Normalization

Helmut Schwichtenberg
Mathematisches Institut, Universität München
Theresienstraße 39, D–8000 München 2, Germany

The aim of this paper is to present a central technique from proof theory, Gentzen's normalization for natural deduction systems, and to discuss some of its applications.

By normalization we mean a collection of algorithms transforming a given derivation into a certain normal form. A derivation is called *normal* if it does not contain any "detour" i.e. an application of an introduction rule immediately followed by an application of an elimination rule. Such normalization algorithms are useful because they allow to "straighten out" complex derivations and in this way extract hidden information.

We will treat many applications which demonstrate this, e.g. the subformula principle, Herbrand's theorem, the interpolation theorem, an exact characterization of the initial cases of transfinite induction provable in arithmetic and a proof that normalization in (the usual finitary) arithmetic is impossible.

From the computer science point of view, an even more interesting field of application for normalization algorithms is the possiblility to extract the constructive content of a maybe complex mathematical argument. Such algorithms can yield verified programs from derivations proving that certain specifications can be fulfilled. Of course, the feasability of programs obtained in this way will depend to a large extent on a good choice of the derivation, which should be done on the basis of a good idea for an algorithm. However, in this approach it is possible to use ordinary mathematical machinery for the development of programs.

Chapter 1 deals with normalization for minimal propositional logic, or more precisely for its implicational fragment. In Section 1.1 it is shown that — by adding stability axioms — classical logic can be embedded in it. In Sections 1.2–1.6 we then treat normalization for this calculus, with special emphasis on complexity questions. In Section 1.7 normalization (for a natural deduction system) is compared with cut–elimination (for a sequent calculus). Section 1.8 discusses a decision procedure for minimal implicational logic.

In Chapter 2 the method of collapsing types developed in (Troelstra and van Dalen 1988) is used to lift these results to minimal first order logic or more precisely to its $\to\forall$–fragment, which again suffices for classical logic. Section 2.4 contains some applications of normalization: the subformula principle, Herbrand's theorem and the interpolation theorem.

The final Chapter 3 treats normalization for arithmetic. Since normalization for finitary arithmetic with the induction rule is impossible (this is proved in Section 3.5), we extend in Section 3.3 the normalization technique to arithmetic with the ω–rule. This is used in Section 3.4 to give an exact characterization of the initial cases of transfinite induction provable in arithmetic as well as in some subsystems of arithmetic obtained by restricting the complexity of the induction formulas.

The expert will certainly note that most of the results and proofs elaborated here are well–known. The only novel points are the following.

- We have based our treatment of normalization in Chapters 1 and 2 on a slight generalization of β–conversion: we not only allow $(\lambda xr)s$ to be converted into $r_x[s]$, but more generally $(\lambda\vec{x}x.r)\vec{s}s$ to be converted into $(\lambda\vec{x}.r_x[s])\vec{s}$. This allows a particularly simple proof of the existence of the normal form (Theorem 1.2.2 below), which also provides an easy estimate of the number of conversion steps needed, and it also makes the results on strong normalization (Section 1.5) and the uniqueness of the normal form (Section 1.3) slightly stronger.
- In Section 1.7 we give an argument that normalization (for a natural deduction system) and cut–elimination (for a sequent calculus) are essentially different, using a recent result of (Hudelmaier 1989).
- In Section 1.8 we present a decision algorithm for implicational logic also due to (Hudelmaier 1989), together with a new proof of its correctness and completeness.
- In Theorem 3.5.1 it is stated that a certain weak form of a normalization theorem does not hold for arithmetic with the induction rule.

1. Normalization for propositional logic

1.1 Minimal implicational logic as a typed λ–calculus

Formulas are built up from propositional variables denoted by P, Q by means of $(\varphi \to \psi)$. We write $\varphi_1, \ldots, \varphi_m \to \psi$ for $(\varphi_1 \to (\varphi_2 \to \ldots (\varphi_m \to \psi) \ldots))$. Derivations are built up from assumption variables u^φ, v^ψ by means of the rule $\to^+$ of implication introduction (or λ–abstraction)

$$(\lambda u^\varphi r^\psi)^{\varphi\to\psi}$$

and the rule $\to^-$ of implication elimination (or application)

$$(t^{\varphi\to\psi} s^\varphi)^\psi.$$

A derivation r^ψ whose free assumption variables are among $u_1^{\varphi_1}, \ldots, u_m^{\varphi_m}$ is also called a derivation of ψ from $\varphi_1, \ldots, \varphi_m$. For readability we often leave out formula superscripts when they are obvious from the context or non–essential.

