

DIE REVISIONBEDÜRFTIGKEIT DER LOGISCHEN SEMANTIK FREGES

CHRISTIAN THIEL

Mit den folgenden Bemerkungen möchte ich die Aufmerksamkeit der Logikhistoriker auf eine neuere Einsicht über das begriffsschriftliche System von Freges «Grundgesetzen der Arithmetik» lenken, die m.E. zu einer völlig neuen Einschätzung desselben zwingt. Es geht dabei um ein wichtiges inhaltliches Problem der Fregeschen Logik und Mengenlehre, wie sie im Hauptwerk der logizistischen Phase Freges, den «Grundgesetzen der Arithmetik» dargestellt ist. Seit Beginn des Jahrhunderts, nämlich seit Russells Entdeckung der heute meist nach ihm benannten Antinomie ist dieses System Freges Gegenstand von Untersuchungen gewesen, die den «eigentlichen» Grund der Inkonsistenz aufdecken sollten. Dass das System verändert werden muss, ist klar, nicht jedoch, in welcher Richtung. Die gegenwärtig bevorzugten Lösungen sind wohl typentheoretischer Art und verwerfen überdies Freges Grundgesetz V, das Abstraktionsprinzip





$$\dot{\epsilon}(\Phi(\epsilon)) = \dot{\alpha}(\Psi(\alpha)) \iff \bigwedge x (\Phi(x) \iff \Psi(x)).$$

Eine andere, auch von mir bevorzugte Lösung akzeptiert das Abstraktionsprinzip wie es steht, und verlangt Einschränkungen für die Bildung der zugelassenen Funktionsnamen $\Phi(\xi)$ und $\Psi(\xi)$ und damit der Wertverlaufsnamen $\dot{\epsilon}(\Phi(\epsilon))$ und $\dot{\alpha}(\Psi(\alpha))$.

Diese hübsche Alternative wurde empfindlich gestört, als Peter Aczel beim Kleene-Symposium 1979 eine Studie «Frege Structures» vortrug, auf die ich erst durch die Kurzfassung in der Frege-Sektion des Kongresses in Hannover 1979 und das mir anschliessend zugänglich gemachte Manuskript für die Kleene-Festschrift aufmerk-

sam wurde. Aczel geht auf die Möglichkeit verschiedener Auswege aus der Inkonsistenz des Fregeschen Systems nicht ein, sondern versucht, sofort einen «haltbaren Teil» desselben zu charakterisieren. Er sieht die Ursache für das Auftreten der Russellschen Antinomie in Freges Einführung des Waagerechten, also der Funktion $\text{—}\xi$, deren Einführung nichts anderes als eine systeminterne Definition des Prädikats «wahr» sei und daher (nach einem bekannten Ergebnis von Tarski) zur Inkonsistenz führe. Die Behauptung, der Waagerechte sei an allem schuld, wurde wohl von vielen Hörern in Hannover als skurril empfunden und mit Unglauben aufgenommen. Auch ich konnte, da Aczel seine Darstellung von den Eigentümlichkeiten der Fregeschen Begriffsschrift weitgehend freihält, seinen Gedankengang nicht voll verstehen und habe mich in diesem Sinne auf mehreren Fregetagungen 1979 geäußert. Heute möchte ich einen Versuch machen, ohne grosse technische Details (deren korrekte Durchführbarkeit ich unterstelle — s. aber unten), die Aczelsche Argumentation zu verdeutlichen.

