Einführung in das mathematische Arbeiten

Harald Woracek

(1.Auflage, druckfehlerkorrigiert 2018, geringfügig ergänzt 2019W und 2022W von M.Goldstern)

Inhaltsverzeichnis

Vorwort

Ich möchte mit zwei Zitaten beginnen, die treffender nicht sein könnten.

Aus: *Kevin Houston, 'How to think like a mathematician'*

> **The power of mathematics**
>
> Mathematics is the most powerful tool we have. It controls our world. We can use it to put men on the moon. We use it to calculate how much insulin a diabetic should take. It is hard to get right.
>
> And yet. And yet ... And yet people who use or like mathematics are considered geeks or nerds (add your own favourite term of abuse for the intelligent but unstylish). And yet mathematics is considered useless by most people – throughout history children at school have whined 'When am I ever going to use this ? '
>
> Why would anyone want to become a mathematician ? As mentioned earlier mathematics is a very powerful tool. Jobs that use mathematics are often well-paid and people do tend to be impressed. There are a number of responses from non-mathematicians when meeting a mathematician, the most common being 'I hated maths at school. I wasn't any good at it ', but another common response is 'You must be really clever. '

Aus: *Daniel J. Velleman, 'How to prove it. A structured approach'*

> What is mathematics ? High school mathematics is concerned mostly with solving equations and computing answers to numerical questions. College mathematics deals with a wider variety of questions, involving not only numbers, but also sets, functions, and other mathematical objects. What ties them together is the use of *deductive reasoning* to find answers to questions. When you solve an equation for x you are using the information given by the equation to *deduce* what the value of x must be. Similarly, when mathematicians solve other kinds of mathematical problems, they always justify their conclusions with deductive reasoning.

Was ist eigentlich Mathematik ?

Mathematik hat viele Gesichter. Sie vermag eine virtuelle Welt zu sein, wo man intellektuelle Herausforderung findet, oder ein Werkzeug um unsere reale Welt zu verstehen und zu verbessern, oder ein Rückzugsort hinter dem Spiegel, oder ein nützliches Mittel zum Zweck, etc. Nicht ohne Grund werden Mathematik und Philosophie, genauso wie Mathematik und Naturwissenschaften, oder Mathematik und Technologie, oder Mathematik und Wirtschaft, oft in einem Atemzug genannt. Es gibt es wohl keine Wissenschaft, wo gar keine mathematischen Methoden Anwendung finden; und auch aus dem täglichen Leben ist mathematisches – logisches und konsequentes – Denken, Argumentieren, und Deduzieren, nicht wegzudenken.

Wollen wir auf den im zweiten Zitat genannten Aspekt näher eingehen. Im Gegensatz zu anderen Wissenschaften, wo empirische Experimente, Interpretationen, und das Aufstellen und Verwerfen von Thesen, an der Tagesordnung sind, stellt Mathematik den Anspruch, eine exakte Wissenschaft zu sein. Ihre Sätze und Theorien sind unmissverständlich und unumstößlich – für alle Ewigkeit nachvollziehbar und wahr.

Wie kann das sein ? Man einigt sich auf gewisse Grundtatsachen, die von allen akzeptiert werden (diese werden *Axiome* genannt werden), und auf gewisse Regeln, nach denen man vorzugehen hat (diese werden *logische*

*Beide genannten Bücher sind höchst empfehlenswert! Ersteres ist auch ein allgemeiner einführender Begleiter, zweiteres ist auf einem ein bisschen höheren Niveau angesiedelt.

Schlussweisen oder *Beweisprinzipien* genannt werden). Alles was dann aus diesen Grundtatsachen mit Hilfe der Regeln abgeleitet wird, *muss* wahr sein, und *muss* auch von allen als wahr akzeptiert werden. Man könnte sagen, es ist wie bei einem Spiel. Bei einem Schachspiel einigt man sich auf eine Grundaufstellung wo die Figuren in ganz bestimmter Weise ihre Plätze einnehmen, und es gibt Regeln wie die Figuren während des Spiels bewegt werden können. Ein Zug kann dann geschickt sein, unbedacht, oder dumm, aber – solange man sich an die Regeln hält – niemals falsch. Genauso kann eine Strategie komplex oder geradlinig sein, kreativ oder bloss die Standarderöffnung, sie kann zum Ziel führen oder in den Abgrund, aber – solange man sich an die Regeln hält – kann sie niemals falsch sein.

Ein anderer Aspekt der Mathematik ist das Streben, die Essenz einer Sache oder die Kausalitäten in einem Prozess glasklar herauszuarbeiten. Meist geht dies mit Verallgemeinerung einher; es ist zielführend, vom Konkreten und Anschaulichen wegzustreben, hin zum Abstrakten und Allgemeinen. Bereinigt man Objekte von all dem Ballast, den eine konkrete Anschauung nun mal mit sich führt, sieht man viel besser was der Kern der Sache ist.

Was braucht es, um Mathematik zu betreiben ?

Um Mathematik zu betreiben braucht man eine Sprache, in der man sich – dem selbst auferlegten Exaktheitsanspruch genüge leistend – ausdrücken kann. Das Vokabular der Sprache der Mathematik ist die Terminologie der Mengenlehre und der Logik. Ihre Grammatik sind die Regeln des logischen Schließens.

Um Mathematik zu betreiben braucht man Dinge über die es sich lohnt zu sprechen. Der Ursprung vieler mathematischer Theorien und Fragestellungen liegt im Verlangen, unsere Welt zu verstehen, und im Bemühen, unsere Fähigkeiten und Möglichkeiten zu verbessern und zu erweitern. Viele betrachtete Objekte haben ihre Wurzel in konkreten Dingen des Lebens. Oft ist es aber auch so, dass man Dinge einfach um ihrer selbst willen studiert, aus purem Interesse und Freude an ihrer Schönheit.

Um Mathematik zu betreiben braucht man Kreativität, Intuition, und pedantische Genauigkeit. Das ist kein Widerspruch ! So wie ein Architekt, der ein innovatives Konzept für Wohnraum erschaffen möchte, einerseits kreativ sein muss und Vorstellungskraft benötigt, wie das Leben in dem in seinem Geiste entstehenden Gebäude ablaufen wird, andererseits aber auch die Fähigkeit haben muss, seine Ideen nachvollziehbar und unmissverständlich in einem Plan darzulegen, an den sich der Baumeister dann genauestens halten muss.

Dieses Skriptum soll . . .

. . . Ihnen einerseits als Unterlage für die Vorlesung „Einführung in das mathematische Arbeiten“ dienen, und andererseits auch ein Begleiter sein, den Sie in den ersten Semestern Ihres Studiums immer wieder konsultieren können. Rein von den mathematischen Inhalten deckt es sich mit dem Stoff der Vorlesung, insgesamt enthält es jedoch viel mehr. Zum Beispiel detaillierte Analysen von Argumenten und Beweisprinzipien, oder auch Beschreibungen von Methoden und Sprechweisen. Alle diese Dinge, die wir hier im Detail durchbesprechen, werden Ihnen im Laufe des gesamten Studiums begegnen, und sind Grundlage für jedes mathematisches Denken und Arbeiten. In höheren Semestern werden sie meistens nicht mehr explizit angesprochen werden, einfach weil erwartet wird, dass Sie das alles dann schon beherrschen.

Tatsächlich habe ich dieses Skriptum mit der Idee geschrieben, dass Sie es dreimal lesen: naturgemäß einmal jetzt gleich, parallel mit der Vorlesung, dann vielleicht in den Weihnachts- oder Semesterferien mit der Erfahrung die Sie bis dorthin gesammelt haben, und dann noch einmal in den ersten Sommerferien mit den Erfahrungen des ganzen ersten Studienjahres. Jetzt, ganz am Anfang Ihres Studiums, ist natürlich alles neu, unbekannt, und vielleicht ein bisschen unheimlich. Ich würde vermuten, dass Sie beim zweiten Mal lesen dann einige Aha-Erlebnisse haben: stimmt, das machen wir die ganze Zeit. Beim dritten Mal lesen wird es Ihnen dann vielleicht schon ein bisschen langweilig erscheinen, diese ganzen Dinge, die ja mittlerweile ohnehin selbstverständlich sind, nochmal durchzuarbeiten. Tun Sie es trotzdem !

> *Besinnen Sie sich immer und immer wieder auf die Grundlagen des mathematischen Denkens und Schließens. So komplex, tiefliegend, oder abgehoben eine Theorie auch sein mag, letzlich passiert nichts anderes als das was hier und jetzt – mit unseren ganz einfachen Inhalten – auch passiert !*

Zum Abschluss möchte ich an ein Sprichwort erinnern: „Ein Lehrer kann Ihnen die Türe öffnen, aber eintreten müssen Sie selbst.“ Man lernt, indem man sich selbst – und auch selbstständig – mit einer Materie beschäftigt. Vorlesungen, Skripten, Bücher, etc., sind notwendige und gute Unterlagen, aber:

> *Denken müssen Sie selbst.*

Kapitel 1

Einleitung

1.1 Logisches Schließen

1.1.1 Aussagen und Aussageformen

Auf der ersten mit mathematischen Inhalten befassten Seite des dtv-Atlas zur Mathematik findet sich der folgenden Text:

> Die Mathematik ist wie jede Wissenschaft darauf angewiesen, ihre Ergebnisse mündlich und schriftlich zu formulieren. Wegen der Vielfalt der Sprachen und der Gefahr von Mißverständnissen beim Gebrauch der Umgangssprache ist man in der Mathematik mehr und mehr dazu übergegangen, die Aussagen in einer künstlichen, formalisierten Sprache wiederzugeben, die nur noch die logisch bedeutsamen Elemente der Umgangssprache enthält. Es beginnt damit, daß man selbst den Begriff der *Aussage* zu präzisieren hat. Im allgemeinen fordert man, daß die Aussagen in die Klasse der wahren und die Klasse der falschen Aussagen eingeteilt werden können (*Prinzip der Zweiwertigkeit*). Eine Aussage ist dann jedes schrift-sprachliche Gebilde, dem entweder der *Wahrheitswert* des Wahren *W* oder der des Falschen *F* zukommt. Dabei spielt es keine Rolle, auf welche Weise der Wahrheitswert festgestellt wird. Bei bis heute nicht bewiesenen Vermutungen in der Mathematik etwa steht der Wahrheitswert gar nicht fest, doch darf bei der üblichen Auffassung angenommen werden, daß sie entweder wahr oder falsch sind.

Obwohl manche der hier genannten Punkte auch oft in Zweifel gezogen werden (z.B. Zweiwertigkeit, Beliebigkeit der Beweismethode, o.ä.), scheint dieser Text ein recht brauchbares Bild zu vermitteln.

Ist A eine Aussage, so schreiben wir $w(A)$ für den *Wahrheitswert* von A. Dieser kann den Wert `W` für „wahr“, oder `F` für „falsch“ annehmen.

Beispiele für Aussagen (mathematischer oder auch nichtmathematischer Natur) wären zum Beispiel

A_1:	Die Rose ist eine Pflanze.	$w(A_1) = \texttt{W}$
A_2:	Ein Affe ist ein Fisch.	$w(A_2) = \texttt{F}$
A_3:	$2 + 4 = 6$.	$w(A_3) = \texttt{W}$
A_4:	Ist dem Bauer kalt im Schuh, steht er in der Tiefkühltruh.	$w(A_4) = \texttt{F}$
A_5:	4 ist eine Primzahl.	$w(A_5) = \texttt{F}$

Es gibt auch Aussagen die eigentlich ganz einfach aussehen, deren Wahrheitswert aber unbekannt ist. Zum Beispiel: Man versteht unter einem Primzahlzwilling ein Paar $(n, n+2)$ wo beide angeschriebenen Zahlen Primzahlen sind; wie $(3, 5)$, $(5, 7)$, $(11, 13)$, $(17, 19)$, $(29, 31)$, etc. Die Aussage ist nun: „Es gibt unendlich viele Primzahlzwillinge“. Wahrheitswert ?...unbekannt.

Natürlich gibt es auch schrift-sprachliche Gebilde die keine Aussage liefern, zum Beispiel „Die Zahl 5 ist größer“, oder „Wenn nicht und rot dann oder grün, dann gelb“, oder ähnliches.

Man kann nun Aussagen in mannigfaltiger Weise miteinander verknüpfen. Manche dieser Operationen treten so häufig auf, dass sie einen eigenen Namen bekommen.

➢ Seien A und B Aussagen. Dann verstehen wir unter

(1) $\neg A$ die Aussage „*nicht A*“(auch „*Negation* von *A*“), welche genau dann wahr ist, wenn A falsch ist.

$\neg A$:

A	$\neg A$
W	F
F	W

(2) $A \vee B$ die Aussage „*A oder B*“(oder „*Disjunktion* von *A* und *B*“), welche genau dann wahr ist, wenn A wahr ist oder B wahr ist (oder beide).

$A \vee B$:

A	B	$A \vee B$
W	W	W
W	F	W
F	W	W
F	F	F

(3) $A \wedge B$ die Aussage „*A und B*“(oder „*Konjunktion* von *A* und *B*“), welche genau dann wahr ist, wenn sowohl A als auch B wahr sind.

$A \wedge B$:

A	B	$A \wedge B$
W	W	W
W	F	F
F	W	F
F	F	F

(4) $A \Rightarrow B$ die Aussage „*wenn A, dann B*“. Diese Aussage ist genau dann wahr, wenn B immer dann wahr ist wenn A wahr ist.

$A \Rightarrow B$:

A	B	$A \Rightarrow B$
W	W	W
W	F	F
F	W	W
F	F	W

Man sagt auch „*A impliziert B*“ oder „*aus A folgt B*“, und bezeichnet „$A \Rightarrow B$“ als *Implikation*. Dabei nennt man A die *Prämisse* der Implikation und B die *Konklusion* der Implikation.

Beachte, dass der einzige Fall wo eine Implikation falsch ist, jener ist wenn ihre Prämisse wahr ist, ihre Konklusion jedoch falsch. Ist die Prämisse einer Implikation falsch, so ist die Implikation – unabhängig vom Wahrheitswert der Konklusion! – wahr. Beachte weiters, dass in der Alltagssprache mit der Aussage „Wenn A, dann B“ meist eine Kausalität verknüpft ist. Beim hier betrachteten Begriff der Implikation geht es hingegen nur um Wahrheitswerte, und um keinen inneren Zusammenhang: „Wenn $\pi = 3$ ist, dann ist 7 eine gerade Zahl“ ist eine wahre Aussage.

(5) $A \Leftrightarrow B$, die Aussage „*A genau dann, wenn B*“. Diese Aussage ist genau dann wahr, wenn A und B entweder gemeinsam wahr oder gemeinsam falsch sind.

$A \Leftrightarrow B$:

A	B	$A \Leftrightarrow B$
W	W	W
W	F	F
F	W	F
F	F	W

Es gibt noch weitere Verknüpfungen wie etwa $A \Leftarrow B$ („*B* folgt aus *A*“) oder A nor B (A nor B, „weder A noch B“), die in der Praxis selten verwendet werden. Statt $A \Leftarrow B$ verwendet man das äquivalente $B \Rightarrow A$, und statt A nor B das äquivalente $\neg(A \vee B)$ oder auch $(\neg A) \wedge (\neg B)$.

Einen Ausdruck in dem Aussagen repräsentierende Variablen, sowie Symbole für logische Verknüpfungen vorkommen nennt man *Aussageform*. Eine sinnvolle Aussageform kann, je nach Belegung der Variablen mit wahren oder falschen Aussagen den Wahrheitswert wahr oder falsch annehmen. Beispiele für sinvolle Aussageformen wären „$(A \vee B) \Rightarrow A$“, oder „$(C \wedge (A \Leftrightarrow B)) \,\mathtt{nor}\, D$“. Man beachte, dass Ausdrücke, die aus Variablen und Symbolen zusammengesetzt sind, auch sinnlos sein können; zum Beispiel macht „$A \,\mathtt{nor}\, (B \Rightarrow)$“ oder „$\neg \wedge A$“ keinen Sinn. Mit diesem syntaktischen Aspekt, wie sinnvolle Aussageformen aufgebaut sind, wollen wir uns hier aber nicht beschäftigen.

1.1.2 Allgemeingültige Aussagenformen — Schlussregeln

Eine wichtige Rolle spielen *allgemeingültige Aussageformen*, auch *Tautologien* genannt. Das sind solche Aussageformen, welche, unabhängig von der Belegung der Variablen mit wahren oder falschen Aussagen, wahr sind. Solche sind von besonderer Bedeutung für das Leben im allgemeinen und die Mathematik im besonderen, da sie logische Schlussweisen repräsentieren die man in Argumentationen verwenden kann. Man spricht daher auch von *Schlussregeln*.

➢ `Wir wollen einige allgemeingültige Aussageformen, welche häufig verwendet werden, zusammenstellen.`

Um Klammern zu vermeiden, verwenden wir im Folgenden immer wieder die Konvention, dass $\neg$ am stärksten bindet, und $\Rightarrow$ bzw. $\Leftrightarrow$ am schwächsten.

(1)	$A \vee \neg A$	(*Satz vom ausgeschlossenen Dritten*)
(2)	$\neg(A \wedge \neg A)$	(*Satz vom Widerspruch*)
(3)	$\neg\neg A \Leftrightarrow A$	(*Satz von der doppelten Verneinung*)
(4)	$(A \wedge B) \Rightarrow A$	(*Konjunktionsbeseitigung*)
(5)	$A \Rightarrow (A \vee B)$	(*Disjunktionseinführung*)
(6)	$(A \Leftrightarrow B) \Rightarrow (A \Rightarrow B)$	
(7)	$\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$ $\neg(A \vee B) \Leftrightarrow (\neg A) \wedge (\neg B)$	(*Sätze von deMorgan*)
J107 (8)	$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$	(*Kontrapositionssatz*)
(9)	$(A \wedge (A \Rightarrow B)) \Rightarrow B$	(*Modus Ponendo Ponens*)
(10)	$(\neg A \wedge (A \vee B)) \Rightarrow B$	(*Modus Tollendo Ponens*)
(11)	$((A \Rightarrow B) \wedge \neg B) \Rightarrow \neg A$	(*Modus Tollendo Tollens*)
(12)	$(\neg(A \wedge B) \wedge B) \Rightarrow \neg A$	(*Modus Ponendo Tollens*)
(13)	$((A \Rightarrow B) \wedge (B \Rightarrow C)) \Rightarrow (A \Rightarrow C)$	(*Modus Barbara*)
(14)	$A \wedge B \Leftrightarrow (B \wedge A)$ $A \vee B \Leftrightarrow (B \vee A)$	(*Kommutativsätze*)
(15)	$A \wedge (B \wedge C) \Leftrightarrow (A \wedge B) \wedge C$ $A \vee (B \vee C) \Leftrightarrow (A \vee B) \vee C$	(*Assoziativsätze*)
(16)	$A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$ $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$	(*Distributivsätze*)

Betrachtet man die eingangs eingeführten Verknüpfungen von Aussagen, so sieht man, dass diese nicht unabhängig voneinander sind. Damit meinen wir, dass man manche von ihnen durch Kombinationen der anderen ausdrücken kann. Zum Beispiel sind die folgenden Aussageformen allgemeingültig:

J108 (17) $(A \Rightarrow B) \Leftrightarrow \neg(A \wedge \neg B)$

Dies bedeutet also: Die Aussage $A \Rightarrow B$ ist genau dann wahr, wenn es unmöglich (sprich, falsch) ist, dass „zwar A aber nicht B“ gilt.

(18) $(A \Leftrightarrow B) \Leftrightarrow (A \Rightarrow B) \wedge (A \Leftarrow B)$

J101 Hier steht „$A \Leftarrow B$“ für „$B \Rightarrow A$“.

Dies bedeutet also: Eine Äquivalenz $A \Leftrightarrow B$ ist genau dann wahr, wenn beide wechselseitigen $A \Rightarrow B$ und $A \Leftarrow B$ gelten.

Um zu überprüfen, dass eine Aussageform allgemeingültig ist, braucht man nur für die in ihr vorkommenden Aussagenvariablen alle möglichen Wahrheitswerte (bei n Variablen sind das 2^n Möglichkeiten) einsetzen und dann den Wahrheitswert der Aussageform ausrechnen.

♧ Wir überprüfen exemplarisch, dass der Modus Ponendo Ponens eine Tautologie ist:

$(A \wedge (A \Rightarrow B)) \Rightarrow B$:

A	B	$A \Rightarrow B$	$A \wedge (A \Rightarrow B)$	$(A \wedge (A \Rightarrow B)) \Rightarrow B$
W	W	W	W	W
W	F	F	F	W
F	W	W	F	W
F	F	W	F	W

♧

Man kann sich vorstellen, dass diese Methode recht schnell mühsam wird, wenn mehr Aussagenvariablen vorkommen. Oft kann man sich diese Arbeit ersparen und die Allgemeingültigkeit einer Aussageform durch anwenden der bereits bekannten Regeln ableiten.

1.1.3 All- und Existenzquantor

Betrachtet man mathematische (oder auch irgendwelche anderen) Aussagen, so stößt man auch auf Redewendungen wie „für alle Dinge gilt...“, oder „es gibt ein Ding für das...“. Um solche Aussagen ebenfalls formal zu erfassen und mit ihnen, auch wenn sie in komplexerem Zusammenhang auftreten, logisch korrekt umgehen zu können, führt man *Quantoren* ein.

➢ Sei $A(x)$ eine von einer Variablen x abhängige Aussage. Dann bezeichnen wir mit

(1) $\forall$ den *Allquantor*. Die Aussage „$\forall x\colon A(x)$“, ist genau dann wahr, wenn für alle x die Aussage $A(x)$ wahr ist.

(2) $\exists$ den *Existenzquantor*. Die Aussage „$\exists\, x\colon A(x)$“, ist genau dann wahr, wenn es (mindestens) ein x gibt sodass die Aussage $A(x)$ wahr ist.

Genauso wie bei den Verknüpfungen der Aussagenlogik, sind auch hier jene Formeln von besonderer Bedeutung, die stets, also unabhängig von der Belegung gewisser Variablen mit gewissen Wahrheitswerten, wahr sind. Der Nachweis solcher Sätze der Prädikatenlogik ist jedoch nicht mehr so einfach wie bei jenen der Aussagenlogik, da Quantoren ja im allgemeinen Aussagen über unendliche Mengen darstellen können. Solche kann man nicht mehr durch Einsetzen aller möglichen Varianten auf Richtigkeit testen.

➢ Wir wollen einige allgemeingültige Formeln, welche häufig verwendet werden, zusammenstellen. Die Formel A darf hier von x bzw von x und y abhängen, die Formel B aber nicht.

(3) $\neg(\forall x\colon A(x)) \Leftrightarrow \exists\, x\colon (\neg A(x))$ (*Verneinungssätze*)

J104 $\neg(\exists\, x\colon A(x)) \Leftrightarrow \forall x\colon (\neg A(x))$

(4) $(\forall x\, \forall y\colon A(x,y)) \Leftrightarrow (\forall y\, \forall x\colon A(x,y))$ (*Vertauschbarkeitssätze*)
$(\exists\, x\, \exists\, y\colon A(x,y)) \Leftrightarrow (\exists\, y\, \exists\, x\colon A(x,y))$
$(\exists\, x\, \forall y\colon A(x,y)) \Rightarrow (\forall y\, \exists\, x\colon A(x,y))$

(5) $\forall x\colon \big((\forall y\colon A(y)) \;\Rightarrow\; A(x)\big)$
$\forall x\colon \big(A(x) \;\Rightarrow\; (\exists y\colon A(y))\big)$

(6) $(\forall x\colon A(x)) \wedge B \;\Leftrightarrow\; \forall x\colon (A(x) \wedge B)$ (*Quantorenverschiebungsgesetze*)
$(\exists x\colon A(x)) \wedge B \;\Leftrightarrow\; \exists x\colon (A(x) \wedge B)$

Man beachte, dass die Formel $\forall y\, \exists x\colon A(x,y) \;\Rightarrow\; \exists x\, \forall y\colon A(x,y)$ *nicht* allgemeingültig ist. Der wesentliche Unterschied zwischen $\forall y \exists x\colon A(x,y)$ und $\exists x \forall y\colon A(x,y)$ ist: Im ersten Fall existiert zwar zu jedem y ein x, dieses darf jedoch von y abhängen. Dagegen muß im zweiten Fall ein x existieren welches für alle y funktioniert, d.h. für jedes y ein und das selbe x.

Weiters beachte man, dass $(\forall x\colon A(x)) \Rightarrow B$ im Allgemeinen **nicht** zu $\forall x\colon (A(x) \Rightarrow B)$ äquivalent ist, sondern zu $\exists x\colon (A(x) \Rightarrow B)$.

➢ `Weitere Schreibweisen, die häufig verwendet werden:`

(7) $\exists! x\colon A(x)$
Das besagt „es *existiert genau ein* x sodass $A(x)$ wahr ist“, oder auch „es gibt ein *eindeutiges* x“
$\big[\; \exists x\colon \big(A(x) \wedge (\forall y\colon A(y) \Rightarrow y = x)\big) \;\big]$

(8) $\nexists x\colon A(x)$
Das ist eine Abkürzung für „$\neg(\exists x\colon A(x))$“, besagt also „es *existiert kein* x sodass $A(x)$ wahr ist“.

(9) Eine weitere Sprechweise ist zu sagen „es *existiert höchstens ein* x sodass $A(x)$ wahr ist“. Dieses heißt, dass entweder kein x existiert oder genau eines.
$\big[\; (\nexists x\colon A(x)) \vee (\exists! x\colon A(x)) \;\big]$
Anders ausgedrückt kann man sagen, dass je zwei Objekte die beide die Eigenschaft $A(\_)$ haben, gleich sein müssen. (Man beachte die scheinbare Paradoxie: um „höchstens **ein** x“ auszudrücken, spricht man über **zwei** Variable x, y.)
$\big[\; \forall x, y\colon (A(x) \wedge A(y)) \Rightarrow x = y \;\big]$
Man schreibt dafür manchmal auch $\exists^{\leq 1} x\colon A(x)$.
Manchmal sieht man auch $\exists^{\geq 2} x\colon A(x)$, und damit ist gemeint, dass $\neg\big(\exists^{\leq 1} x\colon A(x)\big)$, sprich, dass es mehr als ein x mit der Eigenschaft $A(x)$ gibt.

(10) Die folgende Sichtweise auf Existenz und Eindeutigkeit wird oft verwendet: Es existiert genau ein x mit $A(x)$, genau dann wenn es ein Element x gibt sodass $A(x)$ gilt, und wenn je zwei Elemente x, y für die $A(x)$ bzw. $A(y)$ gilt übereinstimmen. Die Formel auf der rechten Seite der Äquivalenz scheint zwar komplizierter zu sein als die auf der linken Seite, aber sie ist oft praktischer, weil man die hier die „schwierige“ Aufgabe (nämlich „es gibt genau ein“) auf zwei einfachere Aufgaben aufgeteilt hat: „es gibt mindestens ein“ und „es gibt höchstens ein“.

J106

$\big[\; (\exists! x\colon A(x)) \Leftrightarrow ((\exists x\colon A(x)) \wedge (\forall x, y\colon A(x) \wedge A(y) \Rightarrow x = y)) \;\big]$

Weitere allgemeingültige Formeln, die also logische Schlussweisen repräsentieren, sind die sogenannten *Syllogismen* der Aristotelischen Logik. Wir führen exemplarisch drei an (es gibt 24 gültige Syllogismen).

(11) $[\forall x\colon A(x) \Rightarrow B(x)] \wedge [\forall x\colon B(x) \Rightarrow C(x)] \;\Rightarrow\; [\forall x\colon A(x) \Rightarrow C(x)]$ (*Modus Barbara*)

(12) $[\forall x\colon A(x) \Rightarrow B(x)] \wedge [\nexists x\colon B(x) \wedge C(x)] \;\Rightarrow\; [\nexists x\colon A(x) \wedge C(x)]$ (*Modus Camestres*)

(13) $[\nexists x\colon A(x) \wedge B(x)] \wedge [\exists x\colon A(x) \wedge C(x)] \;\Rightarrow\; [\exists x\colon C(x) \wedge \neg B(x)]$ (*Modus Ferioque*)

Weitere in der Praxis wichtige Äquivalenzen finden Sie am Ende der Seite `https://dmg.tuwien.ac.at/goldstern/linalg/sprache1.html` .

1.2 Die Sprache der Mengenlehre

1.2.1 Der Cantorsche Mengenbegriff

Die Formulierung mathematischer Aussagen basiert auf dem Begriff der Menge. Eigentlich kann man sagen, dass jede mathematische Theorie – wenn man sie losgelöst von ihren Bedeutung für die reale Welt begreift – nichts anderes ist als eine Ansammlung von Implikationen die Aussagen über gewisse Mengen von Objekten und deren Elemente treffen.

Wir wollen uns in diesem Abschnitt mit der Sprache der Mengenlehre bekannt machen. Die folgende „Definition“ des Begriffes der Menge stammt von Georg Cantor. Sie genügt eigentlich nicht den strengsten logischen Ansprüchen, ist jedoch gut genug um im alltäglichen mathematischen Leben damit das Auslangen zu finden. Ein tiefergehendes Studium dieser Grundlagen würde hier zu weit führen*.

➢ `Der Cantorsche Mengenbegriff:`

Unter einer *Menge* verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten x unserer Anschauung oder unseres Denkens (welche die *Elemente* von M genannt werden) zu einem Ganzen.

Um auszudrücken, dass ein Objekt x Element einer Menge M ist, schreibt man „$x \in M$“. Um auszudrücken, dass x kein Element von M ist, schreibt man „$x \notin M$“.

Eine Menge ist durch ihre Elemente vollständig bestimmt: Sind M_1, M_2 Mengen, so ist M_1 *gleich* M_2, und man schreibt „$M_1 = M_2$“, wenn M_1 und M_2 die selben Elemente enthalten.

$\left[\; M_1 = M_2 \;\Leftrightarrow\; \big(\forall x\colon\, x \in M_1 \Leftrightarrow x \in M_2\big) \;\right]$

1.2.2 Anschreiben von Mengen

➢ `Wie gibt man Mengen an:`

Um eine Menge festzulegen muss man auf irgendeine Weise unmissverständlich klarmachen, welche Objekte Elemente der Menge sind und welche nicht. Dazu gibt es verschiedene Möglichkeiten.

(1) Durch *Aufzählung* aller Elemente. Man schreibt *alle Elemente* der Menge zwischen einem Paar geschweifter Klammern hin. Zum Beispiel ist

$$M = \{a, \hat{z}, 17, \text{Strawberry fields forever}, 117, \heartsuit\}$$

die Menge bestehend aus den sechs Elementen „a“, „$\hat{z}$“, „17“, „Strawberry fields forever“, „117“ und „$\heartsuit$“.

Das geht natürlich nur für Mengen mit sehr wenigen Elementen.

(2) Durch *Angabe einer definierenden Eigenschaft*. Die Menge wird beschrieben als die Menge *aller* Objekte die eine gewisse *Eigenschaft* haben. Diese Methode eine Menge anzugeben ist die universellste, gebräuchlichste und präziseste.

Die Schreibweise sieht wie folgt aus: Sei $A(x)$ die Eigenschaft die die Elemente der Menge charakterisieren soll, dann ist

$$M = \{x :\; A(x)\} \tag{1.2.1}$$

J109

die Menge aller Objekte x für die $A(x)$ wahr ist.

$\left[\; \forall x\colon\, x \in M \Leftrightarrow A(x) \;\right]$

Sprich, wenn $A(x)$ für ein Objekt x wahr ist, dann ist dieses x Element von M, und wenn $A(x)$ für x nicht wahr ist, dann ist $x \notin M$. Umgekehrt gelesen: wenn x in M liegt dann hat x die Eigenschaft $A(x)$, wenn x nicht in M liegt dann hat x die Eigenschaft $A(x)$ nicht.

Es ist auch gebräuchlich†, anstelle des Doppelpunktes in (1.2.1) einen Strich zu schreiben: $M = \{\, x \mid A(x) \,\}$.

*Es sei jedem wärmstens empfohlen sich eine Vorlesung über Mengentheorie und/oder mathematische Logik anzuhören. Allerdings erst in einem höheren Semester, wenn man schon mehr Erfahrung mit mathematischen Inhalten und Vorgangsweisen hat!

Oft findet man auch Schreibweisen wie (hier bezeichnen wir mit $\mathbb{N}$ die Menge aller – aus unserer Anschauung bekannten – nichtnegativen ganzen Zahlen)

$$M_1 = \{x \in \mathbb{N} : \ x \geq 4, \ 3 \text{ ist ein Teiler von } x\} \quad \text{oder} \quad M_2 = \{n^2 : \ n \in \mathbb{N}\}.$$

Ersteres beschreibt die Menge aller Objekte die Element von $\mathbb{N}$ sind und die beiden angeschriebenen Eigenschaften haben

$\left[\ M_1 = \{x : \ A(x)\} \quad \text{wobei} \quad A(x) = (x \in \mathbb{N}) \wedge ((x \geq 4) \wedge (3 \text{ ist ein Teiler von n})) \ \right]$

und zweiteres beschreibt die Menge aller Objekte die von der Gestalt n^2 sind, wobei n ein beliebiges Element von $\mathbb{N}$ ist

$\left[\ M_2 = \{x : \ \exists\, n \in \mathbb{N} \colon \ x = n^2\} \ \right]$

Zum Beispiel enthält die Menge $\{\frac{1}{2x} \mid x \in \mathbb{Q} \ \wedge \ x > 0\}$ alle positiven rationalen Zahlen (denn jede solche Zahl $\frac{p}{q}$ lässt sich mit $x := \frac{q}{2p}$ als $\frac{1}{2x}$ schreiben). Man beachte aber, dass aus der Gleichheit

$$\left\{ \frac{1}{x} \mid x \in \mathbb{Q} \ \wedge \ x > 0 \right\} = \left\{ 2x \mid x \in \mathbb{Q} \ \wedge \ x > 0 \right\}$$

keine Beziehung der Form $\frac{1}{x} = 2x$ folgt; tatsächlich gibt es überhaupt kein $x \in \mathbb{Q}$, welches diese Gleichung erfüllt.

Auch die folgende Methode, eine Menge anzugeben, ist recht gebräuchlich – weil oft praktisch. Allerdings ist sie mit Vorsicht zu behandeln.

(3) Durch *Aufzählung „mit u.s.w."*. Man schreibt exemplarisch einige Elemente der Menge an, und erwartet, dass der Rest selbsterklärend verstanden wird. Zum Beispiel schreibt man die Menge aller geraden positiven ganzen Zahlen manchmal an als

$$\{2, 4, 6, \dots\}.$$

Die, bei solcher Art Notation, stillschweigend getroffene Annahme, dass der Leser die drei Punkte „…" von sich aus richtig interpretiert, ist höchst problematisch. Sie ist Quelle vielfältiger Missverständnisse (schließlich kann sich jeder Leser etwas anderes unter den drei Punkte vorstellen). Was zum Beispiel könnte die Menge $Q = \{1, 2, 3, \dots\}$ sein ?

Diese Art von Notation ist nur in Ausnahmefällen zu verwenden. Wenn es den geringsten Verdacht auf mögliche Missinterpretationen gibt, ist sie zu vermeiden!

P.S.: Natürlich ist jedem sofort völlig klar, dass das die obige Menge Q die Menge aller positiven ganzen Zahlen ist die nur durch 1 und sich selbst teilbar sind. Zugegeben, man hätte vielleicht besser $Q = \{1, 2, 3, 5, \dots\}$ schreiben sollen; dann wäre es wohl wirklich klar, dass Q tatsächlich unser Q ist.

P.P.S.: Oder ist $Q = \{1, 2, 3, 5, \dots\}$ vielleicht doch offensichtlich die Menge aller Folgenglieder der *Fibonacci-Folge* (das ist – aus der Anschauung bekannte – Folge $1, 1, 2, 3, 5, 8, \dots$ wo jedes Folgenglied die Summe der beiden vorhergehenden ist) ?

Eine besondere Rolle spielt die *leere Menge*. Das ist jene Menge die kein Element besitzt.

$\left[\ \forall x \colon \ x \notin \emptyset \ \right]$

Man schreibt die leere Menge an als $\emptyset$, oder auch als $\{\,\}$. (Man beachte, dass die leere Menge keine Elemente enthält, während die Menge $\{\{\}\} = \{\emptyset\}$ genau ein Element enthält.)

Manchmal ergibt sich beim Anschreiben einer Menge, dass ein und das selbe Element der Menge mehrmals aufgelistet wird. Das heißt *nicht* dass dieses Objekt „mehrmals" Element der Menge ist! Soetwas wie „öfters

[†] Inhaltlich gibt es keinen Unterschied zwischen $\{ x \mid A(x) \}$ und $\{ x : A(x) \}$, aber die richtige Wahl des Symbols kann die Lesbarkeit erhöhen:

- Man kürzt die Eigenschaft „f ist eine Funktion von A nach B" (siehe Abschnitt 2.4) gerne mit $f : A \to B$ ab. Die Menge aller Funktionen von A nach B schreibt man dann lieber als $\{ f \mid f : A \to B \}$ an, nicht $\{ f : f : A \to B \}$.
- Aus der Schule kennen Sie vielleicht die Betragsfunktion, zum Beispiel $|5| = |-5| = 5$.
  Die Menge aller Zahlen mit Betrag > 5 schreibt man eher als $\{ x \in \mathbb{R} : |x| > 5 \}$, nicht $\{ x \in \mathbb{R} \mid |x| > 5 \}$.

Element sein“ gibt es nicht; man denke an die Passage „. . . bestimmten *wohlunterschiedenen* Objekten. . . “ in der Cantorschen Definition des Mengenbegriffes. Ein Objekt kann nur entweder Element der Menge sein, oder eben nicht.

Zum Beispiel wäre die Menge $\{a, a, a, a, a\}$ jene Menge die nur das eine Element „a“enthält, oder die Menge $\{(x-4)^2 : x \in \mathbb{N}\}$ jene Menge die die Elemente $0, 1, 4, 9, 16, 25, 36, 49, \ldots$ enthält. Bemerke, dass die Elemente $0, 1, 4, 9, 16$ in der beschreibenden Definition dieser Menge zweimal aufgelistet sind, alle weitern (das sind also $25, 36, 49, 64, \ldots$) dagegen nur einmal.

1.2.3 Operationen mit Mengen

Man kann Mengen miteinander vergleichen.

➢ `Seien` M_1 `und` M_2 `Mengen.`

(1) Man sagt M_1 *ist eine Teilmenge von* M_2 und schreibt $M_1 \subseteq M_2$, wenn jedes Element von M_1 auch Element von M_2 ist.
$\left[M_1 \subseteq M_2 \;\Leftrightarrow\; \big(\forall x\colon x \in M_1 \Rightarrow x \in M_2\big) \right]$

(2) Man sagt M_1 *ist eine Obermenge von* M_2 und schreibt $M_1 \supseteq M_2$, wenn jedes Element von M_2 auch Element von M_1 ist. Sprich, wenn $M_2 \subseteq M_1$.
$\left[M_1 \supseteq M_2 \;\Leftrightarrow\; \big(\forall x\colon x \in M_1 \Leftarrow x \in M_2\big) \right]$

Man spricht von $\subseteq$ und $\supseteq$ auch als *Mengeninklusionen*.

$x \subseteq y$ und $x \in y$ sind im Allgemeinen nicht äquivalent. Zum Beispiel sind die beiden Aussagen $2 \in \{1, 2\}$ und $\{2\} \subseteq \{1, 2\}$ wahr, aber die beiden Aussagen $2 \subseteq \{1, 2\}$ und $\{2\} \in \{1, 2\}$ sind falsch.

Beachte, dass zwei Mengen M_1 und M_2 genau dann gleich sind, wenn beide wechselseitigen Inklusionen $M_1 \subseteq M_2$ und $M_1 \supseteq M_2$ gelten, also jedes Element der einen Menge in der anderen liegt, und umgekehrt.

$\left[M_1 = M_2 \;\Leftrightarrow\; (M_1 \subseteq M_2 \wedge M_1 \supseteq M_2) \right]$

Dies beruht auf der Tautologie (18) aus §1.1.2.

In diesem Zusammenhang sind auch die folgenden Schreibweisen gebräuchlich:

(3) Man sagt M_1 *ist eine echte Teilmenge von* M_2 und schreibt $M_1 \subsetneq M_2$, wenn jedes Element von M_1 auch Element von M_2 ist, es aber (mindestens) ein Element von M_2 gibt welches nicht Element von M_1 ist.
$\left[M_1 \subsetneq M_2 \;\Leftrightarrow\; (M_1 \subseteq M_2 \wedge M_1 \neq M_2) \right]$

(4) Man sagt M_1 *ist eine echte Obermenge von* M_2 und schreibt $M_1 \supsetneq M_2$, wenn jedes Element von M_2 auch Element von M_1 ist, es aber (mindestens) ein Element von M_1 gibt welches nicht Element von M_2 ist. Sprich, wenn $M_2 \subsetneq M_1$.
$\left[M_1 \supsetneq M_2 \;\Leftrightarrow\; (M_1 \supseteq M_2 \wedge M_1 \neq M_2) \right]$

(5) Die Schreibweise $M_1 \not\subseteq M_2$ steht für $\neg(M_1 \subseteq M_2)$
$\left[M_1 \not\subseteq M_2 \;\Leftrightarrow\; \big(\exists x\colon x \in M_1 \wedge x \notin M_2\big) \right]$
und die Schreibweise $M_1 \not\supseteq M_2$ für $\neg(M_1 \supseteq M_2)$
$\left[M_1 \not\supseteq M_2 \;\Leftrightarrow\; \big(\exists x\colon x \in M_2 \wedge x \notin M_1\big) \right]$

Oft findet man auch die Schreibweise „$M_1 \subset M_2$“. Leider *nicht* in einheitlicher Bedeutung! Meist steht sie für „$M_1 \subseteq M_2$“, manchmal aber auch für „$M_1 \subsetneq M_2$“. Es ist darum wohl besser, sie zu vermeiden.
Das Analoge gilt für „$M_1 \supset M_2$“.

Man kann nun Mengen in mannigfaltiger Weise miteinander verknüpfen. Manche dieser Operationen treten so häufig auf, daß sie einen eigenen Namen bekommen.

➢ `Seien` M_1 `und` M_2 `Mengen.`

(6) Die *Vereinigung* von M_1 und M_2 ist die Menge die all jene Elemente enthält, die in einer der Mengen M_1 und M_2 vorkommen (oder in beiden). Man schreibt $M_1 \cup M_2$ für diese Menge.
$\left[M_1 \cup M_2 = \{x : \ x \in M_1 \vee x \in M_2\} \right]$

(7) Der *Durchschnitt* von M_1 und M_2 ist die Menge die all jene Elemente enthält, die in beiden Mengen M_1 und M_2 vorkommen. Man schreibt $M_1 \cap M_2$ für diese Menge.
$\left[M_1 \cap M_2 = \{x : \ x \in M_1 \wedge x \in M_2\} \right]$

(8) Das *kartesische Produkt* von M_1 und M_2 ist die Menge deren Elemente all jene geordneten Paare sind, deren erster Eintrag Element von M_1 ist und deren zweiter Eintrag Element von M_2 ist. Man schreibt $M_1 \times M_2$ für diese Menge.
$\left[M_1 \times M_2 = \{(x, y) : \ x \in M_1 \wedge y \in M_2\} \right]$

`Sei` M `eine Menge.`

(9) Die *Potenzmenge* von M ist die Menge, die alle Teilmengen von M als Elemente hat. Man schreibt $\mathcal{P}(M)$ für diese Menge.
$\left[\mathcal{P}(M) = \{x : \ x \subseteq M\} \right]$
Die Potenzmenge der leeren Menge hat genau ein Element. (Welches?)
Die Potenzmenge der Menge $\{1, 2, 3\}$ hat genau 8 Elemente. (Und zwar?)

Man kann Vereinigung und Durchschnitt auch für beliebige Mengen von Mengen definieren.

➢ `Sei` $\mathcal{M}$ `eine Menge deren Elemente Mengen sind.`

(10) Die *Vereinigung* von $\mathcal{M}$ ist die Menge die alle jene Elemente enthält, die in mindestens einem der Elemente von $\mathcal{M}$ vorkommen. Man schreibt $\bigcup \mathcal{M}$, oder auch $\bigcup_{A \in \mathcal{M}} A$, für diese Menge.
$\left[\bigcup \mathcal{M} = \{x : \ \exists A \colon A \in \mathcal{M} \wedge x \in A\} \right]$

`Sei zusätzlich angenommen, dass` $\mathcal{M}$ `nichtleer ist.`

(11) Der *Durchschnitt* von $\mathcal{M}$ ist die Menge die alle jene Elemente enthält, die in allen Elementen von $\mathcal{M}$ vorkommen. Man schreibt $\bigcap \mathcal{M}$, oder auch $\bigcap_{A \in \mathcal{M}} A$, für diese Menge.
$\left[\bigcap \mathcal{M} = \{x : \ \forall A \colon A \in \mathcal{M} \Rightarrow x \in A\} \right]$

Oft spielt das folgende Auswahlverfahren eine Rolle.

➢ `Sei` $\mathcal{M}$ `eine Menge deren Elemente Mengen sind.`

(12) Sei angenommen, dass jedes Element von $\mathcal{M}$ nichtleer ist, und dass je zwei verschiedene Elemente von $\mathcal{M}$ leeren Durchschnitt haben. Dann existiert eine Menge N welche aus jedem Element von $\mathcal{M}$ genau ein Element enthält. Sprich, wir wählen aus jedem Element von $\mathcal{M}$ ein Element aus, und fassen die so erhaltenen Objekte zu einer Menge zusammen.
$\left[\exists N \colon \forall A \colon (A \in \mathcal{M} \ \Rightarrow \ \exists! \, x \colon x \in A \wedge x \in N \right]$

J105

Diese – anschaulich eigentlich recht naheliegende – Eigenschaft heißt das *Auswahlaxiom* oder *axiom of choice*.

An dieser Stelle wollen wir noch eine Sprechweise erwähnen. Hat man zwei Mengen M_1 und M_2, so sagt man

diese sind *disjunkt*, wenn $M_1 \cap M_2 = \emptyset$. Hat man eine Menge $\mathcal{M}$ deren Elemente Mengen sind, so sagt man die Elemente von $\mathcal{M}$ sind *paarweise disjunkt*, wenn je zwei verschiedene Elemente von $\mathcal{M}$ disjunkt sind.

Zwei weitere Operationen die auch öfters auftreten sind die Folgenden.

➢ Seien M_1 und M_2 Mengen.

