Gentzen's original consistency proof and the Bar Theorem

W. W. Tait[*]

The story of Gentzen's original consistency proof for first-order number theory (Gentzen 1974),[1] as told by Paul Bernays (Gentzen 1974), (Bernays 1970), (Gödel 2003, Letter 69, pp. 76-79), is now familiar: Gentzen sent it off to *Mathematische Annalen* in August of 1935 and then withdrew it in December after receiving criticism and, in particular, the criticism that the proof used the Fan Theorem, a criticism that, as the references just cited seem to indicate, Bernays endorsed or initiated at the time but later rejected. That particular criticism is transparently false, but the argument of the paper remains nevertheless invalid from a constructive standpoint. In a letter to Bernays dated November 4, 1935, Gentzen protested this evaluation; but then, in another letter to him dated December 11, 1935, he admits that the "critical inference in my consistency proof is defective." The defect in question involves the application of proof by induction to certain trees, the 'reduction trees' for sequents (see below and §1), of which it is only given that they are well-founded. No doubt because of his desire to reason 'finitistically', Gentzen nowhere in his paper explicitly speaks of reduction trees, only of reduction rules that would generate such trees; but the requirement of well-foundedness, that every path taken in accordance with the rule terminates, of course makes implicit reference to the tree. Gentzen attempted to avoid the induction; but as he ultimately conceded, the attempt was unsatisfactory.

[1]The paper first appeared in print via an appendix to the translation of (Gentzen 1936) in (Gentzen 1969). A somewhat revised version of it is presented in (Bernays 1970) and the full text, together with an introduction by Bernays, in (Gentzen 1974).

Brouwer's Bar Theorem has generally been cited as what is needed to repair the argument.[2] The Bar Theorem does indeed suffice to close the gap between the well-foundedness of a reduction tree and proof by induction on it, but we will see that Brouwer's argument for the Bar Theorem in the context in question involves an argument for the claim that a proof of the well-foundedness of a reduction tree for a sequent can only be based on having the corresponding *deduction tree* of the sequent. The deduction tree in question is obtained by reading the reduction tree, which is constructed 'bottom-up', as 'top-down'. Deduction trees are built up inductively and so proof by induction on them is valid. Moreover, given a deduction tree for a sequent, the corresponding reduction tree can be constructed; but the converse is constructively invalid. So—and this is the main point of this paper—the gap in Gentzen's argument is filled, not by the Bar Theorem, but by taking as the basic notion that of a deduction tree in the first place rather than that of a reduction tree. These deduction trees are well-known objects, namely *cut-free deductions* in a formalization of first-order number theory in the sequent calculus with the ω-rule.

The formalization of number theory in the original paper as well as in the 1936 paper ultimately takes as the logical constants $\neg, \wedge$ and $\forall$. Deductions are of sequents of the form $\Gamma \vdash A$, where A is a formula and Γ a possibly null sequence of formulas. The rules of inference are the natural deduction rules: the introduction and elimination rules for the logical constants are only for the succedent formula (so that a deduction of the sequent $\Gamma \vdash A$ corresponds to a deduction of A in natural deduction whose assumption formulas are all in Γ). I will refer to this system as *the formal system of first-order number theory in natural deduction*. The precise details don't really concern us, since the non-trivial parts of Gentzen's argument do not really concern these natural deductions.

His consistency argument in the original version aims at showing that a natural deduction of a sequent $\Gamma \vdash A$ is a code for constructing a reduction tree for the sequent. Since there is no reduction tree for the sequent $\vdash 1 = 2$, for example, consistency is implied. As we noted and will see, reduction trees can be replaced everywhere in the argument by the corresponding deduction trees.

In the 1936 version of the consistency proof, the notion of a reduction tree

---

[2]Could Bernays' references to the 'Fan Theorem', all over thirty years later, have been the result of a confusion of the Fan Theorem with the Bar Theorem?

plays no role in the proof of consistency: that proof is obtained by means of the notion of a reduction procedure for *deductions* of sequents in the formal system of first-order number theory in natural deduction and an assignment of ordinals to these deductions such that each reduction step results in a decrease in ordinal. A reduction tree for a deduced sequent, along with an ordinal measure of the height of the subtrees, simply falls out of the proof. Of course, the assignment of ordinals to the subtrees allows the reduction tree to be identified with the corresponding deduction tree (since induction on the tree can be expressed by induction on the ordinals).

In the original paper, on the contrary, the notion of a reduction rule for a sequent plays an essential role: the non-trivial part of the argument—and the source of difficulty—is Gentzen's argument for his Lemma. As we noted, the corresponding deduction trees are cut-free deductions in the formal system of first-order number theory in the sequent calculus with the ω-rule; and the Lemma states (in terms of deduction trees) that cuts in that system can be eliminated. This was finally explicitly proved by Lorenzen (1951) and Schütte (1951). Lorenzen's proof applied to ramified analysis of finite order but does not supply ordinal bounds. Schütte's proof applies to a variant formalization of the sequent calculus and supplies the ordinal bounds. (Tait 1968) contains a unified treatment of Schüttes' result and his later papers on cut-elimination for ramified analysis (1952, 1964) with the ω-rule, using a simplified form of the sequent calculus.[3]

Although the notion of a reduction rule played no part in the consistency proof in (Gentzen 1936), it should be noted that it retained a conceptual/philosophical role. Gentzen not only wanted a proof of consistency, he wanted a way to understand the truth of a sentence of number theory that is in some sense 'finitary' but at the same time supported classical reasoning in number theory. In this respect, the original paper as well as the 1936 paper go beyond the original Hilbert program of finding finitary consistency proof for formal systems. Indeed, Gödel's incompleteness theorems would seem to demand such an extension. Consistency of a particular formal system is of less interest when we know that the system, if consistent, is also incomplete. A 'finitist' interpretation of classical mathematical propositions that guarantees their consistency transcends any particular formal system. Gentzen's

---

[3](Gentzen 1938) contains a version of the consistency proof for number theory framed in the sequent calculus which could quite easily be transformed into a proof of cut-elimination for the sequent calculus with the ω-rule. See (Buchholz 1997) for a detailed description of the relation between (Gentzen 1938) and Schütte's 1951 result.

