeben. Wegen $f^m(y) = 0 \in W$ gibt es eine kleinste natürliche Zahl k mit $f^k(y) \in W$. Wäre diese Zahl k nicht 0, so leiten wir aus

$$f\big(f^{k-1}(y)\big) = f^k(y) \in W$$

und obiger Behauptung her, daß schon $f^{k-1}(y) \in W$ gilt, im Widerspruch zur Minimalität von k. Also ist $k = 0$ und damit $y = f^0(y) \in W$, woraus wie gewünscht $V = W$ folgt. □

Alternativer Beweis von Lemma B2.12: Wir zeigen zunächst mit Induktion nach $n = \dim_K(V)$, daß es eine Basis B gibt, bezüglich derer die Matrixdarstellung für f die Form

$$M_B^B(f) = \bigoplus_{j=1}^{n} \bigoplus_{k=1}^{t_j} J_j(0), \tag{162}$$

wobei t_j gerade die Anzahl der Jordanblöcke der Größe j ist. Für $n = 1$ ist dabei nichts zu zeigen, da dann f der Nullhomomorphismus ist. Sei also $n > 1$.

Nach Voraussetzung ist $f \in \mathrm{End}_K(V)$ ein nilpotenter Endomorphismus mit $\mu_f = t^m$. Wir wählen $0 \neq x \in V$ wie in Lemma B2.20 und betrachten den m-dimensionalen f-invarianten Unterraum $U_{f,x}$ mit Basis

$$B' = \big(f^{m-1}(x), f^{m-2}(x), \ldots, f(x), x\big)$$

bezüglich derer die Matrixdarstellung von $f_{U_{f,x}}$ ein Jordanblock der Größe m ist,

$$M_{B'}^{B'}\big(f_{U_{f,x}}\big) = J_m(0).$$

Falls $\dim_K(U_{f,x}) = m = n = \dim_K(V)$ gilt, so ist

$$U_{f,x} = V = \mathrm{Hau}(f, 0)$$

und $B = B'$ ist die gesuchte Jordanbasis von V mit $J_f = M_B^B(f) = J_n(0)$.

Ist $\dim_K(U_{f,x}) = m < n = \dim_K(V)$, so finden wir mit Lemma B2.20 ein f-invariantes Komplement U zu $U_{f,x}$ und wegen

$$t^n = \chi_f = \chi_{f_U} \cdot \chi_{f_{U_{f,x}}}$$

ist auch $f_{U_{f,x}}$ dann wieder nilpotent. Da die Dimension von U echt kleiner als die von V ist, finden wir nun mit Induktion eine Jordanbasis B'' von U mit

$$M_{B''}^{B''}(f_U) = \bigoplus_{j=1}^{n''} \bigoplus_{k=1}^{t''_j} J_j(0),$$

wobei $n'' = \dim_K(U) < \dim_K(V) = n$. Dann ist

$$B = B' \cup B''$$

eine Basis von V und die Matrixdarstellung $M_B^B(f) = M_{B'}^{B'}(f_{U_{f,x}}) \oplus M_{B''}^{B''}(f_U)$ ist eine Matrix in jordanscher Normalform (162).

Da $M_B^B(f)^m = M_B^B(f^m) = 0$ die Nullmatrix ist, kann $M_B^B(f)$ keinen Jordanblock $J_j(0)$ einer Größe $j > m$ enthalten, so daß die erste Summe in (162) nur bis m geht. Zudem gilt $t_m > 0$, da der Jordanblock zu $f_{U_{f,x}}$ die Größe m hat, und die Größe der Matrix ist die Summe der Größen der Jordanblöcke,

$$n = \sum_{j=1}^{m} j \cdot t_j.$$

Schließlich erniedrigt sich der Rang von f mit jedem Jordanblock um 1, da ein Jordanblock der Größe j gerade Rang $j-1$ hat. Damit gilt aber

$$\operatorname{rang}(f) = n - \#\text{Jordanblöcke in } M_B^B(f) = n - \sum_{j=0}^{m} t_j,$$

und da der Eigenraum von f zum Eigenwert 0 gerade der Kern von f ist, folgt

$$\dim_K \operatorname{Eig}(f,0) = \dim_K \operatorname{Ker}(f) = n - \operatorname{rang}(f) = \sum_{j=0}^{m} t_j.$$

Damit ist Lemma B2.12 gezeigt. □

Aufgaben

Aufgabe B2.21 (Berechnung der Elementarteiler).
Mit den Bezeichnungen aus Satz B2.1 zeige man, für $i = 1, \ldots, r$ und $1 \leq j \leq m_i$ gilt:

$$t_{ij} = \operatorname{rang}\left((f - \lambda_i \operatorname{id}_V)^{j-1}\right) - 2 \cdot \operatorname{rang}\left((f - \lambda_i \operatorname{id}_V)^{j}\right) + \operatorname{rang}\left((f - \lambda_i \operatorname{id}_V)^{j+1}\right)$$

bzw.

$$t_{ij} = \operatorname{rang}\left((A - \lambda_i \mathbb{1}_n)^{j-1}\right) - 2 \cdot \operatorname{rang}\left((A - \lambda_i \mathbb{1}_n)^{j}\right) + \operatorname{rang}\left((A - \lambda_i \mathbb{1}_n)^{j+1}\right).$$

Hinweise: 1. Zeige, $J_j(0)^l = \left(\delta_{\mu+l,\nu}\right)_{\mu,\nu=1,\ldots,j}$ und $\operatorname{rang}\left(J_j(0)^l\right) = \max\{0, j-l\}$ für $l \in \mathbb{N}$. 2. Man betrachte zunächst den Fall $r = 1$ und $\lambda_1 = 0$. 3. Den allgemeinen Fall führe man auf die Abbildungen $g_i := (f - \lambda_i \operatorname{id}_V)_{\operatorname{Hau}(f,\lambda_i)}$ zurück.

Aufgabe B2.22.
Bestimme die Jordansche Normalform und die zugehörige Transformationsmatrix für die Matrix A

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ 0 & -4 & 0 & 0 & -2 \end{pmatrix} \in \operatorname{Mat}_5(\mathbb{Q}).$$

Aufgabe B2.23.
Bestimme die Jordansche Normalform und die zugehörige Transformationsmatrix

$T^{-1} \in \mathrm{Gl}_4(\mathbb{Q})$ für

$$A = \begin{pmatrix} 3 & 6 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 9 & 3 & 0 \\ 0 & 12 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q}).$$

Aufgabe B2.24.
Es sei $A \in \mathrm{Mat}(5, K)$ mit $\chi_A = t(t-1)^4$, $\mu_A = t(t-1)^2$ und $\mathrm{rang}(A - \mathbb{1}_5) = 2$. Bestimme die Jordansche Normalform von A.

Aufgabe B2.25.
Zeige, ist $A \in \mathrm{Mat}_n(K)$ so, daß χ_A über K in Linearfaktoren zerfällt, so sind A und A^t konjugiert.

Aufgabe B2.26.
Beweise oder widerlege die folgende Aussage für zwei Matrizen $A, B \in \mathrm{Mat}_n(K)$:

$$A \text{ ist konjugiert zu } B \iff \chi_A = \chi_B,\ \mu_A = \mu_B \text{ und } \mathrm{rang}(A) = \mathrm{rang}(B).$$

Aufgabe B2.27.
Es sei $f \in \mathrm{End}_K(V)$ und $x \in V$.

a. Zeige, daß die Menge

$$I_{f,x} := \{p \in K[t] \mid p(f)(x) = 0\}$$

ein Ideal in $K[t]$ ist.

b. Zeige, daß die Menge

$$U_{f,x} := \{p(f)(x) \mid p \in K[t]\}$$

ein Unterraum von V ist.

c. Zeige, ist $m \in \mathbb{N}$ minimal mit $f^m(x) = 0$, so gilt

$$I_{f,x} = \{t^m \cdot p \mid p \in K[t]\}$$

und

$$U_{f,x} = \mathrm{Lin}\left(f^{m-1}(x), f^{m-2}(x), \ldots, f(x), x\right)$$

ist ein zyklischer Unterraum wie in Aufgabe 31.36.

§ B3 Spektralsatz für normale und unitäre Endomorphismen

In diesem Abschnitt sei V ein euklidischer oder unitärer Raum der Dimension $1 \leq n < \infty$ mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und euklidischer Norm $|| \cdot ||$.

A) Der Spektralsatz für normale Endomorphismen

Definition B3.1.
Es sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ und $A \in \mathrm{Mat}_n(\mathbb{K})$.

a. f heißt *normal*, falls $f^* \circ f = f \circ f^*$.

b. A heißt *normal*, falls $A^* \circ A = A \circ A^*$.

Bemerkung B3.2.

a. Jede symmetrische Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ und jede hermitesche Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ ist normal, denn wegen $A = A^*$ gilt auch

$$A \circ A^* = A \circ A = A^* \circ A.$$

b. Jede orthogonale Matrix $A \in \mathrm{O}(n)$ und jede unitäre Matrix $A \in \mathrm{U}(n)$ ist normal, denn wegen $A^* = A^{-1}$ gilt auch

$$A \circ A^* = A \circ A^{-1} = \mathbb{1}_n = A^{-1} \circ A = A^* \circ A.$$

Lemma B3.3 (Matrixdarstellung normaler Endomorphismen).
Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ und B eine ONB von V.
Genau dann ist f normal, wenn $M_B^B(f)$ normal ist.

Beweis: Ist f normal, so gilt

$$\begin{aligned} M_B^B(f)^* \circ M_B^B(f) &\overset{37.4}{=} M_B^B(f^*) \circ M_B^B(f) = M_B^B(f^* \circ f) \\ &= M_B^B(f \circ f^*) = M_B^B(f) \circ M_B^B(f^*) \overset{37.4}{=} M_B^B(f) \circ M_B^B(f)^* \end{aligned}$$

und somit ist $M_B^B(f)$ normal. Ist umgekehrt $M_B^B(f)$ normal, so gilt

$$\begin{aligned} M_B^B(f^* \circ f) &= M_B^B(f^*) \circ M_B^B(f) \overset{37.4}{=} M_B^B(f)^* \circ M_B^B(f) \\ &= M_B^B(f) \circ M_B^B(f)^* \overset{37.4}{=} M_B^B(f) \circ M_B^B(f^*) = M_B^B(f \circ f^*). \end{aligned}$$

Dann stimmen aber die Abbildungen $f^* \circ f$ und $f \circ f^*$ überein, und somit ist f normal. □

Beispiel B3.4 (Normale Abbildung).
In Beispiel 37.5 gilt

$$\begin{aligned} M_E^E(f) \circ M_E^E(f)^* &= \begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix} \circ \begin{pmatrix} 2 & -4 \\ 4 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 20 & 0 \\ 0 & 20 \end{pmatrix} = \begin{pmatrix} 2 & -4 \\ 4 & 2 \end{pmatrix} \circ \begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix} = M_E^E(f)^* \circ M_E^E(f). \end{aligned}$$

Somit ist $M_E^E(f)$ normal und da E eine ONB ist, ist dann auch f normal.

Lemma B3.5 (Normale Abbildungen).
Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ normal. Genau dann gilt $x \in \mathrm{Eig}(f, \lambda)$, wenn $x \in \mathrm{Eig}\left(f^*, \overline{\lambda}\right)$.