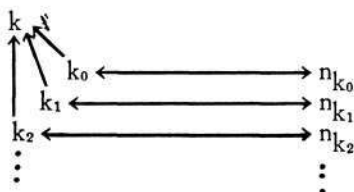
Freges Notation der Begriffsschrift ist bekanntlich zweidimensional, aber sie lässt sich linearisieren, z.B. durch Umwandlung

von «» in «-a», von « Δ » in « (Γ) (Δ)»

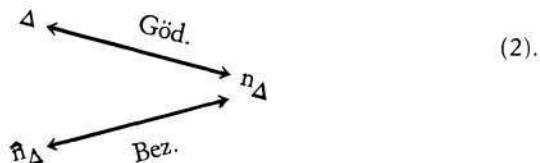
und ähnliche Verfahren. Statt der vielen verschiedenen Buchstaben können wir für metalogische Untersuchungen eine einzige Konstante nehmen und eine gewisse Anzahl von Apostrophen anhängen, z.B. «x» statt «ξ», «x'''» statt «a», «x''''» statt «F» usw. schreiben. Das alles sind bekannte Tricks, und es soll auch nur gezeigt werden, was man ohnehin glaubt, dass nämlich jeder rechtmässig gebildete Ausdruck der Begriffsschrift über einem endlichen Alphabet gebildet ist bzw. durch einen derartigen Ausdruck ersetzt werden kann. Dann ist also die Klasse der Begriffsschriftausdrücke abzählbar, und wir können jedem Begriffsschriftausdruck eineindeutig eine kennzeichnende Zahl als Nummer zuordnen:

$$\Delta \xleftrightarrow{1-1} n_{\Delta} \quad (1)$$

Wir können die Zuordnung darüber hinaus in bekannter Weise als eine «Gödelisierung» vornehmen, d.h. mittels einer (effektiv) entscheidbaren eineindeutigen Abbildung und so, dass Beziehungen wie der Ableitbarkeit eines Ausdrucks aus (einem oder mehreren) anderen, dem Hervorgehen eines Ausdrucks aus einem anderen durch Substitution eines zulässigen Ausdrucks in letzterem etc. entscheidbare Relationen (oder Relationen effektiver elementarer Berechenbarkeit) zwischen den den Ausdrücken zugeordneten Gödelnummern entsprechen. Da Freges System den Aufbau eines Modells der Arithmetik erlaubt, wird jede natürliche Zahl durch gewisse (untereinander bedeutungsgleiche) Begriffsschriftausdrücke bezeichnet:



Wählen wir unter den Ausdrücken, die eine natürliche Zahl k bezeichnen, denjenigen mit der kleinsten Gödelnummer als «Darstellung von k » aus und bezeichnen ihn mit « \hat{k} », so wird die Beziehung eineindeutig. Insbesondere entspricht dann jede Gödelnummer eineindeutig sowohl dem *durch sie numerierten* Begriffsschriftausdruck als auch dem *sie bezeichnenden* Begriffsschriftausdruck:



Ferner entspricht durch *Arithmetisierung* jedem Begriffsausdruck $\Phi(\xi)$ der Begriffsschrift ein Begriffsausdruck $\Psi_1(\xi)$ so, dass $\Phi(\xi)$ von einem für ξ einsetzbaren Begriffsschriftausdruck Δ genau dann erfüllt wird, wenn $\Psi_1(\xi)$ von der Gödelnummer von Δ erfüllt wird. Da Δ in diesem Fall aber sogar *alle* zu $\Psi_1(\xi)$ wertverlaufsgleichen Funktionsnamen erfüllt, wählen wir unter diesen $\Psi_1(\xi)$

wieder denjenigen mit der kleinsten Gödelnummer aus; wir nennen ihn $\Psi(\xi)$, so dass also gilt:

$$\Phi(\Delta) \text{ ist wahr} \iff \Psi(n_\Delta) \text{ ist wahr.} \quad (3).$$

Die erste von Frege in den «Grundgesetzen» eingeführte Funktion ist $-\xi$; « $-\xi$ » ist ein Begriffsausdruck, dem bei der Arithmetisierung ein Begriffsausdruck « $H(\xi)$ » entspricht, so dass

$$-\Delta \text{ ist wahr} \iff H(n_\Delta) \text{ ist wahr.} \quad (4).$$

Nach Freges Erklärung der Funktion $-\xi$ gilt

$$-\Delta \text{ ist wahr} \iff \Delta \text{ ist wahr,} \quad (5).$$

und daher nach seiner Erklärung der Funktion $\neg \xi$ für dasselbe Δ :