candidate for such an interpretation was this: A is true precisely if we can state a 'reduction rule' for $\vdash A$, i.e. a rule for constructing a reduction tree.[4]

Like Gödel, Gentzen had discovered the double negation interpretation of classical first-order number theory in the corresponding intuitionistic version. If from a 'finitist' point of view one were satisfied with the intuitionistic system, i.e. Heyting arithmetic, this result would provide the desired interpretation—and would certainly diminish the significance of a consistency proof for a finitist. But, also like Gödel in (1938a), Gentzen rejected the intuitionistic conception of logic as presented by Heyting as non-finitist. The difficulties they had with it centered on the intuitionistic meaning of implication. The 'circularity' that Gentzen found in Heyting's account of the meaning of $\rightarrow$ in propositions $A \rightarrow B$, where A contains $\rightarrow$ (1936, §11.1), disappears when one adopts the type-theoretic approach of Curry-Howard. But, of course, when A contains $\rightarrow$, its proofs are no longer to be understood as concrete finitary objects on the type-theoretic conception; rather they themselves are already objects of higher type and so it would be a stretch to regard proofs of $A \rightarrow B$, i.e. operations transforming proofs of A into proofs of B, as in any sense 'finitist'.

Gentzen's interpretation of the sentences of arithmetic in terms of reduction rules has a somewhat alien flavor. But it evolved into two different and more homely interpretations of classical reasoning in number theory: the *no-counterexample interpretation* in the hands of Gödel (1938a)[5] and Kreisel (1951, 1952) and a game-theoretic interpretation by Coquand (1995), according to which a reduction rule is a winning strategy in a certain 2-person game. These two interpretations are discussed in (Tait 2001, Tait 2005a) and I will not discuss them here.

1. Reduction Rules. A sequent is of the form $\Gamma \vdash A$, where Γ is a set of formulas and A is a formula. Gentzen defines the notion of a reduction rule for sequents of arbitrary formulas. If free variables occur in formulas in the sequent, the reductions of the sequent begin with replacing one of

---

[4]The rule is, in itself, just a rule for constructing a certain tree—I call it a 'pre-reduction rule' below. It is a reduction rule in virtue of the tree being well-founded. So, to know that A is true would mean, not simply to possess the rule, but also to know that the tree is well-founded. One might have some difficulty in labeling knowledge of this kind as 'finitary'.

[5]That Gödel had anticipated the no-counterexample interpretation in these notes was first noticed by C. Parsons and W. Sieg in their introductory note.

them throughout the sequent by an arbitrarily chosen numeral. If the sequent consists just of sentences and some formula in it contains a closed term $f(t_1, \ldots, t_n)$, then a reduction consists in replacing such a term by the numeral $\bar{k}$ for its value. We can eliminate these reduction steps by considering only sequents of sentences and by identifying the sentence $A(f(t_1, \ldots, t_n))$ with $A(\bar{k})$. Since the reductions of sequents of sentences yields only sequents of sentences, we can also cut down on the number of forms of sentences that must be treated separately by treating $A \wedge B$ and $\forall x A(x)$ as special cases of *conjunctive* sentences $\bigwedge_i A_i$: namely

$$A_0 \wedge A_1 = \bigwedge_{i<2} A_i \qquad \forall x A(x) = \bigwedge_{i<\omega} A(\bar{i}).$$

Concerning the atomic sentences, Gentzen included only decidable sentences, i.e. built up from $\bar{0}$, the successor function constant, constants for other computable functions and decidable relation constants.[6] We will write

$$\perp$$

to denote any false atomic sentence, such as $1 = 2$ (Gentzen's favorite): they are interchangeable.

A sequent $\Gamma \vdash A$ of sentences is called an *axiom sequent* just in case either A is a true atomic sentence or it is a false atomic sentence and Γ contains a false atomic sentence. The rules of inference we will consider are $\bigwedge -R$

$$\frac{\cdots \qquad \Gamma \vdash A_j \qquad \cdots}{\Gamma \vdash \bigwedge_i A_i} \quad (\text{all } j)$$

and $\bigwedge -L$

$$\frac{\Gamma, \bigwedge_i A_i, A_j \vdash \perp}{\Gamma, \bigwedge_i A_i \vdash \perp}$$

for $\bigwedge$; and $\neg - R$

---

[6] When in (Gentzen 1943) he comes to the problem of determining the bound on the provable ordinals, he needs to essentially redo the argument for the case that the atomic formulas also include $t \in V$, where V stands for an indeterminate set of numbers and t a numerical term. But it is obvious how to treat this extension. The atomic sentences must be extended to include the expressions $n \in V$ and the *axiom sets*, defined below, has to be extended to include sequents of the form $\Gamma, n \in V \vdash n \in V$.

$$\frac{\Gamma, A \vdash \bot}{\Gamma \vdash \neg A}$$

and $\neg - L$

$$\frac{\Gamma, \neg A \vdash A}{\Gamma, \neg A \vdash \bot}$$

for $\neg$. The explicitly listed composite sentence in the conclusion of an inference is called its *principal sentence.*

Definition A *pre-reduction rule* R for a sequent $\Gamma \vdash A$ of sentences effectively determines, for each n, the R-admissible sequences $\langle \Gamma_0 \vdash A_0, \ldots, \Gamma_n \vdash A_n \rangle$ of sequents of sentences as follows:

- $\langle \Gamma \vdash A \rangle$ is the only R-admissible sequence of length 1.

- All R-admissible sequences of length $n + 2$ are one-element extensions of R-admissible sequences of length $n + 1$. Let $\langle \Gamma \vdash A, \ldots, \Delta \vdash B \rangle$ be R-admissible. We specify its R-admissible one-element extensions.