Beweis: Für ein beliebiges $x \in V$ gilt

$$\begin{aligned} \langle f^*(x) - \overline{\lambda}x, f^*(x) - \overline{\lambda}x \rangle &= \langle f^*(x), f^*(x) \rangle - \langle f^*(x), \overline{\lambda}x \rangle - \langle \overline{\lambda}x, f^*(x) \rangle + \langle \overline{\lambda}x, \overline{\lambda}x \rangle \\ &= \langle f \circ f^*(x), x \rangle - \overline{\lambda}\langle f^*(x), x \rangle - \lambda \langle x, f^*(x) \rangle + \lambda\overline{\lambda}\langle x, x \rangle \\ &= \langle f^* \circ f(x), x \rangle - \overline{\lambda}\langle x, f(x) \rangle - \lambda \langle f(x), x \rangle + \lambda\overline{\lambda}\langle x, x \rangle \\ &= \langle f(x), f(x) \rangle - \langle \lambda x, f(x) \rangle - \langle f(x), \lambda x \rangle + \langle \lambda x, \lambda x \rangle \\ &= \langle f(x) - \lambda x, f(x) - \lambda x \rangle. \end{aligned}$$

Da das Skalarprodukt positiv definit ist, gilt somit

$$\begin{aligned} x \in \mathrm{Eig}(f, \lambda) &\iff f(x) - \lambda x = 0 \iff \langle f(x) - \lambda x, f(x) - \lambda x \rangle = 0 \\ &\iff \langle f^*(x) - \overline{\lambda}x, f^*(x) - \overline{\lambda}x \rangle = 0 \iff f^*(x) - \overline{\lambda}x = 0 \\ &\iff x \in \mathrm{Eig}\left(f^*, \overline{\lambda}\right). \end{aligned}$$

□

Satz B3.6 (Spektralsatz für normale Endomorphismen).
Für $f \in \mathrm{End}_{\mathbb{K}}(V)$ sind die folgenden beiden Aussagen gleichwertig:

a. f ist normal und χ_f zerfällt über $\mathbb{K}$ in Linearfaktoren.

b. V besitzt eine ONB aus Eigenvektoren von f.

Insbesondere ist f dann bezüglich einer ONB diagonalisierbar.

Beweis:
b. $\Longrightarrow$ a.: Besitzt V eine ONB B aus Eigenvektoren, so zerfällt χ_f nach Satz B1.27

über $\mathbb{K}$ in Linearfaktoren. Zudem ist dann $M_B^B(f)$ eine Diagonalmatrix, und wegen Korollar 37.4 ist dann auch

$$M_B^B(f)^* = \overline{M_B^B(f)}^t$$

eine Diagonalmatrix. Da zwei Diagonalmatrizen stets kommutieren, gilt also

$$M_B^B(f)^* \circ M_B^B(f) = M_B^B(f) \circ M_B^B(f)^*,$$

d.h. $M_B^B(f)$ ist normal. Nach Lemma B3.3 ist dann aber auch f normal.

a. $\Longrightarrow$ b.: Wir führen den Beweis mit Induktion nach $n = \dim_{\mathbb{K}}(V)$, wobei für $n = 1$ der Endomorphismus f für jede ONB $B = (x_1)$ diagonalisierbar ist und zudem x_1 ein Eigenvektor von f ist. Sei also $n > 1$.

Da χ_f über $\mathbb{K}$ in Linearfaktoren zerfällt, besitzt f einen Eigenwert $\lambda \in \mathbb{K}$ und einen zugehörigen Eigenvektor $0 \neq x \in V$. Dann ist $U := \mathrm{Lin}\,(x)$ ein f-invarianter Unterraum der Dimension 1 und es gilt

$$V = U \perp U^\perp = U \oplus U^\perp$$

nach Proposition 36.28.

Wir zeigen nun zunächst, daß auch $U^\perp$ ein f-invarianter Unterraum ist, der dann die Dimension $n-1$ hat. Sei dazu $y \in U^\perp$, dann gilt

$$\langle f(y), x\rangle = \langle y, f^*(x)\rangle \overset{B3.5}{=} \langle y, \overline{\lambda}x\rangle = \overline{\lambda} \cdot \langle y, x\rangle = 0,$$

da $y \perp x$. Damit gilt dann aber auch $f(y) \perp x$, und somit $f(y) \in U^\perp$.

Als nächstes wollen wir zeigen, daß $f_{U^\perp}$ normal ist. Dazu beachten wir zunächst, daß $U^\perp$ auch f^*-invariant ist, da für $y \in U^\perp$ wie oben

$$\langle f^*(y), x\rangle = \langle y, f(x)\rangle = \langle y, \lambda x\rangle = \lambda \cdot \langle y, x\rangle = 0$$

und somit $f^*(y) \perp x$ und $f^*(y) \in U^\perp$ gilt. Aus der definierenden Eigenschaft der adjungierten Abbildung folgt dann aber, daß die adjungierte Abbildung $(f_{U^\perp})^*$ der Einschränkung $f_{U^\perp}$ genau die Einschränkung $(f^*)_{U^\perp}$ der adjungierten Abbildung f^* auf $U^\perp$ ist. Die Normalität von f überträgt sich also direkt auf $f_{U^\perp}$ durch

$$f_{U^\perp} \circ (f_{U^\perp})^* = (f \circ f^*)_{U^\perp} = (f^* \circ f)_{U^\perp} = (f_{U^\perp})^* \circ f_{U^\perp}.$$

Außerdem gilt

$$\chi_f = \chi_{f_U} \cdot \chi_{f_{U^\perp}},$$

da V die direkte Summe der beiden f-invarianten Unterräume U und $U^\perp$ ist, und deshalb zerfällt $\chi_{f_{U^\perp}}$ über $\mathbb{K}$ in Linearfaktoren. Nach Induktion besitzt $U^\perp$ deshalb eine ONB $(x_2, \ldots, x_n)$ aus Eigenvektoren von $f_{U^\perp}$. Dann ist aber $B = (x_1, x_2, \ldots, x_n)$ mit $x_1 = \frac{x}{||x||}$ eine ONB aus Eigenvektoren von f. □

Korollar B3.7 (Spektralsatz für normale Matrizen).
Für eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ sind die folgenden beiden Aussagen gleichwertig:

a. A ist normal und χ_A zerfällt über $\mathbb{K}$ in Linearfaktoren.

b. Es gibt ein T in $\mathrm{O}(n)$ bzw. $\mathrm{U}(n)$, so daß $T^{-1} \circ A \circ T$ eine Diagonalmatrix ist.

Beweis: Wenden wir den Spektralsatz B3.6 auf f_A und $\mathbb{K}^n$ mit dem kanonischen Skalarprodukt an, so enthält die Basistransformationsmatrix $T = T_E^B$ genau die Vektoren der ONB B als Spalten und ist nach Proposition 36.23 daher orthogonal bzw. unitär. □

Der Beweis ist konstruktiv, sofern man die Eigenwerte von A exakt kennt. Man leitet daraus folgenden prinzipiellen Algorithmus zur Bestimmung von T her.

Algorithmus B3.8.
INPUT: $A \in \mathrm{Mat}_n(\mathbb{K})$ normal mit χ_A zerfällt über $\mathbb{K}$.
OUTPUT: T in $\mathrm{O}(n)$ bzw. $\mathrm{U}(n)$, so daß $T^{-1} \circ A \circ T$ Diagonalgestalt hat.

1. **Schritt:** Bestimme die Eigenwerte von A.
2. **Schritt:** Bestimme für jeden Eigenwert von A eine Basis des zugehörigen Eigenraumes.
3. **Schritt:** Orthonormalisiere die Basen der Eigenräume mit dem Orthonormalisierungsverfahren von Gram-Schmidt und schreibe die Vektoren als Spalten in eine Matrix T.
4. **Schritt:** Gib schließlich T zurück.

Beispiel B3.9 (Diagonalisierung einer normalen Abbildung).
Die Abbildung f in Beispiel 37.5 ist nach Beispiel B3.4 normal.

$$\chi_f = \det\left(t \cdot \mathbb{1}_2 - M_E^E(f)\right) = \begin{vmatrix} t-2 & -4 \\ 4 & t-2 \end{vmatrix} = t^2 - 4t + 20 = \big(t-(2+4i)\big) \cdot \big(t-(2-4i)\big)$$

zerfällt über $\mathbb{C}$ in Linearfaktoren. Mithin gibt es wegen des Spektralsatzes eine ONB aus Eigenvektoren von f, so daß

$$M_B^B(f) = \begin{pmatrix} 2+4i & 0 \\ 0 & 2-4i \end{pmatrix}.$$

Um B zu berechnen, berechnen wir zunächst die beiden Eigenräume $\mathrm{Eig}(f, 2+4i)$ und $\mathrm{Eig}(f, 2-4i)$ ausgehend von der Matrixdarstellung $M_E^E(f)$ in Beispiel 37.5.

Für $\mathrm{Eig}(f, 2+4i) = \mathrm{Lös}\left(M_E^E(f) - (2+4i)\cdot \mathbb{1}_2, 0\right)$ liefert unser Algorithmus

$$\begin{pmatrix} -4i & 4 \\ -4 & -4i \end{pmatrix} \overset{\substack{II\mapsto II+i\cdot I \\ I\mapsto \frac{1}{-4i}\cdot I}}{\longmapsto} \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} \overset{\substack{-1\text{'en} \\ \text{ergänzen}}}{\longmapsto} \begin{pmatrix} 1 & i \\ 0 & -1 \end{pmatrix},$$

so daß $x_1 = (i, -1)^t$ eine Basis von $\mathrm{Eig}(f, 2+4i)$ ist. Analog erhalten wir für $\mathrm{Eig}(f, 2-4i) = \mathrm{Lös}\left(M_E^E(f) - (2-4i)\cdot \mathbb{1}_2, 0\right)$

$$\begin{pmatrix} 4i & 4 \\ -4 & 4i \end{pmatrix} \overset{\substack{II\mapsto II-i\cdot I \\ I\mapsto \frac{1}{4i}\cdot I}}{\longmapsto} \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix} \overset{\substack{-1\text{'en} \\ \text{ergänzen}}}{\longmapsto} \begin{pmatrix} 1 & -i \\ 0 & -1 \end{pmatrix},$$

so daß $x_2 = (-i, -1)^t$ eine Basis von $\mathrm{Eig}(f, 2-4i)$ ist. Die Vektoren x_1 und x_2 sind bereits orthogonal zueinander, da

$$\langle x_1, x_2\rangle = \bar{i}\cdot(-i) + \overline{-1}\cdot(-1) = -1 + 1 = 0.$$

Mithin reicht es, sie zu normieren, und wir erhalten die gewünschte ONB

$$B = \left(\tfrac{1}{||x_1||}\cdot x_1, \tfrac{1}{||x_2||}\cdot x_2\right) = \left(\left(\tfrac{i}{\sqrt{2}}, \tfrac{-1}{\sqrt{2}}\right)^t, \left(\tfrac{-i}{\sqrt{2}}, \tfrac{-1}{\sqrt{2}}\right)^t\right).$$

Es ist übrigens kein Zufall, daß die beiden Eigenvektoren im letzten Beispiel orthogonal zueinander standen.