For obvious reasons we will also use the word term for derivations and type for formulas. The possibility to treat derivations as terms and formulas as types has been discovered by H. B. Curry and elaborated by W. A. Howard in (Howard 1980a). This correspondence can easily be shown to be an isomorphism; it is called the Curry–Howard–isomorphism.

More formally, it can be seen easily that a closed derivation (i. e. one without free assumption variables) is determined by

1. a type–free λ–term describing the derivation and
2. the derived formula.

The formulas in the derivation can be left out since they can easily be reconstructed form the given derived formula.

As an example, a derivation of

$$\varphi \to (\psi \to \varphi)$$

is given by $\lambda u \lambda v u$, and a derivation of

$$(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$$

is given by $\lambda u \lambda v \lambda w((uw)(vw))$. Both derivations can be easily written in the more usual tree form.

Note that our minimal implicational logic contains full classical propositional logic, as follows. Choose a particular propositional variable and denote it by $\perp$ (falsity). Associate with any formula φ in the language of classical propositional logic a finite list $\vec{\varphi}$ of formulas in our implicational language, by induction on φ:

$$\begin{aligned} P &\mapsto P \\ \neg\varphi &\mapsto \vec{\varphi} \to \perp \\ \varphi \to \psi &\mapsto \vec{\varphi} \to \psi_1, \ldots \vec{\varphi} \to \psi_m \\ \varphi \wedge \psi &\mapsto \vec{\varphi}, \vec{\psi} \\ \varphi \vee \psi &\mapsto (\vec{\varphi} \to \perp), (\vec{\psi} \to \perp) \to \perp \end{aligned}$$

Then, if φ is a formula in the language of full classical propositional logic and $\varphi_1, \ldots, \varphi_m$ is its associated list, φ is derivable in classical propositional logic iff each φ_i is derivable in minimal implicational logic from stability assumptions $\neg\neg P \to P$ (with $\neg\psi$ denoting $\psi \to \perp$) for all propositional variables P in φ. The essential step in the proof is to show that from the stability of ψ we can infer the stability of $\varphi \to \psi$: a derivation of

$$(\neg\neg\psi \to \psi) \to (\neg\neg(\varphi \to \psi) \to (\varphi \to \psi))$$

is given by

$$\lambda u_1 \lambda u_2 \lambda u_3 (u_1 \lambda u_4 (u_2 \lambda u_5 (u_4 (u_5 u_3)))).$$

1.2. Conversion

We are interested in the following process of simplification of terms:

$$(\lambda \vec{u} u.r)\vec{s}s \quad \text{converts into} \quad (\lambda \vec{u}.r_u[s])\vec{s}.$$

Here $\vec{u}$ and $\vec{s}$ denote finite lists $u_1 \dots u_m$ and $s_1 \dots s_m$, and $\lambda \vec{u} u.r$ denotes the term $\lambda u_1 \dots \lambda u_m \lambda u r$. Terms of the form $(\lambda \vec{u} u.r)\vec{s}s$ are called *convertible*.

Note that converting $(\lambda \vec{u} u.r)\vec{s}s$ into $(\lambda \vec{u}.r_u[s])\vec{s}$ may be viewed as first converting $(\lambda \vec{u} u.r)\vec{s}s$ "permutatively" into $(\lambda \vec{u}(\lambda u r)s)\vec{s}$ and then performing the inner conversion to obtain $(\lambda \vec{u}.r_u[s])\vec{s}$. One may ask why we take this conversion relation as our basis and not the more common $(\lambda u r)s \mapsto r_u[s]$. The reason is that our notion of level is defined with the clause $\mathrm{lev}(\varphi \to \psi) = \max(\mathrm{lev}(\varphi)+1, \mathrm{lev}(\psi))$ and not $= \max(\mathrm{lev}(\varphi), \mathrm{lev}(\psi))+1$; this in turn seems reasonable since then the level of $P_1, \dots, P_m \to Q$ (i. e. of $(P_1 \to (P_2 \to \dots (P_m \to Q) \dots))$) is 1 and hence independent of m. But given this definition of level, and given the need in some arguments (e. g. in Theorem 1.2.1) to perform conversions of highest level first, we must be able to convert $(\lambda u v.r)st$ with u of a low and v of a high level into $(\lambda u.r_v[t])s$. —In addition, since we allow more conversions here, the results on strong normalization and upper bounds for the length of arbitrary reduction sequences get stronger.

