ÜBER EINE BISHER NOCH NICHT BENÜTZTE ERWEITERUNG DES FINITEN STANDPUNKTES

von Kurt GÖDEL, Princeton

P. Bernays hat wiederholt darauf hingewiesen [1], dass angesichts der Tatsache der Unbeweisbarkeit der Widerspruchsfreiheit eines Systems mit geringeren Beweismitteln als denen des Systems selbst eine Überschreitung des Rahmens der im Hilbertschen Sinn finiten Mathematik nötig ist, um die Widerspruchsfreiheit der klassischen Mathematik, ja sogar um die der klassischen Zahlentheorie zu beweisen. Da die finite Mathematik als die der *anschaulichen* Evidenz definiert ist [2], so bedeutet das (wie auch von Bernays in *L'enseignement mathématique*, 34 (1935), p. 62 und 69, explizit formuliert wurde), dass man für den Widerspruchsfreiheitsbeweis der Zahlentheorie gewisse *abstrakte* Begriffe braucht. Dabei sind unter abstrakten (oder nichtanschaulichen) Begriffen solche zu verstehen, die wesentlich von zweiter oder höherer Stufe sind, das heisst, die nicht Eigenschaften oder Relationen *konkreter Objekte* (z. B. von Zeichenkombinationen) beinhalten, sondern sich auf *Denkgebilde* (z. B. Beweise, sinnvolle Aussagen usw.) beziehen, wobei in den Beweisen Einsichten über die letzteren gebraucht werden, die sich nicht aus den kombinatorischen (raumzeitlichen) Eigenschaften der sie darstellenden Zeichenkombinationen, sondern nur aus deren *Sinn* ergeben.

Obwohl in Ermanglung eines präzisen Begriffs der anschaulichen, beziehungsweise abstrakten, Evidenz ein strenger Beweis für die Bernayssche Feststellung nicht vorliegt, so kann doch über ihre Richtigkeit praktisch kein Zweifel bestehen, insbesondere seit dem Gentzenschen Beweis für die Formalisierbarkeit aller Rekur-

[1] Vgl. z. B.: *Entretiens de Zurich* (herausg. von F. GONSETH, 1941), p. 144, 147; ferner: HILBERT-BERNAYS, *Grundlagen der Mathematik*, Bd. 2 (1939), § 5; und: *Rev. Internat. Phil.* Nr. 27-28 (1954), Fasc. 1-2, p. 2.

[2] Vgl. die Hilbertsche Formulierung in *Math. Ann.* 95 (1925), p. 171-173.

sionen nach Ordinalzahlen $< \varepsilon_0$ in der Zahlentheorie. Denn die Gültigkeit des Rekursionsschlusses für ε_0 kann sicher nicht unmittelbar anschaulich gemacht werden, wie das zum Beispiel bei ω^2 möglich ist. Das heisst genauer, man kann die verschiedenen strukturellen Möglichkeiten, die für absteigende Folgen bestehen, nicht mehr übersehen und hat daher keine anschauliche Erkenntnis von der Notwendigkeit des Abbrechens jeder solchen Folge. Insbesondere kann durch schrittweises Übergehen von kleineren zu grösseren Ordinalzahlen eine solche *anschauliche* Erkenntnis nicht realisiert werden, sondern bloss eine abstrakte Erkenntnis mit Hilfe von Begriffen höherer Stufe. Das letztere wird durch den abstrakten Begriff der «Erreichbarkeit[1]» geleistet, welcher durch die inhaltliche Beweisbarkeit der Gültigkeit einer gewissen Schlussweise definiert ist. Auch ist es im Rahmen der für uns anschaulichen Mathematik nicht möglich, den Induktionsschluss nach einer hinreichend grossen Ordinalzahl auf eine Kette anderer Einsichten zurückzuführen. Vielmehr führt jeder Versuch, das zu tun, zu Induktionen von im wesentlichen derselben Ordnung. Ob die Notwendigkeit abstrakter Begriffe bloss durch die praktische Unmöglichkeit, kombinatorisch allzu komplizierte Verhältnisse anschaulich vorzustellen [2], bedingt ist oder prinzipielle Gründe hat,

[1] W. Ackermann erklärt zwar in *Math. Zeits.*, Bd. 53, p. 407, dass « erreichbar » einen anschaulichen Sinn habe, wenn Beweisbarkeit als formale Beweisbarkeit nach gewissen Regeln verstanden wird. Aber darauf ist zu erwidern, dass aus dieser anschaulichen Tatsache die Gültigkeit des Schlusses durch transfinite Induktion für eine vorgelegte Eigenschaft nur mit Hilfe abstrakter Begriffe (oder mit Hilfe transfiniter Induktion in der Metamathematik) folgt. Allerdings ist der Begriff « erreichbar », zumindest für Induktionen bis ε_0, durch schwächere abstrakte Begriffe ersetzbar (vgl. HILBERT-BERNAYS, *Grundlagen der Mathematik*, Bd. 2).