Lemma B3.10.
Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ normal, $x \in \mathrm{Eig}(f, \lambda)$, $y \in \mathrm{Eig}(f, \mu)$ und $\lambda \neq \mu$, dann gilt $x \perp y$.

Beweis: Nach Voraussetzung gilt

$$\lambda\cdot\langle x, y\rangle = \langle\overline{\lambda}x, y\rangle \overset{B3.5}{=} \langle f^*(x), y\rangle \overset{37.2}{=} \langle x, f(y)\rangle = \langle x, \mu y\rangle = \mu\cdot\langle x, y\rangle.$$

Für die Differenz der beiden Seiten erhalten wir dann $(\lambda-\mu)\cdot\langle x, y\rangle = 0$, wobei nach Voraussetzung $\lambda - \mu \neq 0$ gilt. Also ist $\langle x, y\rangle = 0$ und somit $x \perp y$. □

Satz B3.11 (Spektralzerlegung für normale Endomorphismen).
Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ normal mit zerfallendem charakteristischem Polynom und seien $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ die paarweise verschiedenen Eigenwerte von f. Ferner bezeichne

$$\pi_i : V \longrightarrow V$$

die orthogonale Projektion von V auf $\mathrm{Eig}(f, \lambda_i)$.
Dann ist

$$V = \mathrm{Eig}(f, \lambda_1) \perp \ldots \perp \mathrm{Eig}(f, \lambda_r)$$

die *orthogonale Summe* der Eigenräume von f und es gilt

$$f = \lambda_1\cdot\pi_1 + \ldots + \lambda_r\cdot\pi_r.$$

Man nennt dies die *Spektralzerlegung* von f.

Beweis: Nach dem Spektralsatz für normale Abbildungen ist f diagonalisierbar. Deshalb folgt aus Satz B1.27, daß V die direkte Summe der Eigenräume von f ist. Wegen Lemma B3.10 ist diese Summe dann eine orthogonale Summe.

Ist nun $x = x_1 + \ldots + x_r \in V$ mit $x_i \in \mathrm{Eig}(f, \lambda_i)$ gegeben, so gilt

$$\pi_j(x_i) = \delta_{ij} \cdot x_i,$$

da $x_j \in \mathrm{Eig}(f, \lambda_j)$ und $x_i \perp \mathrm{Eig}(f, \lambda_j)$ für $i \neq j$, und somit

$$\pi_j(x) = \pi_j(x_1) + \ldots + \pi_j(x_r) = x_j.$$

Damit erhalten wir dann

$$\begin{aligned}(\lambda_1 \cdot \pi_1 + \ldots + \lambda_r \cdot \pi_r)(x) =& \lambda_1 \cdot \pi_1(x) + \ldots + \lambda_r \cdot \pi_r(x) \\ =& \lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r = f(x_1) + \ldots + f(x_r) = f(x).\end{aligned}$$

□

Bemerkung B3.12 (Spektralzerlegung und Hauptraumzerlegung).
Die Projektionen π_i stimmen übrigens mit den Projektionen $Q_i(f)$ aus Bemerkung B1.25 überein.

B) Orthogonale und unitäre Abbildungen

Wir kommen jetzt zu den strukturerhaltenden Abbildungen, d. h. zu solchen, die mit dem Skalarprodukt verträglich sind. Diese haben einen speziellen Namen.

Definition B3.13 (Orthogonale / unitäre Abbildungen).

a. Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{R}}(V)$ heißt *orthogonal*, falls $f^* \circ f = \mathrm{id}_V$ gilt. $\mathrm{O}(V) := \{f \in \mathrm{End}_{\mathbb{R}}(V) \mid f \text{ ist orthogonal}\}$ heißt die *orthogonale Gruppe* von V.

b. Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{C}}(V)$ heißt *unitär*, falls $f^* \circ f = \mathrm{id}_V$ gilt. Wir nennen $\mathrm{U}(V) := \{f \in \mathrm{End}_{\mathbb{C}}(V) \mid f \text{ ist unitär}\}$ die *unitäre Gruppe* von V.

Proposition B3.14 (Matrixdarstellung orthogonaler / unitärer Abbildungen).
Es sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ und B eine ONB von V.
Genau dann ist f orthogonal bzw. unitär, wenn $M_B^B(f)$ orthogonal bzw. unitär ist.

Beweis: Es gilt:

$$\begin{aligned} f \text{ ist orthogonal bzw. unitär} &\iff f^* \circ f = \mathrm{id}_V \\ &\iff M_B^B(f^*) \circ M_B^B(f) = M_B^B(f^* \circ f) = \mathbb{1}_n \\ &\overset{37.4}{\iff} M_B^B(f)^* \circ M_B^B(f) = \mathbb{1}_n \\ &\iff M_B^B(f) \text{ ist orthogonal bzw. unitär.} \end{aligned}$$

□

Korollar B3.15 (Orthogonal / unitär $\Longrightarrow$ normal).
Orthogonale und unitäre Abbildungen sind normal.

Beweis: Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ orthogonal oder unitär und B eine ONB von V, so ist $M_B^B(f)$ nach Proposition B3.14 orthogonal oder unitär. Nach Bemerkung B3.2 ist dann $M_B^B(f)$ auch normal, und mit Lemma B3.3 ist f deshalb normal. □

Proposition B3.16 (Charakterisierung orthogonaler / unitärer Endomorphismen).
Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ ist genau dann orthogonal bzw. unitär, wenn

$$\langle f(x), f(y)\rangle = \langle x, y\rangle$$

für alle $x, y \in V$ gilt.

Beweis: Ist f orthogonal bzw. unitär und sind $x, y \in V$, so gilt

$$\langle x, y\rangle = \langle x, \mathrm{id}_V(y)\rangle = \langle x, f^* \circ f(y)\rangle = \langle f(x), f(y)\rangle.$$

Ist umgekehrt f mit dem Skalarprodukt verträglich, so gilt

$$\langle x, y\rangle = \langle f(x), f(y)\rangle = \langle (f^* \circ f)(x), y\rangle$$

und somit

$$\langle (f^* \circ f)(x) - \mathrm{id}_V(x), y\rangle = \langle (f^* \circ f)(x), y\rangle - \langle x, y\rangle = 0$$

für alle $x, y \in V$. Für $x \in V$ beliebig setzen wir $y = (f^* \circ f)(x) - \mathrm{id}_V(x)$ und erhalten

$$\langle (f^* \circ f)(x) - \mathrm{id}_V(x), (f^* \circ f)(x) - \mathrm{id}_V(x)\rangle = 0,$$

was $(f^* \circ f)(x) - \mathrm{id}_V(x) = 0$ und $(f^* \circ f)(x) = \mathrm{id}_V(x)$ zur Folge hat. Mithin ist $f^* \circ f = \mathrm{id}_V$ und f ist orthogonal bzw. unitär. □

Proposition B3.17 (Eigenschaften orthogonaler / unitärer Abbildungen).
Es seien $f, g \in \mathrm{End}_{\mathbb{K}}(V)$ orthogonal bzw. unitär und $x, y \in V$

a. $||f(x)|| = ||x||$.

b. $x \perp y$ genau dann, wenn $f(x) \perp f(y)$.

c. Jeder Eigenwert von f hat Betrag 1.

d. f ist bijektiv.

e. f^{-1} und $f \circ g$ sind orthogonal bzw. unitär, d.h. $\mathrm{O}(V)$ und $\mathrm{U}(V)$ sind Gruppen.

Insbesondere, orthogonale und unitäre Abbildungen erhalten Längen und Abstände.

Beweis:

a. $||f(x)||^2 = \langle f(x), f(x)\rangle = \langle x, x\rangle = ||x||^2$.

b. $x \perp y \iff \langle f(x), f(y)\rangle = \langle x, y\rangle = 0 \iff f(x) \perp f(y)$.

c. Ist $0 \neq x \in V$ ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{K}$, so gilt nach a.

$$||x|| = ||f(x)|| = ||\lambda x|| = |\lambda| \cdot ||x||,$$

also $|\lambda| = 1$, da $x \neq 0$.

d. Ist $x \in \mathrm{Ker}(f)$, so gilt nach a. $0 = ||f(x)|| = ||x||$, und somit $x = 0$. Also ist f injektiv, und da V endlich-dimensional ist, ist f somit auch bijektiv.

e. f^{-1} und $f \circ g$ sind orthogonal bzw. unitär, da für $x, y \in V$ gelten

$$\big\langle f^{-1}(x), f^{-1}(y)\big\rangle = \Big\langle f\big(f^{-1}(x)\big), f\big(f^{-1}(y)\big)\Big\rangle = \langle x, y\rangle$$

und

$$\big\langle (f \circ g)(x), (f \circ g)(y)\big\rangle = \Big\langle f\big(g(x)\big), f\big(g(y)\big)\Big\rangle = \big\langle g(x), g(y)\big\rangle = \langle x, y\rangle.$$

□

Bemerkung B3.18 (Orthogonale Abbildungen sind winkeltreu.).
Orthogonale Abbildungen erhalten Winkel, d.h. ist $f \in \mathrm{O}(V)$ und sind $0 \neq x, y \in V$, so gilt

$$\measuredangle\big(f(x), f(y)\big) = \measuredangle(x, y),$$

da f das Skalarprodukt und die euklidische Norm erhält. Es gilt nämlich

$$\measuredangle\big(f(x), f(y)\big) = \arccos\left(\frac{\langle f(x), f(y)\rangle}{||f(x)|| \cdot ||f(y)||}\right) = \arccos\left(\frac{\langle x, y\rangle}{||x|| \cdot ||y||}\right) = \measuredangle(x, y).$$

Satz B3.19 (Spektralsatz für unitäre Abbildungen).
Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{C}}(V)$ ist genau dann unitär, wenn V eine ONB aus Eigenvektoren von f besitzt und alle Eigenwerte von f Betrag 1 haben.

Beweis: Ist f unitär, so ist f nach Korollar B3.15 normal und χ_f zerfällt nach dem Fundamentalsatz der Algebra über $\mathbb{C}$ in Linearfaktoren. Aufgrund des Spektralsatzes für normale Abbildungen B3.6 besitzt V dann eine ONB aus Eigenvektoren von f. Zudem haben die Eigenwerte nach Proposition B3.17 alle Betrag 1.

Besitzt umgekehrt V eine ONB $B = (x_1, \ldots, x_n)$ aus Eigenvektoren von f zu Eigenwerten $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ vom Betrag 1, so gilt

$$M_B^B(f)^* \circ M_B^B(f) = \begin{pmatrix} \overline{\lambda_1} & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \overline{\lambda_n} \end{pmatrix} \circ \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

$$= \begin{pmatrix} \overline{\lambda_1}\lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \overline{\lambda_n}\lambda_n \end{pmatrix} = \begin{pmatrix} |\lambda_1|^2 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & |\lambda_n|^2 \end{pmatrix} = \mathbb{1}_n.$$

Mithin ist $M_B^B(f)$ unitär, und damit ist auch f unitär nach Proposition B3.14. □

Korollar B3.20 (Spektralsatz für unitäre Matrizen).
Ist $A \in \mathrm{U}(n)$, dann gibt es ein $T \in \mathrm{U}(n)$ mit

$$T^{-1} \circ A \circ T = T^* \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}$$

wobei $\lambda_i \in \mathbb{C}$ mit $|\lambda_i| = 1$, $i = 1, \ldots, n$, die Eigenwerte von A sind. Insbesondere ist jede unitäre Matrix diagonalisierbar.

