Introduction to Bernays Text No. 6, “Appendix to Hilbert’s Lecture ‘The Foundations of Mathematics’”

Richard Zach

Hilbert’s 1928 article, to which this piece is an appendix, was presented in July 1927 to the Hamburg Mathematical Seminar. Hilbert had first introduced his program for the foundations of mathematics in the same venue in a series of talks in 1921 (Hilbert, 1922). In the 1927 talk, he presented a mature version of his program, including technical details of the axiomatization of mathematics based on the ε-calculus, Hilbert’s ε-substitution method, as well as a discussion of the finitary standpoint. Bernays’s appendix concerns the ε-substitution method.

The ε-calculus is a version of first-order logic which contains the ε-operator instead of quantifiers. Given a formula $\mathfrak{A}(a)$ with free variable a, the ε-operator can be used to form a term $\varepsilon_a \mathfrak{A}(a)$ (an “ε-functional”). Intuitively (if $\mathfrak{A}$ contains no free variables other than a), it provides a witness for $\mathfrak{A}(a)$, if one exists, and an arbitrary object otherwise. The ε-calculus is the quantifier-free first-order calculus with identity in a language including the ε-operator, and the additional *transfinite axiom*,

$$A(a) \to A(\varepsilon_a A(a)).$$

Note that in the version of the first-order calculus used by Hilbert and Bernays, A is a formula variable, and the rule of substitution (for object and formula variables) allows the derivation of any instance of the transfinite axiom. Such an instance, e.g.,

$$\mathfrak{A}(\mathfrak{a}) \to \mathfrak{A}(\varepsilon_a \mathfrak{A}(a)),$$

where $\mathfrak{a}$ is a term, is called a *critical formula*. Using the ε-operator and the transfinite axiom, quantifiers may be defined by

$$(a)A(a) \leftrightarrow A(\varepsilon_a \overline{A(a)}) \text{ and } (\exists a)A(a) \leftrightarrow A(\varepsilon_a A(a)).$$

The ε-substitution method is a method to remove ε-functionals from proofs in number-theoretic systems formulated in the ε-calculus. The basic idea is that in such a proof, ε-functionals can be replaced by concrete numerical terms; if the end-formula does not contain an ε-operator, this would result in a proof of the same formula which does not use the ε-operator altogether. If the ε-substitution method itself is a finitary operation, then this yields a finitary consistency proof: The formula $0 = 1$ does not contain the ε-operator, so if there were a proof of it in the full system, there would have to be a proof not using ε-operators, and this can be finitarily seen to be impossible.

The ε-substitution method was originally introduced by Hilbert (1923), but only outlined for the case where all critical formulas share the same formula $\mathfrak{A}(a)$. In that case, a two-step procedure suffices: First, replace $\varepsilon_a \mathfrak{A}(a)$ everywhere by 0. If, after this replacement, all critical formulas result in correct formulas (i.e., true equalities and inequalities between numerical terms), we are done. Otherwise, at least one critical formula $\mathfrak{A}(\mathfrak{a}) \to \mathfrak{A}(0)$ is incorrect, in which case $\mathfrak{A}(\mathfrak{a})$ must be correct. In this case, replacing $\varepsilon_a \mathfrak{A}(a)$ by $\mathfrak{a}$ results in a correct proof. However, the general case is significantly more complicated: In general, ε-functionals may be nested, and they may be nested in two ways. First, an ε-functional may be *embedded* (*eingelagert*) in another, in which case no outer ε binds a variable occurring in the embedded functional. Second, ε-functionals may be in *superposition* (*Überordnung*), in which case cross-binding of variables does occur, as in Bernays's example $\varepsilon_a \mathfrak{A}(a, \varepsilon_b \mathfrak{K}(a, b))$. These complications require a more sophisticated substitution procedure, in which a sequence of *total replacements* of numerical terms for the ε-functionals occurring in a proof is computed, which sequence ends in a correct total replacement. It is then also required to give a bound on the length of such a sequence.