    - If $\Delta \vdash B$ is an axiom sequent, then there are no R-admissible extensions.
    - If Δ consists only of true atomic sentences and B is a false atomic sentence, then $\langle \Gamma \vdash A, \ldots, \Delta \vdash B, \Delta \vdash B \rangle$ is its only R-admissible one-element extension.
    - Otherwise R determines an inference with conclusion $\Delta \vdash B$ and $\langle \Gamma \vdash A, \ldots, \Delta \vdash B, \Theta \vdash C \rangle$ is R-admissible for every premise $\Theta \vdash C$ of that inference.

A pre-reduction rule R for $\Gamma \vdash A$ is a *reduction rule* for $\Gamma \vdash A$ iff every ω-sequence $\langle \Gamma \vdash A, \Delta \vdash B, \Theta \vdash C, \ldots \rangle$ of sequents of sentences contains a finite initial segment that is not R-admissible. $\square$

Our definition of a sequent $\Gamma \vdash A$ differs from Gentzen's in that, for him, Γ is a *sequence* of sentences rather than a set. But that makes no difference in the definition of a reduction rule. Notice that the inferences specified above have the property that the antecedent of the conclusion is a subset of the antecedent of each premise. Gentzen also allows another form of both $\forall - L$ and $\neg - L$, namely

$$\frac{\Gamma, A_j \vdash \bot}{\Gamma, \bigwedge_i A_i \vdash \bot}$$

and

$$\frac{\Gamma, \vdash A}{\Gamma, \neg A \vdash \bot}$$

But, since adding a sentence to the antecedent of each premise and the conclusion of any inference in our sense is again an inference, it is clear that a reduction rule for $\Gamma \vdash A$ in the wider sense of Gentzen can be transformed into one in our sense.

Let R be a reduction rule for $\Gamma \vdash A$, where $\Gamma \vdash A$ is not an axiom sequent. Then there is a unique inference

$$\frac{\ldots \qquad \Gamma_i \vdash A_i \qquad \ldots}{\Gamma \vdash A}$$

such that the R-admissible sequences of length 2 are precisely $\langle \Gamma \vdash A, \Gamma_i \vdash A_i \rangle$ of length 2. The principal sentence of this inference is called the *principal sentence of R* and the sequents $\Gamma_i \vdash A_i$ are called the *reducts* of $\Gamma \vdash A$ determined by R.

If R is a pre-reduction rule for $\Gamma \vdash A$, the *reduction tree* $\mathcal{T}_R$ for Γ determined by R has as its nodes the R-admissible sequences, where, for nodes μ and ν, $\mu <_{\mathcal{T}_R} \nu$ means that μ is a proper initial segment of the R-admissible sequence ν. The root of the tree is of course the one-element sequence $\langle \Gamma \vdash A \rangle$. The condition that the pre-reduction rule R be a reduction rule is precisely that $\mathcal{T}_R$ be a well-founded tree.

Let $\langle \Gamma \vdash A, \ldots, \Delta \vdash B \rangle$ be an R-admissible sequence. R determines a reduction rule R' for $\Delta \vdash B$: the R'-admissible sequences are just those $\langle \Delta \vdash B \ldots, \Theta \vdash C \rangle$ such that $\langle \Gamma \vdash A, \ldots, \Delta \vdash B \ldots, \Theta \vdash C \rangle$ is R-admissible. We say in this case that R' is a *reduction sub-rule* of R. Note that $\mathcal{T}_{R'}$ is isomorphic to the subtree $\mathcal{T}_{R,\mu} = \{\nu | \mu \leq_{\mathcal{T}_R} \nu\}$, where $\mu = \langle \Gamma \vdash A, \ldots, \Delta \vdash B \rangle$. We will sometimes confuse the two.

Given the reduction rule R for $\Gamma \vdash A$ and a set Θ of sentences, we may obtain a reduction rule R' for $\Gamma \cup \Theta \vdash A$: the R' admissible sequences are the sequences $\langle \Gamma \cup \Theta \vdash A, \ldots, \Delta \cup \Theta \vdash B \rangle$ such that $\langle \Gamma \vdash A, \ldots, \Delta \vdash B \rangle$

is R-admissible. $\mathcal{T}_{R'}$ is of course isomorphic to $\mathcal{T}_R$ and we will sometimes confuse them.

Let R be a reduction rule for the sequent $\Gamma \vdash A$ of sentences. To each node $\langle \Gamma \vdash A, \ldots, \Delta \vdash B \rangle$ of $\mathcal{T}_R$ assign the set Δ. Then the connection between the assignments to successive nodes is given by the rules of inference. Note that when $\bigwedge_i A_i = \forall x A(x)$ the rule for conjunctions is the ω-rule. Since $\mathcal{T}_R$ is well-founded, every path upward through it terminates in an axiom set. Moreover, the root of $\mathcal{T}_R$, the bottom-most node $\langle \Gamma \vdash A \rangle$, is assigned $\Gamma \vdash A$. Thus, from a *classical* point of view, $\mathcal{T}_R$ is a cut-free deduction of $\Gamma \vdash A$ in a sequent calculus formalization of first-order number theory with the ω-rule.

2. Induction on Trees. From a constructive standpoint the situation is more complicated. Deductions are top-down, starting with axiom sequents and passing from premises to conclusion and finally to the sequent deduced. In this form, proof by induction on the deduction tree (or on its height) is fully justified. But reduction trees for $\Gamma \vdash A$ are built bottom-up, starting with $\Gamma \vdash A$ and passing up from conclusion to premises and finally to axiom sets. They are to be well-founded, but that does not constructively justify the principle of induction applied to them. That is exactly the problem that Gentzen failed to avoid in the original consistency proof.

We say that a property P defined on the nodes of a connected tree $\mathcal{T}$ is $\mathcal{T}$-*inductive* if $P(\nu)$ for every ν immediately above μ implies $P(\mu)$. The *principle of induction* on $\mathcal{T}$ states that every $\mathcal{T}$-inductive property defined on its nodes holds for all of its nodes.