Beweis: Ist A unitär, dann ist f_A unitär und wir finden eine ONB von $\mathbb{C}^n$ aus Eigenvektoren von f_A, also von A, und alle Eigenwerte haben Betrag 1. Schreiben wir die Eigenvektoren als Spalten in eine Matrix T, so ist $T \in \mathrm{U}(n)$ und T transformiert A in eine Diagonalmatrix. □

Beispiel B3.21 (Unitäre Matrix).
Betrachten wir $\mathbb{C}^3$ mit dem kanonischen Skalarprodukt sowie die Matrix

$$A = \frac{1}{9} \cdot \begin{pmatrix} 1 & 8 & -4 \\ -4 & 4 & 7 \\ 8 & 1 & 4 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{C}).$$

Man rechnet sofort nach, daß $A^* \circ A = \mathbb{1}_3$, daß A also orthogonal bzw. unitär ist, mit charakteristischem Polynom

$$\chi_A = t^3 - t^2 + t - 1 = (t-1) \cdot (t-i) \cdot (t+i).$$

Es gibt also eine unitäre Matrix $T \in \mathrm{U}(3)$ mit

$$T^* \circ A \circ T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}.$$

Um T zu bestimmen, berechnen wir zunächst den Eigenraum $\mathrm{Eig}(A, 1)$ und finden $\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)^t$ als ONB. Durch Einsetzen in das Gleichungssystem überzeugt man sich, daß $(4, -1+3i, -1-3i)^t$ eine Lösung von $(A - i\mathbb{1}_3)x = 0$ ist, und durch Normierung erhalten wir dann $\left(\frac{2}{3}, \frac{-1+3i}{6}, \frac{-1-3i}{6}\right)^t$ als ONB von $\mathrm{Eig}(A, i)$. Da A eine reelle Matrix ist, muß somit $-i$ gerade den konjugiert komplexen Vektor als Eigenvektor haben, d. h. $\left(\frac{2}{3}, \frac{-1-3i}{6}, \frac{-1+3i}{6}\right)^t$ ist eine ONB von $\mathrm{Eig}(A, -i)$.

Wir erhalten also

$$T = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{-1+3i}{6} & \frac{-1-3i}{6} \\ \frac{2}{3} & \frac{-1-3i}{6} & \frac{-1+3i}{6} \end{pmatrix} \in \mathrm{U}(3)$$

als Transformationsmatrix mit

$$T^{-1} \circ A \circ T = T^* \circ A \circ T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}.$$

Bemerkung B3.22 (Spektralsatz für orthogonale Abbildungen).
Mit dem gleichen Beweis wie in Satz B3.19 kann man zeigen, daß eine orthogonale Abbildung $f \in \mathrm{O}(V)$, deren charakteristisches Polynom über $\mathbb{R}$ in Linearfaktoren zerfällt, ebenfalls bezüglich einer ONB diagonalisierbar ist.

Orthogonale Abbildungen lassen sich im allgemeinen aber nicht diagonalisieren, insbesondere nicht durch eine ONB. Wir haben in Beispiel 35.13 gesehen, daß die Matrix

$$T(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

die eine Drehung um den Ursprung um den Winkel α beschreibt, i.a. nicht diagonalisierbar ist. Man kann zeigen, daß dieses Beispiel im wesentlichen auch das einzige ist. Es gilt nämlich:

Ist $f \in \mathrm{End}_{\mathbb{R}}(V)$ orthogonal, dann gibt es eindeutig bestimmte Zahlen $r, s, t \in \mathbb{N}$ sowie Winkel $\alpha_1, \ldots, \alpha_t \in (0, 2\pi) \setminus \{\pi\}$ und eine ONB B von V, so daß

$$M_B^B(f) = \mathbb{1}_r \oplus -\mathbb{1}_s \oplus T(\alpha_1) \oplus \ldots \oplus T(\alpha_t).$$

C) Klassifikation der Kegelschnitte in der euklidischen Ebene

Man kann die Überlegungen in 37.17 verallgemeinern, was wir hier im Fall $n = 2$ tun wollen. Dazu betrachten wir die Lösungsmenge einer allgemeinen quadratischen Gleichung in zwei Unbekannten. Dies führt zur Klassifikation der Kegelschnitte in der euklidischen Ebene.

Definition B3.23.
Es sei $\big(V, \langle\cdot,\cdot\rangle\big)$ ein euklidischer Raum.

a. Eine Abbildung $f : V \to V$ heißt eine *affine Abbildung* auf V, falls es ein $y \in V$ gibt und ein $g \in \mathrm{End}_{\mathbb{R}}(V)$ mit $f(x) = y + g(x)$ für alle $x \in V$.

b. Für $y \in V$ nennen wir die affine Abbildung

$$\tau_y : V \to V : x \mapsto x + y$$

die *Translation* um den Vektor y.

c. Eine Abbildung $f : V \to V$ heißt eine *Ähnlichkeit*, wenn es einen Vektor $y \in V$ gibt und eine orthogonale Abbildung $g \in \mathrm{O}(V)$ mit $f = \tau_y \circ g$, d. h.

$$f(x) = \tau_y\big(g(x)\big) = y + g(x) \quad \forall\, x \in V.$$

d. Ist $V = \mathbb{R}^n$ und sei $f = \tau_y \circ g$ mit $g \in \mathrm{End}_{\mathbb{R}}(V)$ eine bijektive affine Abbildung auf V, dann nennen wir die induzierte Abbildung

$$\mathbb{R}[t_1, \ldots, t_n] \to \mathbb{R}[t_1, \ldots, t_n] : p \mapsto p\big(f(t_1, \ldots, t_n)\big)$$

auf der Menge der Polynome in den Veränderlichen $t_1, \ldots, t_n$ einen *affinen Koordinatenwechsel* von $\mathbb{R}[t_1, \ldots, t_n]$.

Bemerkung B3.24.

a. Jede affine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ läßt sich offenbar in eindeutiger Weise schreiben, als $f = \tau_y \circ g$ mit $y = f(0) \in V$ und $g \in \mathrm{End}_{\mathbb{R}}(V)$.

b. Ist $f = \tau_y \circ g : \mathbb{R}^n \to \mathbb{R}^n$ eine affine Abbildung mit $y \in \mathbb{R}^n$ und $g \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^n)$ bijektiv, dann gibt es eine eindeutig bestimmte Matrix $T \in \mathrm{Gl}_n(\mathbb{R})$ mit $g = f_T$. Damit gilt für $p \in \mathbb{R}[t_1, \ldots, t_n]$ und $t = (t_1, \ldots, t_n)^t$

$$p\big(f(t_1, \ldots, t_n)\big) = p(Tt + y).$$

Ist beispielsweise $p = t_1^2 + 3t_2 - 1 \in \mathbb{R}[t_1, t_2]$, $T = T\left(\frac{\pi}{2}\right)$ die Drehung um $90°$ und $y = (2, -2)$, dann ist für $f = \tau_y \circ f_T$

$$p\big(f(t_1, t_2)\big) = p(-t_2 + 2, t_1 - 2) = (-t_2 + 2)^2 + 3(t_1 - 2) - 1.$$

Definition B3.25.
Es sei $p \in \mathbb{R}[t_1, \ldots, t_n]$ dann nennen wir die Menge

$$N(p) = \{\lambda = (\lambda_1, \ldots, \lambda_n)^t \in \mathbb{R}^n \mid p(\lambda) = 0\}$$

eine *algebraische Hyperfläche* von $\mathbb{R}^n$. Ist $\deg(p) = d$, so nennen wir d auch den *Grad* der Hyperfläche. Ist $n = 2$, so sprechen wir auch von *algebraischen Kurven* statt von algebraischen Hyperflächen.

Definition B3.26.
Wir definieren auf $\mathbb{R}[t_1, \ldots, t_n]$ eine Relation durch

$$p \equiv q \quad :\Leftrightarrow \quad \exists\, c \in \mathbb{R} \setminus \{0\} : p = c \cdot q$$

für $p, q \in \mathbb{R}[t_1, \ldots, t_n]$. Wir nennen p und q mit $p \equiv q$ auch *äquivalent*.

Bemerkung B3.27.
Man sieht sofort, daß $\equiv$ eine Äquivalenzrelation auf $\mathbb{R}[t_1, \ldots, t_n]$ definiert.

Ferner gilt offensichtlich, daß für zwei äquivalente Polynome $p, q \in \mathbb{R}[t_1, \ldots, t_n]$ auch $N(p) = N(q)$ gilt. Interessiert man sich also nur für das Nullstellengebilde von p, so kann man p getrost durch ein äquivalentes Polynom ersetzen und somit erreichen, daß der konstante Anteil von p entweder 0 oder -1 ist.

Im Folgenden interessieren wir uns nur noch für algebraische Kurven vom Grad zwei.

Bemerkung B3.28.
Ist $p \in \mathbb{R}[t_1, t_2]$ ein allgemeines Polynom zweiten Grades, dann gibt es reelle Zahlen $\alpha_{11}, \alpha_{12} = \alpha_{21}, \alpha_{22}, \alpha_1, \alpha_2, \alpha \in \mathbb{R}$ so, daß

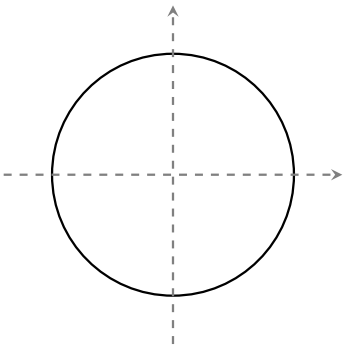
$$p = \alpha_{11}t_1^2 + 2\alpha_{12}t_1t_2 + \alpha_{22}t_2^2 + \alpha_1t_1 + \alpha_2t_2 + \alpha = \langle t, St\rangle + \langle a, t\rangle + \alpha,$$

wobei $t = (t_1, t_2)^t$, $0 \neq S = (\alpha_{ij})_{i,j\in\{1,2\}} \in \mathrm{Mat}_2(\mathbb{R})$ und $a = (\alpha_1, \alpha_2)^t$.

Beispiel B3.29.
Für $S = \mathbb{1}_2$, $a = (0, 0)^t$ und $\alpha = -1$ erhalten wir $p = t_1^2 + t_2^2 - 1$, und die Nullstellenmenge davon,

$$N\big(t_1^2 + t_2^2 - 1\big) = \{\lambda = (\lambda_1, \lambda_2)^t \in \mathbb{R}^2 \mid \lambda_1^2 + \lambda_2^2 = 1\},$$

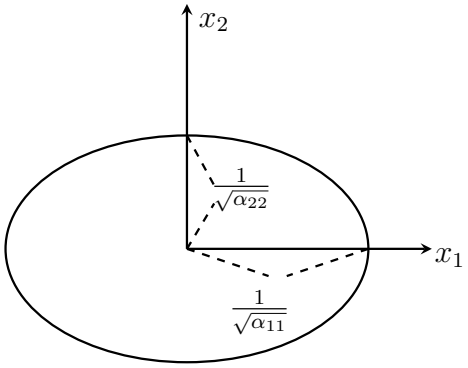
ist offenbar der Einheitskreis.



