PAUL R. HALMOS

NAIVE MENGENLEHRE

VANDENHOECK & RUPRECHT IN GÖTTINGEN

KAPITEL 9

Familien

Gelegentlich wird der Wertebereich einer Funktion für wichtiger gehalten als die Funktion selbst. In einem solchen Falle werden Terminologie und Notation stark verändert. Sei zum Beispiel x eine Funktion von einer Menge I in eine Menge X. (Schon die Wahl der Buchstaben zeigt an, daß etwas Neues kommt.) Wir wollen jetzt ein Element des Definitionsbereichs I einen *Index* und I selbst die *Indexmenge* nennen; der Wertebereich der Funktion x soll *indizierte Menge* und die Funktion selbst *Familie* heißen; der Wert der Funktion an einer Stelle i, *Term* der Familie genannt, wird (anstelle von $x(i)$) nun x_i geschrieben. (Diese Terminologie ist nicht überall üblich, aber doch ziemlich weit verbreitet neben verwandten, unbedeutend abweichenden Sprechweisen; im folgenden wollen wir nur diese eine verwenden.) In der Notation wird der veränderte Aspekte dadurch angezeigt, daß man

$$(x_i)_{i \in I}$$

für die Funktion x schreibt; man liest: die *Familie der x_i mit $i \in I$*. Geht die Indexmenge aus dem Zusammenhang hervor oder ist sie gerade nicht besonders wichtig, so unterdrückt man ihre Angabe und spricht dann kurz von der Familie (x_i); will man betonen, daß die Werte x_i in der Menge X liegen, dann spricht man von einer Familie (x_i) in X oder von einer Familie (x_i) von ..., wobei dann für die Punkte eine für die Elemente von X charakteristische Eigenschaft steht. Ist zum Beispiel $X = \mathfrak{P}(Y)$ für irgendeine Menge Y, so ist eine Familie (A_i) von Teilmengen von Y nichts weiter als eine Funktion A von einer gewissen Indexmenge I in die Potenzmenge $\mathfrak{P}(Y)$; in diesem Falle spricht man der Deutlichkeit halber von einer *Mengenfamilie.*

Sei nun (A_i) eine solche Mengenfamilie. Die Vereinigung des Wertebereichs der Familie, also der zugehörigen indizierten Menge, wird die *Vereinigung der Familie* (A_i), auch Vereinigung der Mengen A_i genannt; man schreibt sie

$$\bigcup_{i \in I} A_i \quad \text{oder} \quad \bigcup_i A_i,$$

je nachdem, ob es wichtig ist oder nicht, die Indexmenge anzugeben.

Aus der Definition der Vereinigung folgt sofort, daß

$$x \in \bigcup_i A_i \textit{ genau dann, wenn } x \in A_i \textit{ für mindestens ein } i.$$

Falls $I = 2$, ist $\{A_0, A_1\}$ der Wertebereich von (A_i), und man hat

$$\bigcup_i A_i = A_0 \cup A_1.$$

Die Vereinigung von Paarmengen läßt sich also als Spezialfall der Vereinigung von Mengenfamilien ansehen. Ebenso kann man die Vereinigung eines beliebigen Mengensystems als Vereinigung einer gewissen Mengenfamilie auffassen, denn jedes Mengensystem ist der Wertebereich einer gewissen Mengenfamilie: Ist $\mathfrak{C}$ ein Mengensystem, so lasse man $\mathfrak{C}$ selbst die Rolle der Indexmenge übernehmen und betrachte die identische Abbildung von $\mathfrak{C}$ als Familie.

Die algebraischen Gesetze der Vereinigungsbildung für Paarmengen können auf beliebige Vereinigungen verallgemeinert werden. Sei beispielsweise $(I_j)_{j \in J}$ eine Mengenfamilie und $K = \bigcup_j I_j$; ferner sei (A_k) eine Mengenfamilie mit dieser Menge K als Indexmenge. Dann ist es nicht schwer, das allgemeine Assoziativgesetz der Vereinigungsbildung zu beweisen, nämlich

$$\bigcup_{k \in K} A_k = \bigcup_{j \in J} \Big(\bigcup_{i \in I_j} A_i \Big).$$

Übungsaufgabe

Man formuliere und beweise ein entsprechend allgemeines Kommutativgesetz.

Die Vereinigung der *leeren* Mengenfamilie, das heißt der Familie mit leerem Indexbereich, ist die leere Menge; den Durchschnitt der leeren Mengenfamilie kann man gar nicht erst bilden, denn er müßte die berüchtigte Allmenge sein. Von dieser Trivialität abgesehen, läuft die Terminologie samt Notation für Durchschnitte von Mengenfamilien der für Vereinigungen völlig parallel. Sei etwa (A_i) eine nichtleere Mengenfamilie, das heißt eine Familie mit nichtleerer Indexmenge. Der Durchschnitt des Wertebereichs der Familie wird der *Durchschnitt der Familie* (A_i), auch Durchschnitt der Mengen A_i genannt; man schreibt ihn

$$\bigcap_{i \in I} A_i \quad \text{oder} \quad \bigcap_i A_i,$$

je nachdem, ob es wichtig ist, die Indexmenge anzugeben oder nicht.

Aus der Definition des Durchschnitts folgt sofort, daß

$$x \in \bigcap_{i \in I} A_i \textit{ genau dann, wenn } x \in A_i \textit{ für alle } i \in I$$

(vorausgesetzt, daß $I \neq \emptyset$).

Das allgemeine Kommutativgesetz und das allgemeine Assoziativgesetz für Durchschnitte können genauso formuliert und bewiesen werden wie die entsprechenden Gesetze für Vereinigungen. Man kann sie aber auch an Hand der De Morganschen Gesetze aus den entsprechenden Tatsachen für Vereinigungen (gemäß dem Dualitätsprinzip) gewinnen. Dies alles ist fast trivial und deshalb keiner langen Diskussion würdig. Von größerem Interesse sind dagegen diejenigen algebraischen Identitäten, in denen sowohl Vereinigungen als auch Durchschnitte vorkommen. Zum Beispiel gelten für eine Familie (A_i) von Teilmengen einer Menge X und eine beliebige Teilmenge B von X die Gleichungen

$$B \cap \bigcup_i A_i = \bigcup_i (B \cap A_i)$$

und

$$B \cup \bigcap_i A_i = \bigcap_i (B \cup A_i);$$

das sind offenbar geringfügige Verallgemeinerungen der Distributivgesetze für Vereinigungen und Durchschnitte von Paarmengen.