[2] Man beachte, dass eine adäquate beweistheoretische Charakterisierung einer durch Absehen von dieser Schranke *idealisierten* anschaulichen Evidenz Schlussweisen enthalten wird, die *für uns* nicht anschaulich sind und die sehr wohl eine Reduktion des induktiven Schlusses auf den einer wesentlich kleineren Ordnung gestatten könnten. Eine andere Möglichkeit, den ursprünglichen finiten Standpunkt zu erweitern, für die dasselbe gilt, besteht darin, dass man abstrakte Begriffe, die auf nichts anderes als auf finite Begriffe und Gegenstände, und zwar in kombinatorisch finiter Weise, Bezug nehmen, mit zur finiten Mathematik rechnet und diesen Prozess iteriert. Solche Begriffe sind zum Beispiel diejenigen, welche in der Reflexion auf den Inhalt schon konstruierter finiter Formalismen involviert sind. Ein

lässt sich nicht ohne weiteres entscheiden. Im zweiten Fall müsste nach Präzisierung der fraglichen Begriffe ein strenger Beweis für das Bestehen jener Notwendigkeit möglich sein.

Jedenfalls lehrt die Bernayssche Bemerkung, zwei Bestandteile in der finiten Einstellung unterscheiden, nämlich erstens das konstruktive Element, welches darin besteht, dass von mathematischen Objekten nur insoweit die Rede sein darf, als man sie aufweisen oder durch Konstruktion tatsächlich herstellen kann; zweitens das spezifisch finitistische Element, welches darüber hinaus fordert, dass die Objekte, über welche man Aussagen macht, mit welchen die Konstruktionen ausgeführt werden und welche man durch sie erhält, « anschaulich » sind, das heisst letzten Endes raum-zeitliche Anordnungen von Elementen, deren Beschaffenheit abgesehen von Gleichheit und Verschiedenheit irrelevant ist. (Im Gegensatz dazu sind jene Objekte in der intuitionistischen Logik sinnvolle Aussagen und Beweise.)

Es ist die zweite Forderung, welche fallen gelassen werden muss. Dieser Tatsache wurde bisher dadurch Rechnung getragen, dass man Teile der intuitionistischen Logik und Ordinalzahltheorie zur finiten Mathematik adjungierte. Im folgenden wird gezeigt, dass man statt dessen für den Widerspruchsfreiheitsbeweis der Zahlentheorie auch den Begriff der berechenbaren Funktion endlichen Typs über den natürlichen Zahlen und gewisse sehr elementare Konstruktionsprinzipien für solche Funktionen verwenden kann. Dabei wird der Begriff « berechenbare Funktion vom Typus t » folgendermassen erklärt: 1. Die berechenbaren Funktionen vom Typus 0 sind die natürlichen Zahlen. 2. Wenn die Begriffe « berechenbare Funktion vom Typus t_0 », « berechenbare Funktion vom Typus t_1 », ... « berechenbare Funktion vom Typus t_k » (wobei $k \geqq 1$) bereits definiert sind, so wird eine berechenbare Funktion vom Typus $(t_0, t_1, \ldots t_k)$ definiert als eine immer ausführbare (und als solche konstruktiv erkennbare) Operation, welche jedem k-tupel

dieser Idee entsprechender Formalismus wurde von G. Kreisel aufgestellt. Vgl. seinen Vortrag auf dem Internationalen Mathematikerkongress in Edinburgh, 1958. Man beachte, dass bei dieser Art der Erweiterung des Finitismus das abstrakte Element in einer wesentlichen schwächeren Form auftritt als bei der weiter unten besprochenen oder in der intuitionistischen Logik.

berechenbarer Funktionen der Typen $t_1, t_2, \ldots t_k$ eine berechenbare Funktion vom Typus t_0 zuordnet. Dieser Begriff [1] ist als unmittelbar verständlich [2] zu betrachten, vorausgesetzt dass man die Begriffe « berechenbare Funktion vom Typus t_i » ($i = 0, 1, \ldots k$) bereits verstanden hat. Indem man dann den Typus t als Variable betrachtet, gelangt man zu dem für den Widerspruchsfreiheitsbeweis benötigten Begriff einer berechenbaren Funktion endlichen Typs t.