Ist S eine Diagonalmatrix mit positiven Diagonaleinträgen, d. h. $\alpha_{11}, \alpha_{22} > 0$ und $\alpha_{12} = \alpha_{21} = 0$, und ist ferner $a = (0,0)^t$ und $\alpha = -1$, dann erhalten wir als Nullstellengebilde von p

$$N\Big(\big(\sqrt{\alpha_{11}}t_1\big)^2 + \big(\sqrt{\alpha_{22}}t_2\big)^2 - 1\Big) = \Big\{(\lambda_1,\lambda_2)^t \in \mathbb{R}^2 \;\Big|\; \big(\sqrt{\alpha_{11}}\lambda_1\big)^2 + \big(\sqrt{\alpha_{22}}\lambda_2\big)^2 = 1\Big\}$$

eine Ellipse.



Satz B3.30.
Es sei

$$p = \langle t, St\rangle + \langle a, t\rangle + \alpha \in \mathbb{R}[t_1, t_2] \tag{163}$$

ein Polynom zweiten Grades mit symmetrischer Matrix $0 \neq S = (\alpha_{ij}) \in \mathrm{Mat}_2(\mathbb{R})$. Dann gibt es eine affine Koordinatentransformation mittels einer Ähnlichkeit $f = \tau_y \circ f_T$ von $\mathbb{R}^2$ mit $T \in \mathrm{SO}(2)$, so daß $q := p\big(f(t_1,t_2)\big)$ äquivalent zu einer der folgenden Normalformen ist:

I: $\det(S) > 0$.

I.1: $\alpha \neq 0$ und $\alpha_{11} > 0$. Dann ist $q \equiv (\lambda_1 t_1)^2 + (\lambda_2 t_2)^2 - 1$ und $N(q)$ ist eine *Ellipse*.

I.2: $\alpha \neq 0$ und $\alpha_{11} < 0$. Dann ist $q \equiv (\lambda_1 t_1)^2 + (\lambda_2 t_2)^2 + 1$ und $N(q)$ ist die leere Menge.

I.3: $\alpha = 0$. Dann ist $q \equiv (\lambda_1 t_1)^2 + (\lambda_2 t_2)^2$ und $N(q)$ ist ein Punkt.

II: $\det(S) < 0$.
- II.1: $\alpha \neq 0$. Dann ist $q \equiv (\lambda_1 t_1)^2 - (\lambda_2 t_2)^2 - 1$ und $N(q)$ ist eine *Hyperbel*.
- II.2: $\alpha = 0$. Dann ist $q \equiv (\lambda_1 t_1)^2 - (\lambda_2 t_2)^2$ und $N(q)$ besteht aus zwei verschiedenen Geraden durch den Ursprung.

III: $\det(S) = 0$, $a \neq (0,0)^t$. Dann ist $q \equiv t_1^2 - \lambda t_2$ und $N(q)$ ist eine *Parabel*.

IV: $\det(S) = 0$, $a = (0,0)^t$.
- IV.1: $\alpha \neq 0$ und S hat einen positiven Eigenwert. Dann ist $q \equiv t_1^2 - \lambda$, $\lambda > 0$, und $N(q)$ besteht aus zwei parallelen Geraden.
- IV.2: $\alpha \neq 0$ und S hat einen negativen Eigenwert. Dann ist $q \equiv t_1^2 + \lambda$, $\lambda > 0$, und $N(q)$ ist die leere Menge.
- IV.3: $\alpha = 0$. Dann ist $q \equiv t_1^2$ und $N(q)$ besteht aus einer *Doppelgraden*, d. h. einer Geraden, die man doppelt zählt.

Bemerkung B3.31.
Dies ist die vollständige Klassifikation der Kurven zweiten Grades. Sie heißen auch *Kegelschnitte*, da alle, bis auf die Fälle I.2, IV.1 und IV.2 als Schnitt des Kreiskegels

$$N\left(t_1^2 + t_2^2 - t_3^2\right) \subset \mathbb{R}^3$$

mit einer geeigneten Ebene im $\mathbb{R}^3$ realisierbar sind (siehe Abbildung 8).

I.1 besagt, daß sich jede Ellipse durch Translation und Drehung so bewegen läßt, daß die Hauptachsen der Ellipse mit den Koordinatenachsen übereinstimmen. Daher kommt der Name Hauptachsentransformation.

Beweis von Satz B3.30:
1. Fall: $a = (0,0)^t$: Wir betrachten zunächst den Fall $a = (0,0)^t$.

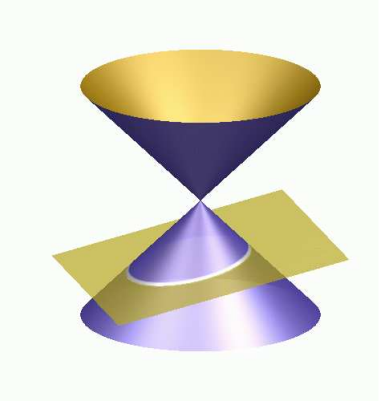
Nach dem Satz über die Hauptachsentransformation 37.10 existiert ein $T \in \mathrm{SO}(2)$, so daß

$$T^t \circ S \circ T = T^{-1} \circ S \circ T = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}.$$

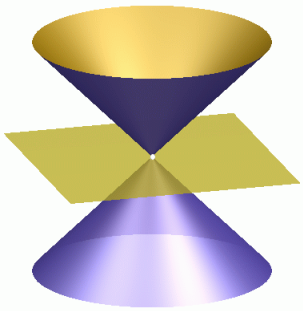
Man beachte noch, daß nicht beide Eigenwerte μ_1 und μ_2 null sein können, da $S \neq 0$. Also können wir o. E. annehmen, daß $\mu_1 \neq 0$ und daß $\mu_1 \geq \mu_2$ gilt, falls $\mu_2 \neq 0$.

Die lineare Abbildung $f_T : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto Tx$ ist eine Drehung und es gilt

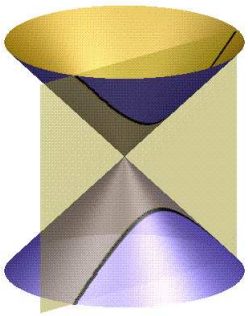
$$\begin{aligned} p(Tt) &= \langle Tt, (S \circ T)t \rangle + \alpha \\ &= \langle t, (T^t \circ S \circ T)t \rangle + \alpha \\ &= \mu_1 t_1^2 + \mu_2 t_2^2 + \alpha. \end{aligned}$$



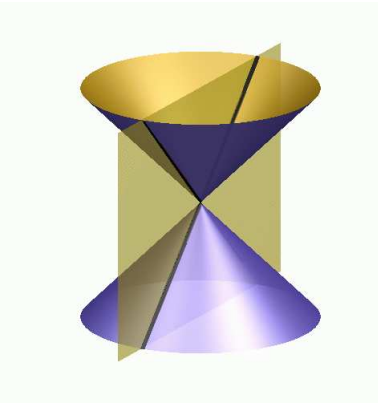
I.1: Ellipse



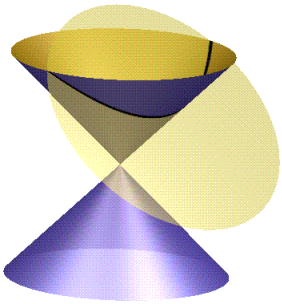
I.3: Punkt



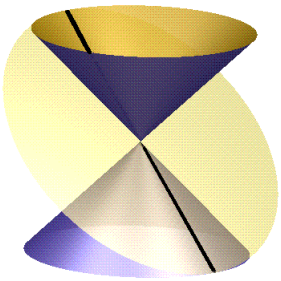
II.1: Hyperbel



II.2: Zwei Geraden mit Schnitt



III: Parabel



IV.3: Doppelgerade

Abbildung 8. Kegelschnitte

Da wir p ohnehin nur bis auf Äquivalenz klassifizieren wollen, können wir o. E. annehmen, daß $\alpha = 0$ oder $\alpha = -1$ gilt. Setzen wir nun noch $\lambda_i = \sqrt{|\mu_i|}$, dann erhalten wir folgende Fälle.

Fall 1.1: $\mu_1, \mu_2 > 0$: Dies ist gleichbedeutend dazu, daß S positiv definit ist, und nach dem Hauptminorenkriterium dazu, daß $\det(S) > 0$ und $\alpha_{11} > 0$. Ist $\alpha = -1$, so sind wir im Fall I.1, und ist $\alpha = 0$, so sind wir Fall I.3.

Fall 1.2: $\mu_1, \mu_2 < 0$: Dies ist gleichbedeutend dazu, daß $-S$ positiv definit ist, daß also $\det(S) = \det(-S) > 0$ und $-\alpha_{11} > 0$. Ist $\alpha = -1$, so sind wir im Fall I.2, und für $\alpha = 0$ wieder im Fall I.3, da wir dann das Polynom nochmals mit -1 multiplizieren können, um ein äquivalentes der gesuchten Form zu erhalten.

Fall 1.3: $\mu_1 > 0, \mu_2 < 0$: Dies ist gleichbedeutend dazu, daß $\mu_1 \cdot \mu_2 = \det(S) < 0$ ist. Im Fall $\alpha = -1$ führt dies zu Fall II.1, und im Fall $\alpha = 0$ führt es zu Fall II.2.

Fall 1.4: $\mu_1 > 0, \mu_2 = 0$ oder $\mu_1 < 0, \mu_2 = 0$: Das ist dann gleichbedeutend dazu, daß $\det(S) = 0$ ist. Für $\mu_1 > 0$ und $\alpha = -1$ erhalten wir Fall IV.1, für $\mu_1 < 0$ und $\alpha = -1$ den Fall IV.2, und für $\alpha = 0$ in den Fall IV.3.

2. Fall: $a \neq (0,0)^t$: Sind wir im Fall $a = (0,0)^t$ noch ohne Translation ausgekommen, so werden wir jetzt doch Translationen betrachten müssen.

Für $c \in \mathbb{R}^2$ bewirkt die Translation $\tau_c : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto x + c$ folgende Koordinatentransformation für p

$$\begin{aligned} p(t+c) &= \langle t+c, St+Sc\rangle + 2\langle a, t+c\rangle + \alpha \\ &= \langle t, St\rangle + 2\langle a+Sc, t\rangle + \langle c, Sc\rangle + 2\langle a, c\rangle + \alpha \\ &= \langle t, St\rangle + 2\langle b, t\rangle + \beta, \end{aligned} \tag{164}$$

wenn wir $b = a + Sc$ und $\beta = \langle c, Sc\rangle + 2\langle a, c\rangle + \alpha$ setzen.

Fall 2.1: $\exists\, c \in \mathbb{R}^2 : b = a + Sc = (0,0)^t$: Dann haben wir p durch $p\big(\tau_c(t)\big)$ auf den ersten Fall "$a = (0,0)^t$" zurückgeführt. Es gibt also ein $T \in \mathrm{SO}(2)$, so daß $q = p\big((\tau_c \circ f_T)(t)\big)$ äquivalent zu einem der Fälle I, II oder IV ist.