Übungsaufgabe

Sind (A_i) und (B_j) Mengenfamilien, so

$$(\bigcup_i A_i) \cap (\bigcup_j B_j) = \bigcup_{i,j} (A_i \cap B_j)$$

und

$$(\bigcap_i A_i) \cup (\bigcap_j B_j) = \bigcap_{i,j} (A_i \cup B_j);$$

dabei soll etwa $\bigcup_{i,j} (A_i \cap B_j)$ eine Akkürzung sein für $\bigcup_{(i,j) \in I \times J} C_{(i,j)}$ mit $C_{(i,j)} = A_i \cap B_j$.

Die Familienschreibweise wird normalerweise bei der Verallgemeinerung des kartesischen Produkts benutzt. Das kartesische Produkt zweier Mengen X und Y war definiert als die Menge aller geordneten Paare (x, y) mit x in X und y in Y. Zwischen dieser Menge $X \times Y$ und einer gewissen Menge von Familien gibt es eine natürliche eineindeutige Korrespondenz: Sei $\{a, b\}$ irgendeine Paarmenge mit $a \neq b$; wir betrachten die Menge Z aller Familien z mit der Indexmenge $\{a, b\}$, für die $z_a \in X$ und $z_b \in Y$. Die durch $f(z) = (z_a, z_b)$ definierte Funktion f von Z in $X \times Y$ ist offenbar die gesuchte eineindeutige Korrespondenz. Der Unterschied zwischen Z und $X \times Y$ mündet also grob gesagt in einer Frage der Bezeichnung. Die Verallgemeinerung des kartesischen Produkts ist genaugenommen eine von Z. (Deshalb stimmt die Terminologie des allgemeinen Falles nicht ganz mit der des Spezialfalles überein. Dagegen können wir nichts machen; der mathematische Sprachgebrauch ist heute eben so.) Die Verallgemeinerung ergibt sich nun von selbst: Das *kartesische Produkt der Mengenfamilie* $(X_i)_{i \in I}$ ist per definitionem die Menge aller Familien $(x_i)_{i \in I}$ mit $x_i \in X_i$ für jedes i in I. Zur Bezeichnung dieses Produkts sind verschiedene Symbole mehr oder weniger üblich, wir werden es in diesem Buch

$$\bigtimes_{i \in I} X_i \quad \text{oder} \quad \bigtimes_i X_i$$

schreiben. Ist jedes X_i gleich ein und derselben Menge X, so ist offenbar $\bigtimes_{i \in I} X_i = X^I$. Ist I eine Paarmenge $\{a, b\}$ mit $a \neq b$, so identifiziert man oft $\bigtimes_{i \in I} X_i$ mit dem kartesischen Produkt $X_a \times X_b$ nach der alten Definition; ist I eine Einermenge $\{a\}$, so identifiziert man oft $\bigtimes_{i \in I} X_i$ mit X_a selbst. *Geordnete Tripel, geordnete Quadrupel* usw. können als Familien definiert werden, deren Indexmengen genau drei bzw. vier usw. Elemente enthalten.

Sei $(X_i)_{i \in I}$ eine Mengenfamilie und X ihr kartesisches Produkt. Ist J eine Teilmenge von I, so entspricht jedem Element von X in natürlicher Weise ein Element des partiellen kartesischen Produkts $\bigtimes_{i \in J} X_i$. Zur Definition dieser Zuordnung sei daran erinnert, daß jedes Element x von X eine Familie (x_i), also eine auf I definierte Funktion ist; das dem x zugeordnete Element, wir nennen es y, erhält man durch Einschränkung der Funktion x auf J; ausführlich können wir $y_i = x_i$ für jedes i in J schreiben. Diese Zuordnung $x \mapsto y$ wird die *Projektion* von X auf $\bigtimes_{i \in J} X_i$ genannt; man bezeichnet sie gelegentlich mit f_J. Ist insbesondere J eine Einermenge $\{j\}$, so wird f_j statt $f_{\{j\}}$ geschrieben. Das Wort „Projektion" wird in vielerlei Bedeutung benutzt; für ein x aus X wird der Wert von f_j an der Stelle x, nämlich x_j, ebenfalls Projektion von x auf X_j genannt oder auch die *j-Koordinate* von x. Eine auf einem kartesischen Produkt definierte Funktion wird Funktion von *mehreren Veränderlichen* genannt; insbesondere heißt eine auf $X_a \times X_b$ definierte Funktion eine Funktion von *zwei Veränderlichen*.

Übungsaufgaben

1. Man beweise, daß für Mengenfamilien (A_i) und (B_j) stets

$$(\bigcup_i A_i) \times (\bigcup_j B_j) = \bigcup_{i,j} (A_i \times B_j)$$

und daß die entsprechende Gleichung für Durchschnitte gilt, falls die beiden Indexmengen nicht leer sind.

2. Man beweise (unter entsprechenden Vorbehalten über leere Familien), daß

$$\bigcap_{i \in I} A_i \subset A_{i_0} \subset \bigcup_{i \in I} A_i$$

für jeden Index i_0 in I und daß Durchschnitt und Vereinigung sogar als Extremallösungen dieser Inklusionen charakterisiert sind. Das heißt, wenn $A_{i_0} \subset X$ für jeden Index i_0 in I, so $\bigcup_{i \in I} A_i \subset X$, und $\bigcup_{i \in I} A_i$ ist die einzige Menge, die diese Minimalbedingung erfüllt. (Man sagt dafür, daß $\bigcup_{i \in I} A_i$ die kleinste Menge ist, die alle A_i als Teilmenge enthält.) Entsprechend ist $\bigcap_{i \in I} A_i$ die größte Menge, die in allen A_i als Teilmenge enthalten ist.