(13) Die *Differenz* „M_1 ohne M_2" von M_1 und M_2 ist die Menge die all jene Elemente enthält, die in M_1 liegen aber nicht in M_2. Man schreibt $M_1 \setminus M_2$ für diese Menge.
$\left[M_1 \setminus M_2 = \{x : \ x \in M_1 \wedge x \notin M_2\} \right]$

(14) Die *symmetrische Differenz* von M_1 und M_2 ist die Menge die all jene Elemente enthält, die in M_1 oder in M_2 liegen aber nicht in beiden. Man schreibt $M_1 \triangle M_2$ für diese Menge.
$\left[M_1 \triangle M_2 = \{x : \ (x \in M_1 \vee x \in M_2) \wedge \neg(x \in M_1 \wedge x \in M_2)\} \right]$

Oft trifft man auch auf den Begriff des Komplementes einer Menge M. Dieser ist für sich alleine stehend sinnlos. Was man dazu braucht ist eine „große Grundmenge" auf die sich das „Komplement" bezieht: Ist X eine Menge, und $M \subseteq X$, so ist das *Komplement von M in X*, geschrieben als M^c die Menge $X \setminus M$. Diese Notation findet dann Verwendung, wenn man stillschweigend angenehmen kann, dass es aus dem Zusammenhang klar ist was die Grundmenge X ist auf die sich das Komplement bezieht.

Man kann Beziehungen zwischen Mengenoperationen recht einfach durch Umformungen logischer Verknüpfungssymbole erhalten. Wir listen einige Beispiele für solche *Rechenregeln* auf, die häufig verwendet werden.

J102 (15) $A \cap B = B \cap A, \quad A \cup B = B \cup A,$

(16) $(A \cap B) \cap C = A \cap (B \cap C), \quad (A \cup B) \cup C = A \cup (B \cup C),$

(17) $A \cap \bigcup_{B \in \mathcal{M}} B = \bigcup_{B \in \mathcal{M}} (A \cap B), \quad A \cup \bigcap_{B \in \mathcal{M}} B = \bigcap_{B \in \mathcal{M}} (A \cup B),$

(18) $A \cap (A \cup B) = A, \quad A \cup (A \cap B) = A, \quad A \cap A = A \cup A = A,$

J103 (19) $X \setminus \bigcup_{A \in \mathcal{M}} A = \bigcap_{A \in \mathcal{M}} (X \setminus A), \quad X \setminus \bigcap_{A \in \mathcal{M}} A = \bigcup_{A \in \mathcal{M}} X \setminus A,$

(20) $A \subseteq B \Leftrightarrow A \cap B = A \Leftrightarrow A \cup B = B.$

Zum Beispiel beruht (15) auf den Kommutativsätzen

$$\begin{aligned} x \in A \cap B &\Leftrightarrow x \in A \wedge x \in B \\ &\Leftrightarrow x \in B \wedge x \in A \quad \Leftrightarrow \quad x \in B \cap A \end{aligned}$$

und (19) auf den Verneinungssätzen

$$\begin{aligned} x \in X \setminus \bigcap_{A \in \mathcal{M}} A &\Leftrightarrow x \in X \wedge x \notin \bigcap_{A \in \mathcal{M}} A \\ &\Leftrightarrow x \in X \wedge \neg \left(\forall A : A \in \mathcal{M} \Rightarrow x \in A \right) \\ &\Leftrightarrow x \in X \wedge \exists A : \neg \left(A \in \mathcal{M} \Rightarrow x \in A \right) \\ &\Leftrightarrow x \in X \wedge \exists A : (A \in \mathcal{M} \wedge x \notin A) \\ &\Leftrightarrow \exists A : x \in X \wedge A \in \mathcal{M} \wedge x \notin A \\ &\Leftrightarrow \exists A : A \in \mathcal{M} \wedge x \in X \setminus A \quad \Leftrightarrow \quad x \in \bigcup_{A \in \mathcal{M}} X \setminus A \end{aligned}$$

Auch sind die angeführten Operationen nicht unabhängig voneinander. Zum Beispiel gilt

$$M_1 \triangle M_2 = (M_1 \setminus M_2) \cup (M_2 \setminus M_1).$$

Dies beruht auf der Formel

$$\begin{aligned} x \in M_1 \triangle M_2 &\Leftrightarrow (x \in M_1 \vee x \in M_2) \wedge \neg(x \in M_1 \wedge x \in M_2) \\ &\Leftrightarrow (x \in M_1 \vee x \in M_2) \wedge (x \notin M_1 \vee x \notin M_2) \\ &\Leftrightarrow (x \in M_1 \wedge x \notin M_1) \vee (x \in M_1 \wedge x \notin M_2) \vee (x \in M_2 \wedge x \notin M_1) \vee (x \in M_2 \wedge x \notin M_2) \\ &\Leftrightarrow (x \in M_1 \wedge x \notin M_2) \vee (x \in M_2 \wedge x \notin M_1) \quad \Leftrightarrow \quad x \in (M_1 \setminus M_2) \cup (M_2 \setminus M_1) \end{aligned}$$

Zum Schluss wollen wir noch ein paar Schreibweisen anführen, die man auch oft vorfindet.

– Es steht $\forall x \in M\colon A(x)$ für „für alle Elemente x von M ist $A(x)$ wahr“.
$\left[\ \forall x\colon \Big(x \in M \Rightarrow A(x)\Big)\ \right]$

– Analog steht $\exists\, x \in M\colon A(x)$ für „es gibt ein Element von M sodass $A(x)$ wahr ist“,
$\left[\ \exists\, x\colon \Big(x \in M \wedge A(x)\Big)\ \right]$

– und $\exists!\, x \in M\colon A(x)$ für $\exists!\, x\colon \Big(x \in M \wedge A(x)\Big)$,

– sowie $\nexists\, x \in M\colon A(x)$ für $\nexists\, x\colon \Big(x \in M \wedge A(x)\Big)$.

– Schreibweisen wie zum Beispiel „$\forall n \in \mathbb{N}, n \geq 2\colon A(x)$“ sind entsprechend zu verstehen. Um Fehler oder Missverständnisse zu vermeiden, empfiehlt sich eine Schreibweise, die explizit den Implikationspfeil enthält, also
$\forall n\colon \Big((n \in \mathbb{N} \wedge n \geq 2) \Rightarrow A(x)\Big)$. oder $\forall n \in \mathbb{N}\colon (n \geq 2 \Rightarrow A(x))$.

Ein Beispiel für eine sehr gebräuchliche – aber noch (mathematisch) umgangssprachlichere – Notation, ist soetwas wie „$A(x),\ x \in M$“. Dieses steht oft für $\forall x \in M\colon A(x)$. Zum Beispiel könnte „$f(x) \leq 1,\ x > 0$“ gut für $\forall x\colon x > 0 \Rightarrow f(x) \leq 1$ stehen.

> *Diese Schreibweise kann auch in anderem Zusammenhang auftreten und dann etwas ganz anderes bedeuten; also immer auf den Kontext achten!*

Zum Beispiel ist die Beistrich-Phrase „$f(x) \leq 1,\ x > 0$“ im Satz „Für x mit $f(x) \leq 1, x > 0$, gilt dann $A(x)$“ wohl eher als „und“ gemeint: $\forall x\colon (f(x) \leq 1 \wedge x > 0) \Rightarrow A(x)$.

Bemühen Sie sich, solche Abkürzungen zu vermeiden. Wenn Sie den Beistrich[‡] durch das richtige logische Symbol ersetzen (meist $\wedge$ oder $\Rightarrow$ oder $\Leftarrow$), drücken Sie damit klarer aus, was Sie meinen.

Wenn nach einer quantifizierten Variable wie $\forall x$ oder $\exists y$ eine kompliziertere Aussage steht – insbesondere: eine Implikation – dann empfiehlt es sich, dieser Aussage durch ein Klammerpaar zusammenzufassen: Statt $\forall x > 0 : x < 1 \Rightarrow f(x) > 0$ lieber $\forall x > 0 : (x < 1 \Rightarrow f(x) > 0)$. Das schaut nach mehr Schreibarbeit[§] aus, macht aber die Struktur der Formel klarer.

[‡] oder das Komma

[§] Umgekehrt kommt es gelegentlich vor, dass man bei der Anwendung eine Funktion auf ein Argument die Klammern weglässt, also einfach Fx statt $F(x)$ schreibt, wenn aus dem Kontext hervorgeht, das F eine Funktion ist und x im Definitionsbereich von F liegt.

Kapitel 2

Relationen und Funktionen

2.1 Relationen

2.1.1 Motivation

Unsere Welt ist voll von Beziehungen und Eigenschaften. Tatsächlich ist es schwierig, an ein Beispiel zu denken, wo keine Beziehung oder Attribut eines Objektes oder Subjektes dahintersteckt. Zum Beispiel sind Menschen in sozialen Netzwerken miteinander verbunden (in Beziehung stehend) oder eben nicht, oder es kann eine Blume rot oder weiß oder verblüht sein (ein Attribut besitzen oder eben nicht).

Wir möchten einen präzisen, d.h. keinen Raum für Missverständnisse oder Auslegungen lassenden Begriff, mit dem sich Beziehungen zwischen Objekten bzw. Besitz von Attributen fassen lassen.

Eine Motivation dafür wäre es zum Beispiel, dass wir unser Wissen und Verständnis über solche Beziehungen erweitern wollen, dass wir neue Erkenntnisse generieren oder Schlüsse aus unserem vorhandenen Wissen ziehen wollen.

2.1.2 Formalisierung

J201. **2.1.1 Definition.** Seien M und N Mengen. Eine Teilmenge von $M \times N$ heißt eine *Relation*.
Spezifischer sagt man auch eine *Relation zwischen M und N*. Ist $M = N$ so sagt man auch *Relation in M*. ◊

> ➢ `Was ist eine Definition:`
>
> Eine *Definition* gibt einem Begriff (einem Wort, einer Phrase, oder einer Sprech- oder Schreibweise) mathematische Bedeutung.
>
> – Der neue Begriff wird aus schon bekannten Begriffen und zusätzlichen Eigenschaften solcher aufgebaut; alles was zur Beschreibung des neuen Begriffes herangezogen wird muss bereits bekannt sein.
>
> – Kein Begriff kann mehr als einmal definiert werden.
>
> – Es muss unmissverständlich klargestellt sein wovon wir sprechen.

`Analyse von Definition` 2.1.1`:`

Die neu eingeführten Begriffe sind *Relation*, sowie *Relation zwischen M und N*, und *Relation in M*.

① Überprüfung formal logischer Richtigkeit:

– Wir wissen was eine Menge ist, wir wissen was das kartesische Produkt zweier Mengen ist, und wir wissen was eine Teilmenge einer Menge ist.

– Das Wort *Relation*, bzw. die Phrasen *Relation zwischen M und N*, und *Relation in M*, haben zum jetzigen Zeitpunkt in unserer mathematischen Welt noch keine Bedeutung.

– Die Definition ist präzise; sie lässt keinerlei Spielraum für Auslegungen oder Interpretationen.

➢ `Wozu wird eine Definition gemacht:`

Motivation dafür eine bestimmte Definition zu geben kann sein

- wichtige/interessante Konzepte explizit zu machen;
- zentrale Eigenschaften die ein Objekt haben kann herauszustreichen;
- Klassen von Objekten zu benennen die viele – spezielle oder „gute“ – Eigenschaften haben.

`Analyse von Definition` 2.1.1 `(Fortsetzung):`

② Motivation ?

- Ist der Begriff der *Relation* ein wichtiges oder interessantes Konzept ?
  Ja; wie sich im Laufe der Zeit erwiesen hat.
- Werden zentrale Eigenschaften aufgezeigt ?
  Nein. Es werden ja gar keine Eigenschaften verlangt außer Teilmenge von $M \times N$ zu sein.
- Wird eine Klasse von Objekten mit vielen Eigenschaften benannt ?
  Nein. Dazu ist der Begriff zu allgemein; es fallen so viele Objekte darunter, dass die Gültigkeit vieler spezieller Eigenschaften nicht zu erwarten ist.

Die Bedeutung des Begriffs der *Relation* liegt also in erster Linie darin einen konzeptuellen Rahmen abzustecken, und nicht so sehr darin weitgehende Schlüsse zu ziehen. Innerhalb dieses Rahmens kann – und wird – man spezielle Relationen, also solche mit gewissen zusätzlichen Eigenschaften betrachten, über welche bzw. mit deren Hilfe man dann mehr aussagen kann.

Sei R eine Relation zwischen M und N. Entsprechend unserer Motivation in §2.1.1 wollen wir R auch auffassen als Beschreibung einer Beziehung zwischen Elementen m aus M und Elementen n aus N. Nämlich jene, dass $(m, n) \in R$. Aus diesem Grund verwendet man oft die – manchmal intuitivere – Schreibweise $m\,R\,n$ um auszudrücken, dass $(m, n) \in R$.

In obigem Absatz wurde auch eine Definition gegeben, nämlich die der Schreibweise mRn.

① Die formalen Anforderungen sind erfüllt.

② Man denke an aus der Anschauung bekannte Beziehungen zwischen (ebenso aus der Anschauung bekannt geglaubten) Zahlen wie „x ist kleiner als y“ oder „x teilt y“. Für solche Beziehungen schreibt man ja traditioneller Weise $x < y$ bzw. $x \mid y$, und nicht $(x, y) \in <$ oder $(x, y) \in |$.

Man sollte an dieser Stelle schon anmerken, dass der Begriff der Relation als abstraktes Konzept entstanden ist, welches solche von jeher verwendete Relationen allgemein fasst.

2.2 Äquivalenzrelationen

2.2.1 Motivation

Erzählung eines Autofahrers:

> Mein Freund Rainer und ich...wir haben das *gleiche* Auto; nämlich einen Mazda 6. Also, eigentlich sind sie *gleich*, haben sogar das *gleiche* Baujahr, nur seiner ist weiß und meiner violett.
>
> Und jetzt für die Pingeligen unter uns: natürlich sind sie nicht *gleich*, weil wir ja nicht car-sharing betreiben. Und überhaupt, sie können ja gar nicht *gleich* sein, weil ja seiner weiß ist und meiner violett.

Der Begriff der Gleichheit tritt hier offenbar in vier verschiedenen Bedeutungen auf, ist je nach Interpretation dann zutreffend oder nicht, und verschiedene Interpretationen werden miteinander in Beziehung gesetzt. Zwei Autos sind gleich, wenn sie:

– die gleiche Marke und Typ haben;

– das gleiche Baujahr haben;

– zusätzlich noch die gleiche Farbe haben;

– tatsächlich ein und dasselbe Objekt sind.

Welche Eigenschaften sind es, die es ausmachen, dass man das Wort *gleich* in „vernünftiger“ Weise verwenden kann ? Anders gefragt, was würde man sich von einem „vernünftigen“ Gleichheitsbegriff erwarten ?

Sicherlich sollten die folgenden drei Eigenschaften erfüllt sein (und jede der vier obigen Gleichheitsinterpretationen für Autos erfüllt diese):

(i) Jedes Objekt unserer Grundmenge ist gleich zu sich selbst.

(ii) Sind x, y zwei Objekte unserer Grundmenge und ist x gleich zu y, dann ist auch y gleich zu x.

(iii) Sind x, y, z drei Objekte unserer Grundmenge und ist x gleich zu y sowie auch y gleich zu z, so muss auch x gleich zu z sein.

2.2.2 Formalisierung

Zuerst wollen wir den oben erwähnten Eigenschaften (i) – (iii) Namen geben.

J205. **2.2.1 Definition.** Sei M eine Menge, und R eine Relation in M. Dann heißt R *reflexiv*, wenn

(Ref) jedes Element $x \in M$ erfüllt, dass xRx,
$\left[\forall x \in M \colon (x, x) \in R \right]$

symmetrisch, wenn

(Sym) für je zwei Elemente $x, y \in M$ mit xRy auch yRx gilt,
$\left[\forall x, y \in M \colon (x, y) \in R \;\Rightarrow\; (y, x) \in R \right]$

und *transitiv*, wenn

(Tra) für $x, y, z \in M$ mit xRy und yRz auch xRz gilt.
$\left[\forall x, y, z \in M \colon \big((x, y) \in R \wedge (y, z) \in R \big) \;\Rightarrow\; (x, z) \in R \right]$

◊

➢ `Was sind Axiome:`

Als *Axiom* bezeichnet man eine, von möglicherweise mehreren, Eigenschaften die ein Objekt haben soll um einen gewissen Namen tragen zu dürfen. Zum Beispiel sagt man, dass eine Relation reflexiv heißt, wenn sie dem Axiom **(Ref)** genügt.

Manchmal denkt man bei *Axiom* eher an *Grundwahrheiten* als an scheinbar beliebige Eigenschaften. Damit meint man Eigenschaften die ein Objekt haben soll, und die anschaulich(!) so außerordentlich naheliegend sind, dass sie sicher von allen akzeptiert werden. Zum Beispiel sagt man, dass die natürlichen Zahlen dem Axiom genügen, dass jede Zahl einen Nachfolger hat.

♧ `Analyse von Definition` 2.2.1:
Die neu eingeführten Begriffe sind *reflexiv*, *symmetrisch*, sowie *transitiv*.

① Überprüfung formal logischer Richtigkeit:

– Wir wissen was eine Menge ist und wir wissen was eine Relation in M ist. Die angeführten Eigenschaften sind schlüssig formuliert; wir haben in den Axiomen logische Formeln angegeben die sinnvolle Gestalt haben.
– Die verwendeten Worte *reflexiv*, *symmetrisch*, sowie *transitiv*, haben zum jetzigen Zeitpunkt in unserer mathematischen Welt noch keine Bedeutung.
– Die Definitionen sind präzise und lassen keinen Spielraum für Auslegungen.

② Motivation ?
– Ist der Begriff *reflexiver* (*symmetrischer*, *transitiver*) Relationen ein wichtiges oder interessantes Konzept ?
Ja. Solche Relationen treten in verschiedensten Zusammenhängen, auch in Kombination mit anderen Eigenschaften, auf.
– Werden zentrale Eigenschaften aufgezeigt ?
Ja.
– Wird eine Klasse von Objekten mit vielen Eigenschaften benannt ?
Ein bisschen. Jeder einzelne Begriff für sich selbst genommen ist aber immer noch recht allgemein.

♲

2.2.2 Bemerkung. Wir sind daran gewöhnt, Relationen als *Prädikate* zu betrachten, mit denen wir Aussagen bilden können. Wenn wir zum Beispiel die Relation „kleiner (<)“ auf den natürlichen Zahlen betrachten, können wir mit ihr die (falsche) Aussage $5 < 3$ bilden.

Aber die Relation $<$ ist auch ein *Objekt* unserer Überlegungen, nämlich eine Menge von Paaren: $\{(0,1),(0,2),\dots,(1,2),(1,3),\dots,\dots\}$. Wir können auch Aussagen *über* diese Relation machen, wie etwa „die Relation $<$ ist transitiv aber nicht symmetrisch“.

J234. **2.2.3 Lemma.** *Sei M eine Menge und R eine Relation in M. Dann ist R genau dann symmetrisch, wenn*

(Sym′) *für je zwei Elemente $x, y \in M$ genau dann xRy gilt, wenn yRx gilt.*
$\left[\ \forall x, y \in M\colon (x,y) \in R \ \Leftrightarrow\ (y,x) \in R\ \right]$

➢ `Was ist ein Lemma:`

Ein *Lemma* ist eine mathematische Aussage deren Wahrheit behauptet wird. Die Wahrheit dieser Aussage muss bewiesen werden !

– In der Aussage dürfen nur Begriffe und Symbole vorkommen die bereits definiert wurden.
– Es muss unmissverständlich klargestellt sein wovon wir sprechen. Also
① was sind die Objekte mit denen wir arbeiten,
② was sind die Voraussetzungen die verlangt werden,
③ was ist die Behauptung die gemacht wird.

Die Bezeichnung *Lemma* steht (in der Mathematik) ursprünglich für „Hilfsaussage“, also eine Aussage deren Wahrheit für den Beweis anderer – komplexerer – Aussagen gebraucht wird. Allerdings hat es sich eingebürgert auch „kleinere/einfachere“ Aussagen als Lemma zu bezeichnen.
Die Mehrzahl von *Lemma* ist übrigens *Lemmata*.

♲ `Analyse von Lemma 2.2.3:`

– Wir wissen was eine Menge ist, was eine Relation in M ist, und wann eine Relation symmetrisch heißt. Wir wissen was Elemente einer Menge sind, und was die Schreibweise xRy bedeutet. Die angeschriebene Formel ist formal sinnvoll.
– Klar ?
① Das Objekt mit dem wir arbeiten ist die Relation R in M.
② Voraussetzungen werden keine verlangt.

③ Die Behauptung ist, dass die Symmetrie von R, sprich die Eigenschaft **(Sym)**, äquivalent zu der angeschriebenen Eigenschaft **(Sym′)** ist.

Wir wollen uns jetzt mit dem Beweis beschäftigen.

➢ `Was ist ein Beweis:`

Ein *Beweis* ist eine Herleitung der Wahrheit einer Aussage mittels logischer Schlüsse aus den zugrundegelegten Axiomen, sowie bereits als wahr erkannten (sprich bewiesenen) Aussagen.

Typischerweise wird es für eine wahre Aussage viele verschiedene Beweise geben. Diese können grundlegend verschieden sein, zum Beispiel in dem Sinn, dass sie unterschiedliche Methoden oder Zugänge verwenden, oder sie können ähnlich sein, nur einer vielleicht kürzer oder eleganter als der andere. Grundsätzlich ist aber jeder Beweis – natürlich sofern er richtig ist – gleichberechtigt. Er zeigt, dass die Aussage wahr ist; Punkt.

Jedenfalls ist es oft interessant für eine Aussage verschiedene Beweise zu kennen; es kann neue Einsichten bringen, neue Ideen motivieren, und vertieft in jedem Fall das Verständnis der Materie.

Zuerst der Beweis wie man ihn wahrscheinlich in der Literatur vorfinden würde; wir werden ihn danach filetieren und um jedes Argument im Detail zu verstehen.

Beweis von Lemma 2.2.3. Für „**(Sym)⇒(Sym′)**“ seien $x, y \in M$. Ist $(x, y) \in R$, so folgt nach **(Sym)** auch $(y, x) \in R$. Vertauscht man die Rollen von x und y, so erhält man die umgekehrte Implikation.

Die Umkehrung ist trivial. □

➢ `Was heißt hier trivial:`

Es ist ein häufig verwendeter Sprachgebrauch eine mathematische Aussage *trivial* zu nennen, wenn ihr Beweis kurz ist, keine neuen Ideen oder Konstruktionen erfordert, und von Leser/Autor sofort ohne viel nachdenken hingeschrieben werden könnte.

Eigentlich ist die Verwendung dieses Terminus ganz schlechter Stil (und alles andere als präzise), weil es offensichtlich *für jeden Leser/Autor* es etwas *ganz anderes* bedeuten kann.

Warum verwendet man diesen Terminus dann trotzdem? Nun, es gibt schon einen gewissen Konsens darüber, welche Argumente/Aussagen „leicht“ und welche „schwer“ sind, und bis zu welchem Grad man es erwarten kann, dass durchschnittliche Leser Argumente selbstständig aus dem Ärmel schütteln können. Der Autor mit Fingerspitzengefühl schafft es die potentielle Zielgruppe so einzuschätzen, dass es für die meisten Leser passen wird.

Ein zweiter Grund für die Verwendung solcher Phrasen ist, dass es dem Leser eine Information darüber gibt was der Autor als essentiell/komplex/innovativ und was als eh-klar/Beiwerk/Standardmethode sieht. Auch das ist eine subjektive Einschätzung, aber auch darüber gibt es meist einen gewissen Konsens.

Das oben Gesagte trifft genauso auf Phrasen zu wie zum Beispiel *eine einfache Rechnung zeigt*, *offensichtlich*, *unmittelbar klar*, oder (wahrscheinlich die am häufigsten als Frechheit empfundene Phrase) *wie man leicht einsieht*.

➢ `Was bedeutet das Symbol □ ?:`

Das Symbol □ (oder ähnliches) wird verwendet um explizit herauszustreichen, dass der eben durchgeführte Beweis vollständig und abgeschlossen ist.

In der älteren Literatur findet man auch oft *q.e.d*. Das ist eine Abkürzung für *quod erat demonstrandum* (d.h. „was zu beweisen war“). Die Verwendung von □ als Markierung für das Ende eines Beweises geht auf den Mathematiker Paul Halmos zurück; man spricht auch – wohl ein bisschen scherzhaft – vom *Halmos tombstone*.

♲ `Analyse des Beweises von Lemma 2.2.3:`
Dieser Beweis ist tatsächlich recht einfach, enthält aber doch ein paar häufig verwendete Prinzipien.

	Wir haben die Äquivalenz zweier Aussagen zu zeigen, nämlich **(Sym)** $\Leftrightarrow$ **(Sym′)**. Nun ist die Wahrheit der Äquivalenz zweier Aussagen äquivalent zur Wahrheit der beiden wechselseitigen Implikationen. Wir erinnern uns an die Tautologie (18) aus §1.1.2: für zwei beliebige Aussage A_1, A_2 gilt $$(A_1 \Leftrightarrow A_2) \;\Leftrightarrow\; \big((A_1 \Rightarrow A_2) \wedge (A_1 \Leftarrow A_2)\big) \qquad (2.2.1)$$ J235 Wir gehen darauf los die beiden Implikationen **(Sym)** $\Rightarrow$ **(Sym′)** und **(Sym)** $\Leftarrow$ **(Sym′)** zu zeigen.
Für „**(Sym)** $\Rightarrow$ **(Sym′)**“	Im ersten Beweisteil gehen wir auf die Implikation **(Sym)** $\Rightarrow$ **(Sym′)** los. ➪ Dazu sei, bis zum Ende dieses Beweisteiles, vorausgesetzt, dass **(Sym)** wahr ist. ⟩①⟩ Wir müssen zeigen, dass **(Sym′)** wahr ist.
seien $x, y \in M$.	Wir müssen eine Aussage zeigen, die mit einem Allquantor anfängt; sprich, dass alle Elemente $x, y \in M$ eine gewisse Eigenschaft haben. $\big[\; \forall x, y \in M\colon (x, y) \in R \;\Leftrightarrow\; (y, x) \in R \;\big]$ ➪ Seien uns beliebige, bis zum Ende dieses Beweisteiles festgehaltene, Elemente $x, y \in M$ vorgegeben. ⟩②⟩
	Wir müssen die Äquivalenz von $(x, y) \in R$ und $(y, x) \in R$ zeigen. Dazu benützen wir wiederum (2.2.1), und versuchen die beiden Implikationen $(x, y) \in R \Rightarrow (y, x) \in R$ und $(x, y) \in R \Leftarrow (y, x) \in R$ zu zeigen.
Ist $(x, y) \in R$,	Im ersten Schritt zeigen wir die in obiger Zeile erstgenannte Implikation. ➪ Sei dazu, bis zum Ende dieses ersten Schrittes, vorausgesetzt, dass $(x, y) \in R$. ⟩③⟩ Wir müssen zeigen, dass $(y, x) \in R$.
so folgt nach **(Sym)** auch $(y, x) \in R$.	Wir wenden die Implikation in **(Sym)**, die ja wegen ① für alle Elemente von M wahr ist, an auf die Elemente x und y. Da die Prämisse wegen ③ wahr ist, erhalten wir, dass auch die Konklusion wahr ist, sprich, dass $(y, x) \in R$.
	Der erste Schritt ist abgeschlossen. ➩ Es ist nicht länger angenommen, dass $(x, y) \in R$. ⟨③⟨
	Im zweiten Schritt zeigen wir die zweitgenannte Implikation. ➪ Sei dazu, bis zum Ende dieses zweiten Schrittes, vorausgesetzt, dass $(y, x) \in R$. ⟩③⟩
	Wir wenden die Implikation in **(Sym)**, die ja wegen ① für alle Elemente von M wahr ist, an auf die Elemente y und x. Da die Prämisse wegen ③ wahr ist, erhalten wir, dass auch die Konklusion wahr ist, sprich, dass $(x, y) \in R$.
	Der zweite Schritt ist abgeschlossen. ➩ Es ist nicht länger angenommen, dass $(y, x) \in R$. ⟨③⟨

Vertauscht man die Rollen von x und y, so erhält man die umgekehrte Implikation.	Was wir in obigem zweiten Schritt getan haben, ist tatsächlich genau das Gleiche wie das was wir im ersten Schritt getan haben. Einzig, dass an Stelle von x und y nun y und x verwendet wurden.
➥ Absatz	Der erste Beweisteil ist abgeschlossen. ⇜ Die Buchstaben x und y sind nicht mehr gebunden (gleich den uns am Anfang des Schrittes vorgegebenen Elementen von M) und wieder frei verfügbar. ⟨ ② ⟨ ⇜ Es ist nicht länger angenommen, dass **(Sym)** wahr ist. ⟨ ① ⟨
Die Umkehrung	Im zweiten Beweisteil gehen wir auf die Implikation **(Sym)** $\Leftarrow$ **(Sym′)** los. Das ist die zu der im ersten Beweisteil behandelten Umgekehrte. ⇝ Dazu sei, bis zum Ende dieses Beweisteiles, vorausgesetzt, dass **(Sym′)** wahr ist. ⟩ ① ⟩ Wir müssen zeigen, dass **(Sym)** wahr ist.
ist trivial.	Warum ist diese Implikation trivial ? *Vor dem weiterlesen zuerst selbst nachdenken!* Wir benützen (2.2.1), aber nicht in voller Stärke sondern nur zur Hälfte (damit meinen wir hier in Kombination mit einer Konjuktionsbeseitigung): für zwei beliebige Aussage A_1, A_2 gilt $$(A_1 \Leftrightarrow A_2) \;\Rightarrow\; (A_1 \Rightarrow A_2)$$ Wenn also für zwei vorgegebene Elemente $x, y \in M$ die Äquivalenz in **(Sym′)** gilt, so gilt für diese Elemente insbesondere auch die Implikation in **(Sym)**.
	Der zweite Beweisteil ist abgeschlossen. ⇜ Es ist nicht länger angenommen, dass **(Sym′)** wahr ist. ⟨ ① ⟨
□	Der Beweis des Lemmas ist abgeschlossen.

♧

♧ `Analyse der Argumentation im Beweis von Lemma` 2.2.3`:`

In diesem Beweis verwenden wir wiederholt ein Argument, welches man wahrscheinlich als die grundlegendste und natürlichste logische Argumentation bezeichnen kann. Nämlich das *Prinzip des direkten Beweises*. Wir sagen zum Beispiel:

> Wir wenden die Implikation in **(Sym)**, die ja wegen ① für alle Elemente von M wahr ist, an auf die Elemente x und y. Da die Prämisse wegen ③ wahr ist, erhalten wir, dass auch die Konklusion wahr ist, sprich, dass $(y, x) \in R$.

Was heißt es eine Implikation „anzuwenden“ ? Hier verwenden wir eigentlich den modus ponendo ponens. Wir *wissen* dass eine Implikation, sagen wir $A \Rightarrow B$, wahr ist, und wir *wissen* dass ihre Prämisse A wahr ist. Daher können wir schließen, dass die Aussage B wahr ist. ♧

➢ `Ein direkter Beweis funktioniert so:`

Wir gehen davon aus, dass die zu beweisende Aussage (zum Beispiel das gerade betrachtete Lemma) bereits in dem Sinne analysiert wurde, dass es klar ist was die zur Verfügung stehenden Voraussetzungen sind, und

was die zu zeigende Behauptung ist. Wollen wir diese Voraussetzungen als A und diese Behauptung als B bezeichnen.

① Man zeigt eine, oder eine Abfolge von, Implikationen die mit A anfängt und zu B führt. Sprich, man findet Zwischenaussagen, sagen wir $C_1, C_2, \ldots, C_n$, und zeigt die Wahrheit der Implikationen

$$A \Rightarrow C_1\,,\ C_1 \Rightarrow C_2\,,\ \ldots\,,\ C_{n-1} \Rightarrow C_n\,,\ C_n \Rightarrow B\,.$$

Was nun folgt, ist die formallogische Argumentation, dass damit der Beweis tatsächlich abgeschlossen ist.

❷ Man wendet wiederholt den modus barbara an, wodurch man $A \Rightarrow B$ erhält.

❸ Man wendet den modus ponendo ponens an, nachdem man aus der Wahrheit von A und der Wahrheit der Implikation $A \Rightarrow B$ auf die Wahrheit von B schließen kann.

Diese Erklärung scheint vielleicht unnötig kompliziert, wohl deswegen weil der modus ponendo ponens ja wirklich ganz unmittelbar aus der Wahrheitstafel der Verknüpfung „$\Rightarrow$“ ersichtlich ist. Sie soll uns aber, abgesehen davon, dass sie die grundlegende Beweismethode des direkten Beweises ganz im Detail aufschlüsselt, auch als Modell dienen, nämlich für Erklärungen anderer – logisch komplizierterer – Beweismethoden die wir später kennenlernen werden.

Jetzt definieren wir, was wir unter einem „vernünftigen Gleichheitsbegriff“ verstehen wollen.

J206. **2.2.4 Definition.** Sei M eine Menge und R eine Relation in M. Dann heißt R eine *Äquivalenzrelation* (spezifischer sagt man auch *Äquivalenzrelation auf M*), wenn R reflexiv, symmetrisch, und transitiv ist. ◊

♧ `Analyse von Definition` 2.2.4`:`
Der neu eingeführte Begriff ist *Äquivalenzrelation*.

① Überprüfung formal logischer Richtigkeit:

– Wir wissen was eine Menge ist, was eine Relation in M ist, und wann eine Relation reflexiv, symmetrisch, bzw. transitiv heißt.

– Das verwendete Wort *Äquivalenzrelation* hat zum jetzigen Zeitpunkt in unserer mathematischen Welt noch keine Bedeutung.

– Die Definition ist präzise und läßt keinen Spielraum für Auslegungen.

② Motivation ?

– Ist der Begriff der *Äquivalenzrelation* ein wichtiges oder interessantes Konzept ?
  Ja. Es wird damit unsere Anschauung von „vergröberter Gleichheit“ modelliert. Die Idee eine Grundmenge einmal mit Vergrößerungsglas und einmal ohne Brille (verschwommen/vergröbert) anzuschauen, spielt in unzähligen Zusammenhängen eine wesentliche Rolle.

– Werden zentrale Eigenschaften aufgezeigt ?
  Ja.

– Wird eine Klasse von Objekten mit vielen Eigenschaften benannt ?
  Ja. Die drei Eigenschaften reflexiv, symmetrisch, und transitiv gemeinsam sind recht stark, und man kann etliche interessante Schlüsse ziehen.

♧

➢ `Wie versteht man eine Definition:`

Bisher haben wir uns damit beschäftigt ob

– eine Definition formal richtig ist, und ob

– eine Idee, Anschauung, oder Konzept hinter ihr steckt, bzw. ob sie eine reichhaltige Struktur beschreibt.

Es ist wichtig, beide diese Schritte zu tun!

Ersterer ist, entsprechend des Exaktheitsanspruches der Mathematik an sich, unabdingbar. Zweiterer ist notwendig, wenn man nicht in einem, zwar logisch soliden aber leeren Gebäude verbleiben möchte, sondern dieses Gebäude mit Inhalt und Leben erfüllen will.

Um einen Begriff zu *verstehen*, kann man sich den folgenden Fragen widmen.

- Gibt es solche Objekte, wie sie in der Definition beschrieben werden, überhaupt ? Normalerweise wird die Antwort „ja“ lauten. Beachte jedoch, dass auch die Antwort „nein“ sehr interessant sein kann.
- Was sind Standardbeispiele für solche Objekte, was sind Trivialbeispiele, was sind Extremfälle ?
- Gibt es viele solche Objekte, oder vielleicht genau eines, oder sind alle die es gibt vielleicht in irgendeinem Sinne ähnlich ?
- Kann man vielleicht in ganz simpler Weise aus gegebenen Objekten andere konstruieren, die auch die in der Definition verlangten Eigenschaften haben ?
- Wie arbeitet der definierte Begriff, was kann man mit ihm machen ?

Meist ist es so, dass man erstmal eine Zeitlang kritiklos mit einem Begriff hantieren muss, bis sich Antworten auf solche oder ähnliche Fragen ergeben, die dann zu tieferem Verständnis beitragen. Antworten auf die genannten Fragen sind oft alles andere als offensichtlich oder schnell und einfach zu finden !

Als Beispiele für Äquivalenzrelationen wollen wir zwei Extremfälle angeben (die auch ein bisschen trivial sind).

Dazu noch eine Bezeichnung: Für eine Menge M bezeichnen wir mit Δ_M jene Teilmenge von $M \times M$ die gegeben ist als

$$\Delta_M := \{(x, y) \in M \times M : \ x = y\}. \tag{2.2.2}$$ J236

Diese Menge heißt die *Diagonale* von M.

♲ Wieder einmal eine Definition.

① Die formalen Anforderungen sind erfüllt.

② Diese Relation modelliert *tatsächliche* Gleichheit. Bezüglich der Namensgebung denke man an das aus der Anschauung bekannte Bild der Diagonale eines Quadrates (in der Ebene) mit Mittelpunkt im Nullpunkt.

♲

J207. *2.2.5 Beispiel.* Sei M eine Menge. Dann sind Δ_M und $M \times M$ Äquivalenzrelationen.

Dabei ist Δ_M eine Äquivalenzrelation weil tatsächliche Gleichheit offenbar reflexiv, symmetrisch, und transitiv ist: Für alle Elemente $x, y, z \in M$ gilt

$$x = x, \qquad x = y \ \Rightarrow \ y = x, \qquad (x = y \wedge y = z) \ \Rightarrow \ x = z.$$

Die Relation $M \times M$ ist eine Äquivalenzrelation, weil die Bedingung **(Ref)**, sowie die Konklusionen der Implikationen in **(Sym)** und **(Tra)** (sogar unabhängig von dem Wahrheitswert der Prämisse) wahr sind, da ja alle Paare zu $M \times M$ gehören. ◊

➢ `Was ist ein Beispiel:`

Ein *Beispiel* kann zweierlei Zweck haben.

Einerseits kann es eine konkrete Situation oder ein konkretes Objekt erläutern, welches in die definierte oder betrachtete Klasse von Objekten fällt. Dann dient es dazu die abstrakten Begriffe oder Aussagen zu illustrieren, und vielleicht ein Gefühl zu geben wie diese arbeiten oder zusammenspielen.

Andererseits kann ein *Beispiel* auch die Rolle eines Beweises übernehmen, nämlich als *Gegenbeispiel*. Möchte man beweisen, dass eine gewisse Aussage vom Typ „Für alle ␣ mit den Eigenschaften ␣ gilt ␣“ (wobei an der Stelle von ␣ gewisse konkrete Eigenschaften stehen) *nicht* wahr ist, so macht man das oft durch Angabe einer konkreten Situation (sprich eines Beispiels) welche in die betrachtete Klasse fällt, in der Voraussetzung

genannte Eigenschaften hat, aber die *behaupteten* Eigenschaften eben *nicht* hat. Diese Vorgangsweise beruht auf dem Verneinungssatz (3) aus §1.1.3, sprich, die Formel

$$\neg \left(\forall x\colon P(x)\right) \;\Leftrightarrow\; \exists x\colon \neg P(x)$$

Ein gutes *Gegenbeispiel* ist oft goldeswert; es kann aufzeigen, woran es scheitert, dass die betrachtete Aussage falsch ist und damit einen Hinweis geben, wie sie zu modifizieren wäre, um eine wahre Aussage zu erhalten. Modifizieren bedeutet hier typischerweise

- einschränken (oder abändern) der betrachteten Klasse von Objekten,
- hinzufügen einer (oder verstärken eines Teils der) Voraussetzung,
- weglassen (oder abschwächen) eines Teils der Behauptung.

Beispiele werden typischerweise gewisse Definitionen enthalten und auch gewisse Aussagen (die dann natürlich eines Beweises bedürfen).

♧ `Analyse von Beispiel` 2.2.5`:`

- Wir wissen was eine Menge ist, wofür Δ_M steht, was $M \times M$ ist und was eine Äquivalenzrelation ist.
- Weitere Definitionen werden nicht gemacht.
- Die konkreten Objekte mit denen wir arbeiten sind Δ_M und $M \times M$.
- Voraussetzungen werden keine verlangt.
- Die Behauptung ist, dass beide Äquivalenzrelationen sind.
- Das angeführte Argument ist (...hoffentlich) klar und einsichtig.

♧

Schließlich wollen wir noch eine einfache Konstruktion angeben wie man aus zwei gegebenen Äquivalenzrelationen eine neue bauen kann.

J208. **2.2.6 Lemma.** *Sei M eine Menge und seien R_1 und R_2 Äquivalenzrelationen auf M. Dann ist auch $R_1 \cap R_2$ eine Äquivalenzrelation.*

♧ `Analyse von Lemma` 2.2.6`:`

- Wir wissen was eine Menge ist und was eine Äquivalenzrelation ist.
- Klar ?

  ① Die Objekte mit denen wir arbeiten sind R_1, R_2, und $R_1 \cap R_2$.
  ② Die Voraussetzung ist, dass R_1 und R_2 beide Äquivalenzrelationen sind.
  ③ Die Behauptung ist, dass $R_1 \cap R_2$ eine Äquivalenzrelation ist.

♧

Beweis von Lemma 2.2.6. Für $x \in M$ gilt $(x, x) \in R_1$ und $(x, x) \in R_2$, also auch $(x, x) \in R_1 \cap R_2$.

Sei $(x, y) \in R_1 \cap R_2$. Dann ist $(x, y) \in R_1$ und auch $(x, y) \in R_2$. Also folgt $(y, x) \in R_1$ und auch $(y, x) \in R_2$, und wir erhalten $(y, x) \in R_1 \cap R_2$.

Die Transitiviät folgt genauso. □

Dieser Beweis ist wirklich *straightforward*.

➢ `Was heißt hier straightforward:`

Ein Beweis (oder ein Argument) wird oft – eigentlich in sprechender Bezeichnungsweise – als *straightforward* bezeichnet, wenn er bloss Standardmethoden benützt, und/oder keinerlei neue Ideen oder Konstruktionen erfordert, und/oder schon nahezu rezeptartig algorithmisch verläuft. Sprich, wenn er zwar nicht ganz trivial ist aber keinerlei wirklichen intellektuellen Herausforderungen beinhaltet.

Die Verwendung dieses Terminus ist wieder einmal alles andere als präzise weil es offensichtlich *für jeden Leser/Autor* etwas *ganz anderes* bedeuten kann. Meist ist er so zu verstehen:

> Wenn man erstmal mit der Materie vertraut ist und schon einige Erfahrung hat, dann ist es wirklich straightforward.

Durch Verwendung dieses Terminus erhält der Leser wieder eine Information darüber was der Autor als essentiell/komplex/innovativ und was als eh-klar/Beiwerk/Standardmethode sieht. Und wieder gibt es meist einen gewissen Konsens darüber, auf welche Argumente dies zutrifft.

Warum beschäftigen wir uns überhaupt mit als *straightforward* erkannten Argumenten ? Gerade deswegen um mit einer Materie vertraut zu werden, und um Erfahrung zu sammeln. Wenn man Routine mit simplen Argumentationen hat, kann man besser die Spreu vom Weizen trennen und in einem komplexen Beweis die Essenz vom Beiwerk unterscheiden.

Es ist wichtig diese Erfahrung und Fingerfertigkeit zu bekommen!

`Analyse des Beweises von Lemma 2.2.6:`

Wir wollen nun den Beweis filetieren um zu lernen, dass er wirklich straightforward ist.

	Wir müssen zeigen, dass $R_1 \cap R_2$ reflexiv, symmetrisch, und transitiv ist. Die Voraussetzung des Lemmas ist, dass beide, R_1 und R_2, diese drei Eigenschaften haben.
Für $x \in M$	Im ersten Schritt gehen wir darauf los die Reflexivität zu zeigen. Das ist eine Eigenschaft die mit einem Allquantor anfängt: $\left[\ \forall x \in M\colon (x,x) \in R_1 \cap R_2\ \right]$ ⇀ Sei uns ein beliebiges, bis zum Ende dieses Schrittes festgehaltenes, Element $x \in M$ vorgegeben. ⟩①⟩
gilt $(x,x) \in R_1$ und $(x,x) \in R_2$,	Nach der Voraussetzung des Lemmas sind die Relationen R_1 und R_2 beide reflexiv.
also auch $(x,x) \in R_1 \cap R_2$.	Ein Objekt ist Element des Durchschnitts zweier Mengen, genau dann wenn es Element beider Mengen ist: für beliebige Mengen M_1, M_2 gilt $$a \in M_1 \cap M_2 \ \Leftrightarrow\ (a \in M_1 \ \wedge\ a \in M_2) \qquad (2.2.3)$$ J237
➥ Absatz	Der Beweis der Reflexivität ist abgeschlossen. ↽ Der Buchstabe x ist nicht mehr gebunden (gleich dem uns am Anfang des Schrittes vorgegebenen Element von M) und wieder frei verfügbar. ⟨①⟨

	Im zweiten Schritt gehen wir darauf los die Symmetrie zu zeigen. Wieder eine Eigenschaft die mit einem Allquantor anfängt: $\left[\ \forall x, y \in M\colon (x,y) \in R_1 \cap R_2 \Rightarrow (y,x) \in R_1 \cap R_2\ \right]$ ⇀ Seien uns beliebige, bis zum Ende dieses Beweisschrittes festgehaltene, Elemente $x, y \in M$ vorgegeben. ⟩①⟩
Sei $(x,y) \in R_1 \cap R_2$.	Wir müssen zeigen, dass für die uns vorgegebenen Elemente x, y eine gewisse Implikation wahr ist. ⇀ Sei dazu, bis zum Ende dieses Schrittes, vorausgesetzt, dass $(x,y) \in R_1 \cap R_2$. ⟩②⟩ Ziel ist zu zeigen, dass die Konklusion der gewünschten Implikation wahr ist.
Dann ist $(x,y) \in R_1$ und auch $(x,y) \in R_2$.	Das ist (2.2.3).
Also folgt $(y,x) \in R_1$ und auch $(y,x) \in R_2$,	Nach der Voraussetzung des Lemmas sind die Relationen R_1 und R_2 beide symmetrisch.
und wir erhalten $(y,x) \in R_1 \cap R_2$.	Wieder (2.2.3).
➥ Absatz	Der Beweis der Symmetrie ist abgeschlossen. ↽ Es ist nicht länger angenommen, dass $(x,y) \in R_1 \cap R_2$. ⟨②⟨ ↽ Die Buchstaben x, y sind nicht mehr gebunden (gleich den uns am Anfang des Schrittes vorgegebenen Elementen von M) und wieder frei verfügbar. ⟨①⟨
Die Transitiviät folgt genauso. *Wir formulieren die „genauso" ablaufende Argumentation in dieser Spalte aus:*	Im dritten Schritt gehen wir darauf los die Transitivität zu zeigen. Wieder eine Eigenschaft die mit einem Allquantor anfängt: $\left[\ \forall x, y, z \in M\colon ((x,y) \in R_1 \cap R_2 \wedge (y,z) \in R_1 \cap R_2) \Rightarrow (x,z) \in R_1 \cap R_2\ \right]$ ⇀ Seien uns beliebige, bis zum Ende dieses Beweisschrittes festgehaltene, Elemente $x, y, z \in M$ vorgegeben. ⟩①⟩
Seien $(x,y), (y,z) \in R_1 \cap R_2$.	Wir müssen zeigen, dass für die gegebenen Elemente x, y, z eine gewisse Implikation wahr ist. ⇀ Sei dazu, bis zum Ende dieses Schrittes, vorausgesetzt, dass $(x,y), (y,z) \in R_1 \cap R_2$. ⟩②⟩ Ziel ist zu zeigen, dass die Konklusion der gewünschten Implikation wahr ist.
Dann sind $(x,y), (y,z) \in R_1$ und auch $(x,y), (y,z) \in R_2$.	Das ist (2.2.3).
Also folgt $(x,z) \in R_1$ und auch $(x,z) \in R_2$,	Nach der Voraussetzung des Lemmas sind die Relationen R_1 und R_2 beide transitiv.

und wir erhalten $(x,z) \in R_1 \cap R_2$.	Wieder (2.2.3).
	Der Beweis der Transitivität ist abgeschlossen. ⇀ Es ist nicht länger angenommen, dass $(x,y),(y,z) \in R_1 \cap R_2$. ⟨②⟨ ⇀ Die Buchstaben x,y,z sind nicht mehr gebunden (gleich den uns am Anfang des Schrittes vorgegebenen Elementen von M) und wieder frei verfügbar. ⟨①⟨
□	Der Beweis des Lemmas ist abgeschlossen.