Fall 2.2: $\forall\, c \in \mathbb{R}^2 : b = a + Sc \neq (0,0)^t$: Aus Lemma B3.32 folgt, daß es ein $c \in \mathbb{R}^2$ gibt mit $Sb = S^2c + Sa = 0$. Setzen wir nun noch $\delta := -\frac{\beta}{2\langle b,b\rangle}$, dann gilt für die Translation $\tau_{c+\delta b}$[1]

$$\begin{aligned} p(t+c+\delta b) =& \langle t, St\rangle + 2\big\langle a + S(c+\delta b), t\big\rangle + \big\langle c+\delta b, S(c+\delta b)\big\rangle + 2\langle a, c+\delta b\rangle + \alpha \\ =& \langle t, St\rangle + 2\langle b + \delta Sb, t\rangle + \delta^2\langle b, Sb\rangle + 2\delta\langle b, b\rangle + \beta \\ =& \langle t, St\rangle + 2\langle b, t\rangle + 2\delta\langle b, b\rangle + \beta \\ =& \langle t, St\rangle + 2\langle b, t\rangle. \end{aligned}$$

Beachtet man, daß, wegen $Sb = 0$, Null auf alle Fälle ein Eigenwert von S ist und daß $S \neq 0$, so folgt aus dem Satz über Hauptachsentransformation 37.10 die Existenz eines $T \in \mathrm{SO}(2)$, so daß

$$D := T^t \circ S \circ T = T^{-1} \circ S \circ T = \begin{pmatrix} \mu_1 & 0 \\ 0 & 0 \end{pmatrix},$$

wobei $\mu_1 \neq 0$. Insbesondere sind wir also in dem Fall $\det(S) = 0$.

Ferner gilt für $T^t b =: (\mu, \lambda)^t$ unter Berücksichtigung, daß $T^t = T^{-1}$,

$$(\mu_1\mu, 0) = \big(T^t \circ S \circ T\big) \circ (T^t b) = T^t \circ (Sb) = 0,$$

und mithin ist $T^t b = (0, \lambda)^t$, wobei $\lambda \neq 0$, da T^t invertierbar und $b \neq (0,0)^t$. Aber dann überführt $t \mapsto Tt$ das Polynom $\langle t, St\rangle + 2\langle b, t\rangle$ in das Polynom

$$\big\langle Tt, (S \circ T)t\big\rangle + 2\big\langle b, Tt\big\rangle = \big\langle t^t, Dt\big\rangle + 2\big\langle T^t b, t\big\rangle = \mu_1 t_1^2 + 2\lambda t_2.$$

[1] Man setze zunächst in der Gleichung (164) für c den Wert $c + \delta b$ ein. Dann ziehe man die Skalarprodukte auseinander und gruppiere sie neu, so daß man $b = a + Sc$, $Sb = 0$ sowie die Definition von β verwenden kann. Man beachte auch, daß S symmetrisch, also selbstadjungiert, ist.

D. h. dann aber, daß

$$q := p\big((\tau_{c+\delta b} \circ f_T)(t)\big) = \mu_1 t_1^2 + 2\lambda t_2,$$

und damit sind wir genau im Fall III. □

Lemma B3.32.
Ist $S \in \mathrm{Mat}_n(\mathbb{R})$ symmetrisch, so gilt für die lineare Abbildung $f_S : \mathbb{R}^n \to \mathbb{R}^n$.

a. $\mathrm{Ker}\left(f_S^2\right) = \mathrm{Ker}(f_S)$ und $\mathrm{Im}\left(f_S^2\right) = \mathrm{Im}(f_S)$.

b. Zu jedem $a \in \mathbb{R}^n$ existiert ein $c \in \mathbb{R}^n$, so daß $S^2c + Sa = 0$.

Beweis: a. Für $x \in \mathrm{Ker}\left(f_S^2\right)$ ergibt sich aus

$$0 = \langle x, S^2x\rangle = \langle Sx, Sx\rangle,$$

also $f_S(x) = Sx = 0$ und $x \in \mathrm{Ker}(f_S)$. Die umgekehrte Inklusion ist klar.
Wir wissen bereits, daß $\mathrm{Im}(f_S) \supseteq \mathrm{Im}\left(f_S^2\right)$ gilt. Da nun ferner

$$\begin{aligned}\dim_{\mathbb{R}}\left(\mathrm{Im}(f_S)\right) &= n - \dim_{\mathbb{R}}\left(\mathrm{Ker}(f_S)\right)\\ &= n - \dim_{\mathbb{R}}\left(\mathrm{Ker}\left(f_S^2\right)\right) = \dim_{\mathbb{R}}\left(\mathrm{Im}\left(f_S^2\right)\right)\end{aligned}$$

gilt, folgt also die Gleichheit.

b. Es gilt für $a \in \mathbb{R}^n$, daß $S(-a) = f_S(-a) \in \mathrm{Im}(f_S) = \mathrm{Im}\left(f_S^2\right)$, also gibt es nach a. ein $c \in \mathbb{R}^n$ mit $S^2c + Sa = f_S^2(c) - f_S(-a) = 0$.

□

Aufgaben

Aufgabe B3.33.
Zeige, wenn für $f \in \mathrm{End}_{\mathbb{K}}(V)$ stets $||f(x)|| = ||x||$ gilt, so ist f orthogonal bzw. unitär.

Aufgabe B3.34 (Lineare Funktionale).
Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum.
Dann gibt es für jedes $g \in \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$ genau ein $y \in V$, so daß für alle $x \in V$ gilt

$$g(x) = \langle y, x\rangle.$$

Aufgabe B3.35 (Die adjungierte Abbildung).
Seien V und W zwei endlich-dimensionale euklidische oder unitäre Räume mit

Skalarprodukten $\langle\cdot,\cdot\rangle_V$ und $\langle\cdot,\cdot\rangle_W$. Dann gibt es zu jeder linearen Abbildung $f : V \longrightarrow W$ genau eine lineare Abbildung $f^* : W \longrightarrow V$, so daß

$$\langle f(x), y\rangle_W = \langle x, f^*(y)\rangle_V \tag{165}$$

für alle $x \in V$ und $y \in W$. Die Abbildung f^* heißt die *adjungierte Abbildung* zu f.

Aufgabe B3.36 (Matrixdarstellung der adjungierten Abbildung).
Seien V und W zwei endlich-dimensionale euklidische oder unitäre Räume mit Orthonormalbasen B bzw. D. Dann gilt für jede $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow W$

$$M_B^D(f^*) = \big(M_D^B(f)\big)^*,$$

d.h. die Matrixdarstellung der adjungierten Abbildung ist die Adjungierte der Matrixdarstellung.

Aufgabe B3.37.
Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum.
Zeige, ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ normal, so gelten

$$\mathrm{Ker}(f) = \mathrm{Ker}(f^*)$$

und

$$V = \mathrm{Ker}(f) \perp \mathrm{Im}(f).$$

Aufgabe B3.38.
Sei V ein endlich-dimensionaler unitärer Raum und $f \in \mathrm{End}_{\mathbb{C}}(V)$ normal.
Zeige, es gibt ein Polynom $p \in \mathbb{C}[t]$ mit $f^* = p(f)$.

Aufgabe B3.39.
Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $f \in \mathrm{End}_{\mathbb{K}}(V)$ bijektiv. Zeige, die folgenden Aussagen sind äquivalent:

a. Für $x, y \in V$ mit $x \perp y$ gilt $f(x) \perp f(y)$.

b. Für $x, y \in V$ mit $||x|| = ||y||$ gilt $||f(x)|| = ||f(y)||$.

c. Es gibt ein $\lambda \in \mathbb{R}_{>0}$ und ein $g \in \mathrm{O}(V)$ bzw. $g \in \mathrm{U}(V)$ mit $f = \lambda g$.

Aufgabe B3.40.
Es sei $V \neq 0$ ein endlich-dimensionaler unitärer Raum und $f \in \mathrm{End}_{\mathbb{C}}(V)$. Zeige, die folgenden Aussagen sind gleichwertig:

a. $f^* = -f$.

b. Für alle $x \in V$ gilt: $\langle f(x), x\rangle \in i\mathbb{R}$.

c. Es gibt eine Orthonormalbasis von V aus Eigenvektoren von f und der Realteil aller Eigenwerte ist Null.

Aufgabe B3.41.

Überprüfe, ob die folgende symmetrische Matrix $A \in \mathrm{Mat}_3(\mathbb{R})$ positiv definit ist:

$$A = \begin{pmatrix} 9 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Aufgabe B3.42.

Bestimme eine orthogonale Matrix $T \in \mathrm{O}(3)$, die die folgende symmetrische Matrix A diagonalisiert:

$$A = \begin{pmatrix} 5 & 2 & 2 \\ 2 & 2 & -4 \\ 2 & -4 & 2 \end{pmatrix}.$$

Ist die Matrix A positiv definit?

§ B4 Lineare Algebra mit SINGULAR

Im vorliegenden Abschnitt wollen wir zeigen, wie ein Computeralgebrasystem eingesetzt werden kann, um Rechnungen in der linearen Algebra durchzuführen. Wir verwenden hierzu das am Fachbereich entwickelte System SINGULAR. Es ist frei erhältlich für die Betriebssysteme Linux, Windows und MacOS von der Webseite:

`http://www.singular.uni-kl.de`

Auf den Linuxrechnern des Fachbereichs startet man SINGULAR einfach durch den Befehl `Singular` von einer einfachen Textkonsole aus. Man erhält dann zunächst einige Informationen zum Programm sowie einen Eingabeprompt `>`:

```
                     SINGULAR                                 /
 A Computer Algebra System for Polynomial Computations   /   version 3-1-1
                                                       0<
     by: G.-M. Greuel, G. Pfister, H. Schoenemann        \   Feb 2010
FB Mathematik der Universitaet, D-67653 Kaiserslautern    \
>
```

Der Eingabeprompt `>` fordert zur Eingabe von SINGULAR-Befehlen auf. Wir wollen hier nur einige kurze Anmerkungen zur allgemeinen Syntax machen und hoffen, daß sich alles weitere aus den im folgenden besprochenen Beispielen erschließt. Unsere Konvention dabei ist, daß SINGULAR-Ein- und Ausgaben im Gegensatz zu begleitenden Erläuterungen stets im `Typewriter`-Stil geschrieben werden.

a. Jede SINGULAR-Sitzung sollte mit dem Befehl

   `ring r=0,t,lp;`

   beginnen. Dadurch wird der Polynomring $\mathbb{Q}[t]$ als Grundring festgelegt und erhält den Namen `r`. Selbst, wenn man nicht vor hat, Polynome zu verwenden, ist dies nötig, um mit den rationalen Zahlen rechnen zu können. Ersetzt man die Zahl `0` in der Definition von `r` durch eine Primzahl p, so verwendet man statt der rationalen Zahlen den Körper $\mathbb{Z}/p\mathbb{Z}$; ersetzt man sie durch `real` oder `complex`, so rechnet man mit reellen oder komplexen Dezimalzahlen, was aber tunlichst vermieden werden sollte, da dann Rundungsfehler auftreten können.

b. Man kann Ergebnisse von Rechnungen sowie Eingaben auch in Variablen speichern. Ein Beispiel dafür ist die Variable `r` in Teil a., in der der Polynomring $\mathbb{Q}[t]$ abgespeichert wurde. Jede Variable in SINGULAR hat einen Namen und einen festgelegten Typen, der sagt, ob es sich um einen Ring (`ring`), ein Polynom (`poly`), ein Körperelement (`number`), eine ganze Zahl (`int`), eine Matrix (`matrix`) oder eine Liste (`list`) von Objekten handelt.