We write $r \to r'$ if r' is obtained from r as follows. Mark some occurrences of convertible subterms in r. Then convert them all simultaneously. Hence new convertible subterms generated by such a conversion can not be converted. More precisely, $r \to r'$ is defined by the following rules

1. $u \to u$.
2. If $r \to r'$, then $\lambda u r \to \lambda u r'$,
3. If $r \to r'$ and $s \to s'$, then $rs \to r's'$.
4. If $r \to r'$, $\vec{s} \to \vec{s'}$ and $s \to s'$, then $(\lambda \vec{u} u.r)\vec{s}s \to (\lambda \vec{u}.r'_u[s'])\vec{s'}$.

As a special case, we take

$$r \to_1 r'$$

to mean that r' is obtained from r by converting exactly one convertible subterm in r. Finally

$$r \to^* r' \qquad (r \text{ reduces to } r')$$

denotes the transitive and reflexive closure of $\to$ (or equivalently of $\to_1$).

A term is said to be in *normal form* if it does not contain a convertible subterm.

We want to show now that any term reduces to a normal form. This can be seen easily if we follow a certain order in our conversions. To define this order we have to make use of the fact that all our terms (i.e. derivations) have types (i.e. formulas).

Define the *level* of a formula by

$$\begin{aligned} \mathrm{lev}(P) &= 0, \\ \mathrm{lev}(\varphi \to \psi) &= \max(\mathrm{lev}(\varphi) + 1, \mathrm{lev}(\psi)) \end{aligned}.$$

A convertible derivation

$$(\lambda \vec{u}^{\vec{\varphi}} u^{\varphi}.r)\vec{s}s$$

is also called a *cut* with *cut–formula* φ. By the level of a cut we mean the level of its cut–formula. The *cut–rank* of a derivation r is the least number bigger than the levels of all cuts in r. Now let t be a derivation of cut–rank $k+1$. Pick a cut

$$(\lambda\vec{u}u.r)\vec{s}s$$

of the maximal level k in t, such that s does not contain another cut of level k. (E.g., pick the rightmost cut of level k.) Then it is easy to see that replacing the picked occurrence of $(\lambda\vec{u}u.r)\vec{s}s$ in t by $(\lambda\vec{u}.r_u[s])\vec{t}$ reduces the number of cuts of the maximal level k in t by 1. Hence

Theorem 1.2.1. *We have an algorithm which reduces any given term into a normal form.*

We now want to give an estimate of the number of conversion steps our algorithm takes until it reaches the normal form. The key observation for this estimate is the obvious fact that replacing one occurrence of

$$(\lambda\vec{u}u.r)\vec{s}s \quad \text{by} \quad (\lambda\vec{u}.r_u[s])\vec{s}$$

in a given term t at most squares the length of t; here the *length* of t is taken to be the number of variables in t (except those immediately following a λ–symbol).

A bound $s_k(l)$ for the number of steps our algorithm takes to reduce the rank of a given term of length l by k can be derived inductively, as follows. Let $s_0(l) := 0$. To obtain $s_{k+1}(l)$, first note that by induction hypothesis it takes $\leq s_k(l)$ steps to reduce the rank by k. The length of the resulting term is $\leq l^{2^s}$ where $s := s_k(l)$ since any step (i.e. conversion) at most squares the length. Now to reduce the rank by one more the number of additional steps is obviously bounded by that length. Hence the total number of steps to reduce the rank by $k+1$ is bounded by

$$s_k(l) + l^{2^{s_k(l)}} =: s_{k+1}(l).$$

Theorem 1.2.2. (Upper bound for the complexity of normalization) *The normalization algorithm given in the proof of Theorem 1.2.1 takes at most $s_k(l)$ steps to reduce a given term of cut–rank k and length l to normal form, where*

$$s_0(l) := 0 \quad \textit{and} \quad s_{k+1}(l) := s_k(l) + l^{2^{s_k(l)}}.$$

1.3. Uniqueness

We shall show that the normal form of a term is uniquely determined; this will be done by an argument which also applies to type–free terms, i.e. terms without formula superscripts. The main idea of the proof (due to J. B. Rosser and W. W. Tait) is to use the relation $r \to r'$ defined in Section 1.2. Its crucial property is given by