♧

J238. *2.2.7 Bemerkung.* Betrachtet man die obige Beweisanalyse, so sieht man, dass wir eigentlich mehr gezeigt haben als Lemma 2.2.6 behauptet. Nämlich haben wir die Gültigkeit jeder einzelner der Implikationen

$$(R_1 \text{ reflexiv } \wedge\ R_2 \text{ reflexiv}) \ \Rightarrow\ R_1 \cap R_2 \text{ reflexiv} \qquad (2.2.4)$$ J245

$$(R_1 \text{ symmetrisch } \wedge\ R_2 \text{ symmetrisch}) \ \Rightarrow\ R_1 \cap R_2 \text{ symmetrisch} \qquad (2.2.5)$$ J246

$$(R_1 \text{ transitiv } \wedge\ R_2 \text{ transitiv}) \ \Rightarrow\ R_1 \cap R_2 \text{ transitiv} \qquad (2.2.6)$$ J247

gezeigt. Es wäre also sicherlich formallogisch stärker, vielleicht didaktisch besser, und jedenfalls ohne zusätzlichen Aufwand möglich, Lemma 2.2.6 in zwei Aussagen aufzuteilen:

1. Ein Lemma: „Die drei oben genannten Implikationen sind wahr."
2. Ein Korollar: „Das was jetzt in Lemma 2.2.6 steht."

◊

➢ `Was ist ein Korollar:`

Ein *Korollar* ist eine mathematische Aussage deren Wahrheit behauptet wird. Die Wahrheit dieser Aussage muss bewiesen werden !

- In der Aussage dürfen nur Begriffe und Symbole vorkommen die bereits definiert wurden.
- Es muss unmissverständlich klargestellt sein wovon wir sprechen. Also

  ① was sind die Objekte mit denen wir arbeiten,
  ② was sind die Voraussetzungen die verlangt werden,
  ③ was ist die Behauptung die gemacht wird.

Das ist eigentlich genau das gleiche wie ein Lemma; warum dann eine andere Bezeichnung ?

Als *Korollar*, synonym auch *Folgerung*, bezeichnet man eine Aussage die unmittelbar durch Kombinieren von bereits bewiesenen Aussagen gezeigt werden kann. Meist werden die kombinierten Aussagen dabei solche sein, die man gerade vorher gezeigt hat.

♧ Warum ist nun die Aussage von Lemma 2.2.6 ein Korollar der Implikationen (2.2.4)–(2.2.6) ? Dazu betrachten wir wieder die Beweisanalyse von Lemma 2.2.6. Wir haben hier wiederholt Konjunktionsbeseitigung verwendet: unsere Voraussetzung ist, dass R_1 und R_2 *alle drei* Eigenschaften reflexiv, symmetrisch, transitiv, haben. Dann schließen wir, dass *insbesondere* R_1 und R_2 beide reflexiv (bzw., im zweiten Schritt symmetrisch, und im dritten Schritt transitiv) sind. Jetzt kommt die Implikation (2.2.4) (bzw., im zweiten Schritt (2.2.5), und im dritten Schritt (2.2.6)) ins Spiel, und wir schließen, dass $R_1 \cap R_2$ reflexiv (bzw., im zweiten Schritt symmetrisch, und im dritten Schritt transitiv) ist.

Also erhalten wir die Aussage von Lemma 2.2.6 tatsächlich ganz unmittelbar durch Kombinieren der Implikationen (2.2.4)–(2.2.6). Formallogisch ausgedrückt verwenden wir die Tautologie

$$\Big((A_1 \Rightarrow B_1) \wedge (A_2 \Rightarrow B_2) \wedge (A_3 \Rightarrow B_3)\Big) \Rightarrow \Big((A_1 \wedge A_2 \wedge A_3) \Rightarrow (B_1 \wedge B_2 \wedge B_3)\Big)$$

Diese kann man (mühsam) durch Erstellen einer Wahrheitstafel nachweisen. Man kann sie aber auch aus den im ersten Kapitel genannten Basisregeln herleiten. ♲

2.3 Partitionen

2.3.1 Motivation

Erzählung eines Bauern:

> Auf meinem Hof habe ich etliche Tiere, nämlich Schweine, Hühner, und Kühe. Diesen Sommer habe ich einen neuen Gehilfen bekommen, dem musste ich zu Beginn einmal den Arbeitsablauf erklären: als erstes *die Schweine* füttern, dann *die Kühe* melken, und erst dann *die Hühner* füttern und die Eier aus dem Gelege herausholen.

Die Gesamtheit aller vorhandenen Tiere am Hof wird hier offenbar in verschiedene Teilmengen zerlegt deren Individuen nicht unterschieden werden:

– *die Schweine*;

– *die Kühe*;

– *die Hühner*.

Es ist oft sinnvoll eine Grundmenge in Teile zu zerlegen. Einerseits weil die Individuen der einzelne Teile vielleicht gemeinsame Eigenschaften haben (zwei Kühe sind wohl verschiedene Persönlichkeiten, aber alle müssen gemolken werden). Andererseits weil die einzelnen – kleineren – Teile vielleicht einfacher zu behandeln oder zu verstehen sind.

Welche Eigenschaften sind es, die es ausmachen, dass man in „vernünftiger“ Weise von einer Zerlegung einer Grundgesamtheit sprechen kann ?

Sicherlich sollten die folgenden zwei Eigenschaften erfüllt sein (und diese sind im obigen Beispiel auch erfüllt):

(i) Jedes Objekt unserer Grundmenge ist in einer der Teilmengen enthalten.

(ii) Je zwei verschiedene der Teilmengen der Zerlegung sind disjunkt, d.h. haben keine gemeinsamen Elemente.

Ein weiterer Aspekt den wir hier hervorheben wollen ist, dass die Elemente einer Menge der Zerlegung in mancher Hinsicht als *gleich* betrachtet werden können.

2.3.2 Formalisierung

J204. **2.3.1 Definition.** Sei M eine Menge. Eine Teilmenge Q der Potenzmenge von M heißt eine *Partition von M*, wenn folgende Axiome gelten.

(Par1) Jedes Element von M ist in einem Element von Q enthalten.
$\Big[\ \forall x \in M\ \exists A \in Q\colon\ x \in A\ \Big]$

(Par2) Zwei Elemente von Q sind entweder gleich oder disjunkt.
$\Big[\ \forall A, B \in Q\colon\ A = B\ \vee\ A \cap B = \emptyset\ \Big]$

(Par3) Jedes Element von Q ist nichtleer.
$\Big[\ \forall A \in Q\colon\ A \neq \emptyset\ \Big]$

◊

Das Axiom **(Par1)** kann auch angeschrieben werden als

(Par1′) $M = \bigcup_{A \in Q} A$

Das Axiom **(Par2)** kann auch angeschrieben werden als

(Par2′) $\forall A, B \in Q\colon A \neq B \Rightarrow A \cap B = \emptyset$

Die Äquivalenz von **(Par2)** und **(Par2′)** geht auf die Tautologie $(P \vee Q) \Leftrightarrow (\neg P \Rightarrow Q)$ zurück.

♲ `Analyse von Definition` 2.3.1`:`

① Wir wissen was eine Menge, ihre Potenzmenge und Teilmengen sind. Wir wollen uns überlegen, dass die angeschriebenen Eigenschaften Sinn machen.

– *Für* **(Par1)***:* ein *Element A* von Q ist ein Element der Potenzmenge von M, also eine *Teilmenge* von M, und daher kann für ein *Element x* von M die Elementbeziehung $x \in A$ gelten (oder eben nicht).
– *Für* **(Par2)***:* Elemente A, B von Q sind Mengen (nämlich Teilmengen von M), und daher wissen wir was deren Gleichheit bedeutet, sowie was deren Durchschnitt ist.

② Dieser Begriff modelliert eine Gesamtheit in kleinere Teile aufzuteilen. Wir werden später (in Satz 2.3.3 und Satz 2.5.14) sehen, dass Partitionen im Wesentlichen das selbe sind wie Äquivalenzrelationen; tatsächlich haben wir hier zwei Sichtweisen auf ein und das selbe Ding. Alles was wir für diese gesagt haben, trifft also auch hier zu.

♲

Ist Q eine Partition einer Menge M, $N \in Q$, und $x \in N$, so sagt man x ist ein *Repräsentant* der Klasse N der Partition Q.

2.3.3 Äquivalenzklassen und Partitionen

J223 **2.3.2 Definition.** Sei M eine Menge und R eine Äquivalenzrelation von M. Für $x \in M$ bezeichne

$$[x]_R = \{y \in M :\ (x, y) \in R\}. \tag{2.3.1}$$ J227

Diese Menge heißt die *Äquivalenzklasse von x bzgl. R*. Die Menge $\{[x]_R :\ x \in M\}$ aller Äquivalenzklassen bzgl. R heißt die *Faktormenge* von M nach R, und man schreibt für sie auch M/R. Die Funktion von M nach M/R, welche einem Element $x \in M$ seine Äquivalenzklasse zuordnet, heißt die *kanonische Projektion*. ◊

J222. **2.3.3 Satz.** *Sei M eine Menge.*

(1) *Ist R eine Äquivalenzrelation auf M, so ist die Menge*

$$Q = \{[x]_R :\ x \in M\} \tag{2.3.2}$$ J224

aller Äquivalenzklassen von R eine Partition von M.

(2) *Ist Q eine Partition von M, so ist die Relation*

$$R = \{(x, y) :\ \exists A \in Q : x \in A \wedge y \in A\} \tag{2.3.3}$$ J228

eine Äquivalenzrelation auf M.

➢ `Was ist ein Satz:`

Ein *Satz* ist eine mathematische Aussage deren Wahrheit behauptet wird. Die Wahrheit dieser Aussage muss bewiesen werden !

– In der Aussage dürfen nur Begriffe und Symbole vorkommen die bereits definiert wurden.
– Es muss unmissverständlich klargestellt sein wovon wir sprechen. Also

① was sind die Objekte mit denen wir arbeiten,
② was sind die Voraussetzungen die verlangt werden,
③ was ist die Behauptung die gemacht wird.

Das ist eigentlich genau das gleiche wie ein Lemma oder Korollar; warum dann eine andere Bezeichnung ?

Eine Aussage als *Satz*, synonym bzw. in der englischsprachigen Literatur (was eigentlich der Standard ist!) auch *Theorem*, zu bezeichnen kann mehrere Gründe haben. Einerseits kann die Aussage für das Folgende sehr wichtig sein, und/oder ein mächtiges Werkzeug sein, und/oder sehr schön sein. Andererseits verwendet man diese Bezeichnung auch oft wenn der Beweis essentielle neue Ideen und/oder Konstruktionen erfordert, und/oder einen komplexen Weg geht (sprich, alles andere als Standard oder trivial ist). Meistens werden beide angeführten Gründe zusammenspielen.

Aus dem Satz schließen wir: Äquivalenzrelationen und Partitionen sind verschiedene Methoden, um dasselbe dahinter stehende Konzept zu beschreiben, nämlich dass gewisse Elemente einer Menge einander näher stehen als andere.

Manchmal ist es praktischer, mit Äquivalenzrelationen zu arbeiten: zum Beispiel ist der Durchschnitt zweier Äquivalenzrelationen wiederum eine Äquivalenzrelatio — zwei Autos sind sich in einem Sinn (R_1) ähnlich, wenn sie die gleiche Marke haben, in einem anderen (R_2), wenn sie die gleiche Farbe haben, und in einem dritten Sinn ($R_1 \cap R_2$), wenn sie sowohl die gleiche Marke als auch die gleiche Marke haben).

Manchmal ist es aber praktischer, die Äquivalenzrelation zu verwenden; zum Beispiel lässt sich eine Partition auf einer kleinen endlichen Menge leichter und übersichtlicher darstellen als die zugehörige Äquivalenzrelation.

Der folgende Beweis von Satz 2.3.3 enthält wohl alle wesentlichen Argumente, ist aber (absichtlich!) ein bisschen kurz gehalten. Dies soll einerseits dazu dienen, das selbstständige Denken anzuregen, andererseits alles hinterfragen zu müssen, und nicht durch suggestive Ausführungen dazu verleitet zu werden, etwas zu *glauben*, was man *argumentieren* muss.

Beweis von Satz 2.3.3. Sei R eine Äquivalenzrelation auf M. Wegen $x \in [x]_R$ gilt $M = \bigcup_{x \in M} [x]_R = \bigcup_{A \in Q} A$. Seien nun $x, y \in M$ mit $[x]_R \cap [y]_R \neq \emptyset$. Wähle $z \in [x]_R \cap [y]_R$, dann gilt also $(x, z) \in R$ und $(y, z) \in R$. Es folgt, dass $(x, y) \in R$. Für beliebiges $z \in [y]_R$, d.h. z mit $(y, z) \in R$, folgt nun auch $(x, z) \in R$ also $z \in [x]_R$. Genauso erhält man $[x]_R \subseteq [y]_R$.

Sei Q eine Partition von M. Zuerst gilt wegen **(Par1)** stets $(x, x) \in R$. Offensichtlich ist R symmetrisch. Seien $(x, y) \in R$ und $(y, z) \in R$. Wähle $A, B \in Q$ mit $x, y \in A$ und $y, z \in B$. Dann ist $y \in A \cap B$, und damit $A = B$, woraus wir $(x, z) \in R$ erhalten. □

`Analyse des Beweises von Satz 2.3.3:`
Wollen wir den Beweis filetieren.

Sei R eine Äquivalenzrelation auf M.

Wir gehen darauf los Teil (1) des Satzes zu beweisen.
⇨ Sei uns eine beliebige, bis zum Ende dieses Beweisteiles festgehaltene, Äquivalenzrelation R vorgegeben. ⟩①⟩
Wir müssen zeigen, dass die in (2.3.2) definiert Menge Q eine Partition ist. Offenbar ist $Q \subseteq \mathcal{P}(M)$, also müssen wir noch zeigen, dass Q die Eigenschaften **(Par1)**, **(Par2)**, und **(Par3)** besitzt.

Wir gehen darauf los **(Par1)** zu zeigen.

$\left[\ \forall x \in M\ \exists A \in Q\colon\ x \in A\ \right]$

Dazu erinnern wir uns, dass **(Par1)**⇔**(Par1′)**, und gehen darauf los **(Par1′)** zu zeigen.

$\left[\ M = \bigcup_{A \in Q} A\ \right]$

⇨ Sei uns ein beliebiges, bis zum Ende des Beweises von **(Par1′)** festgehaltenes, Element $x \in M$ vorgegeben. ⟩②⟩

Wegen $x \in [x]_R$

Da R reflexiv ist gilt $(x, x) \in R$. Nach Definition (2.3.1) der Äquivalenzklassen bzgl. R, haben wir $x \in [x]_R$.

---

gilt
$M = \bigcup_{x \in M} [x]_R = \bigcup_{A \in Q} A$.

Die erste Gleichheit folgt mit dem in der letzten Zeile Gezeigtem, denn

$$M = \bigcup_{x \in M} \{x\} \subseteq \bigcup_{x \in M} [x]_R \subseteq \bigcup_{x \in M} M = M.$$

Die zweite Gleichheit gilt, da die Partition Q nach ihrer Definition (2.3.2) genau aus den Mengen $[x]_R$ mit $x \in M$ besteht.

	Der Beweis von **(Par1′)** ist abgeschlossen. ↬ Der Buchstabe x ist nicht mehr gebunden (gleich dem uns am Anfang vorgegebenen Element von M) und wieder frei verfügbar. ⟨ ② ⟨ Der Beweis von **(Par1)** ist abgeschlossen.
	Wir gehen darauf los **(Par2)** zu zeigen. $\left[\ \forall A, B \in Q\colon\ A = B\ \vee\ A \cap B = \emptyset\ \right]$ Wir müssen also zeigen, dass zwei Elemente von Q disjunkt oder gleich sind. Die Elemente von Q sind genau die Äquivalenzklassen von R, also die Mengen der Gestalt $[x]_R$ wobei x die Menge M durchläuft. Wir müssen also zeigen, dass zwei Mengen *dieser Gestalt* disjunkt oder gleich sind.
Seien nun $x, y \in M$	⇀ Seien uns beliebige, bis zum Ende dieses Beweisschrittes festgehaltene, Elemente $x, y \in M$ vorgegeben. ⟩ ② ⟩
	Wir müssen zeigen, dass $$[x]_R \cap [y]_R = \emptyset\ \vee\ [x]_R = [y]_R. \qquad (2.3.4)$$ J226 Dazu machen wir eine Fallunterscheidung. ✾ *Fall 1;* $[x]_R \cap [y]_R = \emptyset$. ✾ *Fall 2;* $[x]_R \cap [y]_R \neq \emptyset$. Offensichtlich decken diese beiden Fälle alle Möglichkeiten ab (Fall 2 ist genau die Negation von Fall 1).
	Wir betrachten Fall 1. ⇀ Sei dazu, bis zum Ende der Betrachtung dieses Falles, vorausgesetzt, dass $[x]_R \cap [y]_R = \emptyset$. ⟩ ③ ⟩ Die Disjunktion (2.3.4) ist wahr, da die erste in ihr vorkommende Aussage wegen ③ wahr ist. Die Betrachtung von Fall 1 ist abgeschlossen. ↬ Es ist nicht länger angenommen, dass $[x]_R \cap [y]_R = \emptyset$. ⟨ ③ ⟨
mit $[x]_R \cap [y]_R \neq \emptyset$.	Wir betrachten Fall 2. ⇀ Sei dazu, bis zum Ende der Betrachtung dieses Falles, vorausgesetzt, dass $[x]_R \cap [y]_R \neq \emptyset$. ⟩ ③ ⟩
	Wir gehen in einem ersten Schritt darauf los zu zeigen, dass $(x, y) \in R$.
Wähle $z \in [x]_R \cap [y]_R$,	Da der Durchschnitt $[x]_R \cap [y]_R$ nicht leer ist, existieren Elemente in diesem Durchschnitt. ⇀ Wir wählen ein Element z aus dem Durchschnitt, und halten es bis zum Ende dieses Schrittes fest. ⟩ ④ ⟩
dann gilt also $(x, z) \in R$ und $(y, z) \in R$.	Das ist die Definition (2.3.1) der Äquivalenzklassen bzgl. R.

Es folgt, dass $(x, y) \in R$.	Da R symmetrisch ist erhalten wir, dass $(z, y) \in R$, und da R transitiv ist folgt nun $(x, y) \in R$. Der Beweis des ersten Schrittes ist abgeschlossen. ⇀ Der Buchstabe z ist nicht mehr gebunden (gleich dem von uns gewählten Element des Durchschnittes $[x]_R \cap [y]_R$) und wieder frei verfügbar. ⟨ ④ ⟨
	Unsere einzige Chance (2.3.4) zu bekommen, ist $[x]_R = [y]_R$ zu zeigen; wir sind ja im Fall $[x]_R \cap [y]_R \neq \emptyset$, also ist die erste Aussage der Disjunktion (2.3.4) nicht wahr. Nun ist die Gleichheit zweier Mengen äquivalent zur Gültigkeit beider wechselseitiger Inklusionen (hier steckt wieder die Formel (2.2.1) dahinter): für zwei beliebige Menge M_1, M_2 gilt $$(A_1 = A_2) \Leftrightarrow \Big((A_1 \subseteq A_2) \wedge (A_1 \supseteq A_2) \Big)$$
Für beliebiges $z \in [y]_R$,	Wir gehen im zweiten Schritt darauf los die Inklusion $[y]_R \subseteq [x]_R$ zu zeigen $\big[\ \forall z\colon z \in [y]_R \Rightarrow z \in [x]_R\ \big]$ ⇁ Sei uns ein beliebiges, bis zum Ende dieses Schrittes festgehaltenes, Element $z \in [y]_R$ vorgegeben. ⟩ ④ ⟩
d.h. z mit $(y, z) \in R$,	Wieder Definition (2.3.1) der Äquivalenzklassen bzgl. R.
folgt nun auch $(x, z) \in R$	Im ersten Schritt hatten wir gezeigt, dass $(x, y) \in R$; diese Aussage dürfen wir also schon verwenden. Beachte hier, dass die im Beweis des ersten Schrittes gültigen Bezeichnungen und Voraussetzungen (das sind ①, ②, ③) immer noch gültig sind. Da R transitiv ist, erhalten wir $(x, z) \in R$.
also $z \in [x]_R$.	Wieder Definition (2.3.1) der Äquivalenzklassen von R. Der Beweis des zweiten Schrittes ist abgeschlossen. ⇀ Der Buchstabe z ist nicht mehr gebunden (gleich dem uns vorgegebenen Element der Klasse $[y]_R$) und wieder frei verfügbar. ⟨ ④ ⟨
Genauso erhält man $[x]_R \subseteq [y]_R$. *Wir formulieren die „genauso“ ablaufende Argumentation in dieser Spalte aus:*	
Für beliebiges $z \in [x]_R$,	Im dritten Schritt gehen wir darauf los die Inklusion $[x]_R \subseteq [y]_R$ zu zeigen. $\big[\ \forall z\colon z \in [x]_R \Rightarrow z \in [y]_R\ \big]$ ⇁ Sei uns ein beliebiges, bis zum Ende dieses Schrittes festgehaltenes, Element $z \in [x]_R$ vorgegeben. ⟩ ④ ⟩

d.h. z mit $(x, z) \in R$,	Wieder Definition (2.3.1) der Äquivalenzklassen bzgl. R.
folgt nun auch $(y, z) \in R$	Im ersten Schritt hatten wir gezeigt, dass $(x, y) \in R$. Da R symmetrisch ist, gilt auch $(y, x) \in R$. Da R transitiv ist, erhalten wir $(y, z) \in R$.
also $z \in [y]_R$.	Wieder die Definition (2.3.1) der Äquivalenzklassen von R. Der Beweis des dritten Schrittes ist abgeschlossen. ↬ Der Buchstabe z ist nicht mehr gebunden (gleich dem uns vorgegebenen Element der Klasse $[x]_R$) und wieder frei verfügbar. ⟨④⟨
	Wir haben beide Inklusionen $[y]_R \subseteq [x]_R$ und $[x]_R \subseteq [y]_R$, und damit die Gleichheit dieser Mengen gezeigt, also gilt (2.3.4) auch im Fall 2. Die Betrachtung von Fall 2 ist abgeschlossen. ↬ Es ist nicht länger angenommen, dass $[x]_R \cap [y]_R \neq \emptyset$. ⟨③⟨ Der Beweis von **(Par2)** ist abgeschlossen. ↬ Die Buchstaben x, y sind nicht mehr gebunden (gleich den uns vorgegebenen Elementen von M) und wieder frei verfügbar. ⟨②⟨
	Wir gehen darauf los **(Par3)** zu zeigen. $\left[\forall A \in Q\colon A \neq \emptyset \right]$ Die Elemente von Q sind genau die Äquivalenzklassen von R, also die Mengen der Gestalt $[x]_R$ wobei x die Menge M durchläuft. Wir müssen also zeigen, dass jede Menge *dieser Gestalt* nichtleer ist.
	⇀ Sei uns ein beliebiges, bis zum Ende dieses Schrittes festgehaltenes, Element $x \in M$ vorgegeben. ⟩②⟩ Nun haben wir im Beweis von **(Par1)** gezeigt, dass $x \in [x]_R$. Insbesondere ist also $[x]_R \neq \emptyset$. ↬ Der Buchstabe x ist nicht mehr gebunden (gleich dem uns am Anfang vorgegebenen Element von M) und wieder frei verfügbar. ⟨②⟨
➥ Absatz	Der Beweis von Teil (1) des Satzes ist abgeschlossen. ↬ Der Buchstabe R ist nicht mehr gebunden (gleich der uns am Anfang vorgegebenen Äquivalenzrelation) und wieder frei verfügbar. ⟨①⟨
Sei Q eine Partition von M.	Wir gehen darauf los Teil (2) des Satzes zu beweisen. ⇀ Sei uns eine beliebige, bis zum Ende dieses Beweisteiles festgehaltene, Partition Q vorgegeben. ⟩①⟩
	Wir müssen zeigen, dass die in (2.3.3) definierte Menge R eine Äquivalenzrelation ist. Offenbar ist $R \subseteq M \times M$, also müssen wir noch zeigen, dass R reflexiv, symmetrisch und transitiv ist.
	Im ersten Schritt gehen wir darauf los Reflexivität zu zeigen. $\left[\forall x \in M\colon (x, x) \in R \right]$ ⇀ Sei uns ein beliebiges, bis zum Ende dieses Schrittes festgehaltenes, Element $x \in M$ vorgegeben. ⟩②⟩

Zuerst gilt wegen **(Par1)** stets $(x, x) \in R$.	Da Q das Axiom **(Par1)** erfüllt, existiert ein Element $A \in Q$ mit $x \in A$. Damit ist die Konjunktion $x \in A \wedge x \in A$ wahr, und wir erhalten nach der Definition (2.3.3) von R das $(x, x) \in R$.
	Der Beweis der Reflexivität ist abgeschlossen. ↬ Der Buchstabe x ist nicht mehr gebunden (gleich dem uns am Anfang des Schrittes vorgegebenen Element) und wieder frei verfügbar. ⟨ ② ⟨
	Im zweiten Schritt gehen wir darauf los Symmetrie zu zeigen.
Offensichtlich ist R symmetrisch.	Warum ist es „offensichtlich“ dass R symmetrisch ist ? *Vor dem weiterlesen zuerst selbst nachdenken!* Für zwei beliebige Aussage A_1, A_2 gilt $$(A_1 \wedge A_2) \;\Leftrightarrow\; (A_2 \wedge A_1).$$ Also haben wir $$\big(\exists A \in Q\colon x \in A \wedge y \in A\big) \;\Leftrightarrow\; \big(\exists A \in Q\colon y \in A \wedge x \in A\big)$$ Nach der Definition (2.3.3) von R ist die linke Seite äquivalent zu $(x, y) \in R$ und die rechte Seite äquivalent zu $(y, x) \in R$. Der Beweis der Symmetrie ist abgeschlossen.
	Im dritten Schritt gehen wir darauf los Transitivität zu zeigen. $\big[\ \forall x, y, z \in M\colon (\,(x, y) \in R \wedge (y, z) \in R\,) \;\Rightarrow\; (x, z) \in R\ \big]$ ⇀ Seien uns beliebige, bis zum Ende dieses Beweisschrittes festgehaltene, Elemente $x, y, z \in M$ vorgegeben. ⟩ ② ⟩
Seien $(x, y) \in R$ und $(y, z) \in R$.	Wir müssen eine Implikation zeigen, setzen also ihre Prämisse als wahr voraus und versuchen ihre Konklusion herzuleiten. ⇀ Sei dazu, bis zum Ende dieses Schrittes, vorausgesetzt, dass $(x, y), (y, z) \in R$. ⟩ ③ ⟩
Wähle $A, B \in Q$ mit $x, y \in A$ und $y, z \in B$.	Nach der Definition (2.3.3) von R existiert $A \in Q$ mit $x \in A \wedge y \in A$ (wegen $(x, y) \in R$), und $B \in Q$ mit $y \in B \wedge z \in B$ (wegen $(y, z) \in R$). ⇀ Wir wählen Elemente A bzw. B aus Q mit der entsprechenden genannten Eigenschaft und halten sie bis zum Ende dieses Schrittes fest. ⟩ ④ ⟩
Dann ist $y \in A \cap B$,	Es ist $y \in A$ wegen ersterer Konjunktion $(x \in A \wedge y \in A)$, und $y \in B$ wegen zweiterer $(y \in B \wedge z \in B)$.
und damit $A = B$	Der Durchschnitt $A \cap B$ enthält das Element y, ist also nicht leer. Nach **(Par2)** muss daher $A = B$ sein.
woraus wir $(x, z) \in R$ erhalten.	Es gilt $x \in A$ wegen der Konjunktion $x \in A \wedge y \in A$, und $z \in A$ wegen der Konjunktion $y \in B \wedge z \in B$ und der eben gezeigten Tatsache $A = B$. Nach Definition (2.3.3) von R haben wir $(x, z) \in R$.

	Der Beweis der Transitivität ist abgeschlossen. ⇜ Die Buchstaben A, B sind nicht mehr gebunden (gleich den von uns gewählten Elementen von Q) und wieder frei verfügbar. ⟨④⟨ ⇜ Es ist nicht länger angenommen, dass $(x, y), (y, z) \in R$. ⟨③⟨ ⇜ Die Buchstaben x, y, z sind nicht mehr gebunden (gleich den uns am Anfang des Schrittes vorgegebenen Elementen von M) und wieder frei verfügbar. ⟨②⟨
	Der Beweis von Teil (2) des Satzes ist abgeschlossen. ⇜ Der Buchstabe Q ist nicht mehr gebunden (gleich der uns am Anfang vorgegebenen Partition) und wieder frei verfügbar. ⟨①⟨
□	Der Beweis des Satzes ist abgeschlossen.

♲

♲ `Analyse einer Vorgangsweise im Beweis von Satz 2.3.3:`

In diesem Beweis, konkret in dem Schritt wo **(Par2)** gezeigt wird, verwenden wir ein Argument das wir als *Fallunterscheidung* bezeichnet haben. Wir haben zu zeigen, dass für alle x, y die gewünschte Aussage gilt, und sagen:

> Dazu machen wir eine Fallunterscheidung: Fall 1 ist $[x]_R \cap [y]_R = \emptyset$, und Fall 2 ist $[x]_R \cap [y]_R \neq \emptyset$.

Danach zeigen wir, dass die gewünschte Aussage in beiden Fällen wahr ist, und behaupten, dass damit der Beweis (dieser Aussage) abgeschlossen ist. Nun, die beiden genannten Fälle sind Negation voneinander. Wegen dem Prinzip des ausgeschlossenen Dritten decken sie gemeinsam daher alle Möglichkeiten ab, und damit gilt die gewünschte Aussage tatsächlich für alle x, y. Man kann ja sagen:

> Sind x, y beliebig, dann *muss* einer der beiden Fälle eintreten, und damit, wie wir gezeigt haben, die gewünschte Aussage für x, y gelten.

Bemerke, dass die in diesem Beweis auftretenden beiden Fälle sich wegen dem Satz vom Widerspruch sogar gegenseitig ausschließen. Das ist aber für die Beweismethode irrelevant.

Natürlich kann man auch in mehr als zwei Fälle unterscheiden. Einzig wesentlich ist, dass die Fälle gemeinsam alle potentiellen Möglichkeiten abdecken.

Beweise mit Fallunterscheidungen verwendet man einerseits oft wenn man zuerst einmal Trivialfälle loswerden möchte, um sich dann auf die essentielle/generische Situation zu konzentrieren. Genau das ist im hier diskutierten Beweis passiert. Fall 1, dass $[x]_R \cap [y]_R = \emptyset$, ist trivial. Dann gilt **(Par2)** ja ganz unmittelbar durch Disjunktionseinführung. Fall 2 dagegen war ziemlich anstrengend.

Andererseits kommt es auch oft vor, dass man Argumente oder Rechnungen macht die meistens – aber eben nicht immer – zum Ziel führen. Dann ist es auch sinnvoll, eine Fallunterscheidung zu machen: Fall 1 ist, dass alles brav ist und die Rechnung funktioniert; der ist erledigt. Fall 2 ist der Rest, sprich die Sonderfälle, und das ist eben noch gesondert zu betrachten.

Zum Beispiel könnte man im Laufe eines Beweises eine Rechnung durchführen, wo man an einer Stelle durch $x - y$ dividiert; das geht aber nur dann, wenn $x \neq y$ gilt. Hier empfiehlt sich eine Fallunterscheidung: Fall 1: $x \neq y$, dann kann der Beweis wie geplant durchgeführt werden; Fall 2: $x = y$, dann muss man einen anderen Beweis finden (der oft ohnehin leichter ist, weil man sich nun nur mehr mit einer Variable herumschlagen muss). ♲

➢ `Ein Beweis mit Fallunterscheidungen funktioniert so:`

Wir gehen davon aus, dass die zu beweisende Aussage (in obigem Beweis ist das **(Par2)**) bereits in dem Sinne analysiert wurde, dass es klar ist, was die zur Verfügung stehenden Voraussetzungen sind, und was die zu zeigende Behauptung ist. Wollen wir diese Voraussetzungen als A und diese Behauptung als B bezeichnen.

① Man findet eine Familie von Eigenschaften, sagen wir $C_1, C_2, \ldots, C_n$, sodass $A \Rightarrow A \wedge (C_1 \vee C_2 \vee \ldots \vee C_n)$ gilt. Das sind die Fälle in die unterschieden wird, und die Implikation besagt, dass sie gemeinsam alle potentiell auftretenden Möglichkeiten abdecken.

② Nun zeigt man die Wahrheit der Implikationen

$$(A \wedge C_1) \Rightarrow B\,,\ (A \wedge C_2) \Rightarrow B\,,\ \ldots\,,\ (A \wedge C_n) \Rightarrow B\,,$$

sprich, dass die gewünschte Behauptung in jedem der Fälle gilt.

Was nun folgt, ist die formallogische Argumentation, dass damit der Beweis tatsächlich abgeschlossen ist.

❸ Man verwendet die Tautologie: für Aussagen $D_1, \ldots, D_n$ und E ist

$$\big[(D_1 \Rightarrow E) \wedge \ldots \wedge (D_n \Rightarrow E)\big] \;\Leftrightarrow\; \big[(D_1 \vee \ldots \vee D_n) \Rightarrow E\big] \qquad (2.3.5)$$ J250

Diese erhält man durch eine wiederholte Anwendung der folgenden, leicht zu überprüfenden Tautologie: für Aussagen D, E gilt

$$(D \Rightarrow E) \wedge (D \vee E) \;\Leftrightarrow\; E$$

❹ Schließlich leitet man, unter Verwendung des in ① und ② bewiesenen, des Distributivitätssatzes, und oben benannter Tautologie (2.3.5), ab dass

$$A \;\Rightarrow\; A \wedge (C_1 \vee C_2 \vee \ldots \vee C_n) \;\Leftrightarrow\; (A \wedge C_1) \vee \ldots \vee (A \wedge C_n) \;\Rightarrow\; B$$

❺ Man wendet den modus ponendo ponens an, nachdem man aus der Wahrheit von A und der Wahrheit der Implikation $A \Rightarrow B$ auf die Wahrheit von B schließen kann.

2.4 Der Funktionsbegriff

ktionsbegriff

2.4.1 Motivation

Beziehungen sind insbesondere dann interessant, wenn jedem Objekt ein und nur ein Attribut zugeordnet werden kann.

Jedes Verkehrszeichen hat eine – und nur eine – Bedeutung.

2.4.2 Formalisierung

J202.

2.4.1 Definition. Seien M und N Mengen. Eine *Funktion von M nach N*, synonym *Abbildung*, ist eine Relation f zwischen M und N mit den folgenden beiden Eigenschaften.

(Fun1) Für jedes Element $x \in M$ existiert ein Element $y \in N$ mit $(x, y) \in f$.
$\big[\ \forall x \in M\ \exists y \in N\colon (x, y) \in f\ \big]$

(Fun2) Ist $x \in M$, $y_1, y_2 \in N$, und gilt $(x, y_1) \in f$ und $(x, y_2) \in f$, so folgt $y_1 = y_2$.
$\big[\ \forall x \in M\ \forall y_1, y_2 \in N\colon \big((x, y_1) \in f \wedge (x, y_2) \in f\big) \;\Rightarrow\; y_1 = y_2\ \big]$

Die Menge M heißt die *Definitionsmenge* (synonym auch *Definitionsbereich*) von f, und N heißt *Zielmenge* (synonym auch *Wertevorrat*) von f. Ist $M = N$, so spricht man auch von einer *Funktion von M in sich*. ◊

Manchmal werden Definitionsmenge und Zielmenge auch in die Notation mit hineinverpackt und man schreibt eine Funktion als ein Tripel (f, M, N), wobei dann $f \subseteq M \times N$ mit **(Fun1)**, **(Fun2)** ist.

Die beiden Axiome **(Fun1)** und **(Fun2)** kann man zusammenfassen zu:

(Fun) Für jedes $x \in M$ existiert genau ein $y \in N$ mit $(x, y) \in f$.
$\left[\forall x \in M \; \exists! \, y \in N \colon (x, y) \in f \right]$

♲ Um die Äquivalenz ((**Fun1**) $\wedge$ (**Fun2**)) $\Leftrightarrow$ (**Fun**) einzusehen, erinnere man sich was es heißt, dass *genau ein* Element existiert. ♲

J203. *2.4.2 Bemerkung.* Sei f eine Funktion von M nach N. Dann kann man f auffassen als eine Zuordnung die jedem Element x von M ein Element y aus N zuweist, nämlich jenes mit $(x, y) \in f$. Dementsprechend vewendet man meistens die – vielleicht intuitivere – Schreibweise $f(x) = y$ um auszudrücken, dass $(x, y) \in f$, und schreibt $f : M \to N$ um auszudrücken, dass f eine Funktion von M nach N ist.

Diese Sichtweise spiegelt sich bei der üblichen Schreibweise wieder, die man verwendet um konkrete Funktionen anzugeben. Zum Beispiel sieht man oft soetwas wie

$$f : \begin{cases} M & \to & N \\ x & \mapsto & \text{␣} \end{cases}$$

wobei an der Stelle von „␣“ ein gewisser, von x abhängiger konkreter Ausdruck steht. Diese Schreibweise sagt dann, dass f jene Funktion von M nach N ist, die dem Element x das durch den konkreten Ausdruck gegebene Element zuweist. Also: f ist die Menge aller Paare $(x, \text{␣})$ wo x die Menge M durchläuft und wo wieder „␣“ durch den konkreten Ausdruck zu ersetzen ist.

Manchmal ist es auch so, dass man von einer Funktion f als Zuordnung denkt, und f aufgefasst als Relation dann als den *Graphen* von f bezeichnet. ◊

Wieder ein paar Definitionen (von diversen Schreibweisen).

➢ `Was ist eine Bemerkung:`

Eine *Bemerkung* ist eine Tatsache/Argument/Idee o.ä., die aus dem umliegenden Text hervorgehoben werden soll. Dies kann dazu dienen etwas das im Folgenden eine wesentliche Rolle spielen wird herauszustreichen, oder zusätzliche Informationen, die das Rundherum ein bisschen erschließen, zu geben. Eine Bemerkung kann kleinere Definitionen enthalten, zum Beispiel von Schreibweisen o.ä., oder auch kleinere oder unmittelbar klare Aussagen (die dann natürlich eines Beweises bedürfen).

➢ `Definition einer Funktion durch Fallunterscheidung:`

Oft sieht man auch das eine Funktion nicht durch einen, sondern durch *mehrere verschiedene* konkrete Ausdrücke angegeben wird. Zum Beispiel sieht man sowetwas wie: $f : M \to N$ sei die Funktion definiert als

$$f(x) = \begin{cases} \text{␣}_1, & k \in M_1 \\ \text{␣}_2, & k \in M_2 \\ \text{␣}_3, & k \in M_3 \end{cases}$$

wobei an der Stelle von „␣_1“, „␣_2“, und „␣_3“, gewisse, von x abhängige, konkrete Ausdrücke stehen, und wobei M_1, M_2, und M_3, gewisse Teilmengen von M bezeichnen. Damit meint man nun jene Funktion die Werten x aus M_i gerade den entsprechenden konkreten Ausdruck ␣_i zuordnet. Um sicherzustellen, dass eine solche Vorschrift tatsächlich eine Funktion definiert, muss man überprüfen, dass

– $M = \bigcup_i M_i$,

– Für $x \in M_i \cap M_j$ stimmen die entsprechen ausgewerteten konkreten Ausrücke ␣_i und ␣_j überein.

Wenn man das Glück hat (oder die Mengen M_i so schlau gewählt hat), dass die Mengen M_i paarweise disjunkt sind (gemeint ist hier: für alle $i \neq j$ gilt $M_i \cap M_j = \emptyset$), dann ist die zweite Bedingung automatisch erfüllt, weil es ja keine $x \in M_i \cap M_j$ gibt, die uns Schwierigkeiten machen könnten.

J232. *2.4.3 Bemerkung.* Das in §1.2.3 unter dem Punkt (12) beschriebene Auswahlverfahren läßt sich auch mit Hilfe des Funktionsbegriffs formulieren.

(AC) Sei C eine Menge, deren Elemente nichtleere Mengen sind. Dann existiert eine Funktion $f : C \to \bigcup_{M\in C} M$ mit der Eigenschaft

$$\forall M \in C\colon\ f(M) \in M. \tag{2.4.1}$$ J231

Eine äquivalente Formulierung, die oft praktischer anzuwenden ist, lautet wie folgt:

(AC′) Sei I eine Menge, und sei für jedes $i \in I$ eine nichtleere Menge A_i gegeben. Dann existiert eine Funktion $f : I \to \bigcup_{i\in I} A_i$ mit der Eigenschaft

$$\forall i \in I\colon\ f(i) \in A_i. \tag{2.4.2}$$ J233

Man bezeichnet Funktionen mit der Eigenschaft (2.4.1) bzw. (2.4.2) auch als *Auswahlfunktionen*. ◊

2.4.3 Komposition, Urbild, und Einschränkung

Wir wollen nun ein paar Konstruktionen studieren, wie man in natürlicher Weise aus gegebenen Funktionen weitere Funktionen bauen kann.

J210. **2.4.4 Definition.** Seien M, N, L Mengen, und seien $f : M \to N$ und $g : N \to L$ Funktionen. Dann ist die *Hintereinanderausführung von f und g* die Funktion

$$g \circ f : \left\{ \begin{array}{ccc} M & \to & L \\ x & \mapsto & g(f(x)) \end{array} \right.$$

Synonym bezeichnet man $g \circ f$ auch als die Funktion *g nach f*, oder die *Verknüpfung von f und g*, oder die *Komposition von f und g*.

ACHTUNG: Bei der Komposition $g \circ f$ wird „zuerst“ f ausgeführt, erst „danach“ g; genauer: um den Wert von $g \circ f$ an der Stelle x zu berechnen, wird zuerst $f(x)$ bestimmt, und dann auf diesen Wert die Funktion g angewendet.

Sei M eine Menge. Dann ist die *Identität auf M* die Funktion

$$\mathrm{id}_M : \left\{ \begin{array}{ccc} M & \to & M \\ x & \mapsto & x \end{array} \right.$$

◊

Der erste Teil dieser Definition bedarf einer Rechtfertigung, nämlich dass durch den in dieser Definition verwendeten Ausdruck $x \mapsto g(f(x))$ tatsächlich eine Funktion auf ganz M wohldefiniert ist.

Das ist der Fall, denn: Für jedes $x \in M$ können wir das eindeutige Element $f(x) \in N$ betrachtet. Für dieses nehmen wir dann, das ebenfalls eindeutige, Element $g(f(x)) \in L$. Also haben wir tatsächlich *jedem* Element der Definitionsbereiches M in *eindeutiger* Weise ein Element der Zielmenge L zugeordnet.

➢ `Was heißt wohldefiniert:`

Man verwendet den Terminus *wohldefiniert* oft im Zusammenhang mit einer Funktionsdefinition die einer Rechtfertigung in obigem oder ähnlichen Sinne bedarf. Dabei bezieht sich *wohldefiniert* oft darauf, dass jedem Element x dem man etwas zuordnen kann, ein *eindeutiges* Element zugeordnet wird (sprich, das **(Fun2)** gilt). Manchmal meint *wohldefiniert* aber auch das wirklich *jedem* Element ein *eindeutiges* Element zugeordnet wird (sprich, dass **(Fun1)** und **(Fun2)** beide gelten).

Dafür, dass man tatsächlich *jedem* Element des Definitionsbereiches etwas zuordnen kann (sprich, dass **(Fun1)** gilt), sagt man manchmal auch, dass die Funktion *überall definiert* ist. Diese Redewendung ist streng genommen sinnlos, denn eine Funktion hat einen – und nur einen – Definitionsbereich, und diesen Elementen ordnet sie etwas zu; Punkt. Der Ursprung der Phrase ist wohl historisch. Man stelle sich die folgende Situation vor: Man hat eine große Menge X, eine Teilmenge M von X, und man hat ein Prozedere mit dem man Elementen aus M irgendetwas (sagen wir Elemente einer Menge N) zuordnen kann. Dann hat man also eine Funktion von M nach N. Da man nun aber X als Grundmenge betrachtet, sagt man diese Funktion ist nicht überall definiert, sondern „nur“ auf M. Sollte nun $M = X$ sein, sagt man sie ist überall definiert.

♧ Wir wollen drei Situationen herausstellen, wo man Wohldefiniertheit überprüfen muss, die häufig vorkommen.