If a tree $\mathcal{T}$ satisfies the principle of induction, then it is well-founded. (For the property $P(\mu)$ of a node μ of $\mathcal{T}$ that the subtree $\mathcal{T}_\mu = \{\nu \in \mathcal{T} \mid \mu \leq_\mathcal{T} \nu\}$ is well-founded is an inductive property.) Classically, we can easily infer from the well-foundedness of a tree $\mathcal{T}$ that it satisfies the principle of induction. If not, choose an inductive property P which is not possessed by every node of $\mathcal{T}$ and, having defined $\mu_0 < \ldots < \mu_n$ where the μ_i do not have the property P, the inductiveness of P implies that there is a $\mu_{n+1} > \mu_n$ which also fails to have P. Iterating this construction, we obtain an infinite path of nodes up through the tree that do not have the property P. Constructively, however, this argument of course fails: from the fact that μ_n does not have P it follows from the inductiveness of P that it is not the case that all successor nodes μ of μ_n have the property P; but that does not imply that there exists such a successor node which does not have P.

Call a tree *inductive* if it is in the least class $\mathcal{I}$ of connected trees such

that whenever $0 \leq \alpha \leq \omega$ and $\mathcal{T}_i \in \mathcal{I}$ for all $i < \alpha$, then the connected tree with whose immediate subtrees are precisely the $\mathcal{T}_i$ is inductive. Inductive trees obviously satisfy the principle of induction.

One example is the constructive ordinals of the second number class. Let $\mathcal{T}_0$ and $\mathcal{T}_1$ be connected trees. $\mathcal{T}_0 \subseteq \mathcal{T}_1$ means that $\mathcal{T}_0$ is a substructure of $\mathcal{T}_1$, i.e. that each node of $\mathcal{T}_0$ is a node of $\mathcal{T}_1$ and that $\mu <_{\mathcal{T}_0} \nu$ implies $\mu <_{\mathcal{T}_1} \nu$. $\mathcal{T}_0 \prec \mathcal{T}_1$ means that $\mathcal{T}_0 \subseteq (\mathcal{T}_1)_\mu$ for some node μ in $\mathcal{T}_1$ other than its root. 0 is the one-node tree with node $\emptyset$, $\alpha + 1$ is the connected tree with root $\{\alpha\}$ and whose only immediate subtree is α, and when $\alpha_n \prec \alpha_{n+1}$ for each n, $\lim_n \alpha_n$ is the connected tree with root $\{\alpha_n \mid n < \omega\}$ and whose immediate subtrees are the distinct α_n. $\alpha + \beta$ is defined as usual by induction on β:

$$\alpha + 0 = \alpha \qquad \alpha + (\beta + 1) = (\alpha + \beta) + 1 \qquad \alpha + \lim_n \beta_n = \lim_n (\alpha + \beta_n).$$

We can always assign ordinal bounds $|\mathcal{T}|$ on the height of an inductive tree $\mathcal{T}$, with $|\mathcal{T}'| < |\mathcal{T}|$ when $\mathcal{T}'$ is a proper subtree of $\mathcal{T}$: let $\mathcal{T}_i$ for $i < \alpha$ be its immediate subtrees and assume $|\mathcal{T}_i|$ is defined for each $i < \alpha$. If $\alpha < \omega$ set $|\mathcal{T}| = |\mathcal{T}_0 + \cdots + |\mathcal{T}_{\alpha-1}| + 1$. Otherwise, $|\mathcal{T}| = \lim_n(|\mathcal{T}_0| + \cdots + |\mathcal{T}_n| + 1)$. Conversely, let $\mathcal{T}$ be any tree and suppose that we have assigned an ordinal $|\mathcal{T}_\mu|$ to each subtree $\mathcal{T}_\mu$ of $\mathcal{T}$ so that $|\mathcal{T}_\mu| < |\mathcal{T}_\nu|$ when $\mu <_{\mathcal{T}} \nu$. Then $\mathcal{T}$ satisfies the principle of induction, since it can be reduced to the principle of induction on the ordinals.

The immediately relevant example of inductive trees is given by what we are calling the deduction trees. If $\Gamma \vdash A$ is an axiom sequent, then a one-node tree with node $\Gamma \vdash A$ is a deduction tree. If $\Gamma \vdash A$ is the conclusion of an inference and $\mathcal{D}_i$ is a deduction of the ith premise, $i < \alpha$, then the tree with root $\{\mathcal{D}_i \mid i < \alpha\}$ and immediate subtrees $\mathcal{D}_i$ $(i < \alpha)$ is a deduction of $\Gamma \vdash A$.[7]

Corresponding to the notion of a reduction rule, we also have the notion of a *deduction rule* for $\Gamma \vdash A$. Such a rule R determines a deduction tree $\mathcal{D}_R$ for $\Gamma \vdash A$ as follows: if $\Gamma \vdash A$ is an axiom sequent then $\mathcal{D}_R$ is the one-node tree with node $\Gamma \vdash A$. Otherwise, R determines an inference with the conclusion $\Gamma \vdash A$ and a deduction rule R_i for each premise $\Gamma_i \vdash A_i$ of the inference. $\mathcal{D}_R$ is a tree with root $\{\mathcal{D}_{R_i} \mid i < \alpha\}$ and immediate subtrees $\mathcal{D}_{R_i}$.

---

[7]It is not excluded that some premise of an inference is identical with the conclusion. Therefore, we have to distinguish the node of a deduction tree from the sequent attached to it. For simplicity, I have defined the nodes of deduction trees to be in general infinitary objects. The finitist will want to replace these with suitable codes.

A deduction rule R for $\Gamma \vdash A$ determines a reduction rule R' for $\Gamma \vdash A$ as follows. The construction is by induction on $\mathcal{D}_R$. If $\Gamma \vdash A$ is an axiom set, $\langle \Gamma \vdash A \rangle$ is R'-admissible. Otherwise, R' determines an inference with conclusion $\Gamma \vdash A$ and a deduction rule R' for each premise $\Gamma_i \vdash A_i$. The R'-admissible sequences are of the form $\langle \Gamma \vdash A, \Gamma_i \vdash A_i, \ldots, \Delta \vdash B \rangle$ such that $\langle \Gamma_i \vdash A_i, \ldots, \Delta \vdash B \rangle$ is R'_i-admissible. Notice that the converse construction, of a deduction rule R from a reduction rule R', is not constructively valid, since induction on $\mathcal{T}_{R'}$ is not valid.