Lemma 1.3.1. *If $r \to r'$ and $t \to t'$ then $r_v[t] \to r'_v[t']$.*

The proof is by induction on the definition of $r \to r'$. All cases are obvious except possibly Rule 4. So assume $r \to r'$, $\vec{s} \to \vec{s'}$ and $s \to s'$. Then

$$r_v[t] \to r'_v[t'], \quad \vec{s}_v[t] \to \vec{s'}_v[t'] \quad \text{and} \quad s_v[t] \to s'_v[t']$$

by induction hypothesis, and hence

$$\underbrace{(\lambda\vec{u}u.r_v[t])\vec{s}_v[t]s_v[t]}_{((\lambda\vec{u}u.r)\vec{s}s)_v[t]} \to \underbrace{(\lambda\vec{u}.(r'_v[t'])_u[s'_v[t']])\vec{s'}_v[t']}_{((\lambda\vec{u}.r'_u[s'])\vec{s'})_v[t']}.$$

by definition of $\to$. $\square$

Lemma 1.3.2. *Assume $r \to r'$ and $r \to r''$. Then we can find a term r''' such that $r' \to r'''$ and $r'' \to r'''$.*

The proof is by induction on the definition of $r \to r'$. Again all cases are obvious except possibly the situation where either $r \to r'$ or $r \to r''$ is obtained via Rule 4. By symmetry we may assume the former. But then the claim follows from Lemma 1.3.1: If

$$(\lambda\vec{u}u.r)\vec{s}s \to (\lambda\vec{u}.r'_u[s'])\vec{s'}$$

and

$$(\lambda\vec{u}u.r)\vec{s}s \to (\lambda\vec{u}u.r'')\vec{s''}s'',$$

then

$$(\lambda\vec{u}u.r)\vec{s}s \to (\lambda\vec{u}.r'_u[s'])\vec{s'} \to (\lambda\vec{u}.r'''_u[s'''])\vec{s'''}$$

and

$$(\lambda\vec{u}u.r)\vec{s}s \to (\lambda\vec{u}u.r'')\vec{s''}s'' \to (\lambda\vec{u}.r'''_u[s'''])\vec{s'''},$$

and if

$$(\lambda\vec{u}u\vec{v}v.r)\vec{s}s\vec{t}t \to (\lambda\vec{u}\vec{v}v.r'_u[s'])\vec{s'}\vec{t'}t'$$

and

$$(\lambda\vec{u}u\vec{v}v.r)\vec{s}s\vec{t}t \to (\lambda\vec{u}u\vec{v}.r''_v[t''])\vec{s''}s''\vec{t''}$$

then

$$(\lambda\vec{u}u\vec{v}v.r)\vec{s}s\vec{t}t \to (\lambda\vec{u}\vec{v}v.r'_u[s'])\vec{s'}\vec{t'}t' \to (\lambda\vec{u}\vec{v}.r'''_{u,v}[s''',t'''])\vec{s'''}\vec{t'''}$$

and

$$(\lambda\vec{u}u\vec{v}v.r)\vec{s}s\vec{t}t \to (\lambda\vec{u}u\vec{v}.r''_v[t''])\vec{s''}s''\vec{t''} \to (\lambda\vec{u}\vec{v}.r'''_{u,v}[s''',t'''])\vec{s'''}\vec{t'''}.$$

$\square$

Theorem 1.3.3. (Church–Rosser) *Assume $r \to^* r'$ and $r \to^* r''$. Then we can find a term r''' such that $r' \to^* r'''$ and $r'' \to^* r'''$.*

The proof is immediate from Lemma 1.3.2. $\square$

Corollary 1.3.4. (Uniqueness of the normal form) *Assume $r \to^* r'$ and $r \to^* r''$, where both r' and r'' are in normal form. Then r' and r'' are identical.*

1.4. Complexity of normalization: a lower bound

In Theorem 1.2.2 we have obtained an upper bound on the number of conversion steps our particular normalization algorithm of Theorem 1.2.1 takes to reach the normal form. This upper bound was superexponential in the length of the given term. It is tempting to think that by choosing a clever normalization strategy one might be able to reduce that bound significantly. It is the purpose of the present section to show that this is impossible. More precisely, we will construct terms r_n of length $3n$ and show that any normalization algorithm needs at least $2_{n-2} - n$ conversions (with $2_0 := 1, 2_{n+1} := 2^{2_n}$) to reduce r_n to its normal form.