Ackermann (1924) extended the ε-substitution method to cover these more complicated cases of embedding and superposition in proofs in the ε-calculus. Indeed, the system considered by Ackermann also contained second-order ε-operators ε_f, and so was a formalization of a fragment of analysis. However, it soon became clear that the procedure did not work in the generality envisaged by Ackermann. In a paper submitted for publication in 1925, von Neumann (1927) introduced some simplifications of Ackermann's original approach to the substitution method, which Ackermann incorporated into a new substitution procedure for first-order number theory based on the ε-calculus. Ackermann communicated this procedure to Bernays by letter only. It is this procedure which Bernays describes in his appendix to Hilbert's paper, and which is carried out in detail in Hilbert and Bernays (1939, §2). The new procedure, however, breaks down in the presence of the induction axiom, which in the ε-calculus takes the form

$$(\varepsilon_a A(a) = b') \to \overline{A}(b).$$

In particular, Bernays's claim in part 3 of his note is incorrect. This error was, however, not discovered until Gödel's second incompleteness theorem cast into doubt the possibility of a finitary consistency proof, and when von Neumann found a counterexample to the alleged bound on the length of the sequence of total replacements. It was not until after Gentzen (1936) provided a consistency proof for first-order arithmetic that Ackermann (1940) gave a correct ε-substitution procedure, which, like Gentzen's proof, used transfinite induction up to ε_0. For further details on these early consistency proofs, see Zach (2003, 2004).

References

Ackermann, Wilhelm. 1924. Begründung des "tertium non datur" mittels der Hilbertschen Theorie der Widerspruchsfreiheit. *Mathematische Annalen* 93: 1–36.

Ackermann, Wilhelm. 1940. Zur Widerspruchsfreiheit der Zahlentheorie. *Mathematische Annalen* 117: 162–194.

Ewald, William Bragg, ed. 1996. *From Kant to Hilbert. A Source Book in the Foundations of Mathematics*, vol. 2. Oxford and New York: Oxford University Press.

Gentzen, Gerhard. 1936. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen* 112: 493–565. English translation in (Gentzen, 1969, 132–213).

Gentzen, Gerhard. 1969. *The Collected Papers of Gerhard Gentzen.* Amsterdam: North-Holland.

Hilbert, David. 1922. Neubegründung der Mathematik: Erste Mitteilung. *Abhandlungen aus dem Seminar der Hamburgischen Universität* 1: 157–77. English translation in (Mancosu, 1998, 198–214) and (Ewald, 1996, 1115–1134).

Hilbert, David. 1923. Die logischen Grundlagen der Mathematik. *Mathematische Annalen* 88: 151–165. English translation in (Ewald, 1996, 1134–1148).

Hilbert, David. 1928. Die Grundlagen der Mathematik. *Abhandlungen aus dem Seminar der Hamburgischen Universität* 6: 65–85. English translation in (van Heijenoort, 1967, 464–479).

Hilbert, David and Paul Bernays. 1939. *Grundlagen der Mathematik*, vol. 2. Berlin: Springer.

Mancosu, Paolo, ed. 1998. *From Brouwer to Hilbert. The Debate on the Foundations of Mathematics in the 1920s.* Oxford and New York: Oxford University Press.

van Heijenoort, Jean, ed. 1967. *From Frege to Gödel. A Source Book in Mathematical Logic, 1897–1931.* Cambridge, Mass.: Harvard University Press.

von Neumann, Johann. 1927. Zur Hilbertschen Beweistheorie. *Mathematische Zeitschrift* 26: 1–46.

Zach, Richard. 2003. The practice of finitism. Epsilon calculus and consistency proofs in Hilbert's Program. *Synthese* 137: 211–259.

Zach, Richard. 2004. Hilbert's "Verunglückter Beweis," the first epsilon-theorem, and consistency proofs. *History and Philosophy of Logic* 25: 79–94.