Als evidente Axiome sind, neben den Axiomen der Identität (auch für Funktionen [3]), dem 3. und 4. Peanoschen Axiom und der Substitutionsregel für freie Variable, keine anderen nötig als erstens solche, die es gestatten, Funktionen durch Gleichsetzung mit einem aus Variablen und vorher definierten Konstanten aufgebauten Term und durch einfache Induktion nach einer Zahlvariablen zu definieren, und zweitens den Schluss der vollständigen Induktion nach einer Zahlvariablen anzuwenden. Das heisst die Axiome dieses Systems (es werde T genannt) sind formal fast dieselben [4] wie die der primitiv rekursiven Zahlentheorie, nur dass

[1] Man kann darüber im Zweifel sein, ob wir eine genügend deutliche Vorstellung vom Inhalt dieses Begriffs haben, aber nicht darüber, ob die weiter unten angegebenen Axiome für ihn gelten. Derselbe scheinbar paradoxe Sachverhalt besteht auch für den der intuitionistischen Logik zugrunde liegenden Begriff des inhaltlich richtigen Beweises. Wie die nachfolgenden Überlegungen und die intuitionistisch interpretierte Theorie der rekursiven Funktionen und Funktionale zeigen, sind diese beiden Begriffe innerhalb gewisser Grenzen als Grundbegriffe durcheinander ersetzbar. Dabei ist zu beachten, dass, wenn der Begriff der berechenbaren Funktion nicht implizit den Begriff des Beweises enthalten soll, die Ausführbarkeit der Operationen unmittelbar aus der Kette der Definitionen ersichtlich sein muss, wie das für alle Funktionen des weiter unten angegebenen Systems T der Fall ist.

[2] A. M. Turing hat bekanntlich mit Hilfe des Begriffs einer Rechenmaschine eine Definition des Begriffs einer berechenbaren Funktion erster Stufe gegeben. Aber wenn dieser Begriff nicht schon vorher verständlich gewesen wäre, hätte die Frage, ob die Turingsche Definition adäquat ist, keinen Sinn.

[3] Identität zwischen Funktionen ist als intensionale oder Definitionsgleichheit zu verstehen.

[4] Bei der Definition durch Gleichsetzung mit einem Term tritt insofern ein Unterschied auf, als man eine Funktion P höheren Typs auch durch $[P(x_1, x_2, \ldots x_n)](y_1, y_2, \ldots y_m) = E$ definieren kann. Aber dieser Unterschied fällt weg, falls mehrstellige Funktionen in der von A. Church angegebenen Weise durch einstellige ersetzt werden.

die Variablen (ausser denen, auf die Induktion angewendet wird), sowie auch die definierten Konstanten, einen beliebigen endlichen Typus über den natürlichen Zahlen haben können. Der Einfachheit halber wird im folgenden der zweiwertige Aussagenkalkül, angewendet auf Gleichungen, hinzugenommen, obwohl die Wahrheitsfunktionen durch zahlentheoretische Funktionen ersetzbar sind. Gebundene Variable werden nicht zugelassen. Das System T ist von gleicher Beweisstärke wie ein System der rekursiven Zahlentheorie, in dem vollständige Induktion für alle Ordinalzahlen $< \varepsilon_0$ (in der gewöhnlichen Darstellung) zugelassen wird.

Die Zurückführung der Widerspruchsfreiheit der klassischen Zahlentheorie auf die des Systems T gelingt mit Hilfe der folgenden Interpretation der Heytingschen Zahlentheorie, auf welche ja die klassische zurückführbar ist [1]:

Es wird jeder Formel F der intuitionistischen Zahlentheorie [2] (deren freie Variable in ihrer Gesamtheit mit x bezeichnet werden) eine Formel F′ der Gestalt $(\exists y)(z) \, A(y, z, x)$ zugeordnet, wobei y und z endliche Reihen von Variablen irgendwelcher Typen sind, und $A(y, z, x)$ ein quantorenfreier Ausdruck mit keinen andern als den in x, y, z vorkommenden Variablen. Die Variablen der Reihen x, y, z, deren Gliederzahl auch 0 sein kann, sind sämtlich untereinander verschieden. Mit xy wird die aus x und y in dieser Reihenfolge zusammengesetzte Reihe bezeichnet.