c. Nicht alle in SINGULAR im Prinzip verfügbaren Befehle sind schon unmittelbar mit dem Programmstart geladen, viele liegen in sogenannten Bibliotheken vor. Sie sind erst verfügbar, wenn man die entsprechende Bibliothek mit dem Befehl `LIB` eingebunden hat. Wie dies geschieht, werden wir in Beispielen sehen.

d. Jede SINGULAR-Eingabe schließt mit einem Semikolon `;` und dem anschließenden Drücken der Return-Taste ab. Das Semikolon fordert den SINGULAR-Interpreter dazu auf, die Eingabe zu übersetzen und auszuführen. Will man eine Eingabe über mehrere Zeilen strecken, so läßt man das Semikolon am Zeilenende weg und drückt die Return-Taste. Man erhält statt des üblichen Promptzeichens `>` dann einen Punkt `.` als Prompt. Dieser zeigt an, daß die Eingabe noch nicht beendet ist und sich über mehrere Zeilen erstreckt.

e. In den folgenden Beispielen ist alles, was auf einen Prompt `>` oder `.` am Zeilenanfang folgt, eine Eingabe, und jede Zeile, die ohne eines dieser Zeichen beginnt, enthält SINGULAR-Ausgaben. Text, der auf `//` folgt, enthält Kommentare, die beim Ausführen des Kommandos nicht beachtet werden. Will man das Beispiel selbst in SINGULAR nachprüfen, kann man sie getrost weglassen. Sie dienen nur der Erläuterung für den Leser. Ausgaben, die beim Laden von Bibliotheken auftreten, werden wir in den Beispielen weglassen.

f. Man beendet SINGULAR mit dem Befehlt `exit`. Hilfe zur Syntax von SINGULAR findet man im Manual auf der SINGULAR-Webseite oder durch den Befehl `help`.

Beispiel B4.1 (Reduzierte Zeilen-Stufen-Form und Rang einer Matrix).
Wir wollen eine reduzierte Zeilen-Stufenform und damit den Rang der folgenden Matrix berechnen:

$$A = \begin{pmatrix} 1 & 2 & 1 & 3 & 1 \\ 2 & 4 & 7 & 3 & 0 \\ 4 & 8 & 9 & 9 & 2 \\ 3 & 6 & 0 & 2 & 1 \end{pmatrix} \in \mathrm{Mat}(4 \times 5, \mathbb{Q})$$

Dazu benutzen wir die SINGULAR-Befehle `rowred`.

```
> LIB "matrix.lib";
> ring r=0,t,dp;
> matrix A[4][5]=1,2,1,3,1,2,4,7,3,0,4,8,9,9,2,3,6,0,2,1;
> print(rowred(A));  // zeigt die rZSF von A
1,2,0,0,1/11,
0,0,1,0,-2/11,
0,0,0,1,4/11,
0,0,0,0,0
```

Beispiel B4.2 (Kern einer Matrix).
Mit dem Befehl `syz` können wir eine Basis des Kerns der Matrix in Beispiel B4.1 berechnen. Man bezeichnet die Relationen zwischen den Spalten der Matrix, die durch die Vektoren im Kern beschrieben werden, auch als *Syzygien*, und `syz` ist die Abkürzung dieses Begriffs.

```
> LIB "matrix.lib";
> ring r=0,t,dp;
> matrix A[4][5]=1,2,1,3,1,2,4,7,3,0,4,8,9,9,2,3,6,0,2,1;
> matrix B=syz(A);
> print(B);
-2,0,
1, -1,
0, 4,
0, -8,
0, 22
```

Der Kern von A hat also die Basisvektoren $(-2,1,0,0,0)^t$ und $(0,-1,4,-8,22)^t$.

Beispiel B4.3 (Lösung eines linearen Gleichungssystems).
Wir setzen nun $b=(2,2,6,4)^t$ und wollen das lineare Gleichungssystem $Ax=b$ lösen. Der Befehl `concat` hängt zwei Matrizen hintereinander.

```
> matrix b[4][1]=2,2,6,4;
> matrix Ab=concat(A,b); // Bilde die erweiterte Koeffizientenmatrix.
> print(Ab);
1,2,1,3,1,2,
2,4,7,3,0,2,
4,8,9,9,2,6,
3,6,0,2,1,4
> print(syz(Ab));  // Berechne eine Basis des Kerns von Ab.
-2,0, 0,
1, -1,0,
0, 4, -2,
0, -8,4,
0, 22,-12,
0, 0, 1
```

Wir haben nun eine Basis des Kerns der erweiterten Koeffizientenmatrix berechnet. Der Algorithmus stellt sicher, daß es genau dann einen Vektor mit letzter Komponente ungleich null gibt, wenn das Gleichungssystem lösbar ist. Es gibt dann auch nur einen solchen Vektor und das ist der letzte Basisvektor. Dividiert man die ersten fünf Einträge

des Vektors durch das Negative des letzten Eintrags, so erhält man eine spezielle Lösung, hier

$$c = (0, 0, 2, -4, 12)^t.$$

Vergißt man bei den übrigen Vektoren in der berechneten Basis die letzte Komponente, so erhält man eine Basis des homogenen Lösungsraums, hier

$$\text{Lös}(A, 0) = \text{Lin}\left((-2, 1, 0, 0, 0)^t, (0, -1, 4, -8, 22)^t\right),$$

wie wir schon aus Beispiel B4.2 wissen.

Beispiel B4.4 (Eigenwerte einer Matrix).
Wir wollen die Eigenwerte der folgenden Matrix bestimmen

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} \in \text{Mat}_3(\mathbb{Q}).$$

Dazu verwenden wir unter anderem die SINGULAR-Befehle `det` zum Berechnen der Determinante, `unitmat` für die Einheitsmatrix und `factorize` zum Berechnen der Primfaktorzerlegung eines Polynoms.

```
> LIB "matrix.lib";
> ring r=0,t,dp;
> matrix A[3][3]=0,1,1,-1,2,1,-1,1,2;
> poly p=det(t*unitmat(3)-A);
> p;
t3-4t2+5t-2
> short=0;
> p;
t^3-4*t^2+5*t-2
> factorize(p);
[1]:
   _[1]=1
   _[2]=t-1
   _[3]=t-2
[2]:
   1,2,1
```

Das charakteristische Polynom von A ist also

$$\chi_A = t^3 - 4t^2 + 5t - 2 = (t-1)^2 \cdot (t-2),$$

so daß die Eigenwerte $\lambda = 1$ mit algebraischer Vielfachheit 2 sowie $\lambda = 2$ mit algebraischer Vielfachheit 1 sind.

Beispiel B4.5 (Minimalpolynom einer Matrix).
Als nächstes wollen wir das Minimalpolynom der Matrix A in Beispiel B4.4 berechnen. Dazu verwenden wir den Algorithmus B1.21 sowie einige SINGULAR-Befehle. `transpose` transponiert eine Matrix, `flatten` schreibt die Einträge einer Matrix in einen Zeilenvektor und `power` potenziert eine Matrix.

```
> matrix C=transpose(flatten(power(A,0)));
> for (int i=1;i<=3;i++)
. {
.   C=concat(C,transpose(flatten(power(A,i))));
. }
> matrix D=syz(C);
> print(D);
2, 0,
-3,2,
1, -3,
0, 1
> poly mu;
> for (i=1;i<=4;i++){mu=mu+D[1][i]*t^(i-1);}
> mu;
t^2-3*t+2
> factorize(mu);
[1]:
   _[1]=1
   _[2]=t-1
   _[3]=t-2
[2]:
   1,1,1
```

Das Minimalpolynom von A ist also

$$\mu_A = t^2 - 3t + 2 = (t-1)\cdot(t-2),$$

und die Matrix A ist somit diagonalisierbar, da das Minimalpolynom in paarweise verschiedene Linearfaktoren zerfällt.

Die obige Befehlssequenz ist recht lang. Falls man bereits weiß, daß das Minimalpolynom über dem Grundkörper in Linearfaktoren zerfällt, so kann man auch den SINGULAR-Befehl `minipoly` aus der Bibliothek `linalg.lib` verwenden, aber nur dann! Um sicherzustellen, daß das Minimalpolynom zerfällt, kann man zunächst das charakteristische Polynom berechnen und faktorisieren, denn nur wenn dieses zerfällt, zerfällt auch das Minimalpolynom. Für die Matrix A aus unserem Beispiel wissen wir bereits, daß es zerfällt. Wir können also den Befehl `minipoly` anwenden.

```
> LIB "linalg.lib";
> minipoly(A);
[1]:          // das Minimalpolynom hat die zwei Nullstellen 1 und 2
   _[1]=1
   _[2]=2
[2]:          // beide kommen mit Vielfachheit 1 vor
   1,1
```

Beispiel B4.6 (Diagonalisierung einer Matrix).
Wir haben in Beispiel B4.5 gesehen, daß die Matrix A aus Beispiel B4.4 diagonalisierbar ist. Nun wollen wir die zugehörige Transformationsmatrix T bestimmen. Dazu erinnern wir uns, daß A genau die Eigenwerte 1 und 2 besitzt. Zu diesen müssen wir Basen der Eigenräume bestimmen.

```
> matrix T1=syz(A-unitmat(3));
> print(T1);
1,0,
1,-1,
0,1
> matrix T2=syz(A-2*unitmat(3));
> print(T2);
1,
1,
1
> matrix T=concat(T1,T2);
> print(T);
1,0, 1,
1,-1,1,
0,1, 1
> print(inverse(T)*A*T);
1,0,0,
0,1,0,
0,0,2
```

Beispiel B4.7 (Jordansche Normalform).
In diesem Beispiel wollen wir die Jordansche Normalform und die Transformationsmatrix

T für die Matrix

$$A = \begin{pmatrix} 21 & 5 & 1 & 21 & 5 \\ -23 & 4 & 8 & -31 & 1 \\ -2 & -1 & -2 & -1 & -1 \\ -17 & -4 & -1 & -17 & -4 \\ 22 & -2 & -8 & 30 & 1 \end{pmatrix} \in \mathrm{Mat}_5(\mathbb{Q})$$

berechnen. Dazu bestimmen wir zunächst das charakteristische Polynom und das Minimalpolynom von A und faktorisieren diese.

```
> LIB "matrix.lib";
> LIB "linalg.lib";
> ring R=0,t,dp;
> matrix A[5][5]=21, 5, 1, 21, 5,
. -23,4, 8, -31,1,
. -2, -1,-2,-1, -1,
. -17,-4,-1,-17,-4,
. 22, -2,-8,30, 1;
> print(A);
21, 5, 1, 21, 5,
-23,4, 8, -31,1,
-2, -1,-2,-1, -1,
-17,-4,-1,-17,-4,
22, -2,-8,30, 1
> short=0;
> poly chi=det(t*unitmat(5)-A);
> chi;
t^5-7*t^4+10*t^3+18*t^2-27*t-27
> factorize(chi);
[1]:
   _[1]=1
   _[2]=t-3
   _[3]=t+1
[2]:
   1,3,2
> minipoly(A);
[1]:
   _[1]=-1
   _[2]=3
[2]:
   2,2
```