KAPITEL 15

Das Auswahlaxiom

Zur Gewinnung tieferer Erkenntnisse über geordnete Mengen brauchen wir ein neues mengentheoretisches Hilfsmittel; wir unterbrechen deshalb die Entwicklung der Ordnungstheorie, um uns dieses Hilfsmittel zu beschaffen.

Zunächst bemerken wir, daß eine Menge entweder leer ist oder nicht, und wenn sie nicht leer ist, so gibt es in ihr ein Element. Diese Feststellung kann verallgemeinert werden. Wenn von zwei gegebenen Mengen X und Y eine leer ist, so ist auch das kartesische Produkt $X \times Y$ leer. Wenn weder X noch Y leer ist, so gibt es ein Element x in X und ein Element y in Y und damit das geordnete Paar (x, y) im kartesischen Produkt $X \times Y$, also ist $X \times Y$ nicht leer. Diese einfachen Tatsachen sind Spezialfälle (für $n = 1$ und $n = 2$) der folgenden Aussage: Ist $(X_i)_{i \in n}$ eine endliche Folge von Mengen, so ist ihr kartesisches Produkt genau dann leer, wenn mindestens eine von ihnen leer ist. Durch Induktion nach n überzeugt man sich leicht von der Gültigkeit dieser Behauptung. (Der Fall $n = 0$ führt auf ein delikates Argument über die leere Abbildung; der nicht neugierige Leser möge mit der Induktion bei $n = 1$ anfangen.)

Für unendliche Mengenfamilien läßt sich der nichttriviale Teil obiger Behauptung, also, daß das kartesische Produkt einer unendlichen Familie von Mengen nur dann leer ist, wenn mindestens eine von ihnen leer ist, nicht mit den bisher eingeführten Prinzipien beweisen. Wir fügen diesen Teil der Behauptung als weitere Grundannahme unserer Theorie hinzu.

Auswahlaxiom

Das kartesische Produkt einer (nichtleeren) Familie von nichtleeren Mengen ist nichtleer.

Mit anderen Worten: Ist $(X_i)_{i \in I}$ eine Familie nichtleerer Mengen (mit nichtleerer Indexmenge I), so gibt es eine Familie $(x_i)_{i \in I}$ mit $x_i \in X_i$ für alle i in I.

Sei $\mathfrak{C}$ ein (nichtleeres) System von nichtleeren Mengen. Wir können $\mathfrak{C}$ als Familie betrachten, genauer: wir können $\mathfrak{C}$ in eine indizierte Menge verwandeln, indem wir die Gesamtheit $\mathfrak{C}$ selbst als Indexmenge und die identische Abbildung auf $\mathfrak{C}$ als indizierende Abbildung verwenden. Dann besagt das Auswahlaxiom, daß das kartesische Produkt der Mengen aus $\mathfrak{C}$ mindestens ein Element enthält. Ein Element eines solchen kartesischen Produkts ist nach Definition eine Funktion (Familie, indizierte Menge), deren Definitionsbereich die Indexmenge ist (in diesem Falle $\mathfrak{C}$) und deren Wert für einen Index in der Menge mit ebendiesem Index liegt. Daraus schließen wir: Es gibt eine Funktion f mit dem Definitionsbereich $\mathfrak{C}$ derart, daß $f(A) \in A$ für alle A aus $\mathfrak{C}$. Diese Folgerung läßt sich insbesondere auf den Fall anwenden, daß $\mathfrak{C}$ das System aller nichtleeren Teilmengen einer (nichtleeren) Menge X ist; sie lautet dann: Es gibt eine Funktion f mit dem Definitionsbereich $\mathfrak{P}(X) - \{\emptyset\}$ derart, daß $f(A) \in A$ für alle A aus diesem Definitionsbereich. Anschaulich gesprochen kann die Funktion f beschrieben werden als simultane Auswahl je eines Elements aus jeder einzelnen von vielen Mengen; daher der Name „Auswahlaxiom". Eine Funktion, die in diesem Sinne aus jeder nichtleeren Teilmenge einer Menge X ein Element „auswählt", heißt *Auswahlfunktion* für X. Wir haben gesehen, daß wir für endliche Systeme von nichtleeren Mengen die Möglichkeit einer solchen simultanen Auswahl, also die Existenz einer Auswahlfunktion, schon vor der Formulierung des Auswahlaxioms aus den bis dahin eingeführten mengentheoretischen Prinzipien folgern konnten; das Auswahlaxiom soll diese Möglichkeit für den unendlichen Fall garantieren.

Die beiden obigen Folgerungen aus dem Auswahlaxiom (eine für die Potenzmenge einer Menge, die andere für allgemeinere Mengensysteme) sind in Wirklichkeit nur Umformulierungen dieses Axioms. Man hielt es gewöhnlich für wichtig, jede Folgerung aus dem Auswahlaxiom daraufhin zu untersuchen, in welchem Ausmaß das Auswahlaxiom im Beweis dieser Folgerung benötigt wird. Die Entdeckung eines anderen Beweises derselben Behauptung ohne Benutzung des Auswahlaxioms weckte Siegesstimmung; eine Umkehrung des Beweises mit dem Ergebnis, daß die fragliche Behauptung äquivalent zum Auswahlaxiom ist (unter Benutzung der übrigen Axiome der Mengenlehre), galt als ehrenvolle Niederlage. Alles was dazwischenlag, wurde als ärgerlich empfunden. Als Beispiele (und Übungsaufgaben) für die vielen als mit dem Auswahlaxiom äquivalent bekannten Aussagen erwähnen wir, daß jede Relation eine Funktion mit demselben Definitionsbereich enthält

sowie daß es zu jedem System $\mathfrak{C}$ von paarweise disjunkten nichtleeren Mengen eine Menge A gibt derart, daß $A \cap C$ eine Einermenge ist für jedes C in $\mathfrak{C}$.