The fact that there is no elementary algorithm (i.e. whose time is exponentially bounded) to compute the normal form of terms also follows from (Statman 1979), where it is shown more generally that the problem whether two terms r_1 and r_2 have the same normal form is not elementary recursive. The simple example treated here is taken from (Schwichtenberg 1982, p. 455).

The pure types k are defined inductively by $0 := P$ (some fixed propositional variable) and $k+1 = k \to k$. We define iteration terms I_n of pure type $k+2$ by

$$I_n :\equiv \lambda f \lambda u (f(f(\dots f(fu)\dots))),$$

with n occurrences of f after $\lambda f \lambda u$; here f, u are variables of type $k+1, k$, respectively. Let $f \circ g$ be an abbreviation for $\lambda u(f(gu))$, and let $r = s$ mean that r and s have the same normal form. With this notation we can write

$$I_n = \lambda f(\underbrace{f \circ f \circ \dots \circ f}_{n}).$$

The main point of our argument is the following simple lemma, which can be traced back to Rosser (cf. (Church 1941, p. 30))

Lemma 1.4.1.

$$(I_m f) \circ (I_n f) = I_{m+n} f,$$
$$I_m \circ I_n = I_{m \cdot n},$$
$$I_m I_n = I_{n^m}. \square$$

As an immediate consequence we have

$$r_n :\equiv \underbrace{I_2 I_2 \dots I_2}_{n} = I_{2_n}.$$

Now consider any sequence of reduction steps transforming r_n into its normal form, and let s_n denote the total number of reduction steps in this sequence.

Theorem 1.4.2. $s_n \geq 2_{n-2} - n$.

Proof. The length of r_n is $3n$. Note that any conversion step can at most square the length of the original term. Hence we have

$$\begin{aligned} 2_n &< \mathrm{length}(I_{2_n}) \qquad \text{(the normal form of } r_n) \\ &\leq \mathrm{length}(r_n)^{2^{s_n}} \\ &= (3n)^{2^{s_n}} \\ &\leq 2^{2^{n+s_n}} \qquad (\text{since } 3n \leq 2^{2^n}), \end{aligned}$$

and the theorem is proved. $\square$

1.5. Strong normalization

In Section 1.2 we have proved that any term can be reduced to a normal form, and in Section 1.3 we have seen that this normal form is uniquely determined. But it is still conceivable that there might be an odd reduction sequence which does not terminate at all. It is the aim of the present Section to show that this is impossible. This fact is called the strong normalization theorem.

For the proof we employ a powerful method due to W. W. Tait, which is based on so–called strong computability predicates. These are defined by induction on the types (i.e. formulas) as follows.

A term r^φ with φ of level 0 (i. e. a propositional variable) is strongly computable iff r is strongly normalizable, i.e. every reduction sequence starting from r terminates. A term $r^{\varphi\to\psi}$ is strongly computable iff for all strongly computable s^φ also $(rs)^\psi$ is strongly computable.

A term r is *strongly computable under substitution* iff for all strongly computable $\vec{s}$ the result of substituting $\vec{s}$ for all variables free in r is again strongly computable.