Ferner werden folgende Bezeichnungen verwendet:

1. v, w sind endliche Reihen von Variablen irgendwelcher Typen; s, t sind Zahlvariable; u ist eine Reihe von Zahlvariablen.

2. V ist eine Reihe von Variablen, deren Anzahl und Typen dadurch bestimmt sind, dass jede von ihnen auf y als Argumentreihe angewendet werden kann und dass die Reihe der so erhaltenen Werte (welche mit $V(y)$ bezeichnet wird) hinsichtlich der Anzahl und der Typen ihrer Glieder mit der Reihe v übereinstimmt.

[1] Vgl. *Ergebnisse eines math. Kolloquiums*, herausg. von K. Menger, Heft 4 (1933), p. 34.

[2] Die Zahlentheorie soll so formalisiert sein, dass keine Aussagen- oder Funktionsvariable vorkommen. Die Axiome des Aussagenkalküls sind als Schemata für alle möglichen Einsetzungen zu betrachten.

3. Analog wird die Variablenreihe Y (bzw. Z, bzw. $\bar{\bar{Z}}$) hinsichtlich der Anzahl und der Typen ihrer Glieder durch die Argumentreihe s (bzw. yw, bzw. y) und durch die mit der Reihe der Werte gleichtypige Reihe y (bzw. z, bzw. z) bestimmt.

Funktionen mit 0 Leerstellen und Werten vom Typus τ werden mit Objekten vom Typus τ identifiziert, eingliedrige Variablenreihen mit Variablen.

Die Zuordnung von F′ zu F geschieht durch Induktion nach der Anzahl k der in F enthaltenen logischen Operatoren. (Die bei der Wahl der Symbole für die gebundenen Variablen zu beachtenden Bedingungen und die heuristische Begründung der Definitionen werden nach den Formeln gegeben.)

I. Für $k = 0$ sei $\mathrm{F}' = \mathrm{F}$.

II. Es sei $\mathrm{F}' = (\exists y)(z)\,\mathrm{A}(y, z, x)$ und
$\mathrm{G}' = (\exists v)(w)\,\mathrm{B}(v, w, u)$ bereits definiert,

dann ist *per definitionem*:

1. $(\mathrm{F} \wedge \mathrm{G})' = (\exists yv)(zw)\,[\mathrm{A}(y, z, x) \wedge \mathrm{B}(v, w, u)]$.
2. $(\mathrm{F} \vee \mathrm{G})' = (\exists yvt)(zw)\,[t=0 \wedge \mathrm{A}(y, z, x) \cdot\!\vee\!\cdot\, t=1 \wedge \mathrm{B}(v, w, u)]$.
3. $[(s)\,\mathrm{F}]' = (\exists \mathrm{Y})(sz)\,\mathrm{A}(\mathrm{Y}(s), z, x)$.
4. $[(\exists s)\,\mathrm{F}]' = (\exists sy)(z)\,\mathrm{A}(y, z, x)$.
5. $(\mathrm{F} \supset \mathrm{G})' = (\exists \mathrm{VZ})(yw)\,[\mathrm{A}(y, \mathrm{Z}(yw), x) \supset \mathrm{B}(\mathrm{V}(y), w, u)]$.
6. $(\neg\,\mathrm{F})' = (\exists \bar{\bar{\mathrm{Z}}})(y)\,\neg\,\mathrm{A}(y, \bar{\bar{\mathrm{Z}}}(y), x)$.

s ist eine beliebige Zahlvariable. Vor Anwendung der Regeln 1—5 sind nötigenfalls die gebundenen Variablen der Formeln F′ und G′ so umzubenennen, dass sie sämtlich untereinander und von den Variablen der Reihen x, u sowie auch von s verschieden sind. Ferner sind die durch Anwendung der Regeln 2, 3, 5, 6 neu eingeführten gebundenen Variablen der Reihen t, Y, V, Z, $\bar{\bar{\mathrm{Z}}}$ so zu wählen, dass sie untereinander und von den in den betreffenden Formeln schon vorkommenden Variablen verschieden sind.