Zur Illustration der Anwendung des Auswahlaxioms betrachten wir folgende Behauptung: Jede unendliche Menge enthält eine mit ω gleichmächtige Teilmenge. Heuristisch könnten wir so argumentieren: Wenn eine Menge X unendlich ist, so ist sie insbesondere nicht gleichmächtig mit 0, d. h. sie ist nicht leer und enthält deshalb ein Element x_0. Weil X nicht gleichmächtig mit 1 ist, kann die Menge $X - \{x_0\}$ nicht leer sein und enthält deshalb ein Element x_1. Diese Argumentation wiederhole man ad infinitum; der nächste Schritt ist zum Beispiel: Weil X nicht gleichmächtig mit 2 ist, kann die Menge $X - \{x_0, x_1\}$ nicht leer sein und enthält deshalb ein Element x_2. Wir erhalten schließlich eine unendliche Folge $(x_n)_{n \in \omega}$ von lauter verschiedenen Elementen von X, q.e.d. Diese heuristische Überlegung offenbart immerhin deutlich die wichtigste Idee einer korrekteren Argumentation, nämlich daß sich der Vorgang, aus einer nichtleeren Menge ein Element auszuwählen, unendlich oft wiederholt. Der mit der Handhabung des Auswahlaxioms vertraute Mathematiker wird oft in ähnlicher, nicht streng formaler Weise argumentieren; seine Erfahrung befähigt ihn dazu, mit einem Blick zu sehen, wie man den Beweis formal in Ordnung bringen könnte. Für unsere Zwecke ist es ratsam, die Vervollständigung der obigen Beweisskizze ausführlich zu behandeln.

Sei f eine Auswahlfunktion für X; das bedeutet, daß f eine Funktion von der Gesamtheit aller nichtleeren Teilmengen von X in X ist derart, daß $f(A) \in A$ für alle A im Definitionsbereich von f. Sei $\mathfrak{C}$ das System aller endlichen Teilmengen von X. Weil X unendlich ist, kann für ein A in $\mathfrak{C}$ die Menge $X - A$ nicht leer sein und muß deshalb zum Definitionsbereich von f gehören. Wir definieren eine Funktion g von $\mathfrak{C}$ in $\mathfrak{C}$ durch $g(A) = A \cup \{f(X - A)\}$; in Worten: Man erhält $g(A)$, indem man zu A das durch die Funktion f aus $X - A$ ausgewählte Element hinzufügt. Wir wenden das Rekursionstheorem auf die Funktion g an; dabei können wir zum Beispiel mit der Menge $\emptyset$ als Anfangswert beginnen. Das Ergebnis ist eine Funktion U von ω in $\mathfrak{C}$ derart, daß $U(0) = \emptyset$ und $U(n^+) = g(U(n)) = U(n) \cup \{f(X - U(n))\}$ für jede natürliche Zahl n. Wir behaupten, daß die Funktion v mit $v(n) = f(X - U(n))$ eine eineindeutige Abbildung von ω in X ist und daher ω in der Tat mit einer Teilmenge von X (nämlich mit dem Wertevorrat von v)

gleichmächtig ist. Zum Beweis machen wir eine Reihe von elementaren Feststellungen, die man leicht an den Definitionen abliest. Erstens: $v(n) \notin U(n)$ für alle n in ω. Zweitens: $v(n) \in U(n^+)$ für alle n in ω. Drittens: Für natürliche Zahlen n und m mit $n \leqq m$ ist $U(n) \subset U(m)$. Viertens: Für natürliche Zahlen n und m mit $n < m$ ist $v(n) \neq v(m)$ (denn $v(n) \in U(n^+) \subset U(m)$, also $v(n) \in U(m)$, aber $v(m) \notin U(m)$). Mithin werden zwei verschiedene natürliche Zahlen stets auf zwei verschiedene Elemente von X abgebildet, denn von zwei verschiedenen natürlichen Zahlen ist immer eine kleiner als die andere. Damit haben wir die heuristische Betrachtung vollständig präzisiert.

Wir wissen jetzt, daß jede unendliche Menge eine mit ω gleichmächtige Teilmenge enthält. Dieses Ergebnis — hier nicht so sehr um seiner selbst willen denn als Beispiel für die richtige Anwendung des Auswahlaxioms bewiesen — hat ein interessantes Korollar. Wir behaupten, daß eine Menge genau dann unendlich ist, wenn sie mit einer echten Teilmenge von sich selbst gleichmächtig ist. Daß eine Menge nicht endlich sein kann, wenn sie mit einer echten Teilmenge gleichmächtig ist, wissen wir bereits. Sei umgekehrt X eine unendliche Menge und v eine eineindeutige Abbildung von ω in X. Liegt x im Wertevorrat von v, etwa $x = v(n)$, so setzen wir $h(x) = v(n^+)$; für diejenigen x in X, die nicht im Wertevorrat von v liegen, schreiben wir $h(x) = x$. Man prüft leicht nach, daß h eine eineindeutige Abbildung von X in sich ist. Da der Wertevorrat von h eine echte Teilmenge von X ist (er enthält nämlich nicht $v(0)$), ist damit das Korollar bewiesen. Die Aussage dieses Korollars wurde von Dedekind gerade als Definition der Unendlichkeit benutzt.

KAPITEL 16

Das Zornsche Lemma

Ein Existenztheorem behauptet die Existenz eines Objekts, das zu einer gewissen Menge gehört und gewisse Eigenschaften besitzt. Viele Existenztheoreme können so formuliert (oder, wenn es sein muß, umformuliert) werden, daß die zugrunde liegende Menge eine geordnete Menge und die entscheidende Eigenschaft Maximalität ist. Unser nächstes Ziel ist der Beweis des wichtigsten Theorems dieser Art:

Zornsches Lemma

Sei X eine geordnete Menge derart, daß jede Kette in X eine obere Schranke besitzt. Dann enthält X ein maximales Element.

Diskussion: Wir erinnern uns, daß eine Kette eine total geordnete Menge ist. Eine Kette in X ist eine Teilmenge von X, die total geordnet ist durch die von X übernommene, d. h. auf die Teilmenge eingeschränkte Ordnung. Ist A eine Kette in X, so verlangt die Voraussetzung des Zornschen Lemmas die Existenz einer oberen Schranke für A in X; die Existenz einer oberen Schranke für A in A ist nicht erforderlich. Die Konklusion des Zornschen Lemmas besteht in der Existenz eines Elements a in X derart, daß aus $a \leqq x$ stets $a = x$ folgt.