Wir sehen also, daß

$$\chi_A = (t-3)^3 \cdot (t+1)^2$$

und

$$\mu_A = (t-3)^2 \cdot (t+1)^2.$$

Damit ist die Jordansche Normalform von A festgelegt. Sie muß zu den beiden Eigenwerten 3 und -1 je mindestens einen Jordanblock der Größe 2 enthalten, weil sie im Minimalpolynom beide mit Vielfachheit zwei vorkommen. Zudem muß sie den Eigenwert 3 noch ein drittes Mal auf der Diagonalen haben, so daß ein weiterer Jordanblock der Größe eins zum Eigenwert 3 nötig ist. Also gilt

$$J_A = \left(\begin{array}{cc|c|cc} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right).$$

Da die Jordansche Normalform drei Jordanblöcke besitzt, müssen wir letztlich drei Basisvektoren finden, die die zyklischen A-invarianten Unterräume definieren, zu denen die Blöcke gehören. Dabei wenden wir den Algorithmus B2.15 an.

```
> matrix B=syz(A-3*unitmat(5)); // Basis von Ker(A-3*id)
> print(B);
-5,0,
1, -1,
1, 0,
4, 0,
0, 1
> matrix C=syz(power(A-3*unitmat(5),2)); // Basis von Ker((A-3*id)^2)
> print(C);
0,-5,0,
1,0, 0,
0,1, 0,
0,4, 0,
0,0, 1
```

Ein kurzer Blick genügt, um zu sehen, daß der erste und der dritte Basisvektor von $\mathrm{Ker}\left((A-3\cdot\mathbb{1}_5)^2\right)$ nicht im $\mathrm{Ker}(A-3\cdot\mathbb{1}_5)$ liegt. Wir können also jeden der beiden wählen, um den zyklischen Unterraum der Größe zwei zum Eigenwert 3 zu bilden. Wählen wir

$$x_1 = (0,1,0,0,0)^t.$$

```
> // berechne (A-3*unitmat(5)) * erste Spalte von C
. matrix X1[5][1]=C[1..5,1];
```

```
> print((A-3*unitmat(5))*X1);
5,
1,
-1,
-4,
-2
```

Damit hat der zyklische Unterraum der Größe zwei zum Eigenwert 3 die Basisvektoren

$$(A - 3 \cdot \mathbb{1}_5)x_1 = (5, 1, -1, -4, -2)^t \quad \text{und} \quad x_1 = (0, 1, 0, 0, 0)^t,$$

und diese sind die ersten beiden Spalten der Matrix T.

Nun müssen wir noch den Vektor $(A - 3 \cdot \mathbb{1}_5)x_1$ zu einer Basis von $\mathrm{Ker}(A - 3 \cdot \mathbb{1}_5)$ ergänzen. Ein Blick auf die Basis von $\mathrm{Ker}(A-3\cdot\mathbb{1}_5)$ zeigt, daß jeder der beiden Vektoren es tut.

```
> matrix X2[5][1]=B[1..5,2];   // waehle X2
```

Wir wählen deshalb

$$x_2 = (0, -1, 0, 0, 1)^t,$$

und dieser ist die dritte Spalte von T.

```
> print(syz(A+unitmat(5))); // Basis von Ker(A+id)
-1,
0,
1,
1,
0
> matrix D=syz(power(A+unitmat(5),2)); // Basis von Ker((A+id)^2)
> print(D);
-1,0,
0, -2,
1, 1,
1, 0,
0, 2
> matrix X3[5][1]=D[1..5,2];
> print((A+unitmat(5))*X3);   // (A+unitmat(5)) * 2. Spalte von D
1,
0,
-1,
-1,
0
```

Daraus folgt, daß der Vektor

$$x_3 = (0,-2,1,0,2)^t \in \operatorname{Ker}\left(A + \mathbb{1}_5^2\right) \setminus \operatorname{Ker}(A + \mathbb{1}_5)$$

liegt, und daß die letzten beiden Spalten von T die Vektoren

$$(A + \mathbb{1}_5)x_3 = (1,0,-1,1-,0)^t \quad \text{und} \quad x_3 = (0,-2,1,0,2)^t$$

sind.

```
> // bestuecke die Matrix T
. matrix T=(A-3*unitmat(5))*X1;
> T=concat(T,X1);
> T=concat(T,X2);
> T=concat(T,(A+unitmat(5))*X3);
> T=concat(T,X3);
> print(T);
5, 0,0, 1, 0,
1, 1,-1,0, -2,
-1,0,0, -1,1,
-4,0,0, -1,0,
-2,0,1, 0, 2
> // invertiere die Matrix T
. matrix S=inverse(T);
> print(S);
1, 0,0, 1, 0,
1, 1,0, 1, 1,
8, 0,-2,10,1,
-4,0,0, -5,0,
-3,0,1, -4,0
> print(inverse(T)*A*T);
3,1,0,0, 0,
0,3,0,0, 0,
0,0,3,0, 0,
0,0,0,-1,1,
0,0,0,0, -1
```

Wir erhalten also

$$T = \begin{pmatrix} 5 & 0 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & -2 \\ -1 & 0 & 0 & -1 & 1 \\ -4 & 0 & 0 & -1 & 0 \\ -2 & 0 & 1 & 0 & 2 \end{pmatrix} \quad \text{und} \quad T^{-1} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 8 & 0 & -2 & 10 & 1 \\ -4 & 0 & 0 & -5 & 0 \\ -3 & 0 & 1 & -4 & 0 \end{pmatrix}$$

sowie

$$J_A = T^{-1} \circ A \circ T = \left(\begin{array}{cc|c|cc} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right).$$

Es gibt in SINGULAR auch einen schnelleren Weg zur Jordanschen Normalform und der Transformationsmatrix, wenn man den Befehl `jordanbasis` verwendet, was in den SINGULAR-Übungsaufgaben aber nicht gemacht werden soll!

```
> matrix E=jordanbasis(A)[1];
> matrix Z[5][5];
> for (int j=1;j<=5;j++) { Z[1..5,j]=E[1..5,6-j]; }
> print(Z);
-5,5, 0,1, 0,
1, 1, 1,0, -2,
1, -1,0,-1,1,
4, -4,0,-1,0,
0, -2,0,0, 2
> print(inverse(Z)*A*Z);
3,0,0,0, 0,
0,3,1,0, 0,
0,0,3,0, 0,
0,0,0,-1,1,
0,0,0,0, -1
```

Die `for`-Schleife oben kehrt die Reihenfolge der Spalten in der Matrix E um. Die neue Matrix Z ist dann eine zulässige Transformationsmatrix T, wobei die Reihenfolge der Jordanblöcke sich geändert hat. Die Vertauschung der Spalten ist nötig, da die Konvention der Jordanschen Normalform in SINGULAR nicht mit unserer Konvention übereinstimmt. Darauf möchte ich hier aber nicht näher eingehen.

Beispiel B4.8 (Näherungsweise Bestimmung von Eigenwerten).

Eine zufällig ausgewählte Matrix in $\mathrm{Mat}_n(\mathbb{Q})$ wird keine rationalen Eigenwerte haben. Betrachen wir sie aber als Matrix in $\mathrm{Mat}_n(\mathbb{C})$, so zerfällt sie in Linearfaktoren und mit Wahrscheinlichkeit 1 sind diese paarweise verschieden. Exakt berechnen können wir sie aber nicht, da die Zerlegung eines Polynoms in $\mathbb{C}[t]$ in seine Primfaktoren im allgemeinen nicht möglich ist. Wir können die Eigenwerte aber näherungsweise berechnen, und dies reicht häufit aus, um zu sehen, daß sie paarweise verschieden sind.

```
> LIB "matrix.lib";
```

```
> ring S=complex,t,lp;
> matrix M[3][3];
> int i,j;
> for (i=1;i<=3;i++){for (j=1;j<=3;j++){M[i,j]=random(-9,9);}}
> print(M);
8, 5, 0,
-6,-2,3,
9, -9,7
> poly f=det(t*unitmat(3)-M);
> short=0;
> f;
t^3-13*t^2+83*t-449
> LIB "solve.lib";
> solve(f);
[1]:
   9.27119961
[2]:
   (1.86440019-i*6.70474155)
[3]:
   (1.86440019+i*6.70474155)
```

Das charakteristische Polynom der 3×3-Matrix ist ein Polynom vom Grad drei mit reellen Koeffizienten. Wegen des Zwischenwertsatzes muß es eine reelle Nullstelle haben. Wenn es keine weitere reelle Nullstelle besitzt, so müssen die übrigen beiden Nullstellen komplex konjugiert zueinander sein. Unsere Rechnung oben approximiert die Nullstellen mit dem Befehl `solve` aus der Bibliothek `solve.lib`, und wir sehen an den approximierten Nullstellen das geschilderte Phänomen.

Aufgaben

Aufgabe B4.9.
Bestimme mit Hilfe von SINGULAR eine Basis B von $\mathbb{Q}^5$, bezüglich derer die Matrixdarstellung der Abbildung $f : \mathbb{Q}^5 \to \mathbb{Q}^5$ Jordansche Normalform hat, wo:

$$f(x_1,x_2,x_3,x_4,x_5) = (x_1-x_2, x_1+2x_2-x_3, -x_1+3x_3, -x_1-2x_2-2x_3+2x_4-x_5, x_1-x_3+2x_5)^t.$$

Literaturverzeichnis

[BF87] Martin Barner and Friedrich Flohr, *Analysis I*, 3. Auflage ed., Walter de Gruyter, 1987.

[Coh96] Henri Cohen, *A course in computational algebraic number theory*, 3 ed., Graduate Texts in Mathematics, no. 138, Springer, 1996.

[Dec10] Wolfram Decker, *Grundlagen der Mathematik I*, Vorlesungsskript, TU Kaiserslautern, 2010.

[DH76] Whitfield Diffie and Martin E. Hellman, *New directions in cryptography*, IEEE Trans. on Info. Theory **IT-22** (1976), 644–654.

[Ebb92] Heinz-Dieter Ebbinghaus (ed.), *Zahlen*, 3 ed., Springer, 1992.

[Gat08] Andreas Gathmann, *Grundlagen der Mathematik*, Vorlesungsskript WS2007/08, TU Kaiserslautern, 2008.

[Har15] Peter Hartmann, *Mathematik für Informatiker*, Vieweg, 2015.

[Hei19] Matthias Hein, *Mathematik für Informatiker I*, Vorlesungsskript 2018/19, Universität Tübingen, 2019.

[Heu03] Harro Heuser, *Lehrbuch der Analysis, Teil 1*, 15 ed., Teubner, 2003.

[Mül07] Rainer Müller, *Aufgaben zur vollständigen Induktion*, http://www.emath.de, 2007.

[RSA78] Ronald L. Rivest, Adi Shamir, and Leonard M. Adleman, *A method for obtaining digital signatures and public-key cryptosystems*, Communications of the ACM **21** (1978), no. 2, 120–126.

[Sie81] Helmut Siemon, *Anwendungen der elementaren Gruppentheorie: in Zahlentheorie und Kombinatorik*, Klett Studienbücher, Klett, 1981.

[Ver75] Jacobus Verhoeff, *Error detecting decimal codes*, Mathematical Centre Tracts, no. 29, Mathematisch Centrum Amsterdam, 1975.