Lemma 1.5.1. *Let φ be a formula.*

1. *Any strongly computable term r^φ is strongly normalizable.*
2. *u^φ is strongly computable.*

We prove 1 and 2 simultaneously by induction on φ. For φ of level 0 both claims are obvious. Now consider $\varphi \to \psi$. For 1, assume that $r^{\varphi\to\psi}$ is strongly computable. By induction hypothesis 2 and the definition of strong computability we know that $(ru)^\psi$ is strongly computable and hence that any reduction sequence starting with ru terminates (by induction hypothesis 1). But this obviously implies that the same is true for r. For 2, assume that $\vec{r}$ are strongly computable. We have to show that $u\vec{r}$ (which is to be of level 0) is strongly computable, i. e. that any reduction sequence starting with $u\vec{r}$ terminates. But this follows from induction hypothesis 1, which says that any reduction sequence starting from r_i terminates. $\square$

Lemma 1.5.2. *If $r \to_1 r'$ and r is strongly computable, then r' is strongly computable.*

Proof. Let $\vec{s}$ be strongly computable. We have to show that $r'\vec{s}$ is strongly computable, i. e. that any reduction sequence starting from $r'\vec{s}$ terminates. But this is obviously true, because otherwise we would also have an infinite reduction sequence for $r\vec{s}$. $\square$

Lemma 1.5.3. *Any term r is strongly computable under substitution.*

The proof is by induction on the height of r.

Case u. Obvious.

Case rs. Let $\vec{t}$ be strongly computable. We have to show that $r[\vec{t}]s[\vec{t}]$ is strongly computable. But this holds, since by induction hypothesis we know that $r[\vec{t}]$ as well as $s[\vec{t}]$ are strongly computable.

Case $\lambda\vec{u}u.r$. Let $\vec{t}$ be strongly computable. We have to show that $\lambda\vec{u}u.r[\vec{t}]$ is strongly computable. So let $\vec{s}, s$ and $\vec{r}$ be strongly computable. We must show that $(\lambda\vec{u}u.r[\vec{t}])\vec{s}s\vec{r}$ is strongly computable, i.e. that any reduction sequence for it terminates. So assume we have an infinite reduction sequence. Since $r[\vec{t}], \vec{s}, s$ and $\vec{r}$ all are strongly normalizable, there must be a term $(\lambda\vec{u}u.r[\vec{t}]')\vec{s'}s'\vec{r'}$ with $r[\vec{t}] \to^* r[\vec{t}]', \vec{s} \to^* \vec{s'}, s \to^* s'$ and $\vec{r} \to^* \vec{r'}$ in that reduction sequence where a "head conversion" is applied, which we may assume to yield

$$(\lambda\vec{u}.(r[\vec{t}]')[s'])\vec{s'}\vec{r'}.$$

But $r[\vec{t}] \to^* r[\vec{t}]'$ implies $\lambda\vec{u}.r[s,\vec{t}] \to^* \lambda\vec{u}.(r[\vec{t}]')[s']$, and hence the fact that $\lambda\vec{u}.r$ is (by induction hypothesis) strongly computable under substitution together with Lemma 1.5.2 implies that $(\lambda\vec{u}.(r[\vec{t}]')[\vec{s'}]$ is strongly computable. But then, again by Lemma 1.5.2, also $(\lambda\vec{u}.(r[\vec{t}])'[s'])\vec{s'}\vec{r'}$ is strongly computable and therefore strongly normalizable. This contradicts our assumption above that we have an infinite reduction sequence. □

From Lemma 1.5.3 and both parts of Lemma 1.5.1 can conclude immediately

Theorem 1.5.4. *Any term r is strongly normalizable.* □

1.6. Complexity of normalization: an upper bound

By Section 1.5 we already know that the full reduction tree for a given term is finite; hence its height bounds the length of any reduction sequence. But it is not obvious how a reasonable estimate for that height might be obtained.

However, using a technique due to (Howard 1980b) (which in turn is based on (Sanchis 1967) and (Diller 1968)) it can be shown that we have the following superexponential universal bound.

Theorem 1.6.1. *Let r be a term of the typed λ–calculus of level 0. Let m be a bound for the levels of subterms of r and $k \geq 2$ be a bound for the arities of subterms of r. Then the length of an arbitrary reduction sequence for r with respect to $\to_1$ is bounded by*

$$k^{2_m(m+2\cdot\mathrm{height}(r)+2k+2)}.$$

For the proof see (Schwichtenberg 1990). □

1.7. Cut elimination versus normalization

Up to now we have considered pure implicational logic by means of Gentzen's rules of natural deduction. Now it is also common to use another type of logical calculus, the *sequent calculus* also introduced by Gentzen. Instead of formulas it treats *sequents* $\Gamma \Rightarrow \varphi$, where Γ is a finite set of formulas. The rules of the sequent calculus for pure implicational logic are the following; here we write Γ, φ for $\Gamma \cup \{\varphi\}$.