Der Grundgedanke des Beweises ist dem im vorigen Kapitel bei der Diskussion unendlicher Mengen benutzten ganz ähnlich. Da die Menge X nach Voraussetzung nicht leer ist (zur Kette $\emptyset$ in X muß es ja eine obere Schranke geben), enthält sie ein Element x_0. Ist x_0 maximal, sind wir fertig; wenn nicht, dann gibt es ein Element x_1 mit $x_0 < x_1$. Ist x_1 maximal, sind wir fertig; andernfalls machen wir weiter. Diese Argumentation wiederholen wir ad infinitum; müssen wir damit schließlich zu einem maximalen Element kommen?

Diese Frage birgt einige Schwierigkeiten. Es könnte zum Beispiel passieren, daß der unendlich oft angewandte obige Schluß zu einer ganzen unendlichen Folge von nichtmaximalen Elementen führt; was sollen wir dann tun? Die Antwort ist zunächst: Der Wertevorrat einer solchen unendlichen Folge ist eine Kette in X und hat deshalb nach Voraussetzung eine obere Schranke; mit dieser oberen Schranke müssen wir die ganze Argumentation wieder von vorn anfangen. Nun ist es aber, gelinde gesagt, ziemlich undurchsichtig, wann und wie das alles enden soll. Auf diesem Wege kommen wir also nicht weiter; wir müssen uns einen präzisen Beweis beschaffen. Wir geben hier eine Modifikation des ursprünglich von ZERMELO mitgeteilten Beweises.

Beweis: Im ersten Schritt ersetzen wir die abstrakte Ordnung durch die Inklusionsordnung in einer passenden Gesamtheit von Mengen: Die Menge $\mathfrak{X}$ aller Ketten in X ist ein durch die Inklusion geordnetes Mengensystem; ein maximales Element M von $\mathfrak{X}$, d. h. eine bezüglich der Inklusion maximale Kette in X, liefert uns ein maximales Element von X. Denn M hat nach Voraussetzung eine obere Schranke x in X; wäre x nicht maximal in X, könnte M nicht maximal in $\mathfrak{X}$ sein. Die Betrachtung des Systems $\mathfrak{X}$ anstelle der Menge X hat den Vorteil, daß die existentielle Form der Voraussetzung des ZORNschen Lemmas in $\mathfrak{X}$ eine explizite Gestalt annimmt, denn wir können zu jeder Kette $\mathfrak{C}$ in $\mathfrak{X}$ eine obere Schranke in $\mathfrak{X}$ direkt angeben: die Vereinigung von $\mathfrak{C}$ (d. h. $\bigcup_{A \in \mathfrak{C}} A$). (Beweis: Trivialerweise ist $K = \bigcup_{A \in \mathfrak{C}} A$ Teilmenge von X und obere Schranke von $\mathfrak{C}$, und zu x_1, x_2 in K hat man A_1, A_2 in $\mathfrak{C}$ mit $x_1 \in A_1$, $x_1 \in A_2$; weil $\mathfrak{C}$ durch die Inklusion total geordnet ist, sind A_1 und A_2 miteinander vergleichbar, also etwa $A_1 \subset A_2$ und damit $x_1 \in A_2$; als Element von $\mathfrak{X}$ ist A_2 total geordnet, weshalb schließlich x_1 und x_2 miteinander vergleichbar sein müssen.)

Jetzt können wir die gegebene Ordnung in X überhaupt vergessen. Wir werden beweisen, daß jedes System $\mathfrak{X}$ von Teilmengen einer Menge X mit der Eigenschaft, daß die Vereinigung jeder Kette (bezüglich der Inklusion) in $\mathfrak{X}$ wieder zu $\mathfrak{X}$ gehört, ein maximales Element besitzt.

Sei f eine Auswahlfunktion für X, d. h. eine Funktion von dem System aller nichtleeren Teilmengen von X in X derart, daß $f(A) \in A$ für alle A im Definitionsbereich von f. Für ein Element A von $\mathfrak{X}$ sei $\hat{A}$ die Menge aller Elemente von X, deren Hinzunahme zu A wieder ein Element von $\mathfrak{X}$ liefert, d. h. $\hat{A} = \{x \in X : \{A \cup \{x\}\} \in \mathfrak{X}\}$. Wir definieren eine Funktion g von $\mathfrak{X}$ in $\mathfrak{X}$: Wenn $\hat{A} - A \neq \emptyset$, setzen wir $g(A) = A \cup \{f(\hat{A} - A)\}$; wenn $\hat{A} - A = \emptyset$, schreiben wir $g(A) = A$. Aus der Definition von $\hat{A}$ folgt sofort, daß $\hat{A} - A = \emptyset$

genau dann eintritt, wenn A maximal in $\mathfrak{X}$ ist. Wir müssen also beweisen, daß es in $\mathfrak{X}$ ein Element A gibt mit $g(A) = A$. Es wird sich zeigen, daß die entscheidende Eigenschaft der Funktion g darin besteht, daß die Menge $g(A)$ (die immer A umfaßt) höchstens ein Element mehr enthält als A.

Um die Darstellung zu vereinfachen, führen wir einige Definitionen ad hoc ein. Wir nennen ein Teilsystem $\mathfrak{T}$ von $\mathfrak{X}$ einen *Turm*, wenn

(i) $\emptyset \in \mathfrak{T}$,

(ii) *wenn* $A \in \mathfrak{T}$, *so* $g(A) \in \mathfrak{T}$,

(iii) *wenn* $\mathfrak{C}$ *eine Kette in* $\mathfrak{T}$ *ist, so* $\bigcup_{A \in \mathfrak{C}} A \in \mathfrak{T}$.

Es gibt Türme, denn $\mathfrak{X}$ selbst ist einer ($\emptyset$ gehört als Vereinigung der leeren Kette $\emptyset$ nach Voraussetzung zu $\mathfrak{X}$). Da der Durchschnitt eines nichtleeren Systems von Türmen offenbar wieder ein Turm ist, folgt insbesondere, daß der Durchschnitt $\mathfrak{T}_0$ aller Türme der kleinste Turm ist. Wir zeigen zunächst, daß $\mathfrak{T}_0$ eine Kette ist.