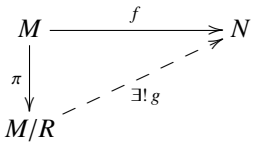
– Als erstes die oben schon betrachtete Definition einer Funktion durch Fallunterscheidung. Gelten die beiden dort genannten Punkte, so sagt man eine Funktion ist durch die Fallunterscheidung wohldefiniert. Offenbar entspricht dabei der erste der Punkte dem Axiom **(Fun1)** (überall definiert), und der zweite dem Axiom **(Fun2)** (wohldefiniert).

– Seien M, N Mengen, f eine Funktion von M nach N, und R eine Äquivalenzrelation auf M. Man sagt, dass durch die Vorschrift $g([x]_R) := f(x)$ eine Funktion g von der Faktormenge M/R nach N wohldefiniert ist, wenn $\forall x, y \in M : (xRy \Rightarrow f(x) = f(y))$ gilt.

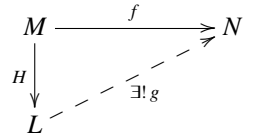
  Beachten Sie, dass Sie beim Beweis dieser Implikation nicht die Notation $g(\ )$ verwenden dürfen, weil Sie ja noch nicht wissen, ob diese Definition tatsächlich eine Funktion g liefert.

  Oft sagt man in dieser Situation auch, dass die Definition von g *von der Wahl des Repräsentanten unabhängig* ist, oder dass f eine *representantenweise definierte Funktion g von M/R nach N induziert*.

  Tatsächlich, wenn oben genannte Eigenschaft erfüllt ist, gibt es genau eine Funktion g von M/R nach N mit der Eigenschaft, dass $f = g \circ \pi$ wobei $\pi : M \to M/R$ die kanonische Projektion bezeichnet.



– Eine etwas allgemeinere, aber analoge, Situation zu der im vorigen Punkt besprochen ist die Folgende: Seien M, N, L Mengen, f Funktion von M nach N, und H Funktion von M nach L. Wenn gilt, dass H surjektiv ist und $\forall x, y \in M$: $H(x) = H(y) \Rightarrow f(x) = f(y)$, dann existiert eine eindeutige Funktion g von L nach N mit $f = g \circ H$.



  Man sagt in dieser Situation, dass eine Funktion g von L nach N durch die Vorschrift $f = g \circ H$ *wohldefiniert* ist.

♧

J239. *2.4.5 Bemerkung.* Wie sieht die Definition der Komposition $g \circ f$ ohne Verwendung der Schreibweise „$f(\_)$“ aus ? Wir behaupten, dass

$$g \circ f = \{(x, z) \in M \times L : \ \exists y \in N : (x, y) \in f \wedge (y, z) \in g\}.$$

Um die Inklusion „$\subseteq$“ zu sehen, erinneren wir uns daran, dass $f(x)$ jenes eindeutige Element $y \in N$ bezeichnet, das $(x, y) \in f$ erfüllt. Weiters ist $g(f(x))$ (was ja gleich $g(y)$ ist) jenes eindeutige Element $z \in L$, das $(y, z) \in g$ erfüllt. Wir haben also ein „verbindendes mittleres“ Element $y \in N$ gefunden. Für die umgekehrte Inklusion, sei angenommen, dass $(x, y) \in f$ und $(y, z) \in g$. Das bedeutet, unter Verwendung der Schreibweise „$f(\_)$“, dass $y = f(x)$ und $z = g(y) = g(f(x)) = (g \circ f)(x)$.

Auch betreffend der Identität id_M wollen wir eine Anmerkung machen. Diese Funktion ist uns nämlich schon einmal begegnet: es ist $\mathrm{id}_M = \Delta_M$. Warum führt man zwei Namen für das selbe Ding ein ? Die ursprüngliche Bezeichnung Δ_M geht auf die tatsächliche Definition der Funktion als Teilmenge von $M \times M$ zurück, dagegen bezieht sich die Bezeichnung id_M auf die Vorstellung einer Funktion als Zuordnung. ◊

J211. **2.4.6 Proposition.** *Es gelten die folgenden beiden Aussagen.*

(1) *Seien M, N, L, K Mengen, und $f : M \to N$, $g : N \to L$, $h : L \to K$ Funktionen. Dann ist*

$$(h \circ g) \circ f = h \circ (g \circ f). \tag{2.4.3}$$ J213

$$\left[\ \forall f : M \to N, g : N \to L, h : L \to K \colon (h \circ g) \circ f = h \circ (g \circ f)\ \right]$$

(2) *Seien M, N, L Mengen, und $f : M \to N$, $g : N \to L$ Funktionen. Dann ist*

$$\mathrm{id}_N \circ f = f, \quad g \circ \mathrm{id}_N = g. \tag{2.4.4}$$ J215

$$\left[\ \forall f : M \to N, g : N \to L \colon\ \mathrm{id}_N \circ f = f \ \wedge\ g \circ \mathrm{id}_N = g\ \right]$$

Die Gleichheit (2.4.3) heißt das *Assoziativgesetz* für die Operation $\circ$. Für die Tatsache, dass die Aussage in (1) wahr ist, sagt man auch, dass die Operation $\circ$ *assoziativ* ist.

➢ `Was ist eine Proposition:`

Eine *Proposition* ist eine mathematische Aussage, deren Wahrheit behauptet wird. Die Wahrheit dieser Aussage muss bewiesen werden !

- In der Aussage dürfen nur Begriffe und Symbole vorkommen die bereits definiert wurden.
- Es muss unmissverständlich klargestellt sein wovon wir sprechen. Also

  ① was sind die Objekte mit denen wir arbeiten,
  ② was sind die Voraussetzungen die verlangt werden,
  ③ was ist die Behauptung die gemacht wird.

Das ist eigentlich genau das gleiche wie ein Lemma, ein Korollar, oder ein Satz; warum dann eine andere Bezeichnung ?

Eine *Proposition* ist etwas zwischen Lemma und Satz. Es ist wohl wichtig und/oder nicht-trivial und/oder von eigenständiger Bedeutung, aber nun auch wieder nicht von so zentraler Bedeutung, dass es verdienen würde Satz zu heißen.

Die Einschätzung, welche Aussagen man wie bezeichnet, ist natürlich höchst subjektiv. Trotzdem gibt es meist einen gewissen Konsens darüber, welche Aussagen man wie benennt. Und nachdem diese Bezeichnungsweisen ja keine formal logischen Konsequenzen haben, sondern nur gewisse Zusatzinformationen tragen (was glaubt man ist wichtig, was glaubt man ist einfach, etc.), ist es letzlich nicht wirklich wichtig (wohl aber für den Leser möglicherweise hilfreich) was wie heißt.

Beweis von Proposition 2.4.6. Für den Beweis der Assoziativität, seien uns Funktionen f, g, h wie in (1) vorgegeben. Weiters sei uns ein Element $x \in M$ vorgegeben. Dann gilt, durch Einsetzen in die Definition der entsprechenden Kompositionen,

$$[(h \circ g) \circ f](x) = (h \circ g)(f(x)) = h(g(f(x)) = h((g \circ f)(x)) = [h \circ (g \circ f)](x).$$

Teil (2) der Proposition ist offensichtlich, da die Identität jedes Element ihrer Definitionsmenge auf sich selbst abbildet:

$$(\mathrm{id}_N \circ f)(x) = \mathrm{id}_N(f(x)) = f(x), \quad (g \circ \mathrm{id}_N)(x) = g(\mathrm{id}_N(x)) = g(x).$$

□

`Analyse des Beweises von Proposition` 2.4.6`:`

In diesem Beweis verwenden wir die folgende Tatsache, die die Definition einer Funktion wiederspiegelt: Zwei Funktionen von M nach N sind genau dann gleich, wenn sie und wenn sie jedem Element ihres (gemeinsamen) Definitionsbereiches M das selbe Element ihrer (gemeinsamen) Zielmenge N zuordnen.

Die Gleichheit von Definitions- bzw. Zielmengen ist in der hier betrachteten Situation (und das ist meistens – aber nicht immer – so) offensichtlich. Was es zu beweisen gilt, ist also, dass jedem Element des Definitionsbereiches der gleiche Wert zugewiesen wird; und das tun die angeschriebenen Rechnungen. ♧

J225. **2.4.7 Definition.** Sei $f : M \to N$ eine Funktion.

(1) Für eine Teilmenge $A \subseteq M$ setzen wir

$$f(A) = \{y \in N : \ \exists\, x \in A : y = f(x)\}.$$

Man spricht von dem *Bild von A unter f*. Speziell nennt man die Menge $f(M)$ das *Bild von f*.

Dadurch ist eine Funktion

$$f : \left\{ \begin{array}{rcl} \mathcal{P}(M) & \to & \mathcal{P}(N) \\ A & \mapsto & f(A) \end{array} \right.$$

gegeben.

(2) Für eine Teilmenge $B \subseteq N$ setzen wir

$$f^{-1}(B) = \{x \in M : \ f(x) \in B\}.$$

Man spricht von dem *vollständigen Urbild von B unter f*.

Dadurch ist eine Funktion

$$f^{-1} : \left\{ \begin{array}{rcl} \mathcal{P}(N) & \to & \mathcal{P}(M) \\ B & \mapsto & f^{-1}(B) \end{array} \right.$$

gegeben.

◊

Teil (1) dieser Definition besteht den formalen Check *nicht*: Das Symbol f *hat schon* eine Bedeutung, nämlich jene der gegebenen Funktion.

Trotzdem hat sich diese Notation eingebürgert, da sie intuitiv ist: $f(A)$ ist jene Menge, die man erhält wenn man die gegebene Funktion f elementweise auf A anwendet, sprich, sie auf jedes Element von A einzeln anwendet und die so entstehenden Elemente zu einer Menge zusammenfasst.

Man muss also immer aufpassen und die Schreibweise $f(\_)$ dem Zusammenhang entsprechend richtig interpretieren. Als kleiner Hinweis: oft – aber nicht immer – gilt die Konvention, dass man Elemente des Definitionsbereiches M mit Kleinbuchstaben, und Teilmengen mit Großbuchstaben bezeichnet.

Manchmal findet man auch die Schreibweise $f[A]$ für das Bild von A unter f. Das ist eine unmissverständliche Notation.

Eine alternative Definition ersetzt $f(x)$ durch $f\,`x$ oder $f'x$ (heute praktisch nicht mehr verwendet), und $f[A]$ durch $f``A$ oder $f''A$.

➢ `Warum macht man etwas, das eigentlich unserem selbst auferlegten Exaktheitsanspruch widerspricht:`

Mehrfachbelegungen von Schreibweisen sind manchmal praktisch (und üblich), wenn es

– aus dem Zusammenhang leicht ersichtlich ist, welche Bedeutung dem Symbol nun wirklich zukommt;

– man sich dadurch zusätzliche Notationen ersparen kann.

Oft ist es so, dass Verwendung absolut präziser Notation die Lesbarkeit derart erschwert, dass die Materie nahezu unlesbar wird und komplizierter wirkt als sie ist.

Abweichungen vom absolut präzisen Weg sind immer eine Gratwanderung!

Generell empfiehlt es sich solche Abweichungen, wenn irgendwie sinnvoll möglich, zu vermeiden. Wie auch immer man vorgeht, man muss zu jedem Zeitpunkt und an jeder Stelle fähig sein, die richtige Bedeutung des Hingeschriebenen zu benennen.

J229. **2.4.8 Lemma.** *Es gelten die folgenden Aussagen.*

(1) *Seien $f : M \to N$ und $g : N \to L$ Funktionen, dann ist*

$$\forall A \in \mathcal{P}(M)\colon (g \circ f)(A) = g(f(A)).$$

Weiters gilt $\forall A \in \mathcal{P}(M)\colon \mathrm{id}_M(A) = A$.

(2) *Seien $f : M \to N$ und $g : N \to L$ Funktionen, dann ist*

$$\forall C \in \mathcal{P}(L)\colon (g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C)).$$

Weiters gilt $\forall A \in \mathcal{P}(M)\colon \mathrm{id}_M^{-1}(A) = A$.

Beweis. Für den Beweis von (1) seien uns Funktionen f, g vorgegeben. Weiters sei ein Element $A \in \mathcal{P}(M)$ vorgegeben. Setzt man die Definition des Bildes einer Menge unter einer Funktion bzw. der Komposition zweier Funktionen ein, so erhält man

$$\begin{aligned}(g \circ f)(A) &= \{z \in L :\ \exists\, x \in A : z = (g \circ f)(x)\} = \{z \in L :\ \exists\, x \in A : z = g(f(x))\},\\ g(f(A)) &= \{z \in L :\ \exists\, y \in f(A) : z = g(y)\}.\end{aligned}$$

Um die Inklusion $(g \circ f)(A) \subseteq g(f(A))$ zu zeigen, sei $z \in L$ und sei angenommen, dass wir $x \in A$ haben mit $z = g(f(x))$. Setze $y = f(x)$, dann ist $y \in f(A)$ und $z = g(y)$. Also ist $z \in g(f(A))$. Für die umgekehrte Inklusion, d.h. $(g \circ f)(A) \supseteq g(f(A))$, sei $z \in L$ und sei angenommen, dass wir $y \in f(A)$ haben mit $z = g(y)$. Wähle $x \in M$ mit $y = f(x)$, dann ist $z = g(y) = g(f(x))$ und wir sehen, dass $z \in (g \circ f)(A)$. Die Aussage, dass $\mathrm{id}_M(A) = A$ ist klar.

Wir kommen zum Beweis von (2). Setzt man die Definition des vollständigen Urbilds einer Funktion ein, erhält man

$$x \in f^{-1}(g^{-1}(C)) \ \Leftrightarrow\ f(x) \in g^{-1}(C) \ \Leftrightarrow\ g(f(x)) \in C \ \Leftrightarrow\ (g \circ f)(x) \in C \ \Leftrightarrow\ x \in (g \circ f)^{-1}(C)$$

Die Aussage, dass $\mathrm{id}_M^{-1}(A) = A$ ist klar. □

Eine weitere Konstruktion, mit der man aus einer Funktion f eine andere bauen kann, ist die Einschränkung.

J241. **2.4.9 Proposition.** *Sei $f : M \to N$ eine Funktion, und seien $\tilde{M} \subseteq M$ und $\tilde{N} \subseteq N$ Teilmengen mit der Eigenschaft, dass $f(\tilde{M}) \subseteq \tilde{N}$. Dann ist $\tilde{f} = f \cap (\tilde{M} \times \tilde{N})$ eine Funktion.*

Beweis. Wir zeigen als erstes, dass $\tilde{f}$ die Eigenschaft **(Fun1)** hat. Dazu sei uns ein Element $x \in \tilde{M}$ vorgegeben. Da f eine Funktion ist, existiert ein Element $y \in N$ mit $(x, y) \in f$. Wegen unserer Voraussetzung, dass $f(\tilde{M}) \subseteq \tilde{N}$, gilt $y \in \tilde{N}$. Daher ist $(x, y) \in \tilde{f}$.

Als nächstes zeigen wir, dass $\tilde{f}$ die Eigenschaft **(Fun2)** hat. Dazu seien uns $x \in \tilde{M}$ und $y_1, y_2 \in \tilde{N}$ vorgegeben, und es sei vorausgesetzt, dass $(x, y_1) \in \tilde{f}$ und $(x, y_2) \in \tilde{f}$. Da $\tilde{f} \subseteq f$, gilt auch $(x, y_1) \in f$ und $(x, y_2) \in f$. Da f eine Funktion ist, folgt $y_1 = y_2$. □

Wir nennen $\tilde{f}$ die *Einschränkungsfunktion* von f bzgl. $\tilde{M}$ und $\tilde{N}$. Im Fall $\tilde{N} = N$, spricht man üblicherweise von der *Einschränkung von f auf $\tilde{M}$*, und schreibt $f|_{\tilde{M}}$ (manchmal auch $f \restriction_{\tilde{M}}$ oder $f \restriction \tilde{M}$).

2.5 Bijektivität

2.5.1 Motivation

Die Identifizierung von Dingen funktioniert üblicherweise via Namen, die den entsprechenden Dingen gegeben werden.

Erzählung eines Gartenbesitzers:

> In meiner Wiese stehen Äpfelbäume und Zwetschkenbäume. Einen Marillenbaum hatte ich auch, aber der hat den letzten Winter leider nicht überlebt. Jetzt habe ich ihn umgeschnitten. Das war ganz schön viel Arbeit...man glaubt ja gar nicht wie tief die Wurzeln von einem Marillenbaum runtergehen.

Woran liegt es, dass der Zuhörer versteht wovon der Erzähler spricht ? Einerseits daran, dass jede Baumart durch ihren Namen eindeutig identifiziert wird, andererseits daran, dass die auftretenden Namen tatsächlich auch bekannte Baumarten beschreiben (wenn der Erzähler in höchsten Tönen von seinem Grrfskbaum schwärmt, wird der Zuhörer wohl erstmal verwirrt sein).

Diese beiden Eigenschaften der *Beziehung* zwischen Baumarten und Namen, eigentlich der *Zuordnung* eines Namens zu jeder Baumart, kann man wie folgt ausdrücken:

(i) Zwei verschiedene Baumarten haben verschiedene Namen.

(ii) Jeder verwendete Name gehört zu einer Baumart.

Die Mengen der Baumarten einerseits und die Menge ihrer Namen andererseits sind also in dem Sinne gleich, als dass ihre Elemente völlig austauschbar – identifizierbar – sind:

$$\text{Baumart} \leftrightarrow \text{Name}$$

Eine solche Identifizierung kann man natürlich in zwei Richtung lesen. In der Rolle des Zuhörers: wenn man einen Namen hört, weiß man wie der Baum aussieht. In der Person des Erzählers: Wenn man einen Baum sieht, weiß man welchen Namen dieser trägt.

$$\text{Name} \leftrightarrow \text{Baumart}$$

Wesentliche Eigenschaften einer solchen Identifizierung sind:

(iii) Steht man vor einem Baum, so weiß man welchen Namen diese Art hat. Schließt man nun die Augen und denkt an einen Baum der Art mit *diesem* Namen, so sieht man vor dem geistigen Auge *genau einen solchen* Baum wie er tatsächlich vor einem steht.

(iv) Umgekehrt funktionert das natürlich genauso. Liest man einen Namen, so denkt man an einen Baum gewisser Art. Geht man nun in den Wald und sieht einen Baum *dieser* Art, so denkt man an *genau jenen* Namen den man vorher gelesen hatte.

2.5.2 Formalisierung

Wir geben den beiden einzelnen oben genannten Eigenschaften (i) und (ii) Namen.

J209. **2.5.1 Definition.** Seien M, N Mengen und $f : M \to N$ eine Funktion. Dann heißt f *injektiv*, wenn

(Inj) verschiedenen Elementen aus M verschiedene Werte zugeordnet werden,
$\left[\forall x_1, x_2 \in M\colon\ x_1 \neq x_2 \ \Rightarrow\ f(x_1) \neq f(x_2) \right]$

und *surjektiv*, wenn

(Sur) es zu jedem Element y aus N ein Element aus M gibt welchem y zugeordnet wird.
$\left[\forall y \in N\ \exists\, x \in M\colon\ y = f(x) \right]$

Weiters nennen wir f *bijektiv*, wenn

(Bij) f injektiv und surjektiv ist.

◊

♧ Man beachte, dass man die Injektivität der Zuordnung f unmittelbar ansieht, wogegen sich die Surjektivität immer auch auf die Zielmenge bezieht. Daher findet man auch oft die Sprechweise *f ist surjektiv auf N*. ♧

Wenn eine Funktion $f\colon M \to N$ surjektiv ist, sagt man statt „f von M *nach* N“ auch gerne „f von M *auf* N“, und schreibt $f\colon M \twoheadrightarrow N$. Im Englischen kann man das Wort „onto“ sogar als Adjektiv verwenden, wenn einem „surjective“ zu kompliziert erscheint.

Statt „injective“ hört man im Englischen auch „one-to-one“; für injektive Funktionen schreibt man auch $f\colon M \rightarrowtail N$.

J249. **2.5.2 Lemma.** *Sei $f : M \to N$ eine Funktion. Dann ist f genau dann injektiv, wenn gilt:*

(Inj′) *Sind die Funktionswerte zweier Elemente gleich, so müssen die beiden Elemente gleich sein.*
$\left[\ \forall x_1, x_2 \in M\colon (f(x_1) = f(x_2) \ \Rightarrow\ x_1 = x_2)\ \right]$

Beweis. Die Implikation in **(Inj′)** ist die Kontraposition von jener in **(Inj)**. □

In diesem Beweis verwenden wir eine sehr einfache, aber extrem häufig verwendete, Schlussweise. Nämlich das Prinzip des *Beweises durch Kontraposition*.

> ➢ `Ein Beweis durch Kontraposition funktioniert so:`
>
> Wir gehen davon aus, dass die zu beweisende Aussage (zum Beispiel das gerade betrachtete Lemma) bereits in dem Sinne analysiert wurde, dass es klar ist was die zur Verfügung stehenden Voraussetzungen sind, und was die zu zeigende Behauptung ist. Wollen wir diese Voraussetzungen als A und diese Behauptung als B bezeichnen.
>
> ① Man zeigt die Wahrheit der Implikation $\neg B \Rightarrow \neg A$.
>
> Was nun folgt, ist die formallogische Argumentation, dass damit der Beweis tatsächlich abgeschlossen ist.
>
> ❷ Man verwendet den Kontrapositionssatz (8) aus §1.1.2. Dies gibt die Wahrheit von $A \Rightarrow B$.
>
> ❸ Man wendet den modus ponendo ponens an, nachdem man aus der Wahrheit von A und der Wahrheit der Implikation $A \Rightarrow B$ auf die Wahrheit von B schließen kann.

2.5.3 Bemerkung.

1. Wenn Sie die Richtung des Pfeils in der Aussage **(Inj′)** umkehren, erhalten Sie eine Aussage, die für *jede* Funktion gilt:
$\left[\ \forall x_1, x_2 \in M\colon (x_1 = x_2 \Rightarrow f(x_1) = f(x_2))\ \right]$
2. Wenn man die Injektivität einer Funktion nachprüfen muss, ist die Formulierung **(Inj)** oft praktischer als die (äquivalente) Formulierung **(Inj′)**, weil wir mehr Übung im Umgang mit Gleichungen als mit Ungleichungen haben.

◊

Schreibt man die Bedingung **(Inj′)** für Injektivität in der ursprünglichen Bedeutung einer Funktion als spezielle Relation, d.h., ohne Verwendung der Schreibweise „$f(\_)$“, so schreibt sie sich als

(Inj″) $\forall y \in N\ \forall x_1, x_2 \in M\colon \Big(\big((x_1, y) \in f \wedge (x_2, y) \in f\big) \ \Rightarrow\ x_1 = x_2\Big)$

Hier sieht man gleich auch noch eine weitere Perspektive auf Injektivität. Nämlich ist f genau dann injektiv, wenn

(Inj‴) es zu jedem Element y aus N höchstens ein Element aus M gibt, welchem y zugeordnet wird.

Die Bedingungen für Surjektivität bzw. Bijektivität kann man auch umformulieren: Eine Funktion ist genau dann surjektiv, wenn

(Sur′) $f(M) = N$,

und sie ist genau dann bijektiv, wenn

(Bij′) es zu jedem Element y von N genau ein Element aus M gibt welchem y zugeordnet wird.
$\left[\ \forall y \in N\ \exists!\, x \in M\colon y = f(x)\ \right]$

Wir haben die Aussagen gemacht, dass für jede Funktion die Äquivalenzen **(Inj′)**⇔**(Inj″)**⇔**(Inj‴)**, **(Sur)**⇔**(Sur′)** und **(Bij)**⇔**(Bij′)** gelten. Dies bedarf natürlich einer Rechtfertigung.

Der erste Teil der ersten genannten Äquivalenz ist nur Verwendung einer anderen Schreibweise, ihr zweiter Teil ist die Definition des Terminus „höchstens ein“. Um die zweite Äquivalenz einzusehen, braucht man sich nur der Definition des Bildes einer Funktion zu erinnern. Für die dritte, erinnere man sich daran was $\exists!$ bedeutet.

J248. *2.5.4 Bemerkung.* Das Zusammenspiel von Injektivität, Surjektivität, und Bijektivität wird von folgender Perspektive besonders deutlich: man vergleiche **(Sur)**, **(Inj‴)**, **(Bij′)**, und erinnere sich an (10) aus §1.1.3. ◊

Wir haben eine Methode kennengelernt wie man aus zwei Funktionen eine neue konstruieren kann, nämlich die Komposition. Es stellt sich die Frage, wie sich die oben definierten Eigenschaften bei dieser Konstruktion verhalten. Die Antwort ist: Brav.

J242. **2.5.5 Proposition.** *Seien $f : M \to N$ und $g : N \to L$ Funktionen. Dann gilt*

(1) *Sind f und g beide injektiv, so ist auch $g \circ f$ injektiv.*

(2) *Sind f und g beide surjektiv, so ist auch $g \circ f$ surjektiv.*

(3) *Sind f und g beide bijektiv, so ist auch $g \circ f$ bijektiv.*

Beweis. Für den Beweis von (1) sei ⟩① vorausgesetzt, dass f, g injektiv sind. Um **(Inj)** für die Funktion $g \circ f$ zu zeigen, ⟩② seien uns $x_1, x_2 \in M$ vorgegeben und sei ⟩③ vorausgesetzt, dass $x_1 \neq x_2$. Da f injektiv ist, folgt, dass $f(x_1) \neq f(x_2)$. Da g injektiv ist, folgt, dass $g(f(x_1)) \neq g(f(x_2))$. Also haben wir $(g \circ f)(x_1) \neq (g \circ f)(x_2)$. ⟨③⟨ ⟨②⟨ ⟨①⟨

Für den Beweis von (2) sei ⟩① vorausgesetzt, dass f, g surjektiv sind. Um **(Sur)** für die Funktion $g \circ f$ zu zeigen, ⟩② sei uns $z \in L$ vorgegeben. Da g surjektiv ist, können wir ⟩③ ein Element $y \in N$ wählen mit $g(y) = z$. Da f surjektiv ist, können wir ⟩④ ein Element $x \in M$ wählen mit $f(x) = y$. Dann gilt $(g \circ f)(x) = g(f(x)) = g(y) = z$. ⟨④⟨ ⟨③⟨ ⟨②⟨ ⟨①⟨

Für den Beweis von (3) sei ⟩① vorausgesetzt, dass f, g bijektiv sind. Dann sind f, g beide injektiv, und nach dem bereits bewiesenen Teil (1) folgt, dass $g \circ f$ injektiv ist. Weiters sind f, g beide surjektiv, und nach dem bereits bewiesenen Teil (2) folgt, dass $g \circ f$ surjektiv ist. Daher ist $g \circ f$ bijektiv. ⟨①⟨ □

Eine weitere Konstruktion, die wir kennengelernt haben, ist die Einschränkung einer Funktion.

J243. **2.5.6 Proposition.** *Seien $f : M \to N$ eine Funktion, seien $\tilde{M} \subseteq M$ und $\tilde{N} \subseteq N$ Teilmengen mit $f(\tilde{M}) \subseteq \tilde{N}$, und sei $\tilde{f} = f \cap (\tilde{M} \times \tilde{N})$ die Einschränkungsfunktion von f.*

(1) *Ist f injektiv, so ist auch $\tilde{f}$ injektiv.*

(2) *$\tilde{f}$ is surjektiv, genau dann wenn $f(\tilde{M}) = \tilde{N}$.*

Beweis. Um (1) zu zeigen, sei ⟩① vorausgesetzt, dass f injektiv ist. ⟩② Seien uns $x_1, x_2 \in \tilde{M}$ vorgegeben, und sei ⟩③ vorausgesetzt, dass $x_1 \neq x_2$. Dann ist $f(x_1) \neq f(x_2)$. Nun gilt für jedes $x \in \tilde{M}$ sicherlich $\tilde{f}(x) = f(x)$, und wir sehen, dass $\tilde{f}(x_1) \neq \tilde{f}(x_2)$. ⟨③⟨ ⟨②⟨ ⟨①⟨

Um (2) einzusehen, genügt es **(Sur′)** zu betrachten, und zu bemerken, dass $\tilde{f}(\tilde{M}) = f(\tilde{M})$. □

2.5.3 Links- und Rechtsinverse

Wir wollen uns im Folgenden mit den in §2.5.1 genannten Eigenschaften (iii) und (iv) beschäftigen. Dazu müssen wir ein bisschen ausholen.

J212. **2.5.7 Definition.** Seien M, N Mengen, und $f : M \to N$ eine Funktion. Eine Funktion $g : N \to M$ heißt

(1) *Linksinverse von f*, wenn $g \circ f = \mathrm{id}_M$;

(2) *Rechtsinverse von f*, wenn $f \circ g = \mathrm{id}_N$;

(3) *Inverse von f*, wenn sie sowohl Links- als auch Rechtsinverse ist.

◊

J214. **2.5.8 Lemma.** *Seien M, N Mengen, und $f : M \to N$ eine Funktion. Sei $g_l : N \to M$ eine Linkinverse von f und $g_r : N \to M$ eine Rechtsinverse. Dann gilt $g_l = g_r$ und diese Funktion ist eine Inverse von f.*

♲ `Analyse von Lemma 2.5.8:`

– Wir wissen, was eine Menge ist, und was eine Äquivalenzrelation ist.

– Klar ?

① Die Objekte, mit denen wir arbeiten, sind die drei Funktionen f, g_l, g_r.
② Voraussetzungen sind, dass g_l Linksinverse von f, und g_r Rechtsinverse von f ist.
③ Die Behauptung ist, dass g_l und g_r übereinstimmen und dass diese (eine) Funktion sogar Inverse von f ist.

Beweis von Lemma 2.5.8. Es gilt

$$g_l = g_l \circ \mathrm{id}_N = g_l \circ (f \circ g_r) = (g_l \circ f) \circ g_r = \mathrm{id}_M \circ g_r = g_r.$$

Damit ist offensichtlich g_l eine Inverse von f. □

 `Analyse des Beweises von Lemma 2.5.8:`
Die angeschriebenen Gleichheiten kommen wie folgt zustande.

$g_l \circ (f \circ g_r) = (g_l \circ f) \circ g_r$	Wir verwenden die Assoziativität (2.4.3) um die dreifach-Hintereinanderausführung $g_l \circ f \circ g_r$ auf zwei verschiedene Arten auszurechnen.
$g_l \circ \mathrm{id}_N = g_l \circ (f \circ g_r)$ und $(g_l \circ f) \circ g_r = \mathrm{id}_M \circ g_r$	Die Voraussetzung, dass g_r Rechtsinverse bzw. g_l Linksinverse ist.
$g_l = g_l \circ \mathrm{id}_N$ und $\mathrm{id}_M \circ g_r = g_r$	Die Eigenschaften (2.4.4) der Identität.

Wir haben in dieser Rechnung alle Voraussetzungen des Lemmas verwendet, und auch die volle Stärke von Proposition 2.4.6 benutzt.

Warum ist es nun „offensichtlich", dass g_l eine Inverse von f ist?

Vor dem Weiterlesen zuerst selbst nachdenken!

Wir müssen zeigen, dass $g_l \circ f = \mathrm{id}_M$ und $f \circ g_l = \mathrm{id}_N$. Ersteres gilt da g_l nach Voraussetzung des Lemmas Linksinverse von f ist. Wegen $g_l = g_r$ und da g_r nach Voraussetzung des Lemmas eine Rechtsinverse von f ist, gilt $f \circ g_l = f \circ g_r = \mathrm{id}_N$.

J217. **2.5.9 Korollar.** *Seien M, N Mengen und $f : M \to N$ eine Funktion. Dann hat f höchstens eine Inverse.*

Beweis. Seien g_1 und g_2 beide Inverse von f. Dann ist g_1 insbesondere eine Linksinverse von f, und g_2 insbesondere eine Rechtsinverse von f. Nach Lemma 2.5.8 folgt $g_1 = g_2$. □

J216. **2.5.10 Satz.** *Seien M und N nichtleere Mengen. Für jede Funktion $f : M \to N$ gelten die folgenden drei Aussagen.*

(1) *f ist injektiv, genau dann wenn f eine Linksinverse besitzt.*

(2) *f ist surjektiv, genau dann wenn f eine Rechtsinverse besitzt.*

(3) *f ist bijektiv, genau dann wenn f eine Inverse besitzt.*

Der Beweis von Satz 2.5.10(1) beruht auf folgender Konstruktion: Für eine Relation $R \subseteq M \times N$, bezeichne mit R^{-1} die Relation

$$R^{-1} = \{(y, x) \in N \times M : \ (x, y) \in R\}.$$

Man spricht von R^{-1} auch als der *inversen Relation* von R.

 Wieder einmal eine Definition die den formalen Check nur teilweise besteht!
Die Notation $\_^{-1}$ *hat schon* eine Bedeutung. Nämlich, ist R sogar eine Funktion, so bezeichnet R^{-1} die vollständige

Urbildfunktion von $\mathcal{P}(N)$ nach $\mathcal{P}(M)$. Man mache sich also bei jedem Auftreten dieser Notation klar, welche Bedeutung gemeint ist.

Vergleicht man die Eigenschaft **(Fun2)** und die alternative Schreibweise **(Inj″)** für Injektivität, so sieht man, dass eine Relation R genau dann das Axiom **(Fun2)** erfüllt, wenn R^{-1} der Bedingung **(Inj″)** genügt.

Wieder einmal eine Aussage!
Zum Beweis dieser Aussage braucht man bloß die Definition von R^{-1} einzusetzen:

$$\begin{aligned} R \text{ erfüllt } \textbf{(Fun2)} &\Leftrightarrow \forall x \in M\ \forall y_1, y_2 \in N\colon \big((x, y_1) \in R \wedge (x, y_2) \in R \big) \ \Rightarrow\ y_1 = y_2 \\ &\Leftrightarrow \forall x \in M\ \forall y_1, y_2 \in N\colon \big((y_1, x) \in R^{-1} \wedge (y_2, x) \in R^{-1} \big) \ \Rightarrow\ y_1 = y_2 \\ &\Leftrightarrow R^{-1} \text{ erfüllt } \textbf{(Inj″)} \end{aligned}$$

Beweis von Satz 2.5.10. Die wesentlichen Beweisteile sind der Beweis der Implikationen „$\Rightarrow$“ in (1) und (2). Der Rest ist dann einfach.

Schritt 1; „$\Rightarrow$“ in (1): Sei vorausgesetzt, dass f injektiv ist. Betrachte die Relation $f^{-1} \subseteq N \times M$. Diese erfüllt nach oben Gesagtem **(Fun2)**.

Wir erweitern sie um eine Funktion zu erhalten: Die Menge M ist nichtleer. Wähle ein Element $x_0 \in M$, und setze

$$g = f^{-1} \cup \big((N \setminus f(M)) \times \{x_0\} \big). \tag{2.5.1}$$ J218

– Als erstes zeigen wir, dass g nun **(Fun1)** erfüllt:
Sei $y \in N$ vorgegeben. Wir machen eine Fallunterscheidung.

  ✾ *Fall 1;* $y \in f(M)$. Wähle $x \in M$ mit $f(x) = y$, dann gilt $(y, x) \in f^{-1} \subseteq g$.
  ✾ *Fall 2;* $y \in N \setminus f(M)$. Nach der Definition von g gilt, dass $(y, x_0) \in g$.

  Offensichtlich decken diese beiden Fälle alle Möglichkeiten ab (Fall 2 ist genau die Negation von Fall 1).

– Als zweites zeigen wir, dass die Relation g immer noch **(Fun2)** erfüllt:
Sei $y \in N$ und $x_1, x_2 \in M$ vorgegeben, mit $(y, x_1), (y, x_2) \in g$. Wir machen eine Fallunterscheidung.

  ✾ *Fall 1;* $y \in f(M)$. Es kann weder (y, x_1) noch (y, x_2) in $(N \setminus f(M)) \times \{x_0\}$ liegen. Daher muss $(y, x_1), (y, x_2) \in f^{-1}$ gelten, und da f^{-1} **(Fun2)** erfüllt folgt $x_1 = x_2$.
  ✾ *Fall 2;* $y \in N \setminus f(M)$. Es existiert kein Element $x \in M$ mit $(x, y) \in f$, also kann weder (y, x_1) noch (y, x_2) in f^{-1} liegen. Daher muss $(y, x_1), (y, x_2) \in (N \setminus f(M)) \times \{x_0\}$ gelten, und wir sehen, dass $x_1 = x_2$, nämlich beide gleich x_0, ist.

  Offensichtlich decken diese beiden Fälle alle Möglichkeiten ab (Fall 2 ist genau die Negation von Fall 1).

Schritt 2; „$\Rightarrow$“ in (2): Für „$\Rightarrow$“ in (2) sei vorausgesetzt, dass f surjektiv ist. Dann ist für jedes $y \in N$ das vollständige Urbild $f^{-1}(\{y\})$ nicht leer. Nach **(AC′)** existiert eine Auswahlfunktion $g : N \to M$ sodass für alle $y \in N$ die Beziehung $g(y) \in f^{-1}(\{y\})$ gilt. Es gilt also für alle $y \in N$ dass $f(g(y)) = y$, und wir haben eine Rechtsinverse von f gefunden.

Schritt 3; „$\Leftarrow$“ in (1): Sei vorausgesetzt, dass f eine Linksinverse hat. Wähle eine solche, sprich, wähle $g : N \to M$ mit $g \circ f = \mathrm{id}_M$. Wir gehen darauf los **(Inj′)** zu zeigen, was ja äquivalent zu Injektivität ist. Seien $x_1, x_2 \in M$ vorgegeben und sei vorausgesetzt, dass $f(x_1) = f(x_2)$ (das ist die Prämisse der Implikation in **(Inj′)**). Dann gilt

$$x_1 = g(f(x_1)) = g(f(x_2)) = x_2,$$

also ist die Konklusion der Implikation in **(Inj′)** wahr.

Schritt 4; „$\Leftarrow$“ in (2): Sei vorausgesetzt, dass f eine Rechtsinverse hat. Wähle eine solche, sprich, wähle $g : N \to M$ mit $f \circ g = \mathrm{id}_N$. Sei $y \in N$ vorgegeben. Es gilt $f(g(y)) = y$, also haben wir ein Element von M gefunden, welches unter f auf y abgebildet wird, nämlich $g(y)$.

Schritt 5; (3): Die Aussage (3) ist nun ein Korollar.

Für „⇒“ sei vorausgesetzt, dass f bijektiv ist. Dann ist f injektiv und surjektiv. Nach den schon bewiesenen Teilen (1) und (2) besitzt f sowohl eine Linksinverse also auch eine Rechtsinverse. Nach Lemma 2.5.8 stimmen diese überein und sind eine Inverse von f.

Umgekehrt: sei vorausgesetzt, dass f eine Inverse hat. Eine Inverse ist insbesondere sowohl eine Links- also auch eine Rechtsinverse. Nach den schon bewiesenen Teilen (1) und (2) ist f injektiv und surjektiv, also bijektiv.

□

♲ Satz 2.5.10 gibt uns eine alternative Möglichkeit Injektivität (bzw. Surjektivität oder Bijektivität) zu *definieren*. Er zeigt, dass man anstelle der Axiome **(Inj)** etc., genausogut mit Existenz von (einseitigen) Inversen hätte starten können. ♲

➢ `Was ist eine alternative Definition`

Oft ist man in der Situation mit einer Eigenschaft zu arbeiten die viele verschiedene äquivalente Erscheinungsformen hat. Hat man zum Beispiel zwei Eigenschaften, sagen wir **(A1)** und **(A2)**, und weiß man, dass **(A1)⇔(A2)**, so hat man freie Wahl ob man lieber sagt:

> *Definition:* Ein Objekt mit der Eigenschaft **(A1)** heißt ein Biffel.
> *Satz:* Ein Objekt ist genau dann ein Biffel, wenn es **(A2)** erfüllt.

oder, anders herum,

> *Definition:* Ein Objekt mit der Eigenschaft **(A2)** heißt ein Biffel.
> *Satz:* Ein Objekt ist genau dann ein Biffel, wenn es **(A1)** erfüllt.

Die Essenz ist in jedem Fall die Tatsache, dass **(A1)⇔(A2)** gilt. Manchmal spricht man nun von so einer Äquivalenz dann als eine *alternative Definition* des Begriffs eines Biffels.

Diese Sprechweise verwendet man eher nur dann wenn die beiden Eigenschaften **(A1)** und **(A2)** ziemlich verschieden aussehen; wie zum Beispiel **(Inj)** einerseits, und die Eigenschaft eine Linksinverse zu haben andererseits. Von **(Inj′)** im Vergleich zu **(Inj)** würde man man wohl nicht von einer alternativen Definition sprechen, sondern, wenn überhaupt, von einer alternativen Schreibweise der Originaldefinition.

d1d2. *2.5.11 Bemerkung.* Wenn Sie zwei äquivalente Definitionen einer Eigenschaft D haben, dann passiert es oft, dass die eine Definition D1 scheinbar mehr als die andere D2 aussagt. Ein Beispiel haben wir schon am Beispiel der Eigenschaften Sym und Sym′ gesehen.

Wenn Sie in so einer Situation beweisen müssen, dass die Eigenschaft D gilt, dass empfiehlt es sich, die scheinbar schwächere Formulierung D2 zu überprüfen, weil das leichter ist. Wenn Sie aber schon wissen, dass D gilt, dürfen Sie D1 verwenden; eine Folgerung aus der scheinbar stärkeren Eigenschaft D1 ist vielleicht leichter zu gewinnen als aus D2. ◊

J219. *2.5.12 Bemerkung.* Sei $f : M \to N$ bijektiv. Wie wir in Schritt 1 des Beweises von Satz 2.5.10 gesehen haben, ist dann die durch (2.5.1) gegebene Funktion eine Linksinverse von f, und, nach dem Argument in Schritt 5 damit sogar Inverse von f. Nun ist, da f ja surjektiv ist, die Menge $N \setminus f(M)$ leer. Damit haben wir $g = f^{-1}$, sprich, die *inverse Funktion* von f ist gleich der *inversen Relation* f^{-1} von f:

$$f^{-1} = \{(y, x) \in N \times M : \ (x, y) \in f\}.$$

◊

Alternativer Beweis von „⇒“ in Satz 2.5.10(3). Da insbesondere f injektiv ist, ist die Relation (2.5.1) eine Funktion und, wie im Beweis von Satz 2.5.10(1) gezeigt, ist g eine Linksinverse von f.

Wir zeigen nun, dass g auch Rechtsinverse von f ist. Sei dazu $y \in N$ gegeben. Wähle $x \in M$ mit $f(x) = y$ Dann ist $(y, x) \in f^{-1} \subseteq g$, und damit $f(g(y)) = f(x) = y$. □

♧ Ist der alternative Beweis von Satz 2.5.10(3) besser ? Eigentlich nicht wirklich, wir erhalten keine weiteren oder stärkeren Erkenntnisse aus ihm. Man könnte höchstens sagen, dass er ästhetischer ist, in dem Sinn, dass er keine unnötig verwinkelten Wege geht.

Was meinen wir damit ? Betrachten wir den erstgeführten Beweis. Dort haben wir (3) unter Verwendung von (1), (2), und Lemma 2.5.8 hergeleitet. Beim Beweis von (2) haben wir eine *reine Existenzaussage* verwendet (nämlich die Eigenschaft **(AC)**). Der Beweis von Lemma 2.5.8 war einfach. Nun ist es naturgemäß so, dass bei Verwendung von reinen Existenzaussagen wohl die Existenz von Objekten als wahr erkannt wird, man aber *keine Ahnung hat* wie diese – existierenden – Objekte nun wirklich aussehen. Bemerkenswerterweise erhalten wir in obigem Beweis dann aus der Anwendung des ganz simplen Lemma 2.5.8 doch eine *explizite* Information darüber wie die inverse Funktion von f aussieht (nämlich, wie schon angemerkt, ist sie die inverse Relation f^{-1})

Diese Tatsache legt es nahe, dass die Verwendung einer reinen Existenzaussage irgendwie unnötig kompliziert war. Und genau das zeigt uns der alternative Beweis: er nimmt ganz konkret die zu f inverse Relation her und rechnet nach, dass diese eine Funktion ist, und daher wirklich Inverse von f ist. ♧

J240. *2.5.13 Bemerkung.* Wir haben in Satz 2.5.10 vorausgesetzt, dass M und N beide nichtleer sind. Ist diese Voraussetzung notwendig gewesen ?

Um dies zu beantworten bemerke dass, falls M oder N leer ist, auch $M \times N$ leer ist. In diesem Fall gibt es also genau eine Relation zwischen M und N, nämlich $\emptyset$. In der folgenden Tabelle listen wir die im jetzigen Zusammenhang interessanten Eigenschaften auf:

M	N	$\exists!$ Relation	**(Fun1)**	**(Fun2)**	**(Inj)**	**(Sur)**	Links-inverse	Rechts-inverse
$=\emptyset$	$=\emptyset$	$\emptyset$	✓	✓	✓	✓	✓	✓
$\neq\emptyset$	$=\emptyset$	$\emptyset$	✗	✓				
$=\emptyset$	$\neq\emptyset$	$\emptyset$	✓	✓	✓	✗	✗	✗

Wir sehen, dass Satz 2.5.10

- im Fall „$M = N = \emptyset$“ *gilt*, denn die einzige Funktion die es gibt ist bijektiv und hat sich selbst als Inverse;
- im Fall „$M \neq \emptyset, N = \emptyset$“ *gilt*, denn es gibt keine Funktion von M nach N.
- im Fall „$M = \emptyset, N \neq \emptyset$“ Teil (1) *nicht* gilt, denn die Funktion $\emptyset$ ist injektiv hat aber keine Linksinverse (weil es ja gar keine Funktion von N nach M gibt), wogegen die Teile (2) und (3) *gelten*, denn es gibt keine surjektive (insbesondere auch keine bijektive) Funktion von M nach N.

◊

Mit Hilfe von Satz 2.5.10 kann man auch einen alternativen Beweis von Proposition 2.5.5 geben. Wir wollen dies am Punkt (1) von Proposition 2.5.5 demonstrieren.

Alternativer Beweis von Proposition 2.5.5(1). Für den Beweis von (1) sei vorausgesetzt, dass f, g injektiv sind. Nach Satz 2.5.10 können wir eine Linksinverse $f' : N \to M$ von f und eine Linksinverse $g' : L \to N$ von g wählen. Dann gilt

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

Wir haben eine Linksinverse von $g \circ f$ gefunden, und Satz 2.5.10 zeigt, dass $g \circ f$ injektiv ist. □

Um den Begriff der bijektiven Funktion und ihrer Inversen an einem komplexeren Beispiel zu illustrieren, erinnern wir uns an den Zusammenhang zwischen Äquivalenzrelationen und Partitionen.