While we are on the subject of induction on trees, let me mention a construction to which we will refer below. Suppose $\mathcal{T}_0$ and $\mathcal{T}_1$ are trees. We define $\mathcal{T} = \mathcal{T}_0 \times \mathcal{T}_1$ as follows: its nodes are $\mu = (\mu_0, \mu_1)$, where μ_e is a node of $\mathcal{T}_e$.

$$\mu <_{\mathcal{T}} \nu \iff [\mu_0 <_{\mathcal{T}_0} \nu_0 \ \& \ \mu_1 \leq_{\mathcal{T}_1} \nu_1] \ or \ [\mu_0 \leq_{\mathcal{T}_0} \nu_0 \ \& \ \mu_1 <_{\mathcal{T}_1} \nu_1.]$$

Observe that, if both $\mathcal{T}_0$ and $\mathcal{T}_1$ satisfy the principle of induction, then so does $\mathcal{T}_0 \times \mathcal{T}_1$.

3. Gentzen's Lemma. The only non-trivial step in Gentzen's demonstration that a reduction rule for a sequent can be extracted from a natural deduction of it is the proof of the following:

Lemma (Gentzen 1974, §14.44, §14.6). If there are reduction rules R_0 for $\Gamma, D \vdash C$ and R_1 for $\Gamma \vdash D$, then there is a reduction rule for $\Gamma \vdash C$.

The proof of the Lemma is familiar if we think of deductions rather than reductions. It proceeds by induction on the *rank* $|D|$ of D, where $|A| = 0$ when A is atomic and

$$|\neg A| = |A| + 1 \qquad |\bigwedge_i A_i| = \sup_i(|A_i| + 1).$$

The inductive assumption is that the Lemma holds for all B with $|B| < |A|$, and then we want to conclude from this that it holds also for A. The proof of this involves an induction within the induction on the rank of A; namely an induction on the tree $\mathcal{T}_{R_0}$. In the original paper, Gentzen tried to avoid this induction, and so I will put each case of its application below in square brackets.

Case 1. $\Gamma, D \vdash C$ is an axiom sequent. Then C is atomic. If it is true, then $\Gamma \vdash C$ is an axiom sequent. So assume C is false. If D is also a false atomic sentence, then the reduction rule R_1' for Γ is obtained replacing D by C in R_1.

Case 2. $\Gamma, D \vdash C$ is not an axiom sequent.

Case 2a. D is not the principal sentence of R_0. Then the reducts of $\Gamma, D \vdash C$ determined by R_0 have the form $\Gamma_i, D \vdash C_i$ with the proper sub-reduction rule R_0^i. [By induction on $\mathcal{T}_{R_0}$, we may assume that there is a reduction rule R^i for $\Gamma_i, \Gamma \vdash C_i$. The R-reduction tree for $\Gamma \vdash C$ has as its immediate sub-trees the $\mathcal{T}_{R^i}$.]

Case 2b. D is the principal sentence of R_0. Thus, $C = \perp$.

Case 2bi. $D = \neg E$. Then the unique reduct is $\Gamma, D \vdash E$. [By the induction hypothesis, we may assume that there is a reduction rule R' for $\Gamma \vdash E$.] R_1 reduces $\Gamma \vdash D$ to $\Gamma, E \vdash C$ with immediate sub-rule R_1'. $|E| < |D|$ and so by induction on $|D|$ applied to R' and R_1', there is a reduction rule R for $\Gamma \vdash C$.

Case 2bii. $D = \bigwedge_i D_i$ and R_0 reduces $\Gamma, D \vdash C$ to $\Gamma, D, D_j \vdash C$ with corresponding sub-reduction rule R_0'. [By induction on $\mathcal{T}_{R_0}$, we may assume that there is a reduction rule R' for $\Gamma, D_j \vdash C$.] The reducts of $\Gamma \vdash D$ determined by R_1 are the $\Gamma \vdash D_i$ for each i, with sub-reduction rules R_1^i. Since $|D_j| < |D|$, the induction hypothesis on rank applied to R' and R_1^j yields R. $\square$

As we indicated, the square-bracketed parts of the argument, explicitly invoking induction on $\mathcal{T}_{R_0}$, do not appear in Gentzen's original paper. His argument rather is as follows: with the rule R_1 for $\Gamma \vdash D$ fixed, we reduce the problem of finding a reduction rule for $\Gamma \vdash C$ to that of finding one for a reduct $\Gamma' \vdash C'$ of $\Gamma, D \vdash C$ as determined by R_0. If $\Gamma', D \vdash C'$ is not an axiom sequent, then we reduce this problem of finding a reduction rule for $\Gamma'' \vdash C''$, where $\Gamma'', D \vdash C''$ is a reduct of $\Gamma', D \vdash C'$ as determined by R_0—and so on:

> Continuing in this way, we must *reach the end in finitely many steps*, i.e. the completion of the proof. (Gentzen 1969, §14.63)

This may sound convincing as a constructive argument until we ask: how many steps? Of course there is no answer to this because it depends upon which path we take. In particular, if we reach stage n and $C^{(n)}$ in $\Gamma^{(n)}, D \vdash C^{(n)}$ is of the form $\bigwedge_i C_i$, then the $n + 1$st stage $\Gamma^{(n)}, D \vdash C_j$ depends on the 'free choice' of j. In the November 4, 1935 letter to Bernays, Gentzen seems to have been arguing that, because the choice of j is *free*, we are really thinking about a generic path $\langle \Gamma, D \vdash C, \Gamma', D \vdash C', \ldots \rangle$ through $\mathcal{T}_{R_0}$ which therefore presumably has a generic finite length x. $\Gamma^{(x)} \vdash C^{(x)}$ has a reduction rule and so, working backward, so has $\Gamma \vdash C$. But as we noted in the inroductory remarks, Gentzen soon gave up on this argument.