Wir wollen ein Element C von $\mathfrak{T}_0$ *komparabel* nennen, wenn es mit jedem Element von $\mathfrak{T}_0$ vergleichbar ist, das bedeutet, daß für A aus $\mathfrak{T}_0$ stets $A \subset C$ oder $C \subset A$ gilt. Die Aussage, daß $\mathfrak{T}_0$ eine Kette ist, kann dann auch so formuliert werden: Jedes A in $\mathfrak{T}_0$ ist komparabel. Es gibt komparable Mengen in $\mathfrak{T}_0$, denn $\emptyset$ ist komparabel. In den beiden folgenden Abschnitten konzentrieren wir unsere Aufmerksamkeit auf eine beliebige, aber für den Augenblick feste komparable Menge C aus $\mathfrak{T}_0$.

Sei $A \in \mathfrak{T}_0$ und C echte Obermenge von A. Dann ist $g(A) \subset C$; denn da C komparabel ist, hat man entweder $g(A) \subset C$, oder C ist echte Teilmenge von $g(A)$; im letzten Fall wäre aber A echte Teilmenge einer echten Teilmenge von $g(A)$, was unmöglich ist, weil $g(A) - A$ höchstens ein Element enthält.

Wir betrachten jetzt die Gesamtheit $\mathfrak{U}$ aller Elemente A von $\mathfrak{T}_0$, für die mindestens eine der beiden Aussagen $A \subset C$ oder $g(C) \subset A$ gilt. Wir zeigen, daß $\mathfrak{U}$ ein Turm ist. Wegen $\emptyset \subset C$ ist die erste Bedingung für Türme erfüllt. Den Beweis der zweiten Bedingung, (d. h. wenn $A \in \mathfrak{U}$, so $g(A) \in \mathfrak{U}$) unterteilen wir in drei Fälle.

Erstens: A ist echte Teilmenge von C. Dann hat man nach dem vorigen Abschnitt $g(A) \subset C$ und damit $g(A) \in \mathfrak{U}$. Zweitens: $A = C$. Dann ist $g(A) = g(C)$, insbesondere also $g(C) \subset g(A)$ und deshalb $g(A) \in \mathfrak{U}$. Drittens: $g(C) \subset A$. Dann ist $g(C) \subset g(A)$ wegen $A \subset g(A)$

und damit $g(A) \in \mathfrak{U}$. Die dritte Bedingung folgt unmittelbar aus der Definition von $\mathfrak{U}$. Also ist $\mathfrak{U}$ ein in $\mathfrak{T}_0$ enthaltener Turm und somit gleich $\mathfrak{T}_0$, denn $\mathfrak{T}_0$ ist der kleinste Turm.

Aus diesen Überlegungen können wir schließen, daß für jede komparable Menge C auch die Menge $g(C)$ komparabel ist. Denn zu C bilde man $\mathfrak{U}$ wie oben; aus $\mathfrak{U} = \mathfrak{T}_0$ folgt für A in $\mathfrak{T}_0$ stets $A \subset C$ (folglich $A \subset g(C)$) oder $g(C) \subset A$.

Wir wissen jetzt, daß $\emptyset$ komparabel ist und daß g komparable Mengen auf komparable abbildet. Da zudem die Vereinigung einer Kette von komparablen Mengen komparabel ist, erkennen wir, daß die komparablen Mengen (in $\mathfrak{T}_0$) einen Turm bilden und somit ganz $\mathfrak{T}_0$ ausschöpfen, denn $\mathfrak{T}_0$ ist der kleinste Turm. Damit ist gezeigt, daß alle Elemente von $\mathfrak{T}_0$ miteinander vergleichbar sind.

Da nun $\mathfrak{T}_0$ eine Kette ist, gehört die Vereinigung A aller Mengen in $\mathfrak{T}_0$ selbst wieder zu $\mathfrak{T}_0$. Weil diese Vereinigung alle Mengen in $\mathfrak{T}_0$ umfaßt, ist insbesondere $g(A) \subset A$; zusammen mit $A \subset g(A)$ folgt schließlich $A = g(A)$, womit das ZORNsche Lemma bewiesen ist.

Übungsaufgaben

Man beweise:

1. Das ZORNsche Lemma ist äquivalent zum Auswahlaxiom. (Hinweis: Zu gegebener Menge X betrachte man Funktionen f derart, daß $\operatorname{dom} f \subset \mathfrak{P}(X)$, $\operatorname{ran} f \subset X$, $f(A) \in A$ für alle A in $\operatorname{dom} f$; man ordne die Menge aller dieser Funktionen durch die Relation der Fortsetzung, finde mit Hilfe des ZORNschen Lemmas eine maximale unter ihnen und beweise, daß für maximales f stets $\operatorname{dom} f = \mathfrak{P}(X) - \{\emptyset\}$.)

2. Jede der folgenden Aussagen ist ebenfalls äquivalent zum Auswahlaxiom:

(i) Jede geordnete Menge enthält eine maximale Kette (d. h. eine Kette, die nicht echte Teilmenge irgendeiner anderen Kette in derselben Menge ist).

(ii) Jede Kette in einer geordneten Menge ist in einer gewissen maximalen Kette (in derselben Menge) enthalten.

(iii) Jede geordnete Menge, in der jede Kette eine obere Grenze besitzt, enthält ein maximales Element.

KAPITEL 17

Wohlordnung

Eine geordnete Menge braucht kein kleinstes Element zu haben, und selbst wenn sie eines hat, ist es durchaus möglich, daß irgendeine Teilmenge keines hat. Man nennt eine geordnete Menge *wohlgeordnet* (und ihre Ordnung *Wohlordnung*), wenn jede ihrer nichtleeren Teilmengen ein kleinstes Element besitzt. Eine unmittelbare Konsequenz dieser Definition, die wir notieren wollen, noch ehe wir uns nach Beispielen und Gegenbeispielen umsehen, ist die Tatsache, daß jede wohlgeordnete Menge total geordnet ist. Sind nämlich x und y Elemente einer wohlgeordneten Menge, so ist $\{x, y\}$ eine nichtleere Teilmenge und besitzt deshalb ein kleinstes Element; je nachdem, ob dieses erste Element x oder y ist, haben wir $x \leqq y$ oder $y \leqq x$.