J221. **2.5.14 Satz** (Verfeinerung von Satz 2.3.3)**.** *Sei M eine Menge. Bezeichne mit $\mathbb{A}$ die Menge aller Äquivalenzrelationen auf M, und mit $\mathbb{P}$ die Menge aller Partitionen von M. Weiters bezeichne f die Funktion*

$$f : \left\{ \begin{array}{ccl} \mathbb{A} & \to & \mathbb{P} \\ R & \mapsto & \{[x]_R : x \in M\} \end{array} \right.$$

Dann ist f bijektiv, und die Inverse von f ist gegeben als

$$g : \left\{ \begin{array}{ccl} \mathbb{P} & \to & \mathbb{A} \\ Q & \mapsto & \{(x,y) : \ \exists A \in Q \colon x \in A \wedge y \in A\} \end{array} \right. \qquad (2.5.2)$$ J220

Beweis. Die Tatsache, dass f eine Funktion von $\mathbb{A}$ nach $\mathbb{P}$ ist, haben wir in Satz 2.3.3(1) gezeigt. Dass die in (2.5.2) angegeben Zuordnung eine Funktion von $\mathbb{P}$ nach $\mathbb{A}$ ist haben wir in Satz 2.3.3(2) gezeigt. Wir haben zu zeigen, dass $g \circ f = \mathrm{id}_{\mathbb{A}}$ und $f \circ g = \mathrm{id}_{\mathbb{P}}$ gilt.

Schritt 1; wir gehen darauf los $g \circ f = \mathrm{id}_{\mathbb{A}}$ zu zeigen:
Zwei Funktionen sind genau dann gleich, wenn sie die gleichen Definitions- und Zielmengen haben, und jedem Element ihrer Definitionsmenge das selbe Element der Zielmenge zuordnen. Die beiden Funktionen $g \circ f$ und $\mathrm{id}_{\mathbb{A}}$ sind beides Funktionen von $\mathbb{A}$ nach $\mathbb{A}$. Wir müssen zeigen, dass sie stets die gleichen Werte ergeben.

⟩① ⟩ Sei uns $R \in \mathbb{A}$ vorgegeben. Wir müssen zeigen, dass $(g \circ f)(R) = R$. Dazu zeigen wir, dass beide wechselseitigen Inklusionen gelten.

– Wir beginnen mit „$(g \circ f)(R) \subseteq R$“:
⟩② ⟩ Sei uns $(x, y) \in g(f(R))$ vorgegeben; man erinnere sich an Lemma 2.4.8(1). Dann existiert, nach der Definition (2.5.2), ein Element A der Partition $f(R)$ welches beide, x und y, enthält. ⟩③ ⟩ Wähle ein solches Element $A \in f(R)$. Die Elemente von $f(R)$ sind genau die Mengen der Gestalt $[z]_R$ mit $z \in M$. ⟩④ ⟩ Wähle $z \in M$ sodass $A = [z]_R$. Dann ist $x, y \in [z]_R$, also $(x, z), (y, z) \in R$, und es folgt, dass auch $(x, y) \in R$. ⟨④⟨ ⟨③⟨ Wir sehen, dass tatsächlich $g(f(R)) \subseteq R$. ⟨②⟨
– Nun kommen wir zu „$(g \circ f)(R) \supseteq R$“:
⟩② ⟩ Sei uns $(x, y) \in R$ vorgegeben. Dann ist $x, y \in [x]_R$, und wir haben eine Menge der Partition $f(R)$ gefunden (nämlich $[x]_R$), welche beide Elemente x und y enthält. Also ist, nach der Definition (2.5.2), (x, y) Element der Relation $g(f(R))$. Wir sehen, dass tatsächlich $g(f(R)) \supseteq R$. ⟨②⟨

Zusammen haben wir $(g \circ f)(R) = R$ gezeigt. ⟨①⟨

Schritt 2; wir gehen darauf los $f \circ g = \mathrm{id}_{\mathbb{P}}$ zu zeigen:
Die beiden Funktionen $f \circ g$ und $\mathrm{id}_{\mathbb{P}}$ haben die gleichen Definitions- und Zielmengen, nämlich $\mathbb{P}$. Wir müssen zeigen, dass sie stets die gleichen Werte ergeben.

⟩① ⟩ Sei uns $Q \in \mathbb{P}$ vorgegeben. Wir müssen zeigen, dass $(f \circ g)(Q) = Q$. Dazu zeigen wir, dass beide wechselseitigen Inklusionen gelten.

– Wir beginnen mit „$(f \circ g)(Q) \subseteq Q$“:
⟩② ⟩ Sei uns $A \in f(g(Q))$ vorgegeben; man erinnere sich an Lemma 2.4.8(1). ⟩③ ⟩ Wähle $z \in M$ mit $A = [z]_{g(Q)}$. ⟩④ ⟩ Wähle $B \in Q$ mit $z \in B$. Da sicher $z \in [z]_{g(Q)}$ ist, haben wir $z \in A \cap B$. Damit ist $A \cap B \neq \emptyset$ und es folgt $A = B$. Also ist $A \in Q$. ⟨④⟨ ⟨③⟨ ⟨②⟨
– Nun kommen wir zu „$(f \circ g)(Q) \supseteq Q$“:
⟩② ⟩ Sei uns $A \in Q$ vorgegeben. Die Menge A ist nichtleer. ⟩③ ⟩ Wähle $z \in A$. Wir gehen darauf los zu zeigen, dass $A = [z]_{g(Q)}$. Dazu zeigen wir, dass beide wechselseitigen Inklusionen gelten. Wir beginnen mit „$A \subseteq [z]_{g(Q)}$“. ⟩④ ⟩ Sei uns $x \in A$ vorgegeben. Da z in A liegt, haben wir $x \in A \wedge z \in A$, und daher $(x, z) \in g(Q)$. Also ist $x \in [z]_{g(Q)}$. Wir sehen, dass tatsächlich $A \subseteq [z]_{g(Q)}$. ⟨④⟨ Nun kommen wir zur Inklusion „$A \supseteq [z]_{g(Q)}$“. ⟩④ ⟩ Sei uns $x \in [z]_{g(Q)}$ vorgegeben. Dann ist $(x, z) \in g(Q)$, also existiert $B \in Q$ mit $x, z \in B$. ⟩⑤ ⟩ Wähle $B \in Q$ mit dieser Eigenschaft. Dann ist $z \in B \cap A$; man erinnere sich, dass z gemäß ③ gewählt wurde. Also ist $B \cap A \neq \emptyset$ und daher $B = A$. Damit haben wir $x \in A$. ⟨⑤⟨ Wir sehen, dass tatsächlich $A \supseteq [z]_{g(Q)}$. ⟨④⟨ Gemeinsam erhalten wir $A = [z]_{g(Q)}$. ⟨③⟨ Damit ist A ein Element von $f(g(Q))$. ⟨②⟨

Zusammen haben wir $(f \circ g)(Q) = Q$ gezeigt. ⟨①⟨

□

Betrachten wir rückblickend nochmal diesen Beweis. Kurz ist er nicht, es gibt doch einiges zu arbeiten. Aber eigentlich ist alles straightforward.

Zum Abschluss wollen wir noch ein Beispiel diskutieren.

J244. *2.5.15 Beispiel.* Sei M eine Menge. Eine bijektive Funktion von M in sich heißt auch *Permutation von M*. Wir bezeichnen die Menge aller Permutationen von M mit $\mathrm{Sym}(M)$.

Diese Menge ist stets nichtleer, sie enthält immer zumindest die Funktion id_M. Ist $M = \emptyset$, oder hat M nur ein Element, ist dies auch die einzige Permutation (sogar die einzige Funktion von M nach M die es überhaupt gibt). Hat M zwei Elemente, so hat $\mathrm{Sym}(M)$ zwei Elemente; um dies zu sehen, wollen wir alle möglichen Funktionen von einer Menge $\{a, b\}$ in sich, das sind vier Stück, aufschreiben:

	f_1	f_2	f_3	f_4
$a \mapsto$	a	a	b	b
$b \mapsto$	a	b	a	b

Die Funktionen f_1 und f_4 sind weder injektiv noch surjektiv, die Funktionen f_2 und f_3 sind bijektiv. Dabei ist $f_2 = \mathrm{id}_M$. Wir haben also $\mathrm{Sym}(M) = \{\mathrm{id}_M, f_3\}$.

Die Komposition zweier bijektiver Funktionen ist wiederum bijektiv, also können wir eine Funktion definieren als

$$\circ : \begin{cases} \mathrm{Sym}(M) \times \mathrm{Sym}(M) & \to & \mathrm{Sym}(M) \\ (f, g) & \mapsto & f \circ g \end{cases}$$

Weiters ist die Inverse einer bijektiven Funktion wiederum bijektiv.

Die Menge $\mathrm{Sym}(M)$ gemeinsam mit der Operation $\circ$ hat die folgenden Eigenschaften:

- Für alle $f, g, h \in \mathrm{Sym}(M)$ gilt $(h \circ g) \circ f = h \circ (g \circ f)$. (*Assoziativgesetz*)
- Für jedes $f \in \mathrm{Sym}(M)$ ist $\mathrm{id}_M \circ f = f \circ \mathrm{id}_M = f$. (*Existenz eines neutralen Elementes*)
- Für jedes $f \in \mathrm{Sym}(M)$ existiert $g \in \mathrm{Sym}(M)$ mit $f \circ g = g \circ f = \mathrm{id}_M$. (*Existenz von Inversen*)

Man nennt $\mathrm{Sym}(M)$ auch die *Symmetrische Gruppe auf M*. Für jede natürliche Zahl n schreibt man auch S_n für die symmetrische Gruppe auf der n-elementigen Menge $\{1, \dots, n\} = \{k \in \mathbb{N} \mid 0 < k \leq n\}$.

Wollen wir uns im Beispiel $M = \{a, b\}$ überlegen wie die Operation $\circ$ agiert und was Inverse sind.

f	g	$f \circ g$
id_M	id_M	id_M
id_M	f_3	f_3
f_3	id_M	f_3
f_3	f_3	id_M

f	f^{-1}
id_M	id_M
f_3	f_3

◊

2.5.16 Definition. Sei M eine Menge, und seien $i, j \in M$ zwei verschiedene Elemente. Dann definieren wir eine Abbildung $\tau^M_{i,j}$ durch Fallunterscheidung so:

$$\tau^M_{i,j}(x) = \begin{cases} i & x = j \\ j & x = i \\ x & \text{sonst} \end{cases}$$

Diese Funktion vertauscht also i und j, aber bildet alle anderen Werte auch sich selbst ab.

Ein Element $t \in \mathrm{Sym}(M)$ heißt *Transposition*, wenn es zwei Elemente $i, j \in M$ gibt mit $i \neq j$, sodass $t = \tau^M_{i,j}$ ist.

Wenn M sich aus dem Kontext ergibt, schreibt man auch oft nur $\tau_{i,j}$ statt $\tau^M_{i,j}$. Überdies verkürzt man $\tau_{i,j}$ auch gerne zu* τ_{ij}. ◊

Wenn man endlich viele Transpositionen hintereinander ausführt, so erhält man jedenfalls eine Permutation; man kann auch umgekehrt beweisen (vielleicht in einer Übungsaufgabe), dass sich jede Permutation einer endlichen Menge (mit mindestens 2 Elementen) als Produkt einer endlichen Anzahl von Transpositionen schreiben lässt. Zum Beispiel ist die durch $f(1) = 2$, $f(2) = 3$, $f(3) = 1$ definierte Abbildung gleich $\tau_{12} \circ \tau_{23}$. (Achtung: nicht $\tau_{23} \circ \tau_{12}$!)

J251.

2.5.17 Bemerkung. Allgemein nennt man ein Tripel $\langle G, \cdot, e\rangle$ eine *Gruppe*, wenn G eine Menge ist, $\cdot$ eine Funktion von $G \times G$ nach G, und e ein Element von G, sodass die folgenden Axiome gelten:

(Gru1) Die Operation $\cdot$ ist assoziativ.

$$\left[\ \forall x, y, z \in G\colon\ (x \cdot y) \cdot z = x \cdot (y \cdot z)\ \right]$$

(Gru2) Das Element e ist neutrales Element für $\cdot$.

$$\left[\ \forall x \in G\colon\ x \cdot e = e \cdot x = x\ \right]$$

*Hier werden i und j nicht multipliziert sondern stehen nur nebeneinander, so wie Längen- und Breitengrad eines Punktes auf der Erde.

(Gru3) Jedes Element von G besitzt ein Inverses bzgl. $\cdot$.

$$\left[\ \forall x \in G\, \exists y \in G\colon\ x \cdot y = y \cdot x = e\ \right]$$

◊

♲ Gruppen sind eine der wichtigsten algebraischen Strukturen in der Mathematik. Einerseits treten sie in unzähligen Zusammenhängen und in allen möglichen Gebieten auf. Andererseits ist es nahezu unglaublich, wieviel man aus diesen drei einfachen Axiomen **(G1)**–**(G3)** bauen und folgern kann. ♲

Kapitel 3

Die natürlichen Zahlen

3.1 Der Zahlbegriff

3.1.1 Motivation

Der Begriff der natürlichen Zahl kommt von der unmittelbar einsichtigen Tätigkeit des Zählens. Das Zählen beginnt mit dem eine Einheit bezeichnenden Zahlwort

`Eins`.

Danach geht man Schritt für Schritt jeweils von einer natürlichen Zahl n zur nächstgrößeren $\hat{n}$ über. Wir verwenden die Zahlworte `Zwei`, `Drei`, `Vier` für die sukzessive gebildeten *Nachfolger* der Einheit `Eins`. Weiters wollen wir das Zahlwort `Null` dafür verwenden, auszudrücken das nichts da ist was man zählen könnte, und die Einheit `Eins` als den Nachfolger von Nichts (der `Null`) verstehen.

Es ergibt sich eine nie endende Folge:

`Null` $\mapsto$ `Eins` $\mapsto$ `Zwei` $\mapsto$ `Drei` $\mapsto$ `Vier` $\mapsto$ ... $\mapsto$ n $\mapsto$ $\hat{n}$ $\mapsto$ ...

Die *unendliche Gesamtheit* aller so erhaltenen Zahlen ist dann die Menge der natürlichen Zahlen.

Das Zählen einer Herde von Schafen funktioniert hervorragend, wenn die Tiere alle schön in einer Reihe stehen. Die einfachste Weise eine natürliche Zahl aufzuschreiben, ist es also die entsprechende Anzahl von Strichen (oder Punkten, oder Schafen, etc.) nebeneinanderzusetzen:

`Eins`: | `Zwei`: || `Drei`: ||| `Vier`: |||| ..... $\hat{n}$: n| .....

Die in diesem Prozess entstehenden Zahlen haben die folgenden unmittelbar einsichtigen Eigenschaften:

(i) Jede Zahl hat genau einen Nachfolger, und verschiedene Zahlen haben auch verschiedene Nachfolger.

(ii) Die `Null` ist nicht Nachfolger irgendeiner Zahl.

(iii) Mit Ausnahme der `Null` ist jede Zahl der Nachfolger einer Zahl.

(iv) Die Gesamtheit aller Zahlen wird durch den schrittweisen Akt des Zählens vollständig ausgeschöpft.

Ebenso einsichtig, wenn auch von seiner Natur her komplexer, ist das *Dominoprinzip* (in mathematischer Ausdrucksweise, das *Induktionsprinzip*).

(v) Man stelle sich die Gesamtheit der natürlichen Zahlen vor als eine Reihe von Dominosteinen die, von der `Null` beginnend, einer nach dem anderen hintereinander aufgestellt werden. Sie seien so knapp beisammen, dass für jeden Dominostein gilt: *wenn* er umfällt, wirft er *auch den nächsten* um.

Wenn nun jemand den allerersten Dominostein (die `Null`) anschubst, fallen alle Steine in der ganzen unendlich langen Reihe um.

3.1.2 Peano Axiome

Wir wollen die in §3.1.1 genannten Eigenschaften (i), (ii), und (v), in Formeln gießen und als charakterisierende Eigenschaften des Zählens – eigentlich, der Zahlen – hernehmen.

J306. **3.1.1 Definition.** Ein Tripel $\langle X, x_0, f\rangle$ heißt *natürliche Zahlen*, wenn X eine Menge ist, x_0 ein Element von X, und f eine Funktion von X nach X, sodaß die folgenden Axiome gelten.

(Nat1) f ist injektiv.

(Nat2) $x_0 \notin f(X)$.

(Nat3) Ist M eine Teilmenge von X mit $x_0 \in M$ und $f(M) \subseteq M$, so ist $M = X$.
$\left[\ \forall M \subseteq X\colon (x_0 \in M \wedge (\forall x\colon x \in M \Rightarrow f(x) \in M)) \ \Rightarrow\ M = X\ \right]$

◊

Das Axiomensystem **(Nat1)** – **(Nat3)** nennt man die *Peano Axiome* für natürliche Zahlen, und das Axiom **(Nat3)** heißt das *Induktionsaxiom*.

♧ `Analyse von Definition` 3.1.1:
Diese Definition ist wieder einmal ein ganz typisches Beispiel für eine *axiomatische Vorgangsweise*. Wir sagen gar nicht was natürliche Zahlen nun wirklich sind, sondern legen nur fest welche Eigenschaften *etwas* haben muss, damit es den Namen „natürliche Zahlen" verdient. Natürlich waren alle unsere Definitionen im vorangegangenen Kapitel ebenfalls genau von diesem Typ. Nur diesmal verwenden wir einen axiomatischen Zugang um etwas zu beschreiben von dem jeder *glaubt* es sowieso gut zu kennen. Aber wer kann *alle* natürlichen Zahlen aufschreiben ?

Man sieht hier auch sehr schön die Bedeutung von Axiomen als *Grundwahrheiten*. Die Eigenschaften **(Nat1)** – **(Nat3)** sind ja nur die – in logische Formeln gegossenen – Eigenschaften (i), (ii), (v), deren Wahrheit im anschaulichen Akt des Zählens wohl niemand bezweifeln wird. Man erkennt ebenso die Macht des axiomatischen Zugangs zu einer Materie. *Jeder* der diese drei Axiome als wahr akzeptiert (und, wie gesagt, das werden wohl nahezu alle sein), und der unsere Prinzipien des logischen Schließens akzeptiert (und das werden wohl auch fast alle sein), der *muss* alles als wahr akzeptieren was wir aus den Axiomen herleiten; selbst dann, wenn es seiner Anschauung, Meinung, Ideologie, oder Hoffnung, widerspricht.

Man muss sich nun den folgenden Fragen stellen.

– Existiert so ein Objekt überhaupt ?

– Wie viele solche Objekte gibt es ? Wenn es mehr als eines gibt, sind alle in irgendeinem Sinn gleich oder ähnlich ?

– Modelliert diese Definition tatsächlich was wir modellieren wollen (nämlich die Zahlen des Zählens) ?

Die erste Frage werden wir hier nicht behandeln, das würde zu weit führen. Wir wollen uns damit begnügen, wegen der Tatsache dass der Akt des Zählens in unserem Leben ständig auftritt, an die Existenz von natürlichen Zahlen zu *glauben*. Bezüglich der zweiten Frage werden wir zeigen (in Satz 3.1.3), dass natürliche Zahlen im wesentlichen eindeutig bestimmt ist.

Die dritte Frage wollen gleich diskutieren. Kurz gesagt, die Anwort ist Ja. Das ausgezeichnete Element wird die Rolle der `Null` spielen, und die Funktion f entspricht gerade der Nachfolgerabbildung $n \mapsto \hat{n}$. Das Axiom **(Nat1)** besagt das zwei Zahlen gleich sein müssen wenn sie den gleichen Nachfolger haben, und **(Nat2)** besagt dass die Zahl `Null` nicht der Nachfolger irgendeiner Zahl ist. Schließlich werden wir sehen (in Satz 3.1.6) dass **(Nat3)** genau dem Prinzip der vollständigen Induktion entspricht. ♧

♤ Oft ist es auch üblich das ausgezeichnete Element von X als Zahl `Eins` zu interpretieren. An dieser Stelle macht das noch keinen Unterschied. Später, wenn man Addition und Multiplikation natürlicher Zahlen betrachtet, dann schon (wenn auch keinen wesentlichen).

Vom Standpunkt der Motivation aus der realen Welt ist die Interpretation als `Eins` wahrscheinlich die passendere: Ist es *natürlich* zu zählen, wenn gar nichts da ist ? Tatsächlich könnte man sagen, dass das Anerkennen von *Nichts* als existentes Objekt ein signifikanter philosophischer Schritt ist.

Warum wollen wir dann unser ausgezeichnetes Element als `Null` interpretieren ? Einerseits weil es oft praktischer ist. Andererseits weil es irgendwie besser mit der üblichen Konstruktion einer Menge natürlicher Zahlen

zusammenpasst — diese Motivation bleibt uns hier allerdings verschlossen, weil wir den Existenzbeweis gar nicht führen (können).

Es mag verwundern, dass wir die in §3.1.1 auch genannten Eigenschaften (iii) und (iv) nicht in unser Axiomensystem für natürliche Zahlen mit aufgenommen haben. Nun ist es so, dass diese Eigenschaft bereits aus den drei Peano Axiomen folgt (wir werden dies in Proposition 3.1.5 beweisen).

Wir hätten also, ohne *den Begriff ansich* zu verändern, die Eigenschaften (iii), (iv) auch als „viertes und fünftes Axiom" fordern können: Für jedes Tripel ist es äquivalent ob es die Peano Axiome erfüllt, oder diese und noch dazu (iii), (iv).

Nun ist es aber sinnvoller und (zugegebenermaßen Geschmackssache) ästhetischer ein Axiomensystem so zu wählen, dass kein Axiom schon aus den anderen hergeleitet werden kann. Sprich, dass kein Axiom unnötig ist oder, anders ausgedrückt, dass jedes Axiom von den anderen unabhängig ist. Wenn man keine unnötigen Axiome mitschleppt, kommt wohl am klarsten heraus was der Kern der Sache ist (und man kann die Allgemeinheit leichter davon überzeugen die Axiome zu akzeptieren).

Der erste Satz über natürliche Zahlen den wir beweisen wollen, ist der *Rekursionssatz*. Er macht eine ganz grundlegende (und ziemlich nicht-triviale) Existenzaussage.

J308. **3.1.2 Satz.** *Seien $\langle X, x_0, f\rangle$ natürliche Zahlen. Weiters sei Y eine Menge, $y_0 \in Y$, und $g : Y \to Y$. Dann existiert genau eine Funktion $\phi : X \to Y$ mit $\phi(x_0) = y_0$ und $\phi \circ f = g \circ \phi$.*

$$\begin{array}{ccc} X & \overset{\phi}{\dashrightarrow} & Y \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle g} \\ X & \underset{\phi}{\dashrightarrow} & Y \end{array}$$

Der Beweis wird in vier Schritten verlaufen.

1. Konstruktion von ϕ.
2. Nachweis dass ϕ eine Funktion ist.
3. Nachweis der verlangten Eigenschaften von ϕ.
4. Beweis der Eindeutigkeit.

Die Konstruktion in Schritt 1 fällt erstmal vom Himmel; warum man ϕ so baut wie man man es tut, werden wir in der nach dem Beweis folgenden Beweisanalyse diskutieren. Die Beweise in den Schritten 2 – 4 werden stets eine Anwendung des Induktionsaxioms **(Nat3)** sein.

> ➢ `Was heißt hier -- umgangssprachlich -- vom Himmel fallen:`
>
> Damit meint man meist eine Konstruktion oder Definition die nicht offensichtlich ist und, scheinbar motivationslos, aus dem Nichts auftaucht.
>
> Liegt eine Konstruktion, ein Beweis, oder auch eine ganze Theorie, erstmal in einer etablierten schriftlichen Form vor, so ist sie meist glattgebügelt und aufpoliert, hoffentlich fehlerbereinigt, und – wenn man ein bisschen zynisch sein möchte – oft eben auch motivationsbefreit. Man sieht nicht mehr alle jene Sackgassen und Irrwege die beschritten wurden, alle Fehler die gemacht wurden, und manchmal eben auch nicht mehr welche Anschauung und Fragen aus der realen oder abstrakt mathematischen Welt zur Entwicklung der Theorie Anlass gegeben haben. Ein intuitives, menschliches, aber eben auch fehleranfälliges, Konstrukt hat sich im Laufe der Zeit gewandelt zu einem abstrakten, formallogischen, und nachweisbar unumstößlichen Gebäude.
>
> Eigentlich ist es auch Aufgabe des Autors mathematischer Literatur dem Leser mitzugeben woher alle die studierten Begriffe kommen, und welche Bedeutung und Konsequenzen sie in der realen Welt, oder in anderen mathematischen Theorien, oder in anderen Wissenschaften, etc., haben. Oft fällt das aber unter den Tisch.
>
> Tipp: Wenn man damit scheitert aus seiner Anschauung und Erfahrung heraus Motivation für eine Theorie zu erkennen, und die Intuition des Autors beim besten Willen nicht nachvollziehen kann, erstmal kritiklos akzeptieren, eine Zeitlang mit den Begriffen arbeiten, und sich in die Welt dieser Theorie hineinleben. Wenn man sich einmal in einer Theorie zu Hause fühlt, kann man sicherlich ihre Ästhetik, wahrscheinlich ihre – unter Umständen weit zurückliegenden – Wurzeln, sehen, und hoffentlich ihre Bedeutung und Konsequenzen für Anderes verstehen.

Beweis von Satz 3.1.2. Wir machen die oben genannten vier Schritte.

Schritt 1; Konstruktion von ϕ: Betrachte die Menge $\mathcal{M}$ aller jener Teilmengen A von $X \times Y$ die die beiden folgenden Eigenschaften haben:

$$(x_0, y_0) \in A. \tag{3.1.1}$$ J330

$$\text{Ist } (x, y) \in A, \text{ so ist auch } (f(x), g(y)) \in A. \tag{3.1.2}$$ J331

Die Menge $\mathcal{M}$ ist nichtleer, denn $X \times Y$ hat offensichtlich die Eigenschaften (3.1.1) und (3.1.2). Wir können daher den Durchschnitt von $\mathcal{M}$ betrachten; bezeichne $\phi = \bigcap_{A \in \mathcal{M}} A$.

Wollen wir uns überlegen, dass die Menge ϕ selbst ebenfalls die beiden Eigenschaften (3.1.1) und (3.1.2) besitzt.

– Wir zeigen (3.1.1):
Das Paar (x_0, y_0) ist Element jedes Elementes von $\mathcal{M}$ und damit auch Element des Durchschnitts.
– Wir zeigen (3.1.2):
Sei $(x, y) \in \phi$. Ist uns $A \in \mathcal{M}$ beliebig vorgegeben, so haben wir wegen $\phi \subseteq A$ dass $(x, y) \in A$ und da A (3.1.2) erfüllt daher auch $(f(x), g(y)) \in A$. Da $A \in \mathcal{M}$ beliebig war folgt $(f(x), g(y)) \in \phi$.

Schritt 2; Nachweis dass ϕ eine Funktion ist: Dazu wollen wir nachweisen dass ϕ die Eigenschaft **(Fun)** erfüllt. Sprich, wir gehen darauf los zu zeigen, dass es für jedes $x \in X$ genau ein Element $y \in Y$ gibt mit $(x, y) \in \phi$. Man erinnere sich an dieser Stelle daran was „es gibt genau ein“ bedeutet. Betrachte die Menge

✧ $M = \{x \in X : \exists!\, y \in Y \colon (x, y) \in \phi\}$.

Unser Ziel ist es also zu zeigen, dass $M = X$. Um dies zu tun werden wir uns des Axioms **(Nat3)** bedienen.

Um die Implikation **(Nat3)** anwenden zu können, müssen wir zeigen dass ihre Prämisse wahr ist. Sprich, wir müssen zeigen, dass $x_0 \in M$ und $\forall x \in X \colon x \in M \Rightarrow f(x) \in M$ gelten.

– *Prämisse von* **(Nat3)** *Teil 1; $x_0 \in M$:*
Zunächst gilt $(x_0, y_0) \in \phi$ wegen (3.1.1). Sei $(x_0, z) \in \phi$; wir müssen zeigen dass $z = y_0$.
Betrachte die Menge $B = \phi \setminus \{(x_0, z)\}$. Als erstes bemerken wir, dass $B \subsetneq \phi$ da $(x_0, z) \in \phi$. Als zweites zeigen wir, dass B die Eigenschaft (3.1.2) erfüllt. Dazu sei $(x, y) \in B$. Dann ist auch $(x, y) \in \phi$, und daher $(f(x), g(y)) \in \phi$. Nun ist wegen **(Nat2)** sicher $f(x) \neq x_0$, insbesondere also $(f(x), g(y)) \neq (x_0, z)$, und wir erhalten $(f(x), g(y)) \in B$.
Für jede Menge $A \in \mathcal{M}$ gilt $\phi \subseteq A$, also ist $B \notin \mathcal{M}$. Da wir schon wissen dass B die Eigenschaft (3.1.2) erfüllt, kann B die Eigenschaft (3.1.1) nicht erfüllen. Also gilt $(x_0, y_0) \in \phi \setminus B = \{(x_0, z)\}$. Wir sehen dass $(x_0, y_0) = (x_0, z)$, und daher $y_0 = z$.
– *Prämisse von* **(Nat3)** *Teil 2; $\forall x \in X \colon x \in M \Rightarrow f(x) \in M$:*
Sei uns $x \in X$ vorgegeben, und sei vorausgesetzt dass $x \in M$, sprich, dass

$$\exists!\, y \in Y \colon (x, y) \in \phi. \tag{3.1.3}$$ J333

Wir müssen zeigen, dass $f(x) \in M$. Als erstes wähle $y \in Y$ mit $(x, y) \in \phi$; wegen (3.1.3) ist das möglich. Da ϕ ja (3.1.2) erfüllt, folgt dass $(f(x), g(y)) \in \phi$, und wir haben eine zweite Komponente für $f(x)$ gefunden sodass das daraus gebildete Paar in ϕ liegt.
Wir müssen zeigen, dass $z = g(y)$ für alle $(f(x), z) \in \phi$. Das machen wir mittels Beweis durch Kontraposition. Sei vorausgesetzt, dass $z \neq g(y)$. Das folgende Argument ist recht ähnlich dem in Teil 1 verwendeten.
Betrachte die Menge $B = \phi \setminus \{(f(x), z)\}$. Als erstes bemerken wir, dass B (3.1.1) erfüllt, denn $(x_0, y_0) \in \phi$ aber $x_0 \neq f(x)$, insbesondere $(x_0, y_0) \neq (f(x), z)$. Als zweites zeigen wir, dass B die Eigenschaft (3.1.2) erfüllt. Dazu sei $(x', y') \in B$. Dann ist auch $(x', y') \in \phi$, und daher $(f(x'), g(y')) \in \phi$. Wir machen eine Fallunterscheidung.

✾ *Fall 1; $x' \neq x$.* In diesem Fall ist wegen **(Nat1)** auch $f(x') \neq f(x)$, und insbesondere $(f(x'), g(y')) \neq (f(x), z)$. Damit erhalten wir $(f(x'), g(y')) \in B$.
✾ *Fall 2; $x' = x$.* In diesem Fall ist $y' = y$ wegen (3.1.3). Damit haben wir $(f(x'), g(y')) = (f(x), g(y))$. Nun ist, nach unserer Voraussetzung, $g(y) \neq z$, insbesondere $(f(x), g(y)) \neq (f(x), z)$. Wieder erhalten wir $(f(x'), g(y')) \in B$.

Wir sehen dass tatsächlich B die Eigenschaften (3.1.1) und (3.1.2) hat. Es ist also $B \in \mathcal{M}$, und damit $B \supseteq \phi$. Nun gilt per definitionem $B \subseteq \phi$, und wir schließen, dass $B = \phi$. Wieder per definitionem gilt $(f(x), z) \notin B$, also ist $(f(x), z) \notin \phi$.

Der Beweis der Prämisse von **(Nat3)** ist abgeschlossen. Wir können also **(Nat3)** anwenden, schließen dass $M = X$, und damit dass **(Fun)** für ϕ gilt.

Schritt 3; Nachweis der verlangten Eigenschaften von ϕ: Da ϕ die Eigenschaft (3.1.1) hat, ist $(x_0, y_0) \in \phi$. Benützt man die Schreibweise einer Funktion als Zuordnung, heißt das gerade $\phi(x_0) = y_0$.

Um die Gleichheit $\phi \circ f = g \circ \phi$ zu zeigen, erinnern wir uns wieder daran das zwei Funktionen mit dem gleichen Definitions- bzw. Zielmengen genau dann gleich sind, wenn sie jedem Element ihres gemeinsamen Definitionsbereiches das selbe Element ihrer gemeinsamen Zielmenge zuordnen. Der Definitionsbereich von $\phi \circ f$ sowie auch von $g \circ \phi$ ist X und die Zielmenge ist Y. Wir gehen nun darauf los zu zeigen, dass für jedes $x \in X$ die Gleichheit $(\phi \circ f)(x) = (g \circ \phi)(x)$ gilt. Dazu werden wir wieder **(Nat3)** verwenden.

Betrachte die Menge

✧ $M = \{x \in X : \ (\phi \circ f)(x) = (g \circ \phi)(x)\}$.

– *Prämisse von* **(Nat3)** *Teil 1; $x_0 \in M$:*
Da ϕ die Eigenschaft (3.1.2) hat, und $(x_0, y_0) \in \phi$ ist wegen (3.1.1), folgt dass auch $(f(x_0), g(y_0)) \in \phi$, sprich, $\phi(f(x_0)) = g(y_0)$. Unter Verwendung der bereits bekannten Eigenschaft $\phi(x_0) = y_0$, erhalten wir

$$(\phi \circ f)(x_0) = \phi(f(x_0)) = g(y_0) = g(\phi(x_0)) = (g \circ \phi)(x_0).$$

Also ist $x_0 \in M$.

– *Prämisse von* **(Nat3)** *Teil 2; $\forall x \in X\colon\ x \in M \Rightarrow f(x) \in M$:*
Sei uns $x \in X$ vorgegeben, und sei vorausgesetzt dass $x \in M$, sprich, dass

$$(\phi \circ f)(x) = (g \circ \phi)(x). \tag{3.1.4}$$ J332

Diese Gleichheit besagt gerade dass $(f(x), (g \circ \phi)(x)) \in \phi$. Da ϕ die Eigenschaft (3.1.2) hat, folgt

$$\big(f(f(x)), g((g \circ \phi)(x))\big) \in \phi,$$

was nichts anderes besagt als dass $\phi(f(f(x))) = g((g \circ \phi)(x))$. Wir erhalten, unter Verwendung der Voraussetzung (3.1.4), dass

$$(\phi \circ f)(f(x)) = \phi(f(f(x))) = g((g \circ \phi)(x)) \overset{(3.1.4)}{=} g((\phi \circ f)(x)) = (g \circ \phi)(f(x)).$$

Also ist $f(x) \in M$.

Der Beweis der Prämisse von **(Nat3)** ist abgeschlossen. Wir können also **(Nat3)** anwenden, schließen dass $M = X$, und damit dass $\phi \circ f = g \circ \phi$.

Schritt 4; Beweis der Eindeutigkeit: Sei uns eine Funktion $\phi' : X \to Y$ gegeben, welche die Eigenschaften $\phi'(x_0) = y_0$ und $\phi' \circ f = g \circ \phi'$ hat. Unser Ziel ist zu zeigen dass $\phi = \phi'$. Offenbar haben ϕ und ϕ' die gleichen Definition- und Zielmengen, es läuft also (wie meistens) darauf hinaus zu zeigen dass $\phi(x) = \phi'(x)$ für alle $x \in X$. Dazu werden wir wieder **(Nat3)** verwenden.

Betrachte die Menge

✧ $M = \{x \in X : \ \phi(x) = \phi'(x)\}$.

– *Prämisse von* **(Nat3)** *Teil 1; $x_0 \in M$:*
Es ist $x_0 \in M$ da ja $\phi(x_0) = y_0 = \phi'(x_0)$.

– *Prämisse von* **(Nat3)** *Teil 2; $\forall x \in X\colon\ x \in M \Rightarrow f(x) \in M$:*
Sei uns $x \in X$ vorgegeben, und sei vorausgesetzt dass $x \in M$, sprich, dass

$$\phi(x) = \phi'(x). \tag{3.1.5}$$ J334

Dann erhalten wir, unter Verwendung von (3.1.5), dass

$$\phi(f(x)) = (\phi \circ f)(x) = (g \circ \phi)(x) = g(\phi(x)) \overset{(3.1.5)}{=} g(\phi'(x)) = (g \circ \phi')(x) = (\phi' \circ f)(x) = \phi'(f(x)),$$

also $f(x) \in M$.

Der Beweis der Prämisse von **(Nat3)** ist abgeschlossen. Wir können also **(Nat3)** anwenden, schließen dass $M = X$, und damit dass $\phi = \phi'$.

□

Dieser Beweis ist recht lang, vollgepackt mit verschiedensten Typen logischer Schlussweisen, und verlangt obendrein Kreativität (bzw. Intuition). Wir wollen daher einige Details diskutieren.

♧ `Analyse des Beweises von Satz 3.1.2, Schritt 1:`
Was in Schritt 1 passiert ist zwar nicht anspruchsvoll, aber dafür umso essentieller: wir werfen die Definition jenes Objektes hin, welches sich als das zu konstruierende herausstellen wird (nämlich von ϕ). Dann rechnen wir noch nach, dass sich die beiden Eigenschaften (3.1.1) und (3.1.2) auf Durchschnitte vererben, aber das ist straightforward. Die interessante Frage ist:

Wie kommt man darauf ϕ so zu definieren wie man es tut ?

Dazu betrachten wir einmal die beiden Eigenschaften (3.1.1) und (3.1.2).

– Das man $(x_0, y_0) \in \phi$, sprich (3.1.1), braucht ist klar. Denn ϕ soll ja die Eigenschaft $\phi(x_0) = y_0$ haben.

– Weiters wollen wir erzwingen, dass $(\phi \circ f)(x) = (g \circ \phi)(x)$ gilt. Es muss also für jedes x gelten, dass $\phi(\, f(x)\,) = g(\, \phi(x)\,)$, umgeschrieben $(f(x), g(\, \phi(x)\,)) \in \phi$.

  Bezeichnet man nun $y = \phi(x)$, so hat man dass $(f(x), g(\, y\,)) \in \phi$ wann immer $y = \phi(x)$, umgeschrieben $(f(x), g(\, y\,)) \in \phi$ wann immer $(x, y) \in \phi$, und das ist gerade die Eigenschaft (3.1.2).

Wir sehen also, dass das *zu konstruierende* ϕ – wenn es denn überhaupt existiert – sicherlich (3.1.1) und (3.1.2) erfüllen *muss*.

Relationen mit (3.1.1) und (3.1.2) gibt es sicher, zum Beispiel $X \times Y$. Nun muss man aber versuchen eine Relation mit (3.1.1) und (3.1.2) zu bauen, die sogar eine Funktion ist (also sicher nicht $X \times Y$ nehmen). Die beiden Eigenschaften die Funktionen unter allen Relationen auszeichnen sind **(Fun1)** und **(Fun2)**. Diese beiden sind in gewissem Sinne gegenläufig.

– **(Fun1)** sagt, dass es *hinreichend viele* Paare in der Relation gibt: Für alle $x \in X$ gibt es eine zweite Komponente $y \in Y$, sodass das Paar (x, y) in der Relation liegt.

– **(Fun2)** sagt, dass es *nicht zu viele* Paare in der Relation geben darf: Es dürfen keine zwei verschiedenen Paare mit der gleichen ersten Komponente in der Relation liegen.

Dabei ist **(Fun2)** die essentielle Eigenschaft. Hat man **(Fun2)**, so kann man immer erweitern und **(Fun1)**, unter Beibehaltung von **(Fun2)**, erzwingen; man erinnere sich an den Beweis von Satz 2.5.10(1).

In der hier vorliegenden Situation brauchen wir uns – anschaulich gesprochen(!) – um **(Fun1)** gar nicht kümmern. Weil, es sollte durch (3.1.1) und (3.1.2) gewährleistet sein. Wenn man an die in §3.3.1 als (iv) genannte Eigenschaft denkt, und wenn tatsächlich alles so sein sollte wie es unsere Intuition verspricht, dann bekommt man aus (x_0, y_0) schrittweise die Paare $(f(x_0), g(y_0))$, $(f(f(x_0)), g(g(y_0)))$, usw., und schöpft damit alle möglichen ersten Komponenten aus.

 Der letzte Absatz ist so richtige Kaffeesudleserei! Wir wissen ja nicht ob es überhaupt eine tatsächlich geltende Entsprechung für (iv) gibt! Aber wir wollen ja „nur“ motivieren warum wir die Definition von ϕ gerade so machen wie wir es machen.

Also . . . glauben wir einmal an das Gute (und hoffen wir dass es auch wirklich durchgeht).

Wir konzentrieren uns daher auf die Eigenschaft **(Fun2)**. Diese sagt ja, dass es nicht zu viele Paare geben darf – je weniger, desto leichter wird es dass **(Fun2)** erfüllt sein könnte. Also:

Nehmen wir für ϕ die kleinstmögliche Relation die (3.1.1) *und* (3.1.2) *erfüllt.*

Gibt es sowas überhaupt ? Ja! Nämlich den Durchschnitt über alle Relationen die (3.1.1) und (3.1.2) erfüllen. ♧

♧ `Analyse des Beweises von Satz 3.1.2, Schritt 2:`
Dieser Beweisteil ist die Crux (und – naturgemäß – der mit den anspruchsvollsten Argumentationen).

Wir müssen zeigen, dass ϕ **(Fun1)** und **(Fun2)** erfüllt. Es ist praktischer gleich auf **(Fun)** loszugehen. Dazu verwenden wir das Induktionsaxiom mit der in offensichtlicher Weise definierten Menge M (nämlich die Menge aller jener Elemente x wo das gilt was wir haben wollen). Nun müssen wir die Prämisse von **(Nat3)**, welche aus zwei Teilen besteht, nachweisen.

❶ *Teil 1:* Die Existenz einer zweiten Komponente für x_0 ist klar; wegen (3.1.1) ist ja $(x_0, y_0) \in \phi$. Interessant ist die Eindeutigkeit der zweiten Komponente. Man bemerke: dies korrespondiert dazu, dass **(Fun2)** die Essenz der Funktionseigenschaft ist.

Was heißt Eindeutigkeit zu zeigen: Wir müssen zeigen dass

$$\forall z\colon (x_0, z) \in \phi \Rightarrow z = y_0$$

Starten wir also mit der

– *Voraussetzung:* $(x_0, z) \in \phi$

und zeigen wir die

– *Konklusio:* $z = y_0$

Dazu betrachten wir die Menge $B = \phi \setminus \{(x_0, y_0)\}$. Wegen $(x_0, z) \in \phi$, ist B offensichtlich eine echte Teilmenge von ϕ. Nun ist ϕ die kleinste Relation mit (3.1.1), (3.1.2), sprich es gilt (für beliebiges A)

$$A \text{ erfüllt (3.1.1)} \wedge A \text{ erfüllt (3.1.2)} \quad \Rightarrow \quad A \supseteq \phi$$

Die Kontraposition davon ist

$$A \text{ erfüllt (3.1.1) nicht} \vee A \text{ erfüllt (3.1.2) nicht} \quad \Leftarrow \quad A \not\supseteq \phi$$

Wegen $A \not\supseteq \phi \Leftarrow A \subsetneq \phi$ haben wir insbesondere (Modus Barabara)

$$A \text{ erfüllt (3.1.1) nicht} \vee A \text{ erfüllt (3.1.2) nicht} \quad \Leftarrow \quad A \subsetneq \phi$$

Wir sehen also, dass (Modus Ponendo Ponens) „B erfüllt (3.1.1) nicht $\vee$ B erfüllt (3.1.2) nicht“.

Was wir dann zeigen, ist „B erfüllt (3.1.2)“. Damit folgt (Modus Tolendo Ponens), dass „B erfüllt (3.1.1) nicht“ gilt. Da von B auf ϕ ja nur ein Element fehlt, nämlich (x_0, z), und (x_0, y_0) in ϕ ist aber nicht in B, erhält man nun unmittelbar $(x_0, z) = (x_0, y_0)$. Also ist $z = y_0$.

❷ *Teil 2:* Die Existenz einer zweiten Komponente für $f(x)$ ist klar; wegen (3.1.2) ist ja $(f(x), g(y)) \in \phi$ wobei y so gwählt wurde dass $(x, y) \in \phi$. Interessant ist – wiederum – die Eindeutigkeit der zweiten Komponente.

Was heißt Eindeutigkeit zu zeigen: Wir müssen zeigen dass

$$\forall z\colon (f(x), z) \in \phi \Rightarrow z = g(y)$$

Die Idee ist die Gleiche wie oben (nämlich die Menge $B = \phi \setminus \{(f(x), z)\}$ zu betrachten), und das tatsächliche Argument ist auch ein ganz ähnlicher Schmäh. Aber der logische Ablauf des Argumentes ist ein anderer, nämlich machen wir Beweis durch Kontraposition.

Wir starten mit der

– *Voraussetzung:* $z \neq g(y)$

und zeigen die

– *Konklusio:* $(f(x), z) \notin \phi$

Dazu betrachten wir die Menge $B = \phi \setminus \{(f(x), z)\}$. Offensichtlich erfüllt B (3.1.1). Was wir dann zeigen, ist dass B auch (3.1.2) erfüllt. Wir erhalten $B \supseteq \phi$, und damit $B = \phi$. Da B das Element $(f(x), z)$ nicht enthält, liegt es also (auch) nicht in ϕ.

♧

♧ `Analyse des Beweises von Satz 3.1.2, Schritt 3:`

Dieser Schritt ist eigentlich nur Routine. Wir müssen nachrechnen, dass die Funktion ϕ die verlangten Eigenschaften hat. Diese sind aber ohnehin schon – salopp gesagt – in die Definition hineingepackt.

Um sie tatsächlich zu beweisen verwenden wir das Induktionsaxiom mit in offensichtlicher Weise definierten Mengen M, nämlich genau den Mengen wo das gilt was wir haben wollen. Auch das Nachprüfen der Prämisse von **(Nat3)** ist straightforward. Was dabei passiert ist nur das hin-und-her Umschreiben von der Notation einer Funktion als Relation bzw. Zuordnung, einsetzten der Definitionen, und einmaliges verwenden der Prämisse der Implikation in der Prämisse von **(Nat3)**. ♧

♧ `Analyse des Beweises von Satz 3.1.2, Schritt 4:`

Um den Eindeutigkeitsbeweis zu führen, erinnern wir uns an die Bedeutung der Schreibweise $\exists!$, nämlich

$$\Big[\ \exists\, x\colon \big(A(x) \wedge (\forall y\colon A(y) \Rightarrow y = x)\big)\ \Big]$$

Wir haben ja schon eine Funktion ϕ gefunden die die verlangten Eigenschaften hat. Es bleibt also zu zeigen, dass jede Funktion ϕ' mit den verlangten Eigenschaften schon gleich ϕ ist. Danach ist auch dieser Schritt Routine. Wir verwenden das Induktionsaxiom in offensichtlicher Weise, wobei das Nachprüfen der Prämisse von **(Nat3)** wieder straightforward ist.