4. The Bar Theorem. Gentzen's reference to 'free choices' seems to point to Brouwer's function theory; but Gentzen seems to have ignored the one feature of Brouwer's theory that would ground his argument: the Bar Theorem. Whether or not he explicitly rejected it, he certainly did not employ it in his argument. The setting for the Bar Theorem is the notion of a *spread law*.

Definition A *spread law* S effectively determines, for each n, the S-admissible sequences $\langle a_0, \ldots a_n \rangle$ of elements of a decidable set M as follows:

- S determines for which $a \in M$ the one-element sequence $\langle a \rangle$ is S-admissible.

- All S-admissible sequences of length $n + 2$ are one-element extensions of S-admissible sequences of length $n + 1$. Given the S-admissible sequence $\langle a_0, \ldots, a_n \rangle$, S determines for which $a \in M$ $\langle a_0, \ldots, a_n, a \rangle$ is S-admissible.

Moreover, it is required that every S-admissible sequence have a proper extension. We will consider only *connected* spread laws S, i.e. such that there is exactly one one-element S-admissible sequence $\langle a_0 \rangle$, called its *root*. $\square$

If S satisfies all the conditions of being a spread law except the condition that every S-admissible sequence have an S-admissible extension, we can turn it into a spread law $S^{\#}$ by the condition that every S-admissible sequence is $S^{\#}$-admissible and, if $\langle a, \ldots, b \rangle$ is either maximal S-admissible or contains $\#$, then $\langle a, \ldots, b, \# \rangle$ is $S^{\#}$-admissible.

Let S be a spread law. An *S-sequence* is an ω-sequence $\langle a, b, \ldots \rangle$ such that each finite initial segment is S-admissible. $[S]$ denotes the set of all S-sequences. A *bar* on S is a set B of S-admissible sequences such each $\theta \in [S]$ has an initial segment in B. When B is decidable, we can assume that the initial segment is unique. $\mathcal{T}_S$ the tree of S-admissible sequences.

Bar Theorem. If

   i. B is a decidable bar on the connected spread S with root $\langle a_0 \rangle$,

   ii. Every element of B has the property P,

   iii. P is inductive on $\mathcal{T}_S$,

then $P(\langle a_0 \rangle)$. $\square$

We can apply the Bar Theorem to validate the induction on $\mathcal{T}_{R_0}$ in the proof of Gentzen's Lemma. The spread law S in this case is $R_0^{\#}$. So we have a connected spread with root $\Gamma, D \vdash C$. The assertion that R_0 is a reduction rule for $\Gamma, D \vdash C$ and in particular is well-founded implies that the set $[R_0]$ of maximal R_0-admissible sequences is a decidable bar on S. (Here we are using the fgact that a reduction rule is to be effective.) Let P be the property of S-admissible sequences σ that, if σ is the R_0-admissible sequence $\langle \Gamma, D \vdash C, \ldots \Gamma', D \vdash C' \rangle$, then there is a reduction rule for $\Gamma', \Gamma \vdash C'$. Every element $\langle \Gamma \vdash C, \ldots \Gamma' \vdash C' \rangle$ of the bar $[R_0]$ has the property P, since $\Gamma' D \vdash C'$ is an axiom sequent. (If C' is a true atomic sentence or Γ' contains a false atomic sentence other than D, $\Gamma', \Gamma \vdash C'$ is an axiom set. If C' and D are both false atomic, a reduction rule for $\Gamma', \Gamma \vdash C'$ is easily obtained from the reduction rule R_1 for $\Gamma \vdash D$.) Gentzen proved that P is inductive on $\mathcal{T}_S$. So by the Bar Theorem, $\langle \Gamma, D \vdash C \rangle$ has the property P, i.e. $\Gamma \vdash C$ has a reduction rule.

But a cynic might wonder at Brouwer's magic: simply by calling $[R_0^{\#}]$ a 'spread' he could conclude from the fact that $\mathcal{T}_{R_0}$ is well-founded that it satisfies the induction principle. But lets look at Brouwer's argument for the Bar Theorem. It begins with the doctrine that a proof of such an implication consists in a method of transforming a proof of the antecedent [conditions i)-iii)] into a proof of the conclusion. (This simply reflects the intuitionistic meaning of implication.) Now consider the condition $i)$. We say that B *bars* the S-admissible sequence σ if each $\theta \in [S]$ of which σ is an initial segment

has an initial segment in B. So condition i) states that B bars $\langle a_0 \rangle$. Brouwer argues—and this is the crucial step—that ultimately the only way to prove that B bars $\langle a_0 \rangle$ is by using the axioms

$$B \ bars \ \sigma$$

for each $\sigma \in B$ and the inferences:

$$\frac{\cdots \qquad B \text{ bars } \sigma_i \qquad \cdots}{B \ bars \ \sigma} \quad (\text{all } i)$$

where the σ_i are all the S-admissible one-element extensions of σ. Thus, to have a proof that B is a bar on S is to have an (in general infinitary) deduction tree $\mathcal{D}$ using just these axioms and inferences. So then, using the conditions ii) and iii) of the Bar Theorem, the proof of its conclusion is obtained by replacing the property 'B bars x' by the property $P(x)$ in the deduction. This is proved by induction on $\mathcal{D}$. This is permissible because deductions, unlike reduction trees, satisfy the principle of induction.

Of course, in our case we concluded i), i.e. that $[R_0]$ is a bar on $S = R_0^{\#}$, not by such a deduction of the statement that $[R_0]$ bars $\langle \Gamma, D \vdash C \rangle$, but from the fact that $\mathcal{T}_{R_0}$ is well-founded. But, whether or not one finds the crucial step in Brouwer's argument for the Bar Theorem convincing in general, *its application in this instance requires that the proof of well-foundedness of $\mathcal{T}_{R_0}$ ultimately be a deduction $\mathcal{D}$ in the above sense that $[R_0]$ is a bar on $R_0^{\#}$.* This is what is presupposed by an application of the Bar Theorem to the proof of Gentzen's Lemma.