Für jede natürliche Zahl n ist die Menge aller Vorgänger von n (also nach unserer Definition die Menge n) bei der Ordnung nach der Größe eine wohlgeordnete Menge; dasselbe gilt für die Menge ω aller natürlichen Zahlen. Die Menge $\omega \times \omega$ mit $(a, b) \leqq (x, y)$ in der Bedeutung von $(2a + 1)\, 2^y \leqq (2x + 1)\, 2^b$ ist nicht wohlgeordnet, denn sie selbst hat kein kleinstes Element, weil für beliebige x, y aus ω stets $(x, y + 1) < (x, y)$. Dennoch haben gewisse Teilmengen von $\omega \times \omega$ kleinste Elemente; man betrachte zum Beispiel die Menge E aller Paare (x, y) in $\omega \times \omega$ mit $(0, 0) \leqq (x, y)$; diese Menge hat $(0, 0)$ als kleinstes Element. Man hüte sich aber vor der Annahme, daß E als geordnete Menge (mit der von $\omega \times \omega$ übernommenen Ordnung) wohlgeordnet sei; denn obwohl E ein kleinstes Element besitzt, gibt es doch Teilmengen von E ohne erstes Element, zum Beispiel die Menge aller von $(0, 0)$ verschiedenen Elemente von E (Beweis als Übungsaufgabe). Ein weiteres Beispiel einer wohlgeordneten Menge ist $\omega \times \omega$ mit der lexikographischen Ordnung.

Eine der angenehmsten Eigenschaften wohlgeordneter Mengen ist die Tatsache, daß man Behauptungen über ihre Elemente durch ein zur vollständigen Induktion ähnliches Verfahren, nämlich das **Prinzip der transfiniten Induktion,** beweisen kann: Hat man in einer wohlgeordneten Menge X eine Teilmenge S mit der Eigenschaft, daß jedes Element x aus X, für welches der Hauptanfang $s(x)$ Teilmenge von S ist, selbst in S liegt, so wird $S = X$ behauptet. Anders ausgedrückt: Wenn die Zugehörigkeit aller strengen Vorgänger eines Elementes x zur Teilmenge S stets die Zugehörigkeit von x selbst zu S nach sich zieht, so muß S alle Elemente von X enthalten.

Bevor wir mit dem Beweis beginnen, sind noch einige Bemerkungen angebracht. Die Formulierung des gewöhnlichen Prinzips der vollständigen Induktion unterscheidet sich von der des Prinzips der transfiniten Induktion in zwei auffallenden Punkten. Erstens darf man bei diesem von der Menge aller Vorgänger des Elementes x auf x schließen, während bei jenem allein vom unmittelbaren Vorgänger von x auf x geschlossen werden muß; zweitens wird bei der transfiniten Induktion keine Annahme über ein Anfangselement (z.B. 0) gemacht.

Der erste Unterschied ist wesentlich: Ein Element einer beliebigen wohlgeordneten Menge braucht gar keinen unmittelbaren Vorgänger zu haben. In der Menge ω, wo jedes Element einen unmittelbaren Vorgänger hat, sind die beiden Induktionsprinzipien zwar äquivalent, wie man sich leicht klarmacht; bei beliebiger wohlgeordneter Menge sind sie aber wesentlich verschieden. Anders ausgedrückt: Die beiden Formulierungen sind im allgemeinen nicht miteinander äquivalent; ihre Gleichwertigkeit in ω ist ein glücklicher, aber sehr spezieller Umstand. Dazu ein Beispiel: Sei $X = \omega^+$ $(= \omega \cup \{\omega\})$. Die Elemente von ω seien der Größe nach geordnet, und es sei $n < \omega$ für alle n in ω. Damit ist X wohlgeordnet, und es gibt eine echte Teilmenge S von X mit $0 \in S$ und $n + 1 \in S$, falls $n \in S$, nämlich $S = \omega$.

Der zweite Unterschied zwischen gewöhnlicher und transfiniter Induktion (daß man für diese kein Anfangselement braucht) ist eher sprachlich als begrifflich. Wenn x_0 das kleinste Element von X ist, so ist $s(x_0)$ leer, daher trivialerweise $s(x_0) \subset S$; also wird auch in der Voraussetzung des transfiniten Induktionsprinzips $x_0 \in S$ gefordert.

Der Beweis des Prinzips der transfiniten Induktion ist beinahe trivial. Wäre $X - S$ nicht leer, so gäbe es darin ein kleinstes Element, etwa x. Dann würde aber jedes Element des Hauptanfangs $s(x)$ zu S gehören, was nach Voraussetzung $x \in S$ nach sich zöge. Das wäre ein Widerspruch zu $x \in X - S$, also muß $X - S$ leer und damit $S = X$ sein.

Wir wollen eine wohlgeordnete Menge B eine (echte) *Fortsetzung* einer wohlgeordneten Menge A nennen, wenn A eine Teilmenge von B und dabei ein Hauptanfang von B ist und zudem die Ordnung der Elemente von A in B dieselbe ist wie in A. Ist zum Beispiel X eine wohlgeordnete Menge und sind a und b Elemente von X mit $a < b$, so ist $s(b)$ eine Fortsetzung von $s(a)$ und natürlich X eine Fortsetzung sowohl von $s(a)$ als auch von $s(b)$.

Ist $\mathfrak{C}$ ein beliebiges System von Hauptanfängen einer wohlgeordneten Menge, so ist $\mathfrak{C}$ eine Kette bezüglich der Relation der Fortsetzung, genauer: Die Relation der Fortsetzung ist eine (strenge) Ordnung in $\mathfrak{C}$, und von je zwei verschiedenen Elementen von $\mathfrak{C}$ ist stets eins eine Fortsetzung des andern. Es gilt auch eine gewisse, oft nützliche Umkehrung dieser Bemerkung: Wenn ein System $\mathfrak{C}$ von wohlgeordneten Mengen eine Kette im obigen Sinne bezüglich der Relation der Fortsetzung ist, so gibt es genau eine Wohlordnung der Vereinigung U des Systems $\mathfrak{C}$ derart, daß U Fortsetzung jeder (von U selbst verschiedenen) Menge im System $\mathfrak{C}$ ist. Grob gesprochen ist demnach die Vereinigung einer Kette wohlgeordneter Mengen wieder wohlgeordnet. (Diese abgekürzte Formulierung ist mißverständlich, weil darin nicht ausgedrückt wird, daß „Kette“ in bezug auf die Relation der Fortsetzung gemeint ist; die Behauptung wird in der Tat falsch, wenn sich „Kette“ nur auf ordnungserhaltende Inklusion bezieht.)