In der durchgeführten Argumentation ist nirgends eingegangen, dass die Funktion ϕ tatsächlich genau die in Schritt 1 konstruierte ist, sondern nur dass sie die verlangten Eigenschaften hat. Wir könnten uns also auch auf die äquivalente Beschreibung von $\exists!$ als

$$\Big[\ (\exists\, x\colon A(x)) \wedge (\forall x, y\colon A(x) \wedge A(y) \Rightarrow x = y)\ \Big]$$

berufen. Der Eindeutigkeitsbeweis verliefe dann also nach dem Plan: Seien ϕ, ϕ' zwei Funktionen mit den verlangten Eigenschaften gegeben, dann müssen wir $\phi = \phi'$ zeigen (und die durchgeführte Induktion tut genau dieses). ♧

In der Situation von Satz 3.1.2 sagt man auch die Funktion ϕ ist durch die Gleichung $\phi \circ f = g \circ \phi$ *rekursiv definiert*.

Mit Hilfe des Rekursionssatzes erhalten wir nun leicht den folgenden *Eindeutigkeitssatz.*

J309. **3.1.3 Satz.** *Seien $\langle X, x_0, f\rangle$ und $\langle Y, y_0, g\rangle$ natürliche Zahlen. Dann existiert eine eindeutige Funktion $\phi : X \to Y$ mit $\phi(x_0) = y_0$ und $\phi \circ f = g \circ \phi$. Diese Funktion ist bijektiv.*

Beweis. Die erste Aussage, dass eine Funktion ϕ mit den beiden genannten Eigenschaften existiert und eindeutig ist, erhält man gerade indem man den Rekursionssatz anwendet auf die natürlichen Zahlen $\langle X, x_0, f\rangle$, und die Menge Y mit ihrem Element y_0 und der Funktion g.

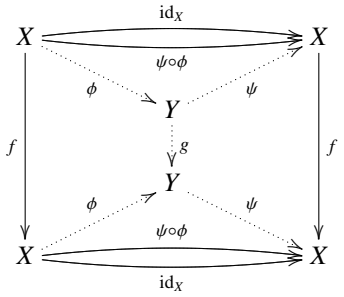
$$\begin{array}{ccc} X & \overset{\phi}{\dashrightarrow} & Y \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle g} \\ X & \underset{\phi}{\dashrightarrow} & Y \end{array}$$

Wir müssen noch zeigen dass ϕ bijektiv ist. Dazu gehen wir darauf los eine Inverse zu ϕ zu konstruieren. Wendet man den Rekursionssatz an auf die natürlichen Zahlen $\langle Y, y_0, g\rangle$, und die Menge X mit ihrem Element x_0 und der Funktion f, so erhält man eine Funktion $\psi : Y \to X$ mit $\psi(y_0) = x_0$ und $\psi \circ g = f \circ \psi$.

$$\begin{array}{ccc} Y & \overset{\psi}{\dashrightarrow} & X \\ {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle f} \\ Y & \underset{\psi}{\dashrightarrow} & X \end{array}$$

Die Funktionen $\mathrm{id}_X : X \to X$ und $\psi \circ \phi : X \to X$ haben die Eigenschaften $\mathrm{id}_X(x_0) = x_0$ und $\mathrm{id}_X \circ f = f \circ \mathrm{id}_X$ bzw.

$(\psi \circ \phi)(x_0) = x_0$ und $(\psi \circ \phi) \circ f = f \circ (\psi \circ \phi)$.



Nach der Eindeutigkeitsaussage im Rekursionssatz, angewendet mit den natürlichen Zahlen $\langle X, x_0, f\rangle$, und der Menge X mit ihrem Element x_0 und der Funktion f, folgt $\mathrm{id}_X = \psi \circ \phi$. Das analoge Argument, angewendet mit id_Y und $\phi \circ \psi$ zeigt $\mathrm{id}_Y = \phi \circ \psi$. Wir haben also tatsächlich eine Inverse von ϕ gefunden, nämlich ψ. □

♲ `Analyse von Satz` 3.1.3`:`

Der Eindeutigkeitssatz besagt dass je zwei Tripel die natürliche Zahlen sind, sich nur durch eine Umbenennung (die bijektive Funktion ϕ) unterscheiden. Diese Umbenennung ist mit den ausgezeichneten Elementen verträglich ($\phi(x_0) = y_0$). Sprich, die `Null` ist hier wie da das Gleiche. Und die Umbenennung ist auch mit den Nachfolgerabbildungen verträglich; es ist egal ob man in X den Nachfolger bildet und dann umbenennt, oder zuerst umbenennt und dann in Y den Nachfolger bildet ($\phi \circ f = g \circ \phi$). Sprich, auch von einer Zahl zur nächsten übergehen ist hier wie da das Gleiche.

Wir können also mit Fug und Recht sagen, dass natürliche Zahlen *im wesentlichen eindeutig* sind. Daher spricht man von *den natürlichen Zahlen*, und führt auch eine eigene Notation ein: $\mathbb{N}$ für die Menge, und 0 für das ausgezeichnete Element. Die Nachfolgerabbildung wollen wir im Folgenden mit S bezeichnen. Weiters bezeichnen wir den Nachfolger $S(0)$ des ausgezeichneten Elementes mit 1, den Nachfolger von 1 mit 2, und den Nachfolger von 2 mit 3; entsprechend den Zahlworten `Eins`, `Zwei`, und `Drei`.

Bemerke dass, wegen **(Nat1)** und **(Nat2)**, die Elemente 0, 1, 2, und 3, paarweise verschieden sind. Es gilt $0 \neq 1, 2, 3$ da $1, 2, 3 \in S(\mathbb{N})$ aber $0 \notin S(\mathbb{N})$. Da S injektiv ist, erhält man der Reihe nach auch $1 \neq 2$, $1 \neq 3$, und $2 \neq 3$.

Im Folgenden werden wir immer diese Notation verwenden. ♲

> ➢ `Was heißt hier im wesentlichen eindeutig`
>
> Man sagt oft ein Objekt mit vorgeschriebenen Eigenschaften ist *im wesentlichen eindeutig*, wenn es durch diese Eigenschaften zwar nicht im strengen Sinne des Wortes eindeutig bestimmt ist, aber jedes andere Objekt mit diesen Eigenschaften sich nicht wirklich davon unterscheiden kann. Meist heißt das, genau so wie in obigem Fall, dass je zwei Objekte durch eine Umbenennung die noch dazu alle wesentlichen Eigenschaften respektiert ineinander transformiert werden können.
>
> Je nach dem Kontext kann die Phrase *im wesentlichen eindeutig* auch etwas ein bisschen anderes bedeuten, aber jedenfalls ist das obige Beispiel ein gutes Modell.

Als eine ganz typische Anwendung des Rekursionssatzes wollen wir erklären was *Iterierte* einer Funktion sind.

J328. *3.1.4 Beispiel.* Sei M eine Menge und f eine Funktion von M in sich. Dann können wir die folgenden Funktionen konstruieren

$$f, \quad f \circ f, \quad f \circ f \circ f, \quad f \circ f \circ f \circ f, \quad \textit{u.s.w.}$$

Diese entstehen also durch fortgesetztes (man sagt auch iteriertes) Verknüpfen mit f.

Was heißt hier „u.s.w." ? Es mag ja offensichtlich *scheinen* wie diese Phrase zu interpretieren ist, aber genauso wie beim Anschreiben einer Menge mit „…", ist es notwendig „u.s.w." zu präzisieren. Und das ist genau was der Rekursionssatz leistet.

Bezeichne mit $\mathbb{F}(M)$ die Menge aller Funktionen von M in sich. Wir wenden den Rekursionssatz an mit den natürlichen Zahlen $\langle \mathbb{N}, 0, S\rangle$, und der Menge $\mathbb{F}(M)$ mit ihrem Element id_M und der Funktion

$$\Gamma : \left\{ \begin{array}{rcl} \mathbb{F}(M) & \to & \mathbb{F}(M) \\ g & \mapsto & f \circ g \end{array} \right.$$

Das gibt uns eine eindeutige Funktion $\phi : \mathbb{N} \to \mathbb{F}(M)$ mit den Eigenschaften $\phi(0) = \mathrm{id}_M$ und $\phi \circ S = \Gamma \circ \phi$. Die zweite Eigenschaft besagt gerade dass $\phi(S(n)) = f \circ \phi(n)$ für alle $n \in \mathbb{N}$, und wir sehen dass

$$\phi(1) = \phi(S(0)) = f \circ \phi(0) = f \circ \mathrm{id}_M = f, \quad \phi(2) = \phi(S(1)) = f \circ \phi(1) = f \circ f, \quad \text{u.s.w.}$$

Also *präzisiert* die Funktion ϕ tatsächlich was wir uns unter den Iterierten von f *vorstellen*.

Man sagt $\phi(n)$ ist die *n-te Iterierte von* f, und schreibt $f^{(n)}$ für diese Funktion. ◊

Die folgende Aussage erklärt sehr gut die Natur des Induktionsaxioms **(Nat3)**. Sie besagt, dass die natürlichen Zahlen durch fortgesetztes bilden des Nachfolgers vollständig ausgeschöpft werden (man erinnere sich an die Eigenschaften (iii), (iv) in §3.1.1).

J329. **3.1.5 Proposition.** *Es gilt*

(1) $\mathbb{N} = \{0\} \cup S(\mathbb{N})$.

(2) *Für jedes* $n \in \mathbb{N}$ *gilt* $n = S^{(n)}(0)$. *Insbesondere ist* $\mathbb{N} = \{S^{(n)}(0) : n \in \mathbb{N}\}$.

Beweis. Wir erhalten beide Behauptungen wieder durch Anwendung von **(Nat3)**.

Zum Beweis von (1)*:* Betrachte die Menge

✧ $M = \{0\} \cup S(\mathbb{N})$.

– *Prämisse von* **(Nat3)** *Teil 1;* $0 \in M$*:*
Nach Definition von M ist $0 \in M$.

– *Prämisse von* **(Nat3)** *Teil 2;* $\forall n \in \mathbb{N}: n \in M \Rightarrow S(n) \in M$*:*
Sei uns $n \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt dass $n \in M$, sprich, dass

$$n \in \{0\} \cup S(\mathbb{N}). \tag{3.1.6}$$ J335

Nun gilt stets $S(n) \in \mathbb{S}(\mathbb{N}) \subseteq \mathbb{N}$. Also ist die gewünschte Konklusion $S(n) \in M$ wahr (wobei wir die Prämisse (3.1.6) nicht einmal verwendet haben).

Der Beweis der Prämisse von **(Nat3)** ist abgeschlossen. Wir können also **(Nat3)** anwenden, schließen dass $M = \mathbb{N}$, und damit dass die Behauptung (1) gilt.

Zum Beweis von (2)*:* Betrachte die Menge

✧ $M = \{n \in \mathbb{N} : n = S^{(n)}(0)\}$.

– *Prämisse von* **(Nat3)** *Teil 1;* $0 \in M$*:*
Es ist $S^{(0)}(0) = \mathrm{id}_{\mathbb{N}}(0) = 0$ nach der Definition der 0-ten Iterierten, also haben wir $0 \in M$.

– *Prämisse von* **(Nat3)** *Teil 2;* $\forall n \in \mathbb{N}: n \in M \Rightarrow S(n) \in M$*:*
Sei uns $n \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt dass $n \in M$, sprich, dass

$$n = S^{(n)}(0). \tag{3.1.7}$$ J336

Dann erhalten wir, unter Verwendung von (3.1.7), dass

$$S(n) \overset{(3.1.7)}{=} S(S^{(n)}(0)) = (S \circ S^{(n)})(0) = S^{(S(n))}(0),$$

also $S(n) \in M$.

Der Beweis der Prämisse von **(Nat3)** ist abgeschlossen. Wir können also **(Nat3)** anwenden, schließen dass $M = X$, und damit dass die Behauptung (2) gilt.

□

3.1.3 Das Prinzip der vollständigen Induktion

Wir wollen nun das *Prinzip der vollständigen Induktion* (kurz *vollständige Induktion*, und synonym auch *Induktionsprinzip*) diskutieren. Es ist eigentlich nichts anderes als eine Umformulierung von **(Nat3)**.

J307. **3.1.6 Satz.** *Für jedes $n \in \mathbb{N}$ sei $A(n)$ eine Aussage. Sei vorausgesetzt dass*

(1) *$A(0)$ wahr ist,*

und dass

(2) *für jedes $n \in \mathbb{N}$ die Implikation $A(n) \Rightarrow A(S(n))$ wahr ist.*

Dann ist $A(n)$ für alle $n \in \mathbb{N}$ wahr.

$$\Big[\ \big(A(0) \wedge \forall n \in \mathbb{N}\colon A(n) \Rightarrow A(S(n))\big) \ \Rightarrow\ \forall n \in \mathbb{N}\colon A(n)\ \Big]$$

Die Voraussetzung (1) nennt man den *Induktionsanfang*, und (2) heißt der *Induktionsschritt*. Die Prämisse der Implikation in (2) nennt man die *Induktionsvoraussetzung*.

Beweis von Satz 3.1.6. Betrachte die Menge $M = \{n \in \mathbb{N} :\ A(n) \text{ ist wahr}\}$. Dann gilt $M \subseteq \mathbb{N}$ nach Definition, $0 \in M$ wegen (1) und $S(M) \subseteq M$ wegen (2). Nach **(Nat3)** folgt $M = X$. Und das heißt gerade dass $A(n)$ für alle $n \in \mathbb{N}$ wahr ist. □

♲ `Analyse von Satz` 3.1.6`:`
Um sich von dem Induktionsprinzip nicht verwirren zu lassen, muss man es schon genau lesen. Auf einen ersten – unachtsamen – Blick sieht diese Beweismethode ja etwas wundersam aus: Wir wollen zeigen dass $A(n)$ wahr ist, und nehmen $A(n)$ dann als Voraussetzung an um irgendetwas abzuleiten. Wie kann da etwas Vernünftiges herauskommen (wenn man als Voraussetzung nimmt was man eigentlich zeigen möchte)?

Dieses scheinbare Rätsel läßt sich rasch auflösen; man muss – wie immer – nur lesen was *wirklich* dort steht.

– Die Aussage die letzlich gefolgert wird ist dass $A(n)$ *für alle* $n \in \mathbb{N}$ gilt.

– Die Aussage die wir im Induktionsschritt als Voraussetzung nehmen ist das $A(n)$ *für das eine* – beliebige, aber uns fest vorgegebene – n wahr ist.

– Wir zeigen im Induktionsschritt nur eine Implikation! *Wenn* das eine $A(n)$ wahr ist, *dann* ist auch $A(S(n))$ wahr. Wir zeigen, oder behaupten, beim Nachprüfen von Satz 3.1.6(2) *nicht* dass $A(n)$ wirklich wahr ist.

Das ist gerade die Stärke des Induktionsprinzips, dass aus der Wahrheit einer Familie von *Implikationen* die *tatsächliche Wahrheit* aller Aussagen folgt, wenn nur der Induktionsanfang auch wahr ist (sprich, wenn nur jemand den ersten Dominostein umwirft). ♲

J337. *3.1.7 Bemerkung.* Warum können wir sagen, dass das Induktionsprinzip *nichts anderes ist* als eine Umformulierung von **(Nat3)**? In Satz 3.1.6 haben wir gesehen dass das Induktionsprinzip folgt wenn man ein Objekt hat das den Peano Axiomen genügt. Nun gilt auch umgekehrt: Hat man ein Objekt mit **(Nat1)**, **(Nat2)**, und sodass die Aussage von Satz 3.1.6 gilt, dann gilt für dieses Objekt auch **(Nat3)**. Um das zu sehen, verwende einfach die Aussage $A(n)$ die besagt „$n \in M$“. ◊

Man erkennt hier die herausragende Bedeutung des Induktionsprinzips: Es ist das einzige Beweismittel welches wir von Anfang an zur Verfügung haben um zu schließen dass eine Aussage *für alle* natürlichen Zahlen gilt! Und wenn man an die Beweise im letzten Abschnitt denkt, dort haben wir ja **(Nat3)** auch ständig verwendet.

Als Beispiel wollen wir ein paar Eigenschaften der Iterierten einer Funktion herleiten.

J344. **3.1.8 Lemma.** *Sei M eine Menge, und $f : M \to M$ eine Funktion.*

(1) *Für alle $k \in \mathbb{N}$ gilt $\mathrm{id}_M{}^{(k)} = \mathrm{id}_M$.*

(2) *Für alle $l, k \in \mathbb{N}$ gilt $f^{(l)} \circ f^{(k)} = f^{\left(S^{(l)}(k)\right)}$.*

(3) *Ist f injektiv, so ist für jedes $k \in \mathbb{N}$ auch $f^{(k)}$ injektiv. Analog: ist f surjektiv, so ist für jedes $k \in \mathbb{N}$ auch $f^{(k)}$ surjektiv.*

In den nächsten beiden Punkten sei $g : M \to M$ eine Funktion mit $f \circ g = g \circ f$.

(4) *Für alle* $l, k \in \mathbb{N}$ *gilt* $f^{(l)} \circ g^{(k)} = g^{(k)} \circ f^{(l)}$.

(5) *Für alle* $k \in \mathbb{N}$ *gilt* $(f \circ g)^{(k)} = f^{(k)} \circ g^{(k)}$.

Wendet man (4) *an mit „* $f = g$ *“, so erhält man insbesondere*

(6) *Für alle* $l, k \in \mathbb{N}$ *gilt* $f^{(l)} \circ f^{(k)} = f^{(k)} \circ f^{(l)}$.

Weitere Eigenschaften sind

(7) *Für alle* $l, k \in \mathbb{N}$ *gilt* $\left[f^{(l)}\right]^{(k)} = \left[f^{(k)}\right]^{(l)}$.

(8) *Für alle* $l, k \in \mathbb{N}$ *gilt* $\left[f^{(l)}\right]^{(k)} = f^{\left([S^{(l)}]^{(k)}(0)\right)}$.

Beweis.

Zum Beweis von (1)*:* Wir machen Induktion nach k. Für $k \in \mathbb{N}$ sei $A(k)$ die Aussage

✧ $A(k)$: $\mathrm{id}_{\mathbb{N}}{}^{(k)} = \mathrm{id}_{\mathbb{N}}$

– *Induktionsanfang;* $A(0)$ *ist wahr:*
Die 0-te Iterierte ist für jede Funktion gleich der Identität, also haben wir auch $\mathrm{id}_{\mathbb{N}}{}^{(0)} = \mathrm{id}_{\mathbb{N}}$.
– *Induktionsschritt;* $\forall k \in \mathbb{N}\colon A(k) \Rightarrow A(S(k))$*:*
Sei $k \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt (die Induktionsvoraussetzung), dass

$$\mathrm{id}_{\mathbb{N}}{}^{(k)} = \mathrm{id}_{\mathbb{N}} \tag{3.1.8}$$ J358

Die gewünschte Konklusion ist $\mathrm{id}_{\mathbb{N}}{}^{(S(k))} = \mathrm{id}_{\mathbb{N}}$.
Wir berechnen, unter Verwendung der Induktionsvoraussetzung (3.1.8) und der Definition der Iterierten,

$$\mathrm{id}_{\mathbb{N}}{}^{(S(k))} = \mathrm{id}_{\mathbb{N}} \circ \mathrm{id}_{\mathbb{N}}{}^{(k)} = \mathrm{id}_{\mathbb{N}}{}^{(k)} \overset{(3.1.8)}{=} \mathrm{id}_{\mathbb{N}}\,.$$

Das Induktionsprinzip Satz 3.1.6 zeigt dass die Aussage $A(k)$ für alle $k \in \mathbb{N}$ gilt.

Zum Beweis von (2)*:* Wir machen Induktion nach l. Für $l \in \mathbb{N}$ sei $A(l)$ die Aussage

✧ $A(l)$: $\forall k \in \mathbb{N}\colon f^{(l)} \circ f^{(k)} = f^{\left(S^{(l)}(k)\right)}$

– *Induktionsanfang;* $A(0)$ *ist wahr:*
Für $l = 0$ ist $f^{(l)} = S^{(l)} = \mathrm{id}_{\mathbb{N}}$. Für beliebiges $k \in \mathbb{N}$ gilt also

$$f^{(0)} \circ f^{(k)} = \mathrm{id}_{\mathbb{N}} \circ f^{(k)} = f^{(k)} = f^{\left(\mathrm{id}_{\mathbb{N}}(k)\right)} = f^{\left(S^{(0)}(k)\right)}.$$

– *Induktionsschritt;* $\forall l \in \mathbb{N}\colon A(l) \Rightarrow A(S(l))$*:*
Sei $l \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt (die Induktionsvoraussetzung), dass

$$\forall k \in \mathbb{N}\colon f^{(l)} \circ f^{(k)} = f^{\left(S^{(l)}(k)\right)} \tag{3.1.9}$$ J345

Die gewünschte Konklusion ist $\forall k \in \mathbb{N}\colon f^{(S(l))} \circ f^{(k)} = f^{\left(S^{(S(l))}(k)\right)}$.
Sei $k \in \mathbb{N}$ vorgegeben. Wir berechnen, unter Verwendung der Induktionsvoraussetzung (3.1.9) und der Definition der Iterierten von f bzw. S,

$$f^{(S(l))} \circ f^{(k)} = [f \circ f^{(l)}] \circ f^{(k)} = f \circ [f^{(l)} \circ f^{(k)}] \overset{(3.1.9)}{=} f \circ f^{\left(S^{(l)}(k)\right)} = f^{\left(S(S^{(l)}(k))\right)} = f^{\left(S^{(S(l))}(k)\right)}.$$

Das Induktionsprinzip Satz 3.1.6 zeigt dass die Aussage $A(l)$ für alle $l \in \mathbb{N}$ gilt.

Zum Beweis von (3)*:* Wir machen Induktion nach k. Für $k \in \mathbb{N}$ sei $A(k)$ die Aussage

✧ $A(k)$: $f^{(k)}$ ist injektiv

– *Induktionsanfang; $A(0)$ ist wahr:*
Für $k = 0$ ist $f^{(k)} = \mathrm{id}_{\mathbb{N}}$, und daher injektiv.
– *Induktionsschritt; $\forall k \in \mathbb{N}\colon A(k) \Rightarrow A(S(k))$:*
Sei $k \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt (die Induktionsvoraussetzung), dass

$$f^{(k)} \text{ ist injektiv} \tag{3.1.10}$$ J350

Die gewünschte Konklusion ist $f^{(S(k))}$ ist injektiv.
Es gilt, nach der Definition der Iterierten, $f^{(S(k))} = f \circ f^{(k)}$. Also ist tatsächlich $f^{(S(k))}$ als Komposition injektiver Funktionen auch injektiv.

Das Induktionsprinzip Satz 3.1.6 zeigt dass die Aussage $A(k)$ für alle $k \in \mathbb{N}$ gilt.

Das für eine surjektive Funktion f alle Iterierten ebenfalls surjektiv sind sieht man genauso.

Zum Beweis von (4)*:* Wir zeigen zuerst die beiden Fälle $l = 0$ und $l = 1$.

✾ *Fall 1; $l = 0$.* Dieser Fall ist trivial, denn für jedes $k \in \mathbb{N}$ gilt

$$f^{(0)} \circ g^{(k)} = \mathrm{id}_{\mathbb{N}} \circ g^{(k)} = g^{(k)} = g^{(k)} \circ \mathrm{id}_{\mathbb{N}} = g^{(k)} \circ f^{(0)}.$$

✾ *Fall 2; $l = 1$.* Dieser Fall ist der essentielle Teil. Wir machen Induktion nach k. Für $k \in \mathbb{N}$ sei $A(k)$ die Aussage

✧ $A(k)\colon \quad f \circ g^{(k)} = g^{(k)} \circ f$

– *Induktionsanfang; $A(0)$ ist wahr:*
Wegen dem schon gezeigten Fall 1, angewendet mit den Rollen von f und g vertauscht und „$k = 1$“, gilt $f \circ g^{(0)} = g^{(0)} \circ f$.
– *Induktionsschritt; $\forall k \in \mathbb{N}\colon A(k) \Rightarrow A(S(k))$:*
Sei $k \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt (die Induktionsvoraussetzung), dass

$$f \circ g^{(k)} = g^{(k)} \circ f. \tag{3.1.11}$$ J348

Die gewünschte Konklusion ist $f \circ g^{(S(k))} = g^{(S(k))} \circ f$.
Wir berechnen, unter Verwendung der Induktionsvoraussetzung (3.1.11) und der Definition der Iterierten,

$$\begin{aligned} f \circ g^{(S(k))} &= f \circ (g \circ g^{(k)}) = (f \circ g) \circ g^{(k)} = (g \circ f) \circ g^{(k)} \\ &= g \circ (f \circ g^{(k)}) \overset{(3.1.11)}{=} g \circ (g^{(k)} \circ f) = (g \circ g^{(k)}) \circ f = g^{(S(k))} \circ f. \end{aligned}$$

Das Induktionsprinzip Satz 3.1.6 zeigt dass die Aussage $A(k)$ für alle $k \in \mathbb{N}$ gilt.

Der allgemeine Fall folgt nun unmittelbar. Seien $l, k \in \mathbb{N}$ vorgegeben. Nach dem bereits gezeigten Fall 2 gilt $[g^{(k)}] \circ f = f \circ [g^{(k)}]$. Wir können daher den bereits gezeigten Fall 2 anwenden mit den Daten „$g^{(k)}, f, l$“ anstelle von „f, g, k“. Dies zeigt $[g^{(k)}] \circ f^{(l)} = f^{(l)} \circ [g^{(k)}]$.

Zum Beweis von (5)*:* Wir machen Induktion nach k. Für $k \in \mathbb{N}$ sei $A(k)$ die Aussage

✧ $A(k)\colon \quad (f \circ g)^{(k)} = f^{(k)} \circ g^{(k)}$

– *Induktionsanfang; $A(0)$ ist wahr:*
Für $k = 0$ ist $f^{(k)} = g^{(k)} = (f \circ g)^{(k)} = \mathrm{id}_{\mathbb{N}}$.
– *Induktionsschritt; $\forall k \in \mathbb{N}\colon A(k) \Rightarrow A(S(k))$:*
Sei $k \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt (die Induktionsvoraussetzung), dass

$$(f \circ g)^{(k)} = f^{(k)} \circ g^{(k)} \tag{3.1.12}$$ J310

Die gewünschte Konklusion ist $(f \circ g)^{(S(k))} = f^{(S(k))} \circ g^{(S(k))}$.
Wir berechnen, unter Verwendung der Induktionsvoraussetzung (3.1.12), der Definition der Iterierten, und der schon bewiesenen Aussage (4),

$$\begin{aligned} (f \circ g)^{(S(k))} &= (f \circ g) \circ (f \circ g)^{(k)} \overset{(3.1.12)}{=} (f \circ g) \circ (f^{(k)} \circ g^{(k)}) \\ &= f \circ (g \circ f^{(k)}) \circ g^{(k)} \overset{(4)}{=} f \circ (f^{(k)} \circ g) \circ g^{(k)} = (f \circ f^{(k)}) \circ (g \circ g^{(k)}) = f^{(S(k))} \circ g^{(S(k))}. \end{aligned}$$

Das Induktionsprinzip Satz 3.1.6 zeigt dass die Aussage $A(k)$ für alle $k \in \mathbb{N}$ gilt.

Zum Beweis von (6)*:* Es gilt trivialerweise $f \circ f = f \circ f$. Also können wir tatsächlich den schon bewiesenen Teil (4) anwenden.

Zum Beweis von (7)*:* Wir machen Induktion nach k. Für $k \in \mathbb{N}$ sei $A(k)$ die Aussage

✧ $A(k)$: $\forall l \in \mathbb{N}\colon [f^{(l)}]^{(k)} = [f^{(k)}]^{(l)}$

– *Induktionsanfang;* $A(0)$ *ist wahr:*
Es ist $[f^{(l)}]^{(0)} = \mathrm{id}_{\mathbb{N}}$ und, nach dem bereits bewiesenen Punkt (1) auch $[f^{(0)}]^{(l)} = \mathrm{id}_{\mathbb{N}}{}^{(l)} = \mathrm{id}_{\mathbb{N}}$.
– *Induktionsschritt;* $\forall k \in \mathbb{N}\colon A(k) \Rightarrow A(S(k))$*:*
Sei $k \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt (die Induktionsvoraussetzung), dass

$$\forall l \in \mathbb{N}\colon [f^{(l)}]^{(k)} = [f^{(k)}]^{(l)} \tag{3.1.13}$$ J359

Die gewünschte Konklusion ist $\forall l \in \mathbb{N}\colon [f^{(l)}]^{(S(k))} = [f^{(S(k))}]^{(l)}$.
Sei $l \in \mathbb{N}$ vorgegeben. Wir berechnen, unter Verwendung der Induktionsvoraussetzung (3.1.13), der Definition der Iterierten, und der schon bewiesenen Aussage (5),

$$[f^{(S(k))}]^{(l)} = [f \circ f^{(k)}]^{(l)} \overset{(5)}{=} f^{(l)} \circ [f^{(k)}]^{(l)} \overset{(3.1.13)}{=} f^{(l)} \circ [f^{(l)}]^{(k)} = [f^{(l)}]^{(S(k))}.$$

Das Induktionsprinzip Satz 3.1.6 zeigt dass die Aussage $A(k)$ für alle $k \in \mathbb{N}$ gilt.

Zum Beweis von (8)*:* Wir machen Induktion nach k. Für $k \in \mathbb{N}$ sei $A(k)$ die Aussage

✧ $A(k)$: $\forall l \in \mathbb{N}\colon [f^{(l)}]^{(k)} = f^{\left([S^{(l)}]^{(k)}(0)\right)}$

– *Induktionsanfang;* $A(0)$ *ist wahr:*
Es ist $[f^{(l)}]^{(0)} = \mathrm{id}_{\mathbb{N}}$ und auch $f^{\left([S^{(l)}]^{(0)}(0)\right)} = f^{(0)} = \mathrm{id}_{\mathbb{N}}$.
– *Induktionsschritt;* $\forall k \in \mathbb{N}\colon A(k) \Rightarrow A(S(k))$*:*
Sei $k \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt (die Induktionsvoraussetzung), dass

$$\forall l \in \mathbb{N}\colon [f^{(l)}]^{(k)} = f^{\left([S^{(l)}]^{(k)}(0)\right)} \tag{3.1.14}$$ J360

Die gewünschte Konklusion ist $\forall l \in \mathbb{N}\colon [f^{(l)}]^{(S(k))} = f^{\left([S^{(l)}]^{(S(k))}(0)\right)}$.
Sei $l \in \mathbb{N}$ vorgegeben. Wir berechnen, unter Verwendung der Induktionsvoraussetzung (3.1.14), der Definition der Iterierten, und der schon bewiesenen Aussage (2),

$$f^{\left([S^{(l)}]^{(S(k))}(0)\right)} = f^{\left((S^{(l)} \circ [S^{(l)}]^{(k)})(0)\right)} = f^{\left(S^{(l)}\left([S^{(l)}]^{(k)}(0)\right)\right)} \overset{(2)}{=} f^{(l)} \circ f^{\left([S^{(l)}]^{(k)}(0)\right)} \overset{(3.1.14)}{=} f^{(l)} \circ \left([f^{(l)}]^{(k)}\right) = [f^{(l)}]^{(S(k))}.$$

Das Induktionsprinzip Satz 3.1.6 zeigt dass die Aussage $A(k)$ für alle $k \in \mathbb{N}$ gilt.

□

Wir werden im wiederholt auch die folgenden Eigenschaften der Nachfolgerabbildung verwenden.

J357. **3.1.9 Lemma.** *Seien* $n, k \in \mathbb{N}$.

(1) *Ist* $n = S^{(k)}(n)$*, so folgt* $k = 0$.

(2) *Ist* $S^{(k)}(n) = 0$*, so folgt* $k = n = 0$.

Im Beweis dieses Lemmas werden wir ein Beweisprinzip kennenlernen das bisher noch nicht aufgetreten ist, nämlich das des *indirekten Beweis* (auch *Widerspruchsbeweis* oder *Reductio ad absurdum*). Diese Schlussweise ist komplexer als die bisher kennengelernten Beweismethoden des direkten Beweises, Beweises durch Fallunterscheidung, und Beweises mittels Kontraposition. Eine ausführliche Diskussion findet sich im Anschluss.

Wieder einmal ist die Terminologie *nicht* uniform. In mancher Literatur wird das was wir „Beweis durch Kontraposition“ genannt haben, als „indirekter Beweis“ bezeichnet. Und für das was wir hier „indirekter Beweis“ nennen werden, sagt man ausschließlich „Widerspruchsbeweis“.

Also, immer aufpassen was gemeint ist! Sorry, aber so ist es nun einmal: man muss einfach immer *verstehen* was *wirklich passiert*, dann können einem solche Notationsverwirrungen nicht beeindrucken.

Wahrscheinlich wäre es besser gleich immer die – deskriptive und uniforme – englische Bezeichnung zu verwenden, wo man eigentlich immer von „proof by contraposition“ bzw. „proof by contradiction“ spricht.

Beweis.

Zum Beweis von (1)*:* Seien uns $n, k \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt, dass $n = S^{(k)}(n)$. Wir berechnen, mit Hilfe von Lemma 3.1.8(6),

$$S^{(n)}(0) = n = S^{(k)}(n) = S^{(k)}(S^{(n)}(0)) = [S^{(k)} \circ S^{(n)}](0) = [S^{(n)} \circ S^{(k)}](0) = S^{(n)}(S^{(k)}(0)) = S^{(n)}(k).$$

Nach Lemma 3.1.8(3) ist die Funktion $S^{(n)}$ injektiv. Also folgt $k = 0$.

Zum Beweis von (2)*:* Sei vorausgesetzt, dass $S^{(k)}(n) = 0$. Wir zeigen zuerst $k = 0$. Dazu sei indirekt angenommen, dass $k \neq 0$. Sei $k' \in \mathbb{N}$ mit $k = S(k')$, dann gilt

$$0 = S^{(k)}(n) = S^{\left(S(k')\right)}(n) = S(S^{(k')}(n)).$$

Insbesondere ist $0 \in S(\mathbb{N})$. Dies ist eine Widerspruch, da wir wissen dass $0 \notin S(\mathbb{N})$. Also muss $k = 0$ sein.

Nun zeigen wir, dass auch $n = 0$. Dazu bemerke dass

$$S^{(k)}(n) = [S^{(k)} \circ S^{(n)}](0) = [S^{(n)} \circ S^{(k)}](0) = S^{(n)}(k).$$

Also bekommen wir aus dem schon gezeigten Beweisteil (angewendet mit n und k anstelle von k und n), dass $n = 0$.

□

`Analyse einer Vorgangsweise im Beweis von Lemma` 3.1.9(2)`:`

Wir haben zu zeigen dass die Implikation „$S^{(k)}(n) = 0 \Rightarrow k = 0$“ gilt. Dazu argumentieren wir wie folgt:

① Sei vorausgesetzt, dass $S^{(k)}(n) = 0$.

② Sei indirekt angenommen, dass $k \neq 0$, und wähle $k' \in \mathbb{N}$ mit $k = S(k')$.

③ Dann gilt, mit Hilfe von Lemma 3.1.8(2),

$$0 \overset{①}{=} S^{(k)}(n) \overset{②}{=} S^{\left(S(k')\right)}(n) \overset{(2)}{=} S(S^{(k')}(0)).$$

Insbesondere ist $0 \in S(\mathbb{N})$.

Danach sagen wir:

④ Dies ist eine Widerspruch, da wir wissen dass $0 \notin S(\mathbb{N})$.

und behaupten dass damit der Beweis von „$k = 0$“ abgeschlossen ist.

Wieso können wir das schließen ? Nun, wir haben eine Implikation gezeigt:

$$S^{(k)}(n) = 0 \wedge k \neq 0 \quad \Rightarrow \quad 0 \in S(\mathbb{N})$$

Die Konlusion dieser Implikation ist, wegen **(Nat2)**, *falsch*. Also muss auch die Prämisse der Implikation *falsch* sein. Schließlich ist „$\neg(S^{(k)}(n) = 0 \wedge k \neq 0)$“ äquivalent zu „$S^{(k)}(n) = 0 \Rightarrow k = 0$“; man erinnere sich an die Tautologie ((17)) aus §1.1.2.

Die formale Grundlage des indirekten Beweises (Widerspruchsbeweises) ist wie folgt.

> ➢ `Ein indirekter Beweis (Widerspruchsbeweis) funktioniert so:`
>
> Wir gehen davon aus, dass die zu beweisende Aussage (zum Beispiel im obigen Schritt 2) bereits in dem Sinne analysiert wurde, dass es klar ist was die zur Verfügung stehenden Voraussetzungen sind, und was die zu zeigende Behauptung ist. Wollen wir diese Voraussetzungen als A und diese Behauptung als B bezeichnen.

① Man zeigt, dass eine Implikation $A \wedge \neg B \Rightarrow F$ wahr ist, wo F einfach *irgendetwas* ist von dem wir *wissen* dass es *falsch* ist.

Was nun folgt, ist die formallogische Argumentation dass damit der Beweis tatsächlich abgeschlossen ist.

❷ Man wendet den Kontrapositionssatz an, wodurch man $\neg F \Rightarrow \neg(A \wedge \neg B)$ erhält.

❸ Man wendet den Modus Ponendo Ponens an, nachdem man aus der Wahrheit von $\neg F$ und der Wahrheit obiger Implikation auf die Wahrheit von $\neg(A \wedge \neg B)$ schließen kann.

❹ Nun erinnert man sich dass $\neg(A \wedge \neg B)$ äquivalent zu $A \Rightarrow B$ ist, cf. (17), Also hat man $A \Rightarrow B$ erhalten.

❺ Man wendet den Modus Ponendo Ponens an, nachdem man aus der Wahrheit von A und der Wahrheit der Implikation $A \Rightarrow B$ auf die Wahrheit von B schließen kann.

Die Stärke eines indirekter Beweises liegt darin, dass

– man die Aussage $A \wedge \neg B$ als Voraussetzung zur Verfügung hat. Und nicht nur A wie bei einem direkten Beweis, bzw. nur $\neg B$ wie bei einem Beweis mittels Kontraposition.

– man völlige Freiheit in der Wahl der falschen Aussage F hat. Sie muss – und wird meistens auch – gar nichts mit den gerade betrachteten Dingen zu tun haben. Ein Klassiker wäre „$0 = 1$“, oder „$a \in M \wedge a \notin M$“ (wobei es ganz egal ist was a und M sind).

  Meistens, wie auch in obigem konkreten Beweis, wird F irgendetwas von der Gestalt $C \wedge \neg C$ sein; was ja nach dem Satz vom Widerspruch ganz sicher falsch ist, ganz unabhängig davon was C ist und welchen Wahrheitswert C hat (sprich, man muss nicht einmal etwas über C wissen).

Die Schwäche eines indirekter Beweises liegt darin, dass

– man oft keine Idee bekommt warum das nun eigentlich wirklich funktioniert.

Führt man einen Existenzbeweis indirekt, so erhält man eine reine Existenzaussage die meist keinen Anhaltspunkt gibt wie ein gewünschtes Objekt wirklich aussieht; im Gegensatz zu einem direkten Beweis einer Existenzaussage, wo man ein gewünschtes Objekt tatsächlich konstruieren muss. Das ist in gewissen Sinn natürlich eine Schwäche des indirekten Beweisprinzipes, andererseits aber auch eine Stärke. Nämlich hat man ein Instrument mit dem man Existenz pathologischer Objekte zeigen kann.

Geht man darauf los einen indirekten Beweis zu führen, so sollte man das immer klar und deutlich ankündigen. Zum Beispiel sagt man: „Sei indirekt angenommen dass B falsch ist.“ (im Englischen sieht man oft die, noch deskriptivere, Formulierung: „Assume, towards a contradiction, that B is false.“). Sagt man nur „sei angenommen dass B falsch ist“, so ist es potentiell unklar ob nun auch vorausgesetzt wird dass A gilt oder nicht, sowie ob man konkret $\neg A$ beweisen muss oder nur irgendetwas wahllos falsches ableiten möchte (sprich, ob man auf einen Beweis durch Kontraposition oder einen indirekten Beweis losgeht).

➢ `Was heißt hier -- umgangssprachlich -- pathologisch:`

Unter einem *pathologischen Objekt* versteht man üblicherweise ein Objekt mit so richtig schlechten Eigenschaften, dass sich typischerweise auch noch jeglicher Vorstellungskraft entzieht, aber gemeinerweise trotzdem existiert.

Zum Beispiel: Man würde sich wünschen dass eine schöne, ins Bild passende und obendrein praktische, Aussage gilt. In allen Beispielen an die man nur denken kann tut sie das auch. Aber leider: es existiert ein komisches, völlig unanschauliches, und ziemlich sicher in allen angedachten Anwendungen nicht auftretendes Objekt, für das die Aussage eben nicht gilt.

3.2 Die Anordnung der natürlichen Zahlen

3.2.1 Motivation

Hat man zwei Schafherden und möchte wissen welche größer ist, so kann man das folgende Verfahren durchführen:

> Man nehme je ein Schaf von jeder der beiden Herden, und führe sie auf einen anderen Platz wo sie dann – wieder getrennt – stehen bleiben. Dies wiederhole man solange bis von einer der beiden Herden keine Schaf mehr übrig ist.

Jene Herde die komplett übersiedelt wurde ist dann die kleinere, und jene von der noch etwas übrig ist die größere. Natürlich kann es auch sein, dass beide Herden komplett übersiedelt wurden, dann sind sie gleich groß.

Nun stellen wir den Originalzustand wieder her, sprich, wir führen die größere Herde wieder zusammen. Was passiert dabei: Wir starten mit dem Teil der Herde der gleich groß zu der kleineren Herde ist, und führen ein Schaf des übriggebliebenen Teiles nach dem anderen dazu, solange bis auch diese Herde komplett übersiedelt ist.

Wir sehen, dass eine Herde größer ist als eine andere, wenn sie durch wiederholtes dazugeben von Schafen zu einem Teil von ihr ensteht, welcher gleich groß zu der anderen Herde ist.

3.2.2 Formalisierung und Eigenschaften

J342. **3.2.1 Definition.** Sei $\leq$ die wie folgt gegebene Relation auf $\mathbb{N}$:

$$\leq = \{(m,n) \in \mathbb{N} \times \mathbb{N} : \ \exists l \in \mathbb{N} \colon n = S^{(l)}(m)\}.$$

Wir schreiben $<$ für die Relation $\leq \setminus \Delta_{\mathbb{N}}$ ◊

Wir werden im Folgenden nahezu ausschließlich die Schreibweise „$m \leq n$“ verwenden. Die Notationen $\geq$ und $>$ sind definiert als

$$n \geq m \ \Leftrightarrow \ m \leq n, \qquad n > m \ \Leftrightarrow \ m < n.$$

♲ `Analyse von Definition` 3.2.1:
Es gilt $m \leq n$ per definitionem genau dann, wenn $\exists l \in \mathbb{N} \colon n = S^{(l)}(m)$. Sprich, wir erhalten die (dann *die größere* heißende) Zahl n durch wiederholtes übergehen zu dem Nachfolger ausgehend von der (dann *die kleinere* heißenden) Zahl m.

Da wir auch $l = 0$ erlauben, ist es hier auch erlaubt dass n und m gleich groß sind. Wollen wir Gleichheit ausschließen, so verwenden wir das eigenständige Symbol $<$; es gilt ja, per definitionem, $m < n$ genau dann, wenn $m \leq n$ und $m \neq n$. ♲

J327. **3.2.2 Satz.** *Die Relation $\leq$ hat die folgenden Eigenschaften.*

(1) *Für alle $a \in \mathbb{N}$ gilt $a \leq a$.* *(*Reflexivität*)*

(2) *Sind $a, b \in \mathbb{N}$ und gilt $a \leq b$ und $b \leq a$, so folgt $a = b$.* *(*Antisymmetrie*)*

(3) *Sind $a, b, c \in \mathbb{N}$ und gilt $a \leq b$ und $b \leq c$, so folgt $a \leq c$.* *(*Transitivität*)*

Sei $M \subseteq \mathbb{N}$. Ein Element x_0 von M heißt kleinstes Element von *M, wenn für jedes Element x von M die Beziehung $x_0 \leq x$ gilt.*

$\left[\ x_0 \in M \wedge \forall x \in M \colon x_0 \leq x\ \right]$

Analog heißt ein Element x_0 von M größtes Element von *M wenn für jedes Element x von M die Beziehung $x \leq x_0$ gilt.*

$\left[\ x_0 \in M \wedge \forall x \in M \colon x \leq x_0\ \right]$

(4) *Jede nichtleere Teilmenge von $\mathbb{N}$ hat ein bezüglich $\leq$ kleinstes Element.* *(*Wohlordnung*)*

(5) *Ist $a \in \mathbb{N} \setminus \{0\}$ so existiert genau eine Zahl $b \in \mathbb{N}$ sodass a das kleinste Element der Menge $\{n \in \mathbb{N} : \ b < n\}$ ist.* *(*$\exists!$ Vorgänger*)*

Für dieses Element b gilt $a = S(b)$, und b ist das größte Element der Menge $\{n \in \mathbb{N} : \ n < a\}$.

Beweis.

Zum Beweis der Reflexivität: Es gilt $S^{(0)} = \mathrm{id}_{\mathbb{N}}$, und damit $n = S^{(0)}(n)$ für jede Zahl $n \in \mathbb{N}$. Also haben wir stets $n \leq n$.

Zum Beweis der Antisymmetrie: Seien $m, n \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt dass $m \leq n$ und $n \leq m$. Wir müssen zeigen, dass $n = m$.

Wähle $l \in \mathbb{N}$ mit $n = S^{(l)}(m)$, und wähle $k \in \mathbb{N}$ mit $m = S^{(k)}(n)$. Dann gilt, mit Lemma 3.1.8(2),

$$n = S^{(l)}(m) = S^{(l)}(S^{(k)}(n))) = S^{\left(S^{(l)}(k)\right)}(n).$$

Nach Lemma 3.1.9(1) folgt $S^{(l)}(k) = 0$, und nach Lemma 3.1.9(2) folgt $l = k = 0$. Damit ist $n = m$.