$$\neg \Delta \text{ ist falsch.}$$

Dann ist aber auch $\neg H(n_\Delta)$ falsch, denn

$$\begin{aligned} \neg H(n_\Delta) \text{ ist wahr} &\implies \neg H(n_\Delta) \text{ ist falsch} \\ &\implies H(n_\Delta) \text{ ist falsch (da «} H(\xi)\text{»} \\ &\quad \text{Begriffsausdruck ist),} \\ &\implies -\Delta \text{ ist falsch (nach (4)),} \end{aligned}$$

im Widerspruch zur Voraussetzung, dass $-\Delta$ wahr sei.

Nun betrachten wir die Klasse K aller Gödelnummern m von Begriffsausdrücken $X(\xi)$, für die $X(m)$ falsch ist. Als Klasse natürlicher Zahlen wird K in der Begriffsschrift durch gewisse Wertverlaufsnamen bezeichnet, von denen wir wieder den mit der kleinsten Gödelnummer auswählen und mit « $\hat{\xi}(F(\epsilon))$ » bezeichnen. Die Elemente von $\hat{\xi}(F(\epsilon))$ und nur diese erfüllen die zugehörige Funktion $F(\xi)$:

$$\bigwedge_x (x \in K \iff F(\hat{x})). \quad (6).$$

t sei die Gödelnummer von « $F(\xi)$ », und unter den t bezeichnenden Begriffsschriftausdrücken sei \hat{t} derjenige mit der kleinsten Gödelnummer. Wir fragen nach dem Wahrheitswert von $F(\hat{t})$.

1. *Annahme*, $F(\mathfrak{f})$ sei wahr. Dann gilt wegen (6) $t \in K$, d.h. t gehört zur Klasse der Gödelnummern von Begriffsausdrücken, die von den Namen der Gödelnummern nicht erfüllt werden,

$$t \in \{m \mid \forall g (m = n_g(\xi) \wedge \neg g(\mathfrak{m}))\}.$$

Also nach der Erklärung der Elementbeziehung:

$$\forall g (t = n_g(\xi) \wedge \neg g(\mathfrak{f})). \quad (7).$$

Da die Gödelisierung eineindeutig ist, kann $g(\xi)$ nur der Begriffsausdruck $F(\xi)$ sein, als dessen Gödelnummer ja t erklärt war. Es gilt also

$$t = n_F(\xi) \wedge \neg F(\mathfrak{f}),$$

letzteres im Widerspruch zu unserer ursprünglichen Annahme, dass $F(\mathfrak{f})$ wahr sei. $F(\mathfrak{f})$ kann also nicht wahr sein und wir kommen zu der gegenteiligen

2. *Annahme*, $F(\mathfrak{f})$ sei falsch. Aber dann gehört t zu K , das gerade als Klasse derjenigen Gödelnummern *erklärt* war, die Nummern von Begriffsausdrücken $X(\xi)$ sind, die bei Einsetzung des Namens der betr. Gödelnummer in ihre Leerstelle eine falsche Aussage liefern. Aber nach (6) gilt dann mit $t \in K$ auch $F(\mathfrak{f})$, es wäre $F(\mathfrak{f})$ wahr im Widerspruch zur 2. Annahme.

Da die 1. und die 2. Annahme zusammen einen Widerspruch ergeben, muss eine ihrer Voraussetzungen fehlerhaft sein. Aczel lokalisiert den Fehler in der Annahme, $-\xi$ und $\neg\xi$ seien in der Begriffsschrift als Prädikate beliebiger Gegenstände zulässig: Freges Erklärung

$$\langle\langle -\Delta \text{ ist wahr} \iff \Delta \text{ ist wahr} \rangle\rangle$$

wäre ja eine begriffsschriftinterne Definition der *Wahrheit*, die nach dem Satz von Tarski nicht möglich ist. Auch die Fregesche Reduktion von $-\xi$ auf das wertverlaufsgleiche $\xi = (\xi = \xi)$ bringt keine Rettung, da die Begriffsschrift, wiederum fehlerhaft, zwischen Identität und extensionaler Gleichheit nicht unterscheidet.