Der Beweis ist einfach: Sind a und b in U, so gibt es Mengen A und B in $\mathfrak{C}$ mit $a \in A$, $b \in B$. Da entweder $A = B$ oder A Fortsetzung von B oder B Fortsetzung von A ist, gibt es jedenfalls in $\mathfrak{C}$ eine Menge, zu der a und b gleichzeitig gehören; aus demselben Grund sind die Elemente a, b in allen Mengen aus $\mathfrak{C}$, in denen sie beide gleichzeitig vorkommen, in derselben Weise angeordnet. Daher können wir jedes Paar $\{a, b\}$ in U per definitionem so ordnen, wie es in allen Mengen aus $\mathfrak{C}$ geordnet ist, zu denen a und b gleichzeitig gehören. (Wenn man beachtet, daß die Ordnungen in den Mengen aus $\mathfrak{C}$ Mengen von geordneten Paaren sind, so kann man die gewünschte Ordnung in U auch einfach als Vereinigung aller Ordnungen der Mengen in $\mathfrak{C}$ definieren.)

Man prüft direkt nach, daß die so definierte Relation eine Ordnung in U ist und daß die Konstruktion bei jedem Schritt zwangsläufig war (d. h. die in U konstruierte Ordnung ist durch die gegebenen Ordnungen eindeutig bestimmt). Bleibt zu beweisen, daß es sich um eine Wohlordnung handelt. Sei V eine nichtleere Teilmenge von U und etwa $a \in V$; dann gibt es eine Menge A in $\mathfrak{C}$ mit $a \in A$. Der Durchschnitt $V \cap A$ hat ein kleinstes Element a_0 in A, welches gleichzeitig das kleinste Element von V ist. Denn ist $b \in V - A$ und etwa $b \in B \in \mathfrak{C}$, so muß B Fortsetzung von A sein, daher A Hauptanfang von B und somit $a_0 < b$.

Übungsaufgabe

Eine Teilmenge A einer geordneten Menge X heißt *kofinal* in X, wenn es zu jedem Element x aus X ein Element a in A gibt mit $x \leqq a$. Man beweise, daß jede total geordnete Menge eine kofinale wohlgeordnete Teilmenge enthält.

Die Wichtigkeit des Begriffs der Wohlordnung rührt von dem folgenden Satz her, aus dem wir unter anderem schließen können, daß das Prinzip der transfiniten Induktion sehr viel mehr Anwendungsmöglichkeiten hat, als man auf den ersten Blick glauben könnte.

Wohlordnungssatz

Jede Menge kann wohlgeordnet werden.

Diskussion: Eine bessere (aber weniger übliche) Formulierung ist folgende: Zu jeder Menge X gibt es eine Wohlordnung mit dem Definitionsbereich X. Es sei vor der Annahme gewarnt, die versprochene Wohlordnung für X habe irgend etwas mit irgendeiner auf X schon gegebenen Struktur zu tun. Wenn der Leser zum Beispiel irgendeine geordnete oder total geordnete Menge kennt, deren Ordnung ganz bestimmt keine Wohlordnung ist, so darf er nicht den Schluß ziehen, er habe eine Paradoxie entdeckt. Die einzige erlaubte Folgerung daraus ist die, daß es auf einer Menge viele Ordnungen gibt, von denen einige möglicherweise Wohlordnungen sind, andere nicht; das wußten wir aber schon vorher.

Beweis: Wir wenden das Zornsche Lemma an. Für die gegebene Menge X betrachten wir das System $\mathfrak{W}$ aller wohlgeordneten Teilmengen von X; ausführlich: Ein Element von $\mathfrak{W}$ ist eine Teilmenge A von X zusammen mit irgendeiner Wohlordnung auf A. Wir ordnen $\mathfrak{W}$ durch die Relation der Fortsetzung.

Das System $\mathfrak{W}$ ist nicht leer, weil zum Beispiel $\emptyset \in \mathfrak{W}$. Wenn $X \neq \emptyset$, kann man auch bessere Elemente von $\mathfrak{W}$ aufweisen, zum Beispiel die Menge $\{x\}$ mit der Ordnung $\{(x, x)\}$ für irgendein Element x von X. Für eine Kette $\mathfrak{C}$ in $\mathfrak{W}$ hat die Vereinigung U aller Mengen in $\mathfrak{C}$ eine eindeutig bestimmte Wohlordnung, die U „größer" oder gleich (bei der Relation der Fortsetzung) als jede Menge in $\mathfrak{C}$ macht; das war gerade das Ergebnis unserer obigen Diskussion der Fortsetzungsrelation. Also ist die Voraussetzung des Zornschen Lemmas erfüllt, und wir erhalten eine maximale wohlgeordnete Menge M in $\mathfrak{W}$. Diese Menge M muß gleich der ganzen Menge X sein; denn wäre $x \in X - M$, so könnte man M durch Hinzunahme von x als letztes Element vergrößern. Die strenge Formulierung dieser unzweideutigen, aber nicht streng formalen Beschreibung sei dem Leser als Übungsaufgabe überlassen. Damit ist dann der Beweis des Wohlordnungssatzes vollständig.

Übungsaufgaben

Man beweise:

1. Eine total geordnete Menge ist wohlgeordnet genau dann, wenn die Menge aller strengen Vorgänger eines jeden Elementes wohlgeordnet ist. Kann man eine ähnliche Bedingung für teilweise geordnete Mengen angeben?
2. Der Wohlordnungssatz impliziert das Auswahlaxiom (und ist daher zu diesem Axiom und zum Zornschen Lemma äquivalent).
3. Ist R eine Ordnung auf einer Menge X, so gibt es eine Totalordnung S in X mit $R \subset S$; mit anderen Worten: Jede Ordnung kann ohne Vergrößerung des Definitionsbereichs zu einer totalen Ordnung erweitert werden.