Zum Beweis der Transitivität: Seien $m, n, k \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt dass $m \leq n$ and $n \leq k$. Wähle $l \in \mathbb{N}$ mit $n = S^{(l)}(m)$ und $j \in \mathbb{N}$ mit $k = S^{(j)}(n)$. Dann gilt, nach Lemma 3.1.8(2),

$$k = S^{(j)}(n) = S^{(j)}(S^{(l)}(m)) = S^{\left(S^{(j)}(l)\right)}(m),$$

und wir sehen dass $m \leq k$.

Zum Beweis der Wohlordnungseigenschaft: Eine Teilmenge M von $\mathbb{N}$ ist nichtleer, genau dann wenn es eine Zahl $n \in \mathbb{N}$ gibt mit $n \in M$. Die Wohlordnungseigenschaft

$\left[\ \forall M \subseteq \mathbb{N}, M \neq \emptyset\, \exists\, x_0 \in M\, \forall x \in M\colon\ x_0 \leq x\ \right]$

ist daher äquivalent zu der folgenden Eigenschaft:

(4′) Für jede Zahl $n \in \mathbb{N}$ und jede Teilmenge M von $\mathbb{N}$ die diese Zahl n enthält, hat die Menge M ein kleinstes Element.
$\left[\ \forall n \in \mathbb{N}\, \forall M \subseteq \mathbb{N}\colon n \in M \Rightarrow \left(\exists\, x_0 \in M\, \forall x \in M\colon\ x_0 \leq x\right)\ \right]$

Man kann sich hier eine Fallunterscheidung in die folgenden, sich überlappenden aber gemeinsam alle potentiellen Möglichkeiten abdeckenden, Fälle vorstellen:

- ✾ *Fall 1;* $0 \in M$.
- ✾ *Fall 2;* $1 \in M$.
- ✾ *Fall 3;* $2 \in M$.

  ⋮

Der Vorteil der äquivalenten Formulierung (4′) ist, dass wir nun eine Aussage zeigen müssen die mit einem Allquantor „$\forall n \in \mathbb{N}$“ anfängt, und wir daher das Werkzeug der vollständigen Induktion einsetzen können.

Wir machen Induktion nach n. Für $n \in \mathbb{N}$ sei $A(n)$ die Aussage

✧ $A(n)\colon\ \ \forall M \subseteq \mathbb{N}\colon n \in M \Rightarrow \left(\exists\, x_0 \in M\, \forall x \in M\colon\ x_0 \leq x\right)$

– *Induktionsanfang; $A(0)$ ist wahr:*
Wir zeigen, dass 0 kleinstes Element von $\mathbb{N}$ ist. Sei $m \in \mathbb{N}$ vorgegeben. Nach Proposition 3.1.5(2) gilt $m = S^{(m)}(0)$, also $0 \leq m$.
Insbesondere folgt nun: Ist $M \subseteq \mathbb{N}$ mit $0 \in \mathbb{N}$, so ist 0 kleinstes Element von M.

– *Induktionsschritt; $\forall n \in \mathbb{N}\colon A(n) \Rightarrow A(S(n))$:*
Sei uns $n \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt (die Induktionsvoraussetzung), dass

$$\forall M \subseteq \mathbb{N}\colon n \in M \Rightarrow \left(\exists\, x_0 \in M\, \forall x \in M\colon\ x_0 \leq x\right) \tag{3.2.1}$$ J346

Die gewünschte Konklusion ist $\forall M \subseteq \mathbb{N}\colon S(n) \in M \Rightarrow \left(\exists\, x_0 \in M\, \forall x \in M\colon\ x_0 \leq x\right)$.
Sei uns $M \subseteq \mathbb{N}$ mit $S(n) \in M$ vorgegeben. Wir machen eine Fallunterscheidung.

✾ *Fall 1;* $0 \in M$. Ist $0 \in M$, dann ist, wie wir im Beweis des Induktionsanfanges schon gesehen haben, 0 kleinstes Element von M.

✾ *Fall 2;* $0 \notin M$. Wir betrachten die Menge $S^{-1}(M)$. Es gilt wegen Proposition 3.1.5(1) dass $M \subseteq S(\mathbb{N})$, und damit

$$M = S(S^{-1}(M)). \tag{3.2.2}$$ J347

Wegen $S(n) \in M$ haben wir $n \in S^{-1}(M)$. Nach der Induktionsvoraussetzung (3.2.1), angewendet auf die Menge $S^{-1}(M)$, hat $S^{-1}(M)$ ein kleinstes Element; nennen wir es x_0.
Wir zeigen nun, dass $S(x_0)$ kleinstes Element von M ist. Zuerst gilt wegen (3.2.2) dass $S(x_0) \in M$. Sei $x \in M$ vorgegeben. Wieder wegen (3.2.2) können wir ein Element $y \in S^{-1}(M)$ wählen, sodass $x = S(y)$. Nun gilt $x_0 \leq y$, sprich, es existiert $l \in \mathbb{N}$ mit $y = S^{(l)}(x_0)$. Damit ist, unter Verwendung von Lemma 3.1.8(6),

$$x = S(y) = S(S^{(l)}(x_0)) = [S \circ S^{(l)}](x_0) = [S^{(l)} \circ S](x_0) = S^{(l)}(S(x_0)),$$

also haben wir $S(x_0) \leq x$.

Das Induktionsprinzip Satz 3.1.6, zeigt dass die Aussage $A(n)$ für alle $n \in \mathbb{N}$ wahr ist, und damit dass die Wohlordnungseigenschaft gilt.

Zum Beweis der Existenz und Eindeutigkeit eines Vorgängers: Sei $a \in \mathbb{N} \setminus \{0\}$ vorgegeben. Dann ist $a \in S(\mathbb{N})$, und daher existiert eine Zahl $b \in \mathbb{N}$ mit $S(b) = a$. Da S injektiv ist, ist b eindeutig durch diese Eigenschaft bestimmt (Achtung: das ist nicht die Eindeutigkeitsaussage in der Behauptung des Satzes!).

Wir zeigen, dass a das kleinste Element der Menge $M_b^+ = \{n \in \mathbb{N} : \ b < n\}$ ist. Zuerst ist $a = S(b)$, also $b \leq a$. Nach Lemma 3.1.9(1) ist $a \neq b$ und wir sehen dass $b < a$. Sei nun $n \in M_b^+$ vorgegeben. Sei $l \in \mathbb{N}$ mit $n = S^{(l)}(b)$. Da $n \neq b$, haben wir $l \neq 0$. Daher finden wir $l' \in \mathbb{N}$ mit $l = S(l')$. Es folgt

$$n = S^{(l)}(b) = S^{\left(S(l')\right)}(b) = [S \circ S^{(l')}](b) = S^{(l')}(S(b)) = S^{(l')}(a),$$

also $a \leq n$.

Betrachte nun die Menge $M_a^- = \{n \in \mathbb{N} : \ n < a\}$. Wie wir schon wissen, gilt $b < a$. Also ist $b \in M_a^-$. Sei nun $n \in M_a^-$ vorgegeben, und sei $l \in \mathbb{N}$ mit $a = S^{(l)}(n)$. Da $n \neq a$, haben wir $l \neq 0$. Sei $l' \in \mathbb{N}$ mit $l = S(l')$. Dann ist $S(b) = a = S^{(l)}(n) = S^{(S(l'))}(n) = S(S^{(l')}(n))$. Da S injektiv ist, folgt $b = S^{(l')}(n)$, und damit $n \leq b$.

Es ist noch die Eindeutigkeit zu zeigen. Sei $b' \in \mathbb{N}$ eine Zahl, sodass a das kleinste Element der Menge $M_{b'}^+ = \{n \in \mathbb{N} : \ b' < n\}$ ist. Klarerweise ist $b' \leq S(b')$. Nach Lemma 3.1.9(1) ist $b' \neq S(b')$, und wir erhalten $S(b') \in M$. Daher folgt $a \leq S(b')$. Andererseits ist $a = S^{(l)}(b')$ mit einem $l \in \mathbb{N} \setminus \{0\}$. Mit $l' \in \mathbb{N}$ sodass $l = S(l')$ gilt also $a = S^{(S(l'))}(b') = S(S^{(l')}(b')) = S^{(l')}(S(b'))$, und damit $S(b') \leq a$. Wir schließen, dass $a = S(b')$. Da auch $a = S(b)$, folgt damit dass $b' = b$.

□

♲ `Analyse von Satz` 3.2.2(5)`:`
In Satz 3.2.2(5) haben wir drei Eigenschaften einer Zahl $b \in \mathbb{N}$.

(A) b ist so, dass a das kleinste Element der Menge $M_b^+ = \{n \in \mathbb{N} : \ b < n\}$ ist.

(B) $a = S(b)$.

(C) b ist das größte Element der Menge $M_a^- = \{n \in \mathbb{N} : \ n < a\}$.

Die Behauptung von Satz 3.2.2(5) ist:

(i) Es gibt eine Zahl mit der Eigenschaft (A).

(ii) Je zwei Zahlen mit der Eigenschaft (A) sind gleich.

(iii) Die Zahl mit der Eigenschaft (A) hat auch die Eigenschaft (B).

(iv) Die Zahl mit der Eigenschaft (A) hat auch die Eigenschaft (C).

Was wir dann zeigen ist:

– Es gibt eine Zahl mit der Eigenschaft (B).

– Je zwei Zahlen mit der Eigenschaft (B) sind gleich.

– Die Zahl mit der Eigenschaft (B) hat auch die Eigenschaft (A).

– Die Zahl mit der Eigenschaft (B) hat auch die Eigenschaft (C).

Bemerke: Da eine Menge höchstens ein größtes Element haben kann, sind auch je zwei Zahlen mit der Eigenschaft (C) gleich.

Damit haben wir bewiesen, dass es eine Zahl gibt die alle drei Eigenschaften (A), (B), (C) hat, und dass je zwei Zahlen mit (B) bzw. auch je zwei Zahlen mit (C) gleich sind. Diese Eindeutigkeitsaussagen sind gut zu wissen, aber nicht das was in (ii) behauptet wurde.

Was wir dann zeigen ist:

– Jede Zahl mit der Eigenschaft (A) hat auch die Eigenschaft (B).

Damit haben also je zwei Zahlen mit der Eigenschaft (A) beide auch die Eigenschaft (B), und sind daher gleich. Wir haben die Eindeutigkeitsaussage des Satzes, sprich für (A), also über den Umweg der Eindeutigkeitsaussage für (B) gezeigt.

J343. *3.2.3 Bemerkung.* Allgemein nennt man eine Relation $\preceq$ auf einer Menge M eine *Ordnungsrelation*, wenn sie reflexiv, antisymmetrisch, und transitiv ist.

Eine Ordnungsrelation $\preceq$ auf einer Menge M heißt eine *Totalordnung*, wenn

(TOrd) Für alle $a, b \in M$ gilt $a \preceq b$ oder $b \preceq a$,

und sie heißt eine *Wohlordnung*, wenn

(WOrd) Jede nichtleere Teilmenge von M hat ein bezüglich $\preceq$ kleinstes Element.

◊

Manchmal spricht man von einer Ordnungsrelation auch als einer *partiellen Ordnung*, und von einer Totalordnung als einer *Ordnung*. Das kann Anlass grober Verwirrung sein! Also, immer aufpassen was der Autor nun wirklich meint.

Der Ursprung dieser Gepflogenheit liegt wohl in der Geschichte begründet. Die Ordnungsrelationen auf den bekannten Zahlen sind Totalordnungen, und man hat von jeher von *Ordnungen* gesprochen. Irgendwann ist man dann Relationen begegnet die auch irgendwie eine Ordnung ausdrücken, aber keine Totalordnung sind. Daher hat man dann von partiellen Ordnungen gesprochen.

Ordnungrelationen sind, neben Äquivalenzrelationen, ein sehr wichtiger Typ spezieller Relationen die oft auftreten, sehr viele Eigenschaften haben, und zu etlichen interessanten weiteren Begriffsbildungen und Erkenntnissen führen. Der Begriff der Wohlordnung spielt zum Beispiel in der Mengentheorie eine herausragende Rolle.

J351. *3.2.4 Bemerkung.* Eine Form der Wohlordnungseigenschaft die häufig verwendet wird ist die sogenannte *descending chain condition*, kurz, **(DCC)**. Sie besagt:
Sei $f : \mathbb{N} \to \mathbb{N}$ eine Funktion mit der Eigenschaft dass

$$\forall n \in \mathbb{N}\colon\ f(S(n)) \leq f(n)$$

Dann existiert $n_0 \in \mathbb{N}$ mit $f(n) = f(n_0)$ für alle $n \in \mathbb{N}$, $n \geq n_0$.

Dies sieht man mit einem induktiven Argument ein. Die Essenz des Argumentes ist das Folgende: Betrachte die Menge $M = \{f(n) : \ n \in \mathbb{N}\}$. Diese Menge ist nichtleer, hat also ein kleinstes Element; nennen wir es x_0. Da $x_0 \in M$, finden wir $n_0 \in \mathbb{N}$ mit $x_0 = f(n_0)$. Nun gilt, da $x_0 \leq f(n)$ für alle $n \in \mathbb{N}$, dass

$$x_0 \leq f(S(n_0)) \leq f(n_0) = x_0,$$

und damit $f(S(n_0)) = x_0 = f(n_0)$. ◊

3.2.3 Das starke Induktionsprinzip

Wir wollen nun eine Variante des Prinzips der vollständigen Induktion angeben, die manchmal auch *starkes Induktionsprinzip* genannt wird.

J314. **3.2.5 Satz.** *Für jedes $n \in \mathbb{N}$ sei $A(n)$ eine Aussage. Wenn*

(1) *für jedes $n \in \mathbb{N}$ die Implikation $(\forall m \in \mathbb{N}, m < n\colon A(m)) \Rightarrow A(n)$ wahr ist,*

dann ist $A(n)$ für alle $n \in \mathbb{N}$ wahr.

$$\Big[\Big(\forall n \in \mathbb{N}\colon (\forall m \in \mathbb{N}, m < n\colon A(m)) \Rightarrow A(n)\Big) \;\Rightarrow\; \forall n \in \mathbb{N}\colon A(n) \Big]$$

Beweis. Wir verwenden Beweis durch Kontraposition.

Bezeichne mit M die Menge $M = \{n \in \mathbb{N} :\ A(n) \text{ ist falsch}\}$, und sei vorausgesetzt dass $M \neq \emptyset$; das ist gerade die Negation von $\forall n \in \mathbb{N}\colon A(n)$. Da $\mathbb{N}$ wohlgeordnet ist, hat M ein kleinstes Element, nennen wir es n_0. Dann ist also $A(n_0)$ falsch, aber für alle $m < n_0$ die Aussage $A(m)$ nicht falsch, sprich, wahr. Wir haben eine Zahl gefunden, nämlich n_0, sodass die Prämisse der Implikation in (1) wahr ist, ihre Konklusion jedoch falsch. Also ist die Implikation in (1) für dieses n_0 falsch. Also ist (1) falsch. □

♲ `Analyse von Satz 3.2.5:`

Als erstes wollen wir anmerken, dass der Induktionsanfang Satz 3.1.6(1) in dieser Formulierung nur *scheinbar* verschwunden ist. Nämlich muss ja die Implikation in Satz 3.2.5(1) auch für das kleinste Element von $\mathbb{N}$, sprich für 0, gelten. Für $n = 0$ ist die Prämisse der Implikation stets wahr (da es ja gar keine m mit $m < n$ gibt). Man könnte die Voraussetzung Satz 3.2.5(1) also ersetzen durch das dazu äquivalente Paar von Bedingungen

(1) $A(0)$ ist wahr;

(2) für jedes $n \in \mathbb{N} \setminus \{0\}$ ist die Implikation $(\forall m \in \mathbb{N}, m < n\colon A(m)) \Rightarrow A(n)$ wahr.

Die Formulierung Satz 3.2.5 des Induktionsprinzips ist aus zwei Gründen von Interesse. Einerseits gilt diese Formulierung, wie man durch betrachten des Beweises sieht, sogar für *jede* wohlgeordnete Menge anstelle von $\mathbb{N}$. Von dieser Verallgemeinerung auf beliebige wohlgeordnete Mengen spricht man auch als *transfinite Induktion*.

Andererseits ist es „leichter“ den Induktionsschritt zu beweisen; deswegen spricht man auch vom starken Induktionsprinzip. Das Werkzeug Satz 3.2.5 ist also mächtiger als Satz 3.1.6.

Woran liegt das ? Wir haben in Satz 3.2.5 als Induktionsvoraussetzung, sprich als Prämisse der Implikation in Satz 3.2.5(1), dass die Aussage $A(m)$ *für alle* $m < n$ gilt. Den Induktionsschritt Satz 3.1.6(2) könnte man auch so formulieren:

(2′) Für $n \in \mathbb{N} \setminus \{0\}$ bezeichne mit $T(n)$ den Vorgänger von n. Dann ist für jedes $n \in \mathbb{N} \setminus \{0\}$ die Implikation $A(T(n)) \Rightarrow A(n)$ wahr.

Wir haben als Prämisse also nur dass $A(m)$ *für die eine* unmittelbar vor n liegende Zahl gilt. Und ist die Prämisse schwächer und die Konklusion gleich, so ist die Implikation stärker, also potentiell auch schwerer zu zeigen.

Man erinnere sich hier an den Modus Barbara: Die Prämisse in Satz 3.2.5(1), nennen wir sie A, impliziert die Prämisse in Satz 3.1.6(2), nennen wir sie B. Wenn wir die Implikation Satz 3.1.6(2), das ist also $B \Rightarrow A(n)$, beweisen können, dann wissen wir nach dem Modus Barbara dass auch $A \Rightarrow A(n)$ wahr ist. Sprich, wenn wir es schaffen Satz 3.1.6 anzuwenden, dann hätten wir es auch geschafft Satz 3.2.5 anzuwenden (bedenke, wie oben angemerkt, dass Satz 3.2.5(1) ja auch Satz 3.1.6(1) impliziert).

Man sollte hier schon anmerken das, letzlich, Satz 3.1.6 und Satz 3.2.5 äquivalent sind. Die Implikation „Satz 3.2.5 $\Rightarrow$ Satz 3.1.6“ ist dabei klar (wir haben sie im letzten Absatz ja erklärt). Für die umgekehrte Implikation muss man etwas mehr investieren. ♲

> ➢ `Was heißt hier stärker:`
>
> Man sagt oft ein Satz ist *stärker* als ein anderer, wenn er diesen anderen in recht einfacher Weise impliziert.
>
> Hat man zwei äquivalente Sätze, sagen wir Satz 1 und Satz 2, dann sagt man manchmal Satz 1 ist der stärkere der beiden, wenn die Implikation „Satz 1 $\Rightarrow$ Satz 2“ recht einfach zu beweisen geht, die umgekehrte Implikation „Satz 1 $\Leftarrow$ Satz 2“ aber schwieriger ist.

3.3 Algebraische Operationen

3.3.1 Motivation

Die Aufgabe; Teil 1:

Von unserer Schafherde ist ein Teil im Laufe des Tages auf eine Nachbaralm gewandert. Wir wollen sie jetzt wieder zusammenführen und dabei gleich nachprüfen ob noch alle da sind. Natürlich könnte man den ausgerissenen Teil der Herde einfach herübertreiben, dabei verliert man aber leicht den Überblick wieviele es wirklich sind.

Wir gehen so vor: wir stellen den Teil der zuhause geblieben ist schön in einer Reihe auf, und führen dann von der Nachbaralm eines nach dem anderen herüber und stellen es dazu. Wenn wir damit fertig sind, ist die Reihe hoffentlich so lang wie es der Gesamtgröße der Herde entspricht.

(i) Wir beginnen mit einer gewissen Anzahl, und gehen, einer gewissen anderen Anzahl entsprechend oft, schrittweise immer wieder zum Nachfolger über, solange bis wir unsere gesamt vorhanden Anzahl erreicht haben.

Die Aufgabe; Teil 2:

Heute kommen Gewitter; unsere Herde muss auf die untere Alm. Nachdem wir sie heruntergetrieben haben, stellt sich die gleich Frage wie vorher: Sind noch alle da ? Die untere Alm ist wohl geschützter, aber leider auch kleiner. Darum geht es sich nicht aus, in bewährter Weise alle in eine Reihe zu stellen.

Wir gehen so vor: Wir stellen einmal ein paar in eine Reihe, so dass es sich gemütlich ausgeht. Dann machen wir parallel dazu eine zweite Reihe, wo wir genauso viele hinstellen. Dann eine dritte, und so weiter bis alle einen Platz haben. Zufälligerweise geht es sich gerade aus, dass auch die letzte Reihe voll ist, und wir haben ein volles Rechteck. Nun wissen wir wieviele in jeder Reihe stehen, und es ist leicht zu zählen wieviele Reihen es gibt.

(ii) Die Herde ist in eine gewisse Anzahl von gleich großen Teilen aufgeteilt. Wir wissen wie groß die Teile sind, und wieviele Teile es sind. Wie bekommen wir daraus die Gesamtanzahl ?

3.3.2 Die Addition

Es ist einfach, dass im obigen §3.3.1 genannte Prozedere (i) zu formalisieren.

J352. **3.3.1 Definition.** Wir definieren eine Funktion $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ als

$$+ : \left\{ \begin{array}{ccc} \mathbb{N} \times \mathbb{N} & \to & \mathbb{N} \\ (m,n) & \mapsto & S^{(n)}(m) \end{array} \right.$$

Wir bezeichnen + als die *Addition auf* $\mathbb{N}$ und schreiben $m + n$ für $+(m,n)$. ◊

♲ Die Nachfolgerabbildung S schreibt sich mit Hilfe der Operation + nun per definitionem als $S(n) = n + 1$. Also modelliert die Addition tatsächlich das n-fache schrittweise Übergehen zu Nachfolgern von m. ♲

Die Funktion + hat etliche Eigenschaften die regeln wie man mit ihr umgehen kann. Man spricht von *Rechenregeln für die Addition* oder auch *Rechengesetzen*.

J354. **3.3.2 Satz.** *Die Operation + hat die folgenden Eigenschaften.*

- *Für alle* $a, b, c \in \mathbb{N}$ *gilt* $(a + b) + c = a + (b + c)$. *(*Assoziativgesetz*)*
- *Für alle* $a, b \in \mathbb{N}$ *gilt* $a + b = b + a$. *(*Kommutativgesetz*)*
- *Sind* $a, b, c \in \mathbb{N}$ *und gilt* $a + c = b + c$*, so folgt* $a = b$. *(*Kürzungsregel*)*
- *Für jedes* $a \in \mathbb{N}$ *ist* $a + 0 = 0 + a = a$. *(*Existenz eines neutralen Elementes*)*

Beweis. Als erstes bemerken wir, dass nach Proposition 3.1.5(2) für je zwei Zahlen $m, n \in \mathbb{N}$

$$m + n = S^{(n)}(m) = S^{(n)}(S^{(m)}(0)) = [S^{(n)} \circ S^{(m)}](0). \tag{3.3.1}$$ J356

Damit lassen sich die genannten Rechenregeln nun leicht herleiten.

Zum Beweis der Assoziativität: Seien $a, b, c \in \mathbb{N}$ vorgegeben. Dann haben wir, mit (3.3.1) und Lemma 3.1.8(2),

$$\begin{aligned}(a+b)+c &= [S^{(c)} \circ S^{(a+b)}](0) = [S^{(c)} \circ S^{\left(S^{(b)}(S^{(a)}(0))\right)}](0)\\ &= [S^{(c)} \circ S^{(b)} \circ S^{\left(S^{(a)}(0)\right)}](0) = [S^{(c)} \circ S^{(b)} \circ S^{(a)} \circ S^{(0)}](0) = [S^{(c)} \circ S^{(b)} \circ S^{(a)}](0)\\ a+(b+c) &= [S^{(b+c)} \circ S^{(a)}](0) = [S^{\left(S^{(c)}(S^{(b)}(0))\right)} \circ S^{(a)}](0)\\ &= [S^{(c)} \circ S^{\left(S^{(b)}(0)\right)} \circ S^{(a)}](0) = [S^{(c)} \circ S^{(b)} \circ S^{(0)} \circ S^{(a)}](0) = [S^{(c)} \circ S^{(b)} \circ S^{(a)}](0)\end{aligned}$$

Zum Beweis der Kommutativität: Seien $a, b \in \mathbb{N}$ vorgegeben. Dann folgt, mit (3.3.1) und Lemma 3.1.8(6),

$$a + b = [S^{(b)} \circ S^{(a)}](0) = [S^{(a)} \circ S^{(b)}](0) = b + a.$$

Zum Beweis der Kürzungsregel: Sei vorausgesetzt dass $a + c = b + c$. Dann ist, nach Definition der Operation $+$ also $S^{(c)}(a) = S^{(c)}(b)$. Wegen der Injektivität, man erinnere sich an Lemma 3.1.8(3), folgt $a = b$.

Zum Beweis der Existenz eines neutralen Elementes: Wegen $S^{(0)} = \mathrm{id}_{\mathbb{N}}$ gilt

$$a + 0 = S^{(0)}(a) = \mathrm{id}_{\mathbb{N}}(a) = a.$$

Nun folgt, mit Hilfe der schon bewiesenen Kommutativität, dass auch $0 + a = a + 0 = a$.

□

J355. *3.3.3 Bemerkung.* Eine algebraischen Operation (wie bei uns $+$) hat höchstens ein neutrales Element. Denn sind e_1 und e_2 neutrale Elemente, so gilt $e_1 = e_1 + e_2 = e_2$. ◊

3.3.3 Die Multiplikation

Nun wollen wir das im obigen §3.3.1 genannte Prozedere (ii) formalisieren.

J312. **3.3.4 Definition.** Wir definieren eine Funktion $\cdot : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ als

$$\cdot : \left\{ \begin{array}{ccc} \mathbb{N} \times \mathbb{N} & \to & \mathbb{N} \\ (m,n) & \mapsto & (S^{(m)})^{(n)}(0) \end{array} \right.$$

Wir bezeichnen $\cdot$ als die *Multiplikation auf* $\mathbb{N}$ und schreiben $m \cdot n$ für $\cdot(m,n)$, ◊

♲ Die m-te Interierte der Nachfolgerabbildung S entspricht dem Dazugeben von m Stück. Ihre n-te Iterierte modelliert also das n-fache Dazugeben der selben Anzahl m. ♲

Man spricht von Addition und Multiplikation auch als den *algebraischen Operationen auf* $\mathbb{N}$.

Auch die Multiplikation erfüllt einige *Rechenregeln*.

J311. **3.3.5 Satz.** *Die Operation $\cdot$ hat die folgenden Eigenschaften.*

- *Für alle $a, b, c \in \mathbb{N}$ gilt $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.* (Assoziativgesetz)
- *Für alle $a, b \in \mathbb{N}$ gilt $a \cdot b = b \cdot a$.* (Kommutativgesetz)
- *Sind $a, b \in \mathbb{N}$, $c \in \mathbb{N} \setminus \{0\}$ und gilt $a \cdot c = b \cdot c$, so folgt $a = b$.* (Kürzungsregel)
- *Für jedes $a \in \mathbb{N}$ ist $a \cdot 1 = 1 \cdot a = a$.* (Existenz eines neutralen Elementes)
- *Für jedes $a \in \mathbb{N}$ ist $a \cdot 0 = 0 \cdot a = 0$.* (Multiplikation mit 0)

Beweis.

Zum Beweis der Assoziativität: Dies folgt leicht mit Lemma 3.1.8(8). Seien $a, b, c \in \mathbb{N}$ vorgegeben. Dann haben wir

$$\begin{aligned}(a \cdot b) \cdot c &= \left[(S^{(a)})^{(b)}(0)\right] \cdot c = \left[S^{\left((S^{(a)})^{(b)}(0)\right)}\right]^{(c)}(0) \overset{(8)}{=} \left[(S^{(a)})^{(b)}\right]^{(c)}(0)\\ a \cdot (b \cdot c) &= (S^{(a)})^{(b \cdot c)}(0) = (S^{(a)})^{\left([S^{(b)}]^{(c)}(0)\right)}(0) \overset{(8)}{=} \left[(S^{(a)})^{(b)}\right]^{(c)}(0)\end{aligned}$$

Zum Beweis der Kommutativität: Dies ist gerade Lemma 3.1.8(7): Seien $a, b \in \mathbb{N}$ vorgegeben, dann ist

$$a \cdot b = (S^{(a)})^{(b)}(0) \overset{(7)}{=} (S^{(b)})^{(a)}(0) = b \cdot a.$$

Zum Beweis der Kürzungsregel: Sei vorausgesetzt dass $a \cdot c = b \cdot c$. Nach der bereits bewiesenen Kommutativität ist dann auch $c \cdot a = c \cdot b$. Nach Definition der Operation $\cdot$, ist also $[S^{(c)}]^{(a)}(0) = [S^{(c)}]^{(b)}(0)$. Wir machen eine Fallunterscheidung.

✾ *Fall 1; $a \leq b$.* Sei $k \in \mathbb{N}$ mit $b = S^{(k)}(a)$. Wir berechnen, mit Lemma 3.1.8(2) und (6),

$$\begin{aligned}[S^{(c)}]^{(a)}(0) &= [S^{(c)}]^{(b)}(0) = [S^{(c)}]^{(S^{(k)}(a))}(0) \\ &\overset{(2)}{=} ([S^{(c)}]^{(k)} \circ [S^{(c)}]^{(a)})(0) \overset{(6)}{=} ([S^{(c)}]^{(a)} \circ [S^{(c)}]^{(k)})(0) = [S^{(c)}]^{(a)}\Big([S^{(c)}]^{(k)}(0)\Big).\end{aligned}$$

Da S injektiv ist, ist nach Lemma 3.1.8(3) auch $S^{(c)}$ injektiv, und in Folge auch $[S^{(c)}]^{(a)}$ injektiv. Daher erhalten wir $0 = [S^{(c)}]^{(k)}(0)$. Sei $c' \in \mathbb{N}$ mit $c = S(c')$, dann können wir berechnen

$$0 = [S^{(c)}]^{(k)}(0) = [S^{(S(c'))}]^{(k)}(0) = [S \circ S^{(c')}]^{(k)}(0) = (S^{(k)} \circ [S^{(c')}]^{(k)})(0) = S^{(k)}\Big([S^{(c')}]^{(k)}(0)\Big).$$

Nach Lemma 3.1.9(2) folgt $k = 0$. Also haben wir $a = b$.

✾ *Fall 2; $b \leq a$.* Wir wenden den bereits bewiesenen Fall 1 an mit „b und a“ anstelle von „a und b“. Dies zeigt, dass wiederum $b = a$.

Da $\leq$ eine Totalordnung ist, decken diese beiden Fälle alle Möglichkeiten ab.

Zum Beweis der Existenz eines neutralen Elementes: Für jede Funktion f gilt $f^{(1)} = f$, also haben wir

$$a \cdot 1 = (S^{(a)})^{(1)}(0) = S^{(a)}(0) = a.$$

Nun folgt, mit Hilfe der schon bewiesenen Kommutativität, dass auch $1 \cdot a = a \cdot 1 = a$.

Zum Beweis der Regel für Multiplikation mit 0*:* Für jede Funktion f gilt $f^{(0)} = \mathrm{id}$, also haben wir

$$a \cdot 0 = (S^{(a)})^{(0)}(0) = \mathrm{id}_{\mathbb{N}}(0) = 0.$$

Nun folgt, mit Hilfe der schon bewiesenen Kommutativität, dass auch $0 \cdot a = a \cdot 0 = 0$.

□

3.3.4 Zusammenspiel der definierten Operationen

Die algebraischen Operatonen der Addition und Multiplikation, sowie die Ordnung der natürlichen Zahlen, hängen über einige weitere *Rechenregeln* zusammen.

J313. **3.3.6 Satz.** *Die Operationen + und $\cdot$, und die Relation $\leq$, haben die folgenden Eigenschaften.*

- *Für alle $a, b, c \in \mathbb{N}$ gilt $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.* (Distributivgesetz)
- *Seien $a, b \in \mathbb{N}$ und $c \in \mathbb{N}$. Dann gilt $a < b$, genau dann wenn $a + c < b + c$.* (Monotonieeigenschaft von $+$)
  Insbesondere gilt, für $a, b \in \mathbb{N}$ und $c \in \mathbb{N}$, dass $a \leq b$ genau dann wenn $a + c \leq b + c$.
- *Sind $a, b \in \mathbb{N}$ und $c \in \mathbb{N} \setminus \{0\}$, dann gilt $a < b$ genau dann wenn $a \cdot c < b \cdot c$.* (Monotonieeigenschaft von $\cdot$)
  Insbesondere gilt, für $a, b \in \mathbb{N}$ und $c \in \mathbb{N} \setminus \{0\}$, dass $a \leq b$ genau dann wenn $a \cdot c \leq b \cdot c$.

Beweis.

Zum Beweis der Distributivität: Seien $a, b, c \in \mathbb{N}$ vorgegeben. Wir berechnen, mit Lemma 3.1.8(2),(8),

$$\begin{aligned}a \cdot (b + c) &= (S^{(a)})^{(b+c)}(0) = (S^{(a)})^{\left(S^{(c)}(b)\right)}(0) \overset{(2)}{=} [(S^{(a)})^{(c)} \circ (S^{(a)})^{(b)}](0) \\ &= (S^{(a)})^{(c)}\Big((S^{(a)})^{(b)}(0)\Big) \overset{(8)}{=} S^{\left((S^{(a)})^{(c)}(0)\right)}\Big((S^{(a)})^{(b)}(0)\Big) = S^{(a \cdot c)}(a \cdot b) = a \cdot b + a \cdot c.\end{aligned}$$

Zum Beweis der Monotonieeigenschaft der Addition: Seien $a, b \in \mathbb{N}$ und $c \in \mathbb{N}$ vorgegeben. Zuerst sei vorausgesetzt, dass $a < b$. Sei $k \in \mathbb{N}$ mit $b = S^{(k)}(a)$. Dann folgt, mit Lemma 3.1.8(6),

$$b + c = S^{(c)}(b) = S^{(c)}(S^{(k)}(a)) = [S^{(c)} \circ S^{(k)}](a) \overset{(6)}{=} [S^{(k)} \circ S^{(c)}](a) = S^{(k)}(S^{(c)}(a)) = S^{(k)}(a + c).$$

Wir sehen, dass $a + c \leq b + c$. Da $a \neq b$, folgt nach der Kürzungsregel Satz 3.3.2, dass $a + c \neq b + c$. Also haben wir $a + c < b + c$.

Für die umgekehrte Implikation verwenden wir Beweis durch Kontraposition. Sei vorausgesetzt, dass $\neg(a < b)$. Da $\leq$ eine Totalordnung ist, folgt $b \leq a$. Mit der bereits bewiesenen Implikation erhalten wir $b + c \leq a + c$. Da $\leq$ antisymmetrisch ist, folgt $\neg(a + c < b + c)$.

Zum Beweis der Monotonieeigenschaft der Multiplikation: Seien $a, b \in \mathbb{N}$ und $c \in \mathbb{N} \setminus \{0\}$ vorgegeben. Zuerst sei vorausgesetzt, dass $a < b$. Sei $k \in \mathbb{N}$ mit $b = S^{(k)}(a)$. Es folgt, mit Lemma 3.1.8(2),(5),(8),

$$\begin{aligned} b \cdot c &= (S^{(b)})^{(c)}(0) = \left(S^{(S^{(k)}(a))}\right)^{(c)}(0) \overset{(2)}{=} (S^{(k)} \circ S^{(a)})^{(c)}(0) \overset{(5)}{=} [(S^{(k)})^{(c)} \circ (S^{(a)})^{(c)}](0) \\ &= (S^{(k)})^{(c)}\left((S^{(a)})^{(c)}(0)\right) = (S^{(k)})^{(c)}(a \cdot c) \overset{(8)}{=} S^{\left((S^{(k)})^{(c)}(0)\right)}(a \cdot c). \end{aligned}$$

Wir sehen, dass $a \cdot c \leq b \cdot c$. Da $c \neq 0$ ist und $a \neq b$, folgt nach der Kürzungsregel Satz 3.3.5, dass $a \cdot c \neq b \cdot c$. Also haben wir $a \cdot c < b \cdot c$.

Für die umgekehrte Implikation verwenden wir Beweis durch Kontraposition. Sei vorausgesetzt, dass $\neg(a < b)$. Da $\leq$ eine Totalordnung ist, folgt $b \leq a$. Mit der bereits bewiesenen Implikation erhalten wir $b \cdot c \leq a \cdot c$. Da $\leq$ antisymmetrisch ist, folgt $\neg(a \cdot c < b \cdot c)$.

□

3.3.5 Einige weitere Varianten des Induktionsprinzips

In den folgenden Korollaren wollen wir zwei weitere häufig verwendete Varianten des Prinzips der vollständigen Induktion angeben. Für den Vergleich mit Satz 3.1.6, erinnere man sich, dass $S(n) = n + 1$.

J316. **3.3.7 Korollar.** *Sei $k \in \mathbb{N}$. Für jedes $n \in \mathbb{N}$, $n \geq k$, sei $A(n)$ eine Aussage. Wenn*

(1) *$A(k)$ ist wahr,*

(2) *für jedes $n \in \mathbb{N}$, $n \geq k$, ist die Implikation $A(n) \Rightarrow A(n + 1)$ wahr,*

dann ist $A(n)$ für alle $n \in \mathbb{N}$, $n \geq k$, wahr.

Beweis. Für jedes $n \in \mathbb{N}$ sei $B(n)$ gleich der Aussage $A(n + k)$.

✧ $B(n)$: $A(n + k)$

Dann sind die obigen Voraussetzungen (1) und (2) gerade

(1) $B(0)$ ist wahr;

(2) für jedes $n \in \mathbb{N}$ ist die Implikation $B(n) \Rightarrow B(n + 1)$ wahr.

Daher können wir Satz 3.1.6 anwenden, und schließen dass $B(n)$ für alle $n \in \mathbb{N}$ wahr ist. Das bedeutet gerade, dass $A(n)$ für alle $n \geq k$ wahr ist. □

J315. **3.3.8 Korollar.** *Für jedes $n \in \mathbb{N}$ sei $A(n)$ eine Aussage. Weiters sei $k \in \mathbb{N}$. Wenn*

(1) *$A(0), \ldots, A(k)$ sind wahr,*

(2) *für jedes $n \in \mathbb{N}$ ist die Implikation $(A(n) \wedge \ldots \wedge A(n + k)) \Rightarrow A(n + k + 1)$ wahr,*

dann ist $A(n)$ für alle $n \in \mathbb{N}$ wahr.

Beweis. Wir gehen darauf los Satz 3.2.5 anzuwenden. Um Satz 3.2.5(1) zu zeigen, sei uns $n \in \mathbb{N}$ vorgegeben. Sei vorausgesetzt dass $A(m)$ wahr ist für alle $m < n$ (die Induktionsvoraussetzung, sprich, die Prämisse in Satz 3.2.5(1)). Um die Konklusion in Satz 3.2.5(1) zu zeigen, machen wir eine Fallunterscheidung.

✾ *Fall 1;* $n \leq k$. Die Aussage $A(n)$ ist wahr wegen (1).

✾ *Fall 2;* $n > k$. Es ist $n \geq k+1$, und daher existiert $n' \in \mathbb{N}$ mit $n = n' + k + 1$. Nun ist $n', \ldots, n' + k < n$, und damit sind insbesondere alle Aussagen $A(n'), \ldots, A(n' + k)$ wahr. Wegen (2) folgt dass $A(n)$ wahr ist.

Also ist Satz 3.2.5(1) erfüllt, wir können Satz 3.2.5 anwenden, und schließen dass $A(n)$ für alle $n \in \mathbb{N}$ wahr ist. □

Schließlich geben wir noch eine Mischform der Typen Satz 3.2.5 und Korollar 3.3.7 des Induktionsprinzips.

J362. **3.3.9 Korollar.** *Sei $k \in \mathbb{N}$. Für jedes $n \in \mathbb{N}$, $n \geq k$, sei $A(n)$ eine Aussage. Wenn*

(1) *$A(k)$ ist wahr,*

(2) *für jedes $n \in \mathbb{N}$, $n \geq k+1$, ist die Implikation $(\forall m \in \mathbb{N}, k \leq m < n\colon A(m)) \Rightarrow A(n)$ wahr,*

dann ist $A(n)$ für alle $n \in \mathbb{N}$, $n \geq k$, wahr.

Beweis. Für jedes $n \in \mathbb{N}$ sei $B(n)$ gleich der Aussage $A(n+k)$.

✧ $B(n)\colon \quad A(n+k)$

Dann sind die obigen Voraussetzungen (1) und (2) gerade

(1) $B(0)$ ist wahr;

(2) für jedes $n \in \mathbb{N} \setminus \{0\}$ ist die Implikation $(\forall m \in \mathbb{N}, m < n\colon B(m)) \Rightarrow B(n)$ wahr.

Daher können wir Satz 3.2.5 anwenden, und schließen dass $B(n)$ für alle $n \in \mathbb{N}$ wahr ist. Das bedeutet gerade, dass $A(n)$ für alle $n \geq k$ wahr ist. □

3.4 Das Schubfachprinzip

3.4.1 Motivation

Die folgende intuitive, und fast lächerlich trivial anmutende, Feststellung ist in Wirklichkeit ein mächtiges und oft verwendetes Beweisprinzip. Es wird *Schubfachprinzip* (oder, synonym, *pigeon hole principle*) genannt.

> Hat man einen Kasten mit einer gewissen Anzahl n von Laden, und möchte mehr als n Paare Socken einräumen, so muss es mindestens eine Lade geben wo mehr als ein Paar Socken landet.

Die Funktion die jedem Paar Socken seine Lade zuordnet ist also nicht injektiv.

3.4.2 Formulierung und Beweis

J320. **3.4.1 Satz.** *Seien $n, m \in \mathbb{N}$ mit $m > n$. Dann gibt es keine injektive Funktion f von $\{k \in \mathbb{N} : k < m\}$ nach $\{k \in \mathbb{N} : k < n\}$.*

Beweis. Der Beweis verläuft in vier Schritten.

1. Eine vorbereitende Bemerkung über die Mengen $M_n = \{k \in \mathbb{N} : k < n\}$.
2. Eine vorbereitende Konstruktion bijektiver Funktionen von M_n in sich.
3. Beweis des Schubfachprinzips im Fall $m = n + 1$.
4. Ableitung des allgemeinen Falles.

Schritt 1: Sei $k, n \in \mathbb{N}$. Dann gilt $k < n+1$ genau dann wenn entweder $k < n$ oder $k = n$. Die Menge M_{n+1} enthält also genau ein Element mehr als M_n; wir können schreiben $M_{n+1} = M_n \cup \{n\}$ wobei wir wissen dass $n \notin M_n$.

Weiters wollen wir anmerken, dass $M_0 = \emptyset$, $M_1 = \{0\}$, und $M_2 = \{0, 1\}$ ist.

Schritt 2: Sei $n \in \mathbb{N}$, $n \geq 2$, und seien $a, b \in M_n$, $a \neq b$. Betrachte die Funktion $\sigma_{a,b} : M_n \to M_n$ die agiert als

$$\sigma_{a,b}(x) = \begin{cases} x, & x \in M_n \setminus \{a, b\}, \\ b, & x = a, \\ a, & x = b. \end{cases} \tag{3.4.1}$$ J323

Das ist also die Funktion die die beiden Elemente a und b vertauscht und alle anderen festläßt. Diese Funktion ist bijektiv, denn $\sigma_{a,b} \circ \sigma_{a,b} = \mathrm{id}_{M_n}$.

Schritt 3: Wir betrachten nun den Fall dass $m = n + 1$.

✾ *Fall 1; $n = 0$.* Zuerst erledigen wir den Sonderfall dass $n = 0$. Dann ist $M_1 = \{0\} \neq \emptyset$ und $M_0 = \emptyset$. Also gibt es keine Funktion von M_1 nach M_0. Insbesondere gibt es keine injektive Funktion von M_1 nach M_0.

✾ *Fall 2; $n \geq 1$.* Wir machen Induktion nach n. Für $n \geq 1$ sei $A(n)$ die Aussage

✧ $A(n)$: $\nexists f : M_{n+1} \to M_n : f$ injektiv

– *Induktionsanfang; $A(1)$ ist wahr:*
Sei $f : M_2 \to M_1$ eine Funktion. Da M_1 nur ein Element hat, nämlich 0, gilt $f(0) = 0$ und auch $f(1) = 0$. Nun ist $0 \neq 1$ (da 1 im Bild der Nachfolgerabbildung liegt, 0 jedoch nicht), also ist f nicht injektiv.

– *Induktionsschritt; $\forall n \in \mathbb{N}, n \geq 1 : A(n) \Rightarrow A(n+1)$:*
Sei $n \in \mathbb{N}$, $n \geq 1$, vorgegeben, und sei vorausgesetzt (die Induktionsvoraussetzung), dass

$$\nexists f : M_{n+1} \to M_n : f \text{ injektiv} \tag{3.4.2}$$ J338

Die gewünschte Konklusion ist $\nexists f : M_{(n+1)+1} \to M_{n+1} : f$ injektiv.
Wir verwenden Beweis durch Kontraposition. Sei vorausgesetzt dass es eine injektive Funktion von $M_{(n+1)+1}$ nach M_{n+1} gibt. Wähle eine solche, sagen wir f. Die Funktion $g = \sigma_{n,f(n+1)} \circ f$ ist injektiv, da beide, $\sigma_{n,f(n+1)}$ und f, injektiv sind; man erinnere sich an Proposition 2.5.5(1). Nun gilt $g(n+1) = \sigma_{n,f(n+1)}(f(n+1)) = n$, und wegen der Injektivität von g daher $g(x) \in M_{n+1} \setminus \{n\}$ für alle $x \in M_{(n+1)+1} \setminus \{n+1\}$. Das besagt gerade, dass $g(M_{n+1}) \subseteq M_n$ gilt. Daher induziert g eine Einschränkungsfunktion $\tilde{g} : M_{n+1} \to M_n$, man erinnere sich an Proposition 2.4.9, und nach Proposition 2.5.6(1) ist $\tilde{g}$ injektiv.

Das Induktionsprinzip, angewendet in der Form von Korollar 3.3.7, zeigt dass die Aussage $A(n)$ für alle $n \geq 1$ gilt, sprich, dass es für kein $n \geq 1$ eine injektive Funktion von M_{n+1} nach M_n geben kann.

Schritt 4: Sei nun $m > n$ beliebig, und sei $f : M_m \to M_n$ eine Funktion. Da $m > n$ ist, gilt $m \geq n + 1$, und daher ist $M_{n+1} \subseteq M_m$. Die Einschränkung $f|_{M_{n+1}}$ ist dann eine Funktion von M_{n+1} nach M_n, und, nach dem in Schritt 3 gezeigten, nicht injektiv. Wegen Proposition 2.5.6(1) (de facto der Kontraposition davon) ist auch f nicht injektiv.