But now let $\mathcal{D}'$ be obtained by replacing

$$[R_0] \ bars \ \langle \Gamma, D \vdash C, \dots, \Gamma', D \vdash C' \rangle$$

throughout $\mathcal{D}$ by

$$\Gamma', D \vdash C'$$

Then $\mathcal{D}'$ is a deduction tree for Γ, DC. Indeed, it is just the R_0-reduction

tree $\mathcal{T}_{R_0}$ read top-down. So *the application of the Bar Theorem already presupposes a deduction tree for $\Gamma, D \vdash C$.* Thus:

It was not the Bar Theorem that Gentzen needed; it was the switch from the basic notion of a reduction tree to that of a deduction tree. Reading the above

14

proof of the Lemma with R_0 and R_1 understood as rules for deduction rather than of reduction, the square bracketed inductions are valid.

5. Cut-elimination. Staying within the domain of deduction trees then, rather than reduction trees, Gentzen's Lemma is constructively valid. If we add to the rules of inference given in §1 the *cut rule*

$$\frac{\Delta, D \vdash C \qquad \Delta \vdash D}{\Delta \vdash C}$$

then we have one formalization of the rules of inference for first-order number theory with the ω-rule. D is called the *cut-formula* of this cut. Call deductions in this system *deduction trees with cuts.* By Gentzen's Lemma, every deduction tree with cuts can be reduced to one without cuts. Simply iterate the operation of eliminating the top-most cuts.

In terms of ordinal bounds on the height of the trees, this is not the most efficient way to eliminate cuts. The more efficient method is essentially the transfinite version of Gentzen's *Hauptsatz* for first-order logic in the sequent calculus. The *cut-degree* of a deduction is the least ordinal greater than the rank $|D|$ of D for all cut-formulas D in it (where $|A| = 0$ for atomic A, $|\neg A| = |A| + 1$, and $|\bigwedge_i A_i| = lub_i(|A_i| + 1)$). If α is the height of the given deduction of $\Gamma \vdash A$ and its cut-degree is some $m < \omega$, the bound we get on the height of the cut-free deduction is 2_m^α, where $2_0^\alpha = \alpha$ and $2_{n+1}^\alpha = 2^{2_n^\alpha}$. The efficient proof proceeds by eliminating cuts of maximum rank $m + 1$, replacing them with cuts of maximum rank m at the cost of increasing the height of the deduction from α to 2^α. (If the cut-degree is ω, then the bound is ϵ_α.) The bound we get from Gentzen's Lemma is much higher.

6. The Sequent Calculus and the Set Calculus. If we were to admit an arbitrary sentence B in the rule of $\bigwedge - L$

$$\frac{\Gamma, \bigwedge_i A_i, A_j \vdash B}{\Gamma, \bigwedge_i A_i \vdash B} \quad (\text{some } j)$$

instead of restricting it to $B = \perp$, the rules of inference would be the natural cut-free rules for first-order number theory with the ω-rule with the logical constants $\forall, \wedge, \neg$. This would of course imply a greater freedom for a reduction rule R for $\Gamma \vdash A$. Namely the reducts of $\Delta \vdash B$ could be of the form $\Delta' \vdash B$ even when B is composite, and not just of the form $\Delta \vdash B'$. So, faced with an admissible $\langle \Gamma \vdash A, \ldots, \Delta \vdash B \rangle$, where B is composite, R could

choose the one-element extension $\langle \Gamma \vdash A, \ldots, \Delta \vdash B, \Delta' \vdash B \rangle$ rather than being restricted to one of the form $\langle \Gamma \vdash A, \ldots, \Delta \vdash B, \Delta \vdash B' \rangle$. By induction on the sentence B, it is easy to see that for all B, if there is a reduction rule for $\Gamma, \bigwedge_i A_i, A_j \vdash B$, then there is one for $\Gamma, \bigwedge_i, A_i \vdash B$. So the general case of the inference is derivable from the special case.

But why did Gentzen restrict reductions to the case $B = \perp$? The answer is that it leads to a simpler proof of his main lemma—or more accurately, in view of the gap in his argument, it seems fair to answer rather that it made it easier for him to convince himself that there was no gap. Let $D = \bigwedge_i D_i$, R_0 be a reduction rule for $\Gamma, D \vdash C$ and R_1 a reduction rule for $\Delta \vdash D$ and suppose that R_0 chooses the reduct $\Gamma, D, D_j \vdash C$. By induction (which we have now justified by replacing the reduction trees by the corresponding deduction trees) we obtain a reduction rule for $\Gamma, D_j, \Delta \vdash C$. Now, on Gentzen's more restricted notion of reduction, the only possible reductions of $\Delta \vdash D$ are all the sequents $\Delta \vdash D_i$, including the case of $\Delta \vdash D_j$. So now the cut is reduced to the simpler cut-formula D_j. But with the more liberal notion of a reduction, the reduct of $\Delta \vdash D$ that R_1 chooses might be of the form $\Delta' \vdash D$. Clearly, in this case, the proof that cuts can be eliminated involves more symmetry between R_0, the deduction of $\Gamma, D \vdash C$ and R_1, the deduction of $\Delta \vdash D$. In fact it requires induction, not on $\mathcal{T}_{R_0}$, but on $\mathcal{T}_{R_0} \times \mathcal{T}_{R_1}$. Gentzen's attempt to avoid/disguise the induction would certainly have been even less plausible in this case.

The symmetry that is revealed by taking the more general form of $\bigwedge$-elimination is made even more evident with two changes in the formalization. One is a change in the logical constants, replacing negation $\neg$ by disjunction $\vee$ and existential quantifier $\exists$, except we will admit negation of atomic formulas. Classically at least, if A is atomic, the choice between A and $\neg A$ as more basic is arbitrary. We will refer to them both as *prime sentences*. Similarly, there is no ground for treating $\neg \bigwedge_i A_i$ as more logically complex than $\bigwedge_i A_i$: for it is expressed by $\bigvee_i \neg A_i$, where

$$A_0 \vee A_1 := \bigvee_{i<2} A_i \qquad \exists x A(x) := \bigvee_{i<\omega} A(\bar{i}).$$

So henceforth, sentences will be built up from prime sentences by means of $\bigvee$ and $\bigwedge$. The *complement* $\bar{A}$ of a sentence A is defined by

$$\bar{A} := \neg A \qquad \overline{\neg A} := A$$

16

if A is atomic and

$$\overline{\bigvee_i A_i} := \bigwedge_i \overline{A_i} \qquad \overline{\bigwedge_i A_i} := \bigvee_i \overline{A_i}$$

So negation is no longer a logical constant: it is only used to arbitrarily mark one of two complementary prime sentences.