Man beachte, dass 6. aus 5. folgt, falls $\neg\, p$ durch $p \supset \cdot\, 0 = 1$ definiert wird. Zu 5. gelangt man, indem man (für die auftretenden Spezialfälle) die Aussage $(\exists x)\,\mathrm{H}(x) \supset (\exists y)\,\mathrm{R}(y)$ (bzw. $(y)\,\mathrm{R}(y) \supset$

$(x)\ H\ (x))$ mit der Existenz von für alle Argumentreihen vom Typus der Variablenreihe x definierten berechenbaren Funktionen identifiziert, welche jedem Beispiel für das Implicans (bzw. Gegenbeispiel für das Implicatum) ein Beispiel für das Implicatum (bzw. Gegenbeispiel für das Implicans) zuordnen.

Selbstverständlich wird nicht behauptet, dass die Definitionen 1—6 den Sinn der von Brouwer und Heyting eingeführten logischen Partikel wiedergeben. Wieweit sie diese ersetzen können, bedarf einer näheren Untersuchung. Man zeigt leicht, dass, *wenn* F *im Heytingschen System* Z *der Zahlentheorie beweisbar ist, Funktionen* Q *in* T *definiert werden können, für welche* A (Q (x), z, x) *in* T *beweisbar ist.* Es ist nämlich leicht nachprüfbar, dass diese Behauptung für die Axiome von Z gilt und ihre Richtigkeit sich bei Anwendung der Schlussregeln von Z von den Prämissen auf die Konklusion überträgt.

Die Verifikation wird besonders einfach, wenn man folgendes Axiomensystem der intuitionistischen Logik zugrunde legt [1]:

Axiome: Taut, Add, Perm, die zu diesen dualen Axiome für $\wedge$, $0 = 1 \cdot \supset p$ ($\neg\, p$ wird durch $p \supset \cdot 0 = 1$ definiert).

Schlussregeln: Modus ponens, Einsetzungsregel für freie Zahlvariable, Syll (mit zwei Prämissen), Sum, Exp, Imp, die Regeln über das Hinzufügen und Weglassen eines All- (bzw. Existenz-) Zeichens im Implicatum (bzw. Implicans) einer bewiesenen Implikation.

Für den Widerspruchsfreiheitsbeweis der klassischen Zahlentheorie können die $\vee$ und die $\exists$ enthaltenden Axiome und Schlussregeln weggelassen werden. Bei allen auf Sum folgenden Regeln stellt sich heraus, dass die in T zu beweisende Aussage im wesentlichen dieselbe ist wie die auf Grund der Prämisse bereits bewiesene.

Es ist klar, dass man, von demselben Grundgedanken ausgehend, auch viel stärkere Systeme als T konstruieren kann, zum Beispiel durch Zulassung transfiniter Typen oder der von Brouwer für den Beweis des « Fan-Theorems [2] » benutzten Schlussweise.

[1] Bezüglich der Bezeichnungen, vgl. *Principia Mathematica*, 2. Aufl., p. XII. Dieselben Bezeichnungen werden auch für die den Formeln entsprechenden Schlussregeln verwendet.

[2] Vgl. A. HEYTING, *Intuitionism*, 1956, p. 42.

Zusammenfassung

P. Bernays hat darauf hingewiesen, dass man, um die Widerspruchsfreiheit der klassischen Zahlentheorie zu beweisen, den Hilbertschen finiten Standpunkt dadurch erweitern muss, dass man neben den auf Symbole sich beziehenden kombinatorischen Begriffen gewisse abstrakte Begriffe zulässt. Die abstrakten Begriffe, die bisher für diesen Zweck verwendet wurden, sind die der konstruktiven Ordinalzahltheorie und die der intuitionistischen Logik. Es wird gezeigt, dass man statt dessen den Begriff einer berechenbaren Funktion endlichen einfachen Typs über den natürlichen Zahlen benutzen kann, wobei keine anderen Konstruktionsverfahren für solche Funktionen nötig sind, als einfache Rekursion nach einer Zahlvariablen und Einsetzung von Funktionen ineinander (mit trivialen Funktionen als Ausgangspunkt).

Abstract

P. Bernays has pointed out that, in order to prove the consistency of classical number theory, it is necessary to extend Hilbert's finitary standpoint by admitting certain abstract concepts in addition to the combinatorial concepts referring to symbols. The abstract concepts that so far have been used for this purpose are those of the constructive theory of ordinals and those of intuitionistic logic. It is shown that the concept of a computable function of finite simple type over the integers can be used instead, where no other procedures of constructing such functions are necessary except simple recursion by an integral variable and substitution of functions in each other (starting with trivial functions).