□

J353. *3.4.2 Beispiel.* Wir (und damit ist gemeint der Erzähler und seine Frau) sind in ein neues Haus gezogen. Es hat drei Stockwerke. Dann ist zu jeder Zeit in mindestens einem Stock niemand.

In diesem Beispiel spielen die Personen die Rolle der Laden, und die Stockwerke die Rolle der Socken! ◊

3.4.3 Endliche Mengen

J317. **3.4.3 Definition.** Sei M eine Menge. Dann heißt M *endlich*, wenn es eine Zahl $n \in \mathbb{N}$ und eine Funktion $f : M \to M_n$ gibt (wobei wieder $M_n = \{k \in \mathbb{N} : k < n\}$), sodass f bijektiv ist.

$\left[\exists n \in \mathbb{N} \, \exists f : M \to M_n : f \text{ bijektiv} \right]$

Eine Menge die nicht endlich ist heißt *unendlich*. ◊

♲ Diese Definition modelliert die Anschauung, dass man mit den natürlichen Zahlen alles endliche auch wirklich zählen kann. Die endlichen Mengen sind damit auch genau jene die man (wenn man genug Papier und Bleistift hat) in aufzählender Schreibweise angeben kann; ist M endlich, so können wir M anschreiben als ($T(k)$ bezeichne den Vorgänger von k)

$$M = \{m_1, m_2, \ldots, m_n\} \quad \text{mit} \quad m_k = f(T(k)),$$

wo f wie in Definition 3.4.3 ist. Umgekehrt ist jede Menge dieser Gestalt endlich, denn die Funktion $f : k \mapsto m_{S(k)}$ ist eine Bijektion von M_n auf M. ♧

Nach dem Schubfachprinzip kann es für $n \neq m$ keine bijektive Funktion von M_n nach M_m geben. Für eine endliche Menge M ist daher jene Zahl n, für die es eine bijektive Funktion von M nach M_n gibt, eindeutig bestimmt. Man nennt sie die *Anzahl der Elemente* von M, oder auch ihre *Kardinalität*, und schreibt $\#M$ oder $|M|$.

J319. *3.4.4 Beispiel.* Die Menge $\mathbb{N}$ aller natürlichen Zahlen ist unendlich.

Um dies zu sehen, betrachten wir eine beliebige Zahl $n \in \mathbb{N}$ und eine beliebige Funktion $f : \mathbb{N} \to M_n$. Die Einschränkung $f|_{M_{n+1}}$ ist eine Funktion von M_{n+1} nach M_n, und kann nach dem Schubfachprinzip daher nicht injektiv sein. Nach Proposition 2.5.6(1) ist auch f nicht injektiv. In Formeln angeschrieben:

$$\begin{aligned} \forall n \in \mathbb{N}\, \forall f\colon \mathbb{N} \to M_n\colon \neg(f \text{ injektiv}) &\Leftrightarrow \neg\big(\exists n \in \mathbb{N}\, \exists f\colon \mathbb{N} \to M_n\colon f \text{ injektiv}\big) \\ &\Rightarrow \neg\big(\exists n \in \mathbb{N}\, \exists f\colon \mathbb{N} \to M_n\colon f \text{ bijektiv}\big) \end{aligned}$$

♢

Als ein komplexeres Beispiel wollen wir die Permutationsgruppe $\mathrm{Sym}(M)$ einer endlichen Menge betrachten.

In Beispiel 2.5.15 haben wir $\mathrm{Sym}(M)$ für die Mengen $M_0 = \emptyset$, $M_1 = \{0\}$, und $M_2 = \{0, 1\}$, explizit bestimmt. Dabei haben wir festgestellt, dass eine Funktion $f : M_n \to M_n$ (wobei $n \in \{0, 1, 2\}$) genau dann injektiv ist, wenn sie surjektiv ist. Für beliebige Mengen muss das nicht gelten; man denke an die Nachfolgerabbildung auf $\mathbb{N}$. Es ist eine bemerkenswerte Tatsache, dass die genannte Äquivalenz für alle endlichen Mengen gilt (tatsächlich charakterisiert sie Endlichkeit, der Beweis dieser Tatsache würde hier aber zu weit führen).

J321. **3.4.5 Satz.** *Sei M eine endliche Menge und $f : M \to M$. Dann ist f genau dann injektiv, wenn f surjektiv ist.*

Beweis. Der Beweis verläuft in drei Schritten.

1. Eine Reduktion: o.B.d.A. kann man sich auf Mengen ganz spezieller Gestalt beschränken, nämlich auf die Mengen $M_n = \{k \in \mathbb{N} : k < n\}$ wobei $n \in \mathbb{N}$.
2. Beweis der Implikation „f injektiv $\Rightarrow$ f surjektiv“ für $f : M_n \to M_n$.
3. Beweis der Implikation „f injektiv $\Leftarrow$ f surjektiv“ für $f : M_n \to M_n$.

Schritt 1; Reduktion: Wir zeigen dass es o.B.d.A. genügt den Fall zu betrachten dass $M = M_n$ für ein $n \in \mathbb{N}$.

Sei vorausgesetzt, dass der folgende Spezialfall des Satzes gilt: die Aussage des Satzes ist für alle Mengen M_n mit $n \in \mathbb{N}$ richtig. Sei uns eine beliebige endliche Menge M vorgegeben. Wir gehen darauf los zu zeigen, dass die Aussage des Satzes auch für M gilt.

Da M endlich ist, existiert $n \in \mathbb{N}$ und eine bijektive Funktion $\phi : M \to M_n$. Sei g eine Funktion von M nach M, dann ist $\tilde{g} = \phi^{-1} \circ g \circ \phi$ eine Funktion von M_n nach M_n. Da ϕ und auch ϕ^{-1} bijektiv sind, ist die Funktion g genau dann injektiv (bzw. surjektiv), wenn $\phi^{-1} \circ g \circ \phi$ injektiv (bzw. surjektiv) ist.

Betrachten wir nun unsere gegebene Funktion $f : M \to M$. Ist f injektiv, so ist, nach dem oben Angemerkten, $\tilde{f} = \phi^{-1} \circ f \circ \phi : M_n \to M_n$ auch injektiv. Nach dem vorausgesetzten Spezialfall des Satzes ist $\tilde{f}$ surjektiv. Damit ist f, nach dem oben Angemerkten, ebenfalls surjektiv.

Die Implikation „f injektiv $\Leftarrow$ f surjektiv“ sieht man genauso.

Schritt 2; „f injektiv $\Rightarrow$ f surjektiv“ für $f : M_n \to M_n$: Wir machen eine Fallunterscheidung.

✾ *Fall 1; $n = 0$.* Ist $n = 0$ so wissen wir dass die Implikation gilt; es gibt ja nur die eine Funktion id_{M_0}, und diese ist bijektiv.

✾ *Fall 2; $n \geq 1$.* Sei im Folgenden angenommen dass $n \geq 1$. Sei m der Vorgänger von n, dann ist also $M_n = M_m \cup \{m\}$ und $m \notin M_m$. Wir verwenden Beweis durch Kontraposition. Sei vorausgesetzt dass die uns vorgegebene Funktion $f : M_n \to M_n$ nicht surjektiv ist. Dann ist also $f(M_n) \subsetneq M_n$. Wähle $x_0 \in M_n \setminus f(M_n)$. Sei $\sigma_{x_0,m}$ die Permutation von M_n die die beiden Elemente x_0 und m vertauscht, und alle anderen festläßt, vgl. (3.4.1). Betrachte die Funktion $g = \sigma_{x_0,m} \circ f$. Für diese gilt $m \notin g(M_n)$, sprich, $g(M_n) \subseteq M_n \setminus \{m\} = M_m$. Ihre Einschränkungsfunktion $\tilde{g} : M_n \to M_m$ ist nach dem Schubfachprinzip Satz 3.4.1 nicht injektiv. Nach (der Kontraposition von) Proposition 2.5.6(1) ist f nicht injektiv.

Schritt 3; „f injektiv ⇐ f surjektiv" für $f : M_n \to M_n$: Sei vorausgesetzt dass f surjektiv ist. Nach Satz 2.5.10(2) existiert eine Rechtsinverse von f. Wähle eine solche; wollen wir sie g nennen. Nach Satz 2.5.10(1) ist $g : M_n \to M_n$ injektiv, und nach der in Schritt 2 bereits gezeigten Implikation daher auch surjektiv, sprich, g ist bijektiv. Daher hat g eine Inverse, und diese ist ebenfalls bijektiv (wieder – zwei mal – Satz 2.5.10). Nun gilt

$$g^{-1} = \mathrm{id}_{M_n} \circ g^{-1} = (f \circ g) \circ g^{-1} = f \circ (g \circ g^{-1}) = f \circ \mathrm{id}_{M_n} = f,$$

insbesondere ist f bijektiv und damit auch injektiv.

□

Hat man eine Reduktion wie in Schritt 1 dieses Beweises, so sagt man auch dass man *ohne Beschränkung der Allgemeinheit* annehmen kann dass $M = M_n$ für ein $n \in \mathbb{N}$. Abgekürzt schreibt man auch *o.B.d.A.*, bzw. in der englischsprachigen Literatur *w.l.o.g.* für *without loss of generality*.

➢ `Was heißt hier o.B.d.A.:`

Man stelle sich vor man hat einen Satz der behauptet dass $\forall x \in \mathcal{M}\colon A(x)$. Manchmal ist es so, dass man eine Teilklasse $\mathcal{M}' \subseteq \mathcal{M}$ findet sodass die Implikation

$$\big(\forall x \in \mathcal{M}'\colon A(x)\big) \Rightarrow \big(\forall x \in \mathcal{M}\colon A(x)\big) \tag{3.4.3}$$ J322

gilt. Sprich, schafft man es die Wahrheit der Aussage $A(x)$ für alle $x \in \mathcal{M}'$ zu zeigen, so folgt dann schon dass $A(x)$ für alle $x \in \mathcal{M}$ wahr ist. Der Beweis des Satzes, sprich von $\forall x \in \mathcal{M}\colon A(x)$, kann also in zwei Schritte aufgespalten werden:

① Man zeigt die Implikation (3.4.3); das ist die Reduktion.

② Man zeigt $\forall x \in \mathcal{M}'\colon A(x)$ (und wendet den Modus Ponendo Ponens an um auf die Aussage des Satzes zu schließen).

Wenn man in dieser Weise vorgeht sagt man oft „o.B.d.A. kann man annehmen dass $x \in \mathcal{M}'$".

Einen Beweis in zwei Schritte der genannten Art aufzuteilen, kann eine wirklich signifikante Vereinfachung bringen. Zum Beispiel wenn die Teilklasse $\mathcal{M}'$ wesentlich kleiner ist, und/oder nur Objekte enthält die wesentlich konkretere Gestalt haben und damit besser angreifbar sind, und/oder nur Objekte enthält die wesentlich einfachere Eigenschaften haben und damit leichter zu behandeln sind.

In dem unten folgenden Satz 3.4.7 verwenden wir die Funktion *n-Faktorielle*; das ist eine spezielle Funktion von $\mathbb{N}$ in sich, welche in den verschiedensten Zusammenhängen auftritt.

J324. **3.4.6 Lemma.** *Es existiert eine eindeutige Funktion von $\mathbb{N}$ nach $\mathbb{N}$, man schreibt sie als $n \mapsto n!$, mit den Eigenschaften*

$$0! = 1, \qquad \forall n \in \mathbb{N}\colon (n+1)! = (n+1) \cdot n! \tag{3.4.4}$$ J325

Beweis. Der Beweis ist eine Anwendung des Rekursionssatzes (das war aufgrund der rekursiven Gestalt von (3.4.4) zu erwarten; die Anwendung selbst ist jedoch ein bisschen trickreich).

Sei $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ die Funktion die agiert als $g(x, y) = (x + 1, (x + 1) \cdot y)$. Nach dem Rekursionssatz, angewendet mit der Menge $Y = \mathbb{N} \times \mathbb{N}$, ihrem Element $y_0 = (0, 1)$, und der Funktion $g : Y \to Y$, existiert eine eindeutige Funktion $\phi : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ mit den Eigenschaften

$$\phi(0) = (0, 1), \qquad \phi \circ S = g \circ \phi.$$

Wir bezeichnen mit $\pi_1 : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ bzw. $\pi_2 : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ jene Funktionen die einem Paar (x, y) seine erste Komponente x bzw. seine zweite Komponente y zuordnen. Man spricht von diesen Funktionen als die *kanonischen Projektionen* auf die erste bzw. zweite Komponente des Produktes. Bemerke, dass man ein Paar (x, y) aus den Werten $\pi_1(x, y)$ und $\pi_2(x, y)$ rekonstruieren kann, nämlich als $(x, y) = (\pi_1(x, y), \pi_2(x, y))$.

Schritt 1: Als erstes zeigen wir, dass $(\pi_1 \circ \phi)(n) = n$ für alle $n \in \mathbb{N}$ gilt.

Wir machen Induktion nach n. Für $n \in \mathbb{N}$ sei $A(n)$ die Aussage

✧ $A(n)$: $(\pi_1 \circ \phi)(n) = n$

– *Induktionsanfang; $A(0)$ ist wahr:*
Es gilt $(\pi_1 \circ \phi)(0) = \pi_1(0,1) = 0$.
– *Induktionsschritt; $\forall n \in \mathbb{N}\colon A(n) \Rightarrow A(n+1)$:*
Sei $n \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt (die Induktionsvoraussetzung), dass

$$(\pi_1 \circ \phi)(n) = n \tag{3.4.5}$$ J339

Die gewünschte Konklusion ist $(\pi_1 \circ \phi)(n+1) = n+1$.

Wir berechnen, unter Verwendung der Induktionsvoraussetzung (3.4.5),

$$\begin{aligned}(\pi_1 \circ \phi)(n+1) &= (\pi_1 \circ \phi)(S(n)) = \pi_1\big((\phi \circ S)(n)\big) = \pi_1\big((g \circ \phi)(n)\big) = \pi_1\big(g(\phi(n))\big)\\ &= \pi_1\big(g((\pi_1 \circ \phi)(n), (\pi_2 \circ \phi)(n))\big) \overset{(3.4.5)}{=} \pi_1\big(g(n, (\pi_2 \circ \phi)(n))\big)\\ &= \pi_1\big(n+1, (n+1)\cdot(\pi_2 \circ \phi)(n)\big) = n+1.\end{aligned}$$

Das Induktionsprinzip Satz 3.1.6 zeigt dass die Aussage $A(n)$ für alle $n \in \mathbb{N}$ gilt, sprich, dass $(\pi_1 \circ \phi)(n) = n$ für alle $n \in \mathbb{N}$ gilt.

Schritt 2: Nun setzen wir $n! = (\pi_2 \circ \phi)(n)$ für alle $n \in \mathbb{N}$, sprich, $n \mapsto n!$ ist bloss ein anderer Name für die Funktion $\pi_2 \circ \phi$. Dann gilt jedenfalls einmal $0! = (\pi_2 \circ \phi)(0) = \pi_2(0,1) = 1$. Um die zweite Formel in (3.4.4) zu zeigen, sei uns $n \in \mathbb{N}$ vorgegeben. Wir wiederholen die obige Rechnung, unter Verwendung der schon gezeigten Beziehung für die ersten Komponenten, um die zweiten Komponenten zu bestimmen:

$$\begin{aligned}(n+1)! &= (\pi_2 \circ \phi)(n+1) = (\pi_2 \circ \phi)(S(n)) = \pi_2\big((\phi \circ S)(n)\big) = \pi_2\big((g \circ \phi)(n)\big) = \pi_2\big(g(\phi(n))\big)\\ &= \pi_2\big(g((\pi_1 \circ \phi)(n), (\pi_2 \circ \phi)(n))\big) = \pi_2\big(g(n, (\pi_2 \circ \phi)(n))\big) = \pi_2\big(n+1, (n+1)\cdot(\pi_2 \circ \phi)(n)\big)\\ &= (n+1)\cdot(\pi_2 \circ \phi)(n) = (n+1)\cdot n!\,.\end{aligned}$$

□

J318. **3.4.7 Satz.** *Sei $n \in \mathbb{N}$. Dann ist $|\mathrm{Sym}(M_n)| = n!$.*

In diesem Beweis werden wir die folgende – anschaulich sehr einsichtige – Aussage verwenden:
Sei M eine endliche Menge, und sei $\mathcal{Q} = \{A_0, \ldots, A_n\}$ eine Partition von M. Dann gilt $|M| = \sum_{k=0}^{n} |A_k|$.
Diese ist nicht allzu schwer zu beweisen, der Beweis würde jedoch hier zu weit führen, und wir wollen uns daher damit begnügen sie an dieser Stelle als wahr zu akzeptieren.

Das Symbol $\sum_{k=0}^{n} |A_k|$ bezeichnet die Summe aller Zahlen $|A_k|$ wo der Index k von 0 bis n läuft. Sprich, es ist

$$\sum_{k=0}^{n} |A_k| = |A_0| + \ldots + |A_n|.$$

Wie immer wenn wo drei Punkte auftauchen muss man das Hingeschriebene präzisieren. Auch in diesem Fall geht das durch eine Anwendung des Rekursionssatzes. Die Anwendung ist ähnlich wie bei der Definition von n-Faktorielle.

Sei $A : \mathbb{N} \to \mathbb{N}$ eine Funktion. Sei $g : \mathbb{N}\times\mathbb{N} \to \mathbb{N}\times\mathbb{N}$ die Funktion die agiert als $g(x,y) = (x+1, A(x+1)+y)$. Nach dem Rekursionssatz, angewendet mit der Menge $Y = \mathbb{N}\times\mathbb{N}$, ihrem Element $y_0 = (0, A(0))$, und der Funktion $g : Y \to Y$, existiert eine eindeutige Funktion $\phi : \mathbb{N} \to \mathbb{N}\times\mathbb{N}$ mit den Eigenschaften

$$\phi(0) = (0, A(0)), \qquad \phi \circ S = g \circ \phi.$$

Wir bezeichnen wieder mit $\pi_1 : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ bzw. $\pi_2 : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ die kanonischen Projektionen.

Genauso wie im Beweis von Lemma 3.4.6 sieht man, dass $(\pi_1 \circ \phi)(n) = n$ für alle $n \in \mathbb{N}$ gilt. Nun setzen wir $\sum_{k=0}^{n} A(k) = (\pi_2 \circ \phi)(n)$ für alle $n \in \mathbb{N}$.

Dann gilt $\sum_{k=0}^{0} A(k) = (\pi_2 \circ \phi)(0) = \pi_2(0, A(0)) = A(0)$. Sei uns nun $n \in \mathbb{N}$ vorgegeben. Wir berechnen:

$$\begin{aligned}\sum_{k=0}^{n+1} A(k) &= (\pi_2 \circ \phi)(n+1) = (\pi_2 \circ \phi)(S(n)) = \pi_2\big((\phi \circ S)(n)\big) = \pi_2\big((g \circ \phi)(n)\big) = \pi_2\big(g(\phi(n))\big) \\ &= \pi_2\big(g\big((\pi_1 \circ \phi)(n), (\pi_2 \circ \phi)(n)\big)\big) = \pi_2\big(g\big(n, (\pi_2 \circ \phi)(n)\big)\big) = \pi_2\big(n+1, A(n+1) + (\pi_2 \circ \phi)(n)\big) \\ &= A(n+1) + (\pi_2 \circ \phi)(n) = A(n+1) + \sum_{k=0}^{n} A(k)\,.\end{aligned}$$

Also modelliert diese Funktion tatsächlich was man anschaulich unter der Summe verstehen würde.

In diesem Kontext sei es auch erwähnt, dass man eine Summe über eine leere Indexmenge (z.B. etwas wie $\sum_{k=1}^{0} a_k$) per definitionem als 0 versteht. ♧

Beweis von Satz 3.4.7. Wir machen (wer hätte wohl damit gerechnet) vollständige Induktion. Der Induktionsschritt ist allerdings diesmal nicht straightforward; er wird eine Idee erfordern. Für $n \in \mathbb{N}$ sei $A(n)$ die Aussage

✧ $A(n)$: $|\operatorname{Sym}(M_n)| = n!$

– *Induktionsanfang; $A(0)$ ist wahr:*
Wie wir wissen gilt $\operatorname{Sym}(M_0) = \{\mathrm{id}_{M_0}\}$, also ist $|\operatorname{Sym}(M_0)| = 1 = 0!$.

– *Induktionsschritt; $\forall n \in \mathbb{N}$: $A(n) \Rightarrow A(n+1)$:*
Sei $n \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt (die Induktionsvoraussetzung), dass

$$|\operatorname{Sym}(M_n)| = n!\,. \tag{3.4.6}$$ J340

Die gewünschte Konklusion ist $|\operatorname{Sym}(M_{n+1})| = (n+1)!$.

Dazu konstruieren wir eine Partition von $\operatorname{Sym}(M_{n+1})$. Nämlich, für $k \in M_{n+1}$ bezeichne

$$A_k = \{f \in \operatorname{Sym}(M_{n+1}) : \ f(n) = k\}.$$

Wegen **(Fun)** existiert für jede Funktion $f : M_{n+1} \to M_{n+1}$ genau eine Zahl $k \in M_{n+1}$ mit der Eigenschaft dass $f(n) = k$. Daher ist $\mathcal{Q} = \{A_k : \ k \in M_{n+1}\}$ eine Partition von $\operatorname{Sym}(M_{n+1})$.

Wegen der oben angeführten Aussage (an die wir ja vereinbarungsgemäß *glauben*), gilt also

$$|\operatorname{Sym}(M_{n+1})| = \sum_{k=0}^{n} |A_k|. \tag{3.4.7}$$ J326

Sei $k \in M_{n+1}$. Wir gehen darauf los die Anzahl der Elemente von A_k zu berechnen. Dazu machen wir eine Fallunterscheidung. Die Crux ist es die Situation für ein k zu verstehen; daraus kann man dann leicht die gleiche Formel für alle anderen k herleiten. Wir nehmen, aus praktischen Gründen, $k = n$ her.

✾ *Fall 1; $k = n$.* Die Menge A_n ist die Menge aller $f \in \operatorname{Sym}(M_{n+1})$ mit $f(n) = n$. Da f bijektiv ist, folgt daraus $f(M_n) = M_n$, und wir können die Einschränkungsfunktion $\mathcal{E}(f) : M_n \to M_n$ betrachten. Wegen Proposition 2.5.6 ist diese bijektiv. Wir haben also eine Funktion, den Einschränkungsoperator,

$$\mathcal{E} : \left\{ \begin{array}{ccc} A_n & \to & \operatorname{Sym}(M_n) \\ f & \mapsto & \mathcal{E}(f) \end{array} \right.$$

Wir gehen nun darauf los eine Inverse zu $\mathcal{E}$ zu konstruieren; den Fortsetzungsoperator. Sei uns $g \in \operatorname{Sym}(M_n)$ vorgegeben. Dann definiere eine Funktion $\mathcal{F}(g) : M_{n+1} \to M_{n+1}$ als

$$\mathcal{F}(g)(k) = \begin{cases} g(k)\,, & k \in M_n \\ n \quad , & k = n \end{cases}$$

Die Funktion $\mathcal{F}(g)$ ist surjektiv, denn

$$[\mathcal{F}(g)](M_{n+1}) \supseteq g(M_n) \cup \{n\} = M_n \cup \{n\} = M_{n+1}.$$

Wir gehen darauf los zu zeigen, dass die Funktion $\mathcal{F}(g)$ auch injektiv ist. Seien uns $k_1, k_2 \in M_{n+1}$ vorgegeben, und sei vorausgesetzt dass $k_1 \neq k_2$. Wir machen eine Fallunterscheidung.

✾ *Fall 1;* $k_1, k_2 \in M_n$. Wenn $k_1, k_2 \in M_n$, dann gilt $[\mathcal{F}(g)](k_1) = g(k_1) \neq g(k_2) = [\mathcal{F}(g)](k_2)$.
✾ *Fall 2;* $k_1 \in M_n$ *und* $k_2 = n$. Wenn $k_1 \in M_n$ und $k_2 = n$, dann gilt $[\mathcal{F}(g)](k_1) = g(k_1) \in M_n$ und $[\mathcal{F}(g)](k_2) = n$, insbesondere $[\mathcal{F}(g)](k_1) \neq [\mathcal{F}(g)](k_2)$.
✾ *Fall 3;* $k_1 = n$ *und* $k_2 \in M_n$. Wenn $k_1 = n$ und $k_2 \in M_n$, dann gilt $[\mathcal{F}(g)](k_1) = n$ und $[\mathcal{F}(g)](k_2) = g(k_2) \in M_n$, insbesondere $[\mathcal{F}(g)](k_1) \neq [\mathcal{F}(g)](k_2)$.

Diese drei Fälle decken tatsächlich alle Möglichkeiten ab, da ja $k_1 \neq k_2$, und insbesondere nicht beide gleichzeitig gleich n sein können. Wir sehen dass tatsächlich $\mathcal{F}(g)$ injektiv ist, und mithin bijektiv.

Also haben wir eine Funktion, den Fortsetzungsoperator

$$\mathcal{F} : \left\{ \begin{array}{rcl} \mathrm{Sym}(M_n) & \to & A_n \\ g & \mapsto & \mathcal{F}(g) \end{array} \right.$$

Nun gehen wir darauf los zu zeigen dass $\mathcal{E} \circ \mathcal{F} = \mathrm{id}_{\mathrm{Sym}(M_n)}$. Sei uns $g \in \mathrm{Sym}(M_n)$, und $k \in M_n$ vorgegeben. Dann gilt, nach den entsprechenden Definitionen,

$$[(\mathcal{E} \circ \mathcal{F})(g)](k) = [\mathcal{E}(\mathcal{F}(g))](k) = [\mathcal{F}(g)](k) = g(k).$$

Also haben wir tatsächlich $(\mathcal{E} \circ \mathcal{F})(g) = g$.

Nun gehen wir darauf los zu zeigen dass $\mathcal{F} \circ \mathcal{E} = \mathrm{id}_{A_n}$. Sei uns $f \in A_n$, und $k \in M_{n+1}$ vorgegeben.

Wir machen eine Fallunterscheidung.

✾ *Fall 1:* $k = n$. Wenn $k = n$ ist, dann gilt einerseits $[(\mathcal{F} \circ \mathcal{E})(f)](k) = [(\mathcal{F} \circ \mathcal{E})(f)](n) = n$ nach der Definition von $\mathcal{F}$, und andererseits $f(k) = f(n) = n$ da $f \in A_n$.
✾ *Fall 2:* $k \neq n$. Wenn $k \neq n$ ist, dann ist $k \in M_n$ und damit haben wir

$$[(\mathcal{F} \circ \mathcal{E})(f)](k) = [\mathcal{F}(\mathcal{E}(f))](k) = [\mathcal{E}(f)](k) = f(k).$$

Also haben wir tatsächlich $(\mathcal{F} \circ \mathcal{E})(f) = f$.

Insgesamt haben wir eine bijektive Funktion von A_n nach $\mathrm{Sym}(M_n)$, und damit $|A_n| = |\mathrm{Sym}(M_n)|$.

✾ *Fall 2:* $k \neq n$. Wir gehen darauf los eine bijektive Funktion von A_k nach A_n zu konstruieren. Bezeichne $\sigma_{k,n}$ wieder die Permutation die k mit n vertauscht und alle anderen Elemente von M_{n+1} festläßt. Ist $f \in A_k$, so ist $\sigma_{k,n} \circ f \in A_n$, denn als Zusammensetzung bijektiver Funktionen ist diese Funktion ebenfalls bijektiv, und es gilt $(\sigma_{k,n} \circ f)(n) = \sigma_{k,n}(f(n)) = \sigma_{k,n}(k) = n$. Das Gleiche funktioniert auch umgekehrt. Ist $f \in A_n$, so ist $\sigma_{k,n} \circ f \in A_k$, denn als Zusammensetzung bijektiver Funktionen ist diese Funktion ebenfalls bijektiv, und es gilt $(\sigma_{k,n} \circ f)(n) = \sigma_{k,n}(f(n)) = \sigma_{k,n}(n) = k$. Damit haben wir also zwei Funktionen

$$\Phi : \left\{ \begin{array}{rcl} A_k & \to & A_n \\ f & \mapsto & \sigma_{k,n} \circ f \end{array} \right. \qquad \Psi : \left\{ \begin{array}{rcl} A_n & \to & A_k \\ f & \mapsto & \sigma_{k,n} \circ f \end{array} \right.$$

Wegen $\sigma_{k,n} \circ \sigma_{k,n} = \mathrm{id}_{M_{n+1}}$ haben wir

$$(\Psi \circ \Phi)(f) = \Psi(\sigma_{k,n} \circ f) = \sigma_{k,n} \circ (\sigma_{k,n} \circ f) = (\sigma_{k,n} \circ \sigma_{k,n}) \circ f = \mathrm{id}_{M_{n+1}} \circ f = f,$$

$$(\Phi \circ \Psi)(f) = \Phi(\sigma_{k,n} \circ f) = \sigma_{k,n} \circ (\sigma_{k,n} \circ f) = (\sigma_{k,n} \circ \sigma_{k,n}) \circ f = \mathrm{id}_{M_{n+1}} \circ f = f.$$

Also sind tatsächlich Φ und Ψ Inverse voneinander. Wir schließen dass $|A_k| = |A_n|$.

Nachdem wir nun für alle Mengen A_k die Anzahl ihrer Elemente bestimmt haben, können wir in die Formel (3.4.7) einsetzen. Unter Verwendung der Induktionsvoraussetzung und der Eigenschaft (3.4.4) erhalten wir

$$|\mathrm{Sym}(M_{n+1})| = \sum_{k=0}^{n} |A_k| = \sum_{k=0}^{n} |\mathrm{Sym}(M_n)| = (n+1) \cdot |\mathrm{Sym}(M_n)| = (n+1) \cdot n! = (n+1)!$$

Das Induktionsprinzip Satz 3.1.6 zeigt dass die Aussage $A(n)$ für alle $n \in \mathbb{N}$ gilt, sprich, dass $|\mathrm{Sym}(M_n)| = n!$ für alle $n \in \mathbb{N}$. □

In der letzten Zeile des Beweises haben wir verwendet, dass die Summe von $(n+1)$ gleichen Zahlen gleich dem Produkt von $(n+1)$ mit eben dieser Zahl ist. Den Beweis dieser Tatsache führt man mit vollständiger Induktion unter Verwendung des Distributivgesetzes.

3.5 Teilbarkeit und Primzahlen

3.5.1 Motivation

Betrachtet man die additive Struktur der natürlichen Zahlen, so kann man alle Zahlen (ausser der 0) aus einem einzigen Baustein erzeugen: der Zahl 1. Nämlich durch fortgesetztes addieren von 1 zu sich selbst; man erinnere sich an Proposition 3.1.5(2).

> Gibt es einen analogen Baustein – oder Bausteine – wenn man die multiplikative Struktur von $\mathbb{N}$ betrachtet ?

Die Antwort ist Ja; und das wollen wir uns in diesem Abschnitt überlegen.

3.5.2 Produktdarstellung natürlicher Zahlen

Die multiplikative Struktur von $\mathbb{N}$ ist *wesentlich* komplizierter als die Additive. Ein Grund, dafür ist dass man viel weniger „Inverse“ hat. Wie wir wissen – eigentlich per definitionem – findet man zu zwei gegebenen Zahlen $n, m \in \mathbb{N}$ eine dritte, l, mit $n = m + l$, genau dann wenn $n \geq m$. Das multiplikative Analogon ist der Begriff der Teilbarkeit.

J301. **3.5.1 Definition.** Seien $n, m \in \mathbb{N} \setminus \{0\}$. Wir sagen *m teilt n*, und schreiben $m \mid n$, wenn es eine Zahl $l \in \mathbb{N}$ gibt mit $n = m \cdot l$.

$\left[\, m \mid n \;\Leftrightarrow\; \exists l \in \mathbb{N}\colon n = m \cdot l \,\right]$

Eine Zahl m mit $m \mid n$ heißt ein *Teiler* von n. ◊

Zum Beispiel wäre, das kleine 1-mal-1 und die Dezimalschreibweise für natürliche Zahlen aus der Anschauung übernehmend, die Menge aller Teiler von 12 gegeben als $\{1, 2, 3, 4, 6, 12\}$.

Wir wollen anmerken, dass 1 und n stets Teiler von n sind, denn $n = n \cdot 1 = 1 \cdot n$. Weiters ist jeder Teiler m von n nicht größer als n, denn für jedes $l \geq 1$ gilt $m \cdot l \geq m$.

Was sind nun „multiplikative Bausteine“ ? Um dies zu herauszufinden, kehren wir zuerst wieder zu der einfacheren additiven Struktur zurück: den additiven Baustein 1 kann man, von anderer Perspektive gesehen, auch auffassen als Zahl $n \geq 1$ die nur in trivialer Weise additiv zerlegt werden kann, nämlich als $n = n + 0 = 0 + n$. Diese Sichtweise läßt nun ein unmittelbares multiplikatives Analogon zu.

J302. **3.5.2 Definition.** Eine Zahl $p \in \mathbb{N}$ heißt *Primzahl*, wenn $p \geq 2$ und wenn p keine Teiler ausser 1 und sich selbst hat. Sprich, wenn p nur in trivialer Weise multiplikativ zerlegt werden kann, nämlich als $p = p \cdot 1 = 1 \cdot p$.

Wir bezeichnen die Menge aller Primzahlen mit $\mathbb{P}$. ◊

Im nächsten Satz zeigen wir, dass Primzahlen tatsächlich multiplikative Bausteine der natürlichen Zahlen sind.

Natürlich kann man sich bei dem Studium multiplikativer Zerlegungen auf Zahlen $n \geq 2$ beschränken; für $n = 0$ bzw. $n = 1$ ist die Situation ja unmittelbar einsichtig. Nämlich läßt sich 0 zerlegen als $0 = 0 \cdot n = n \cdot 0$ für beliebiges $n \in \mathbb{N}$ (und das sind alle möglichen Zerlegungen), und 1 läßt sich zerlegen als $1 = 1 \cdot 1$ (und das ist die einzige Zerlegung).

J303. **3.5.3 Satz.** *Sei $n \in \mathbb{N}$, $n \geq 2$. Dann existieren $k \in \mathbb{N}$ und (nicht notwendig verschiedene) Primzahlen $p_0, \ldots, p_k$, sodass*

$$n = \prod_{i=0}^{k} p_i. \tag{3.5.1}$$ J304

♲ Das Symbol $\prod_{i=0}^{k} p_i$ bezeichnet das Produkt aller Zahlen p_i wo der Index i von 0 bis k läuft. Sprich, es ist

$$\prod_{k=0}^{k} p_i = p_0 \cdot \ldots \cdot p_k.$$

Wie immer wenn wo drei Punkte auftauchen muss man das Hingeschriebene präzisieren. Auch in diesem Fall geht das durch eine Anwendung des Rekursionssatzes. Die Anwendung ist analog bei der Definition von n-Faktorielle oder der Summe.

Sei $A : \mathbb{N} \to \mathbb{N}$ eine Funktion. Sei $g : \mathbb{N}\times\mathbb{N} \to \mathbb{N}\times\mathbb{N}$ die Funktion die agiert als $g(x, y) = (x + 1, A(x + 1) \cdot y)$. Nach dem Rekursionssatz, angewendet mit der Menge $Y = \mathbb{N}\times\mathbb{N}$, ihrem Element $y_0 = (0, A(0))$, und der Funktion $g : Y \to Y$, existiert eine eindeutige Funktion $\phi : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ mit den Eigenschaften

$$\phi(0) = (0, A(0)), \qquad \phi \circ S = g \circ \phi.$$

Wir bezeichnen wieder mit $\pi_1 : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ bzw. $\pi_2 : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ die kanonischen Projektionen.

Genauso wie im Beweis von Lemma 3.4.6 sieht man, dass $(\pi_1 \circ \phi)(n) = n$ für alle $n \in \mathbb{N}$ gilt. Nun setzen wir $\prod_{k=0}^{n} A(k) = (\pi_2 \circ \phi)(n)$ für alle $n \in \mathbb{N}$.

Dann gilt $\prod_{k=0}^{0} A(k) = (\pi_2 \circ \phi)(0) = \pi_2(0, A(0)) = A(0)$. Sei uns nun $n \in \mathbb{N}$ vorgegeben. Wir berechnen:

$$\begin{aligned}\prod_{k=0}^{n+1} A(k) &= (\pi_2 \circ \phi)(n + 1) = (\pi_2 \circ \phi)(S(n)) = \pi_2\big((\phi \circ S)(n)\big) = \pi_2\big((g \circ \phi)(n)\big) = \pi_2\big(g(\phi(n))\big)\\ &= \pi_2\big(g\big((\pi_1 \circ \phi)(n), (\pi_2 \circ \phi)(n)\big)\big) = \pi_2\big(g\big(n, (\pi_2 \circ \phi)(n)\big)\big) = \pi_2\big(n + 1, A(n + 1) \cdot (\pi_2 \circ \phi)(n)\big)\\ &= A(n + 1) \cdot (\pi_2 \circ \phi)(n) = A(n + 1) \cdot \prod_{k=0}^{n} A(k)\,.\end{aligned}$$

Also modelliert diese Funktion tatsächlich was man anschaulich unter einem Produkt verstehen würde.

In diesem Kontext sei es auch erwähnt, dass man eine Summe über eine leere Indexmenge (z.B. etwas wie $\prod_{k=1}^{0} a_k$) per definitionem als 1 versteht. ♲

Beweis von Satz 3.5.3. Wir machen Induktion nach n. Für $n \in \mathbb{N}$, $n \geq 2$, sei $A(n)$ die Aussage

$$\text{✧ } A(n): \quad \exists\, k \in \mathbb{N}\, \exists\, p_0, \ldots, p_k \in \mathbb{P}\colon n = \prod_{i=0}^{k} p_i$$

– *Induktionsanfang; $A(2)$ ist wahr:*
Die Zahl 2 ist selbst eine Primzahl. Denn ein Teiler von 2 kann nicht größer als 2 sein, und daher sind die einzigen möglichen Teiler von 2 die trivialen Teiler 1 und 2. Die Darstellung $2 = 2$ ist daher eine Darstellung der Form (3.5.1) von 2 als Produkt von Primzahlen (ein Produkt mit nur dem einen Faktor 2 sprich, $k = 0$ und $p_0 = 2$). Also hat 2 eine Darstellung der gewünschten Form.

– *Induktionsschritt;* $\forall n \in \mathbb{N}\colon (\forall m \in \mathbb{N}, m < n\colon A(m)) \Rightarrow A(n)$:
Sei $n \in \mathbb{N}$ vorgegeben, und sei vorausgesetzt (die Induktionsvoraussetzung), dass

$$\text{Alle Zahlen } m \text{ mit } 2 \leq m < n \text{ haben eine Darstellung der Gestalt (3.5.1).} \quad \boxed{\texttt{J341}} \qquad (3.5.2) \quad \boxed{\texttt{J341}}$$

Die gewünschte Konklusion ist zu zeigen, dass auch n eine solche Darstellung hat.

Um dies zu sehen, machen wir eine Fallunterscheidung.

✾ *Fall 1; n ist Primzahl.* Die Darstellung $n = n$ eine Darstellung der Form (3.5.1) von n als Produkt von Primzahlen. Also ist $n \in M$.

✾ *Fall 2; n ist keine Primzahl.* In diesem Fall hat n einen Teiler m der verschieden von 1 und n ist. Schreibe $n = m \cdot l$, dann ist auch l ein Teiler von n der verschieden von 1 und n ist. Denn aus $l = 1$ folgt $m = n$ bzw. aus $l = n$ folgt $m = 1$. Damit gilt $2 \leq m < n$ und $2 \leq l < n$. Nach der Induktionsvoraussetzung (3.5.2) haben m und l beide Darstellungen der Form (3.5.1) als Produkt von Primzahlen. Ihr Produkt $m \cdot l$, das ja gleich n ist, hat daher auch eine Darstellung als Produkt von Primzahlen. Also ist $n \in M$.

Das Induktionsprinzip in der Form von Korollar 3.3.9 zeigt, dass die Aussage $A(n)$ für alle $n \in \mathbb{N}$, $n \geq 2$, gilt, sprich, dass jede Zahl $n \geq 2$ eine Darstellung der Form (3.1.6) hat. □

Wir wollen uns nun der Frage zuwenden wieviele multiplikative Bausteine man braucht um *alle* natürlichen Zahlen zu erzeugen. Nach Satz 3.5.3 *genügt es* Primzahlen als Bausteine herzunehmen. Und weil jede Primzahl p nur die eine multiplikative Zerlegung $p = p$ hat, braucht man auch wirklich *alle* Primzahlen dazu. Die obige Frage ist also nichts anderes wie: Wieviele Primzahlen gibt es ?

J305. **3.5.4 Satz.** *Es gibt unendlich viele Primzahlen.*

Beweis. Sei indirekt angenommen, dass es nur endlich viele Primzahlen gibt; sagen wir $\mathbb{P} = \{p_1, \ldots, p_N\}$ mit einem gewissen $N \in \mathbb{N}$. Betrachte die Zahl

$$n = 1 + \prod_{i=1}^{N} p_i.$$

Es ist sicherlich $n \geq 2$, denn wir wissen dass 2 eine Primzahl ist, und damit haben wir sogar $n \geq 1 + 2 > 2$. Nach Satz 3.5.3 läßt sich n darstellen als Produkt von gewissen Primzahlen. Insbesondere gibt es eine Primzahl welche n teilt. Wähle eine solche, sagen wir p_{i_0} mit geeignetem $i_0 \in \{1, \ldots, N\}$. Dann ist also $n = p_{i_0} \cdot l$ mit einem gewissen $l \in \mathbb{N}$, $l \geq 1$. Wegen $n > \prod_{i=1}^{N} p_i$ ist auch $l > \prod_{\substack{i=1 \\ i \neq i_0}}^{N} p_i$, und wir finden $l' \in \mathbb{N}$, $l' \geq 1$, mit $l = \prod_{\substack{i=1 \\ i \neq i_0}}^{N} p_i + l'$. Damit erhalten wir

$$1 + \prod_{i=1}^{N} p_i = n = p_{i_0} \cdot l = p_{i_0} \cdot \Big(\prod_{\substack{i=1 \\ i \neq i_0}}^{N} p_i + l' \Big) = \prod_{i=1}^{N} p_i + p_{i_0} \cdot l',$$

und daher $p_{i_0} \cdot l' = 1$. Nun ist aber $p_{i_0} \geq 2$ und $l' \geq 1$, und daher auch $p_{i_0} \cdot l' \geq 2$, also sicher $p_{i_0} \cdot l' \neq 1$.

Wir haben einen Widerspruch erhalten, also muss unsere indirekte Annahme – dass es nur endlich viele Primzahlen gibt – falsch sein. Und damit muss die Aussage des Satzes wahr sein. □

♲ `Analyse des Beweises von Satz 3.5.4:`

Dieses Argument ist wieder eine typische Variante eines *indirekten Beweises*.

Wir wollen zeigen, dass die Aussage des Satzes

S : Es gibt unendlich viele Primzahlen.

wahr ist. Dazu betrachten wir ihre Negation $A = \neg S$

A : Es gibt nur endlich viele Primzahlen.

Was wir dann zeigen ist dass eine Implikation $A \Rightarrow F$ wahr ist, wo F einfach *irgendetwas* ist von dem wir *wissen* dass es *falsch* ist. In unserem Fall ist dieses „irgendetwas“ die Aussage $C \wedge \neg C$ mit

C : $p_{i_0} \cdot l' = 1$

Wegen dem Satz vom Widerspruch ist eine Aussage der Gestalt $C \wedge \neg C$ sicher falsch; unabhängig davon was C ist und welchen Wahrheitswert C hat.

Damit wissen wir dass $(A \Rightarrow F) \wedge \neg F$ wahr ist. Nach dem Modus Tollendo Tollens muss $\neg A$ wahr sein. Das zeigt, nach dem Satz von der doppelten Verneinung, dass S wahr ist. ♲

Notation

Index

Ein paar Hinweise

Das wichtigste Konzept, das Sie in diesem Skriptum kennengelernt haben, ist der Begriff des *Beweises*.

Die Erfahrung zeigt, dass wegen des großen Umfangs des in kurzer Zeit behandelten Stoffs manche kleinere Ideen, Hinweise und Bemerkungen oft übersehen werden. Daher wiederholen wir einige dieser Punkte.

(Von Seite 8):
Beachte, dass zwei Mengen M_1 und M_2 genau dann gleich sind, wenn beide wechselseitigen Inklusionen $M_1 \subseteq M_2$ und $M_1 \supseteq M_2$ gelten, also jedes Element der einen Menge in der anderen liegt, und umgekehrt.

(Von Seite 28):
Äquivalenzrelationen und Partitionen sind verschiedene Methoden, um dasselbe dahinter stehende Konzept zu beschreiben, nämlich dass gewisse Elemente einer Menge einander näher stehen als andere.

(Von Seite 38):
Seien M, N Mengen, f eine Funktion von M nach N, und R eine Äquivalenzrelation auf M. Man sagt, dass durch die Vorschrift $g([x]_R) := f(x)$ eine Funktion g von der Faktormenge M/R nach N wohldefiniert ist, wenn $\forall x, y \in M : (xRy \Rightarrow f(x) = f(y))$ gilt.

Beachten Sie, dass Sie beim Beweis dieser Implikation nicht die Notation $g(\ )$ verwenden dürfen, weil Sie ja noch nicht wissen, ob diese Definition tatsächlich eine Funktion g liefert.

(Von Seite 39):
Zwei Funktionen von M nach N sind genau dann gleich, wenn sie und wenn sie jedem Element ihres (gemeinsamen) Definitionsbereiches M das selbe Element ihrer (gemeinsamen) Zielmenge N zuordnen.

(Von Seite 47):
Wenn Sie zwei äquivalente Definitionen einer Eigenschaft D haben, dann passiert es oft, dass die eine Definition D1 scheinbar mehr als die andere D2 aussagt. Ein Beispiel haben wir schon am Beispiel der Eigenschaften Sym und Sym′ gesehen.

Wenn Sie in so einer Situation beweisen müssen, dass die Eigenschaft D gilt, dass empfiehlt es sich, die scheinbar schwächere Formulierung D2 zu überprüfen, weil das leichter ist. Wenn Sie aber schon wissen, dass D gilt, dürfen Sie D1 verwenden; eine Folgerung aus der scheinbar stärkeren Eigenschaft D1 ist vielleicht leichter zu gewinnen als aus D2.