Wenn ich Aczels Überlegung in etwas engerer Anlehnung an Freges «Grundgesetze» richtig wiedergegeben habe, so kann ich seiner Argumentation nur zustimmen. Verwunderlich scheint nachträglich nur, weshalb sowohl die Fregekenner als auch die Mengen-

theoretiker, die ja beide in Frege den Ahnherrn der modernen modelltheoretischen Semantik sehen, diesen wunden und — wie sich jetzt zeigt, fatalen — Punkt so lange übersehen konnten. Eine der Ursachen mag die (auch in der vorliegenden sehr skizzenhaften Darstellung nicht behobene) Ungeklärtheit der hier relevanten Eigenschaften des Fregeschen Systems der «Grundgesetze» sein. Um Tarskis Überlegung (man vgl. zu dieser etwa Jean Ladrière 1957, chap. VIII, sect. 2: «Les théorèmes de Tarski») für die Untersuchung dieses Systems anzuwenden, müssten ja zunächst einmal die Darstellbarkeit der (rekursiven) Arithmetik gezeigt sowie die Voraussetzungen der verwendeten mengentheoretischen Metasprache (z.B. die Existenz einer Menge, deren Elemente einem Modell der «Grundgesetze» zugrundegelegt werden können) nachgeliefert werden, abgesehen davon, dass die Rolle der bekannten Inkonsistenz der «Grundgesetze» berücksichtigt werden muss. All dies ist m.W. bis heute nicht bearbeitet, und auch meine obigen Hinweise enthalten keine Schritte in dieser Richtung — sie sollen erst einmal ihre logikhistorische Bedeutsamkeit in den Blick bringen.

Die korrekte Durchführbarkeit dieser Vorklärungen einmal unterstellt, ist es eine zweite Frage, ob man für die Reparatur des Fregeschen Systems typentheoretischen oder prädikativen Ansätzen oder aber dem Vorschlag Aczels folgen sollte. Nach dem letzteren sind die von ihm als «Prädikation» bezeichnete Elementbeziehung in der Form

$$a \text{ Gegenstand, } b \text{ Menge} \implies a \in b \text{ Aussage}$$

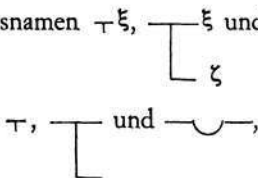
sowie die Komprehension

$$f \text{ einstellige Aussageform 1.Stufe, } a \text{ Gegenstand} \\ \implies b \iff \{x \mid f(x)\} \text{ Menge und } a \in b \iff f(a)$$

intuitiv korrekt und durch Angabe klassischer Modelle als widerspruchsfrei gerechtfertigt.

Die Unzulässigkeit von $\text{—}\xi$ stellt sich logikhistorisch als einschneidend dar, denn das Notationssystem der Fregeschen Begriffsschrift scheint mit dem «Waagerechten» zu stehen und zu fallen: der Urteilsstrich «|-» darf nur vor einen Ausdruck treten, der mit einem Waagerechten beginnt, und es ist für den Algorithmus der Begriffsschrift wesentlich, dass die konstitutiven Teile der ur-

sprünglichen Funktionsnamen $\neg \xi$, ξ und $\neg a \rightarrow \varphi(a)$,



sich als \neg , ξ und $\neg a \rightarrow \varphi(a)$ auffassen

lassen. Es scheint demnach, dass das so originelle und einmalige Notationssystem Freges von Grund auf verfehlt ist und daher aufgegeben werden muss — für viele Kenner und Freunde Freges sicher ein Schock und Anlass zu tiefem Bedauern.