The other change is this: there is no reason in the classical sequence calculus to restrict sequents to one succedent; but moreover, there is no reason to retain sequents at all. The sequent $A_1, \ldots, A_m \vdash B_1, \ldots, B_n$ has the same classical meaning as $\vdash \overline{A_1}, \ldots, \overline{A_m}, B_1, \ldots, B_n$, and so we may as well just take as the units of deduction the corresponding sets $\{\overline{A_1}, \ldots, \overline{A_m}, B_1, \ldots, B_n\}$ the set $\{\overline{A_1}, \ldots, \overline{A_m}, B_1, \ldots, B_n\}$ understood as expressing the disjunction of its elements. In place of axiom sequents, we now have *axiom sets*, i.e. sets of sentences containing a true prime sentence.

The rules of inference now take the simple form:

$$\frac{\Gamma_0, \bigvee_i A_i, A_j}{\Gamma, \bigvee_i A_i} \quad (\text{some } j)$$

and

$$\frac{\cdots \quad \Gamma_j, \bigwedge_i A_i, A_j \quad \cdots}{\Gamma, \bigwedge_i A_i} \quad (\text{all } j)$$

and the cut-rule takes the form

$$\frac{\Gamma, A \qquad \Gamma, \bar{A}}{\Gamma}$$

where A is called the *cut-formula*. Gentzen's Lemma now takes the form that, given cut-free deduction $\mathcal{D}_0$ of Γ, A and $\mathcal{D}_1$ of $\Gamma, \bar{A}$, there is a deduction of Γ involving only cuts with cut-formulas of rank $< |A|$. The argument is again essentially Gentzen's, except that symmetry demands again that the proof must be by induction on $\mathcal{D}_0 \times \mathcal{D}_1$ rather than on just one of the trees $\mathcal{D}_e$. Indeed, the proof is just an extension of Gentzen's proof of his *Hauptsatz* for propositional logic in the framework of the sequent calculus to the case of infinite disjunctions and conjunctions.

References

Bernays, P. (1970). On the original Gentzen consistency proof for number theory, *in* A. Kino, J. Myhill & R. Vesley (eds), *Intuitionism and Proof Theory*, Amsterdam: North-Holland, pp. 409–418.

Buchholz, W. (1997). Explaining Gentzen's consistency proof within infinitary proof theory, *Computational Logic and Proof Theory. Fifth Kurt Gödel Colloquium, KGC '97*. Lecture Notes in Computer Science, Vol. 1298. G. Gottlob, A. Leitsch, D. Mundici editors.

Coquand, T. (1995). A semantics of evidence for classical arithmetic, *The Journal of Symbolic Logic* **60**: 325–337.

Gentzen, G. (1935). Untersuchungen über das logische Schliessen I, II, *Mathematisce Zeitschrift* **39**: 176–210,405–431. English translation in (Gentzen 1969, pp. 68-131).

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

Gentzen, G. (1938). Nue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie, *Forschungen zur Logik und der Grundlagen der exakten Wissenschaften, New Series, No. 4* pp. 5–18. English translation in (Gentzen 1969, pp. 252-286).

Gentzen, G. (1943). Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie, *Mathematische Annalen* **119**: 140–161. English translation in (Gentzen 1969, pp. 287-308).

Gentzen, G. (1969). *The Collected Papers of Gerhard Gentzen*, Amsterdam: North-Holland Publishing Company. translated and edited by M. E. Szabo.

Gentzen, G. (1974). Der erste Widerspruchsfreiheitsbeweis für die klassische Zahlentheorie, *Archiv für Mathematische Logik und Grundlagenforschung* **16**: 97–118. Edited with a brief introduction by P. Bernays.

Gödel, K. (1938a). Lecture at Zilsel's, *Collected Works, Vol. III*, Oxford: Oxford University Press, pp. 87–113.

Gödel, K. (2003). *Collected Works, Vol. IV*, Oxford: Oxford University Press.

Kreisel, G. (1951). On the interpretation of non-finitist proofs—part i, *The Journal of Symbolic Logic* **16**: 241–267.

Kreisel, G. (1952). On the interpretation of non-finitist proofs—part ii. interpretations of number theory, applications, *The Journal of Symbolic Logic* **17**: 43–58.

Lorenzen, P. (1951). Algebraische und logistische Untersuchungen über freie Verbände, *Journal of Symbolic Logic* **16**: 81–106.

Schütte, K. (1951). Beweisetheoretische Erfassung der unendlichen Induktion in der Zahlentheorie, *Mathematische Annalen* **122**: 369–389.

Schütte, K. (1952). Beweistheoretische Untersuchung der verzweigten Analysis, *Mathematische Annalen* **124**: 123–147.

Schütte, K. (1964). Eine Grenze für die Beweisbarkeit der transfinitien Induction in der verzweigten Typenlogik, *Archiv für mathematische Logik und Grundlagenforschung* **7**: 45–60.

Tait, W. (1968). Normal derivability in classical logic, *in* J. Barwise (ed.), *The Syntax and Semantics of Infinitary Languages*, Berlin: Springer-Verlag, pp. 204–236.

Tait, W. (2001). Gödel's unpublished papers on foundations of mathematics, *Philosophia Mathematica* **9**: 87–126. Reprinted in (Tait 2005b), pp. 276–313.

Tait, W. (2005a). Gödel's reformulation of Gentzen's first consistency proof for arithmetic, *Bulletin of Symbolic Logic* **11**(2): 225–238.

Tait, W. (2005b). *The Provenance of Pure Reason: Essays in the Philosophy of Mathematics and Its History*, Oxford: Oxford University Press.