Ich möchte freilich erst einmal zur Diskussion stellen, ob diese katastrophale Folge unausweichlich ist. Aczel selbst hat *nicht* von der Unhaltbarkeit des Notationssystems der «Grundgesetze» gesprochen, sondern von der Unzulässigkeit des Waagerechten, so wie er bei Frege erklärt wird — und er könnte dafür einen guten Grund gehabt haben. Denn was die Erklärung « $\neg \Delta$ ist wahr $\iff \Delta$ ist wahr» zu einer begriffsschriftinternen Definition der Wahrheit macht, ist ja nicht der Waagerechte als solcher, sondern seine Anwendung auf Namen *beliebiger Gegenstände*, wobei sich jedesmal ein Name eines Wahrheitswertes ergibt, und der letztere dabei selbst als *Gegenstand* betrachtet wird. Dass dies etwas anderes ist, wird deutlich, wenn wir statt $\neg \xi$ einmal $\neg \xi$ betrachten. Lassen wir zur Ersetzung von ξ nur *Aussagen* zu, so ist « $\neg A$ ist wahr $\iff A$ ist falsch» nur eine andere Schreibweise für die Wahrheitstafel der Negation:

A	$\neg A$
w	f
f	w

Und ganz analog könnten wir jetzt eine «Position» $\neg A$ durch

A	$\neg A$
w	w
f	f

erklären, wie dies ja tatsächlich geschieht, z.B. in Bochenski-Mennes «Grundriss der Logistik» (Paderborn 1954, 1973) im § 3.22. Dann ist aber $\neg\xi$ ebenso wie $\neg\neg\xi$ nur noch Aussagenfunktion, nicht mehr eine Funktion, die für beliebige Gegenstände als Argumente erklärt ist. Ihre Werte brauchen genausowenig wie ihre Argumente *Namen* des Wahren oder des Falschen zu sein; sie sind wahre oder falsche Aussagen, zu deren Erklärung es möglicherweise auch noch ganz andere Wege als den semantisch-modelltheoretischen gibt.

Wenn ich die neue Situation richtig sehe, fällt nicht der Waagerechte dem von Aczel entdeckten Unglück zum Opfer, sondern seine Verwendung im Rahmen einer naiven Semantik. Das eigentliche Opfer ist nicht der (wie es jetzt scheint, lediglich missbrauchte) Waagerechte, sondern *die Fregesche Ontologie* mit ihrer vollständigen Einteilung des «Mobiliars der Welt» in Gegenstände einerseits, Funktionen (der verschiedenen Stufen) andererseits. Für die Philosophiehistoriker und die Philosophen unter den Logikhistorikern mag dies ein noch weit schmerzlicherer Verlust sein als es der des Waagerechten gewesen wäre (der immerhin als einstelliger Funktor überleben kann). Das Schicksal der Fregeschen Ontologie freilich ist ein Thema, das ich unserem Rahmen nicht mehr erörtern möchte.

LITERATURANGABEN

- ACZEL, Peter: *The Structure of the Formal Language of Frege's Grundgesetze*. 6th International Congress of Logic, Methodology and Philosophy of Science, Hannover August 22-August 29, 1979. Abstracts, Sections 13/14, pp. 16-17.
- *Frege Structures and the Notions of Proposition, Truth and Set*. In: J. Barwise, H. J. Keisler, K. Kunen (eds.), *The Kleene Symposium* (North-Holland Publishing Company: Amsterdam, New York, Oxford 1980), pp. 31-59.
- GRATTAN-GUINNESS, Ivor: *Rezension von Gottlob Frege, Wissenschaftlicher Briefwechsel*, ed. Gabriel, H. Hermes, F. Kambartel, C. Thiel, A. Veraart (Felix Meiner: Hamburg 1976), *Annals of Science* 34 (1977), pp. 214-215.

DIE LOGISCHE SEMANTIK FREGES

LADRIÈRE, Jean: *Les Limitations Internes des Formalismes*. Etude sur la signification du théorème de Gödel et des théorèmes apparentés dans la théorie des fondements des mathématiques. Louvain/Paris 1957.

THIEL, Christian: *Frege und die moderne Grundlagenforschung*. Symposium, gehalten in Bad Homburg im Dezember 1973 (ed.) Anton Hain: Meisenheim am Glan 1975.