Axiom. $\vdash \Gamma, \varphi \Rightarrow \varphi$ for φ atomic.

$\to$*–right.* If $\vdash \Gamma, \varphi \Rightarrow \psi$, then $\vdash \Gamma \Rightarrow \varphi \to \psi$.

$\to$*–left.* If $\vdash \Gamma, \varphi \to \psi \Rightarrow \varphi$ and $\vdash \Gamma, \varphi \to \psi, \psi \Rightarrow \chi$, then $\vdash \Gamma, \varphi \to \psi \Rightarrow \chi$.

Cut. If $\vdash \Gamma \Rightarrow \chi$ and $\vdash \Gamma, \chi \Rightarrow \varphi$, then $\vdash \Gamma \Rightarrow \varphi$.

It is easy to see that the sequent calculus is equivalent to natural deduction, in the sense that $\vdash \Gamma \Rightarrow \varphi$ iff from Γ we can derive φ by means of the rules $\to^+$ and $\to^-$ and the assumption rule.

Now a normal derivation of φ from Γ has the property that all formulas occurring in this derivation are subformulas of either φ or a formula in Γ. The same property holds for derivations of $\Gamma \Rightarrow \varphi$ in the sequent calculus which do not use the cut rule. Hence it is of interest to know that the cut rule can always be eliminated from derivations in the sequent calculus.

Gentzen proved this Cut Elimination Theorem in his thesis. Here we prove it in such a way that we also obtain a good bound on the length of the resulting cut–free derivation, in the form $2^{j(d)} \cdot l(d)$, where $l(d)$ is the length of the original derivation and $j(d)$ is the maximum taken over all paths in d of the sum of the degrees of all cut formulas on the path. The notion of degree used here is rather peculiar. Its crucial property is

$$\deg(\varphi \to \psi) + \deg(\psi \to \chi) < \deg((\varphi \to \psi) \to \chi). \tag{1.1}$$

This can be achieved if we define

- $\deg(\varphi) = 2$ for φ atomic,
- $\deg(\varphi \to \psi) = 1 + \deg(\varphi) \cdot \deg(\psi)$.

For then we have, writing $a := \deg(\varphi)$, $b := \deg(\psi)$ and $c := \deg(C)$, $1 + ab + 1 + bc = 2 + (a + c)b \leq 2 + abc < 1 + c + abc = 1 + (1 + ab)c$ and hence 1.1.

More formally, we define the relation $\vdash^{\alpha}_{m} \Gamma \Rightarrow \varphi$ (to be read: $\Gamma \Rightarrow \varphi$ is derivable with height $\leq \alpha$ and cut–rank $\leq m$) with α, m natural numbers inductively by the following rules.

Axiom. $\vdash^{\alpha}_{m} \Gamma, \varphi \Rightarrow \varphi$ for φ atomic.

$\to$*–right.* If $\vdash^{\alpha}_{m} \Gamma, \varphi \Rightarrow \psi$, then $\vdash^{\alpha+1}_{m} \Gamma \Rightarrow \varphi \to \psi$.

$\to$*–left.* If $\vdash^{\alpha}_{m} \Gamma, \varphi \to \psi \Rightarrow \varphi$ and $\vdash^{\alpha}_{m} \Gamma, \varphi \to \psi, \psi \Rightarrow \chi$, then $\vdash^{\alpha+1}_{m} \Gamma, \varphi \to \psi \Rightarrow \chi$.

Cut. If $\vdash^{\alpha}_{m} \Gamma \Rightarrow \chi$ and $\vdash^{\alpha}_{m} \Gamma, \chi \Rightarrow \varphi$, then $\vdash^{\alpha+1}_{m+\deg(\chi)} \Gamma \Rightarrow \varphi$.

Then the bound mentioned above is a consequence of the following

Theorem 1.7.1. (Cut Elimination Theorem) *If* $\vdash^{\alpha}_{m+1} \Gamma \Rightarrow \varphi$, *then* $\vdash^{2\alpha}_{m} \Gamma \Rightarrow \varphi$.

This theorem is due to (Hudelmaier 1989); its present formulation and proof is the result of Buchholz' analysis (Buchholz 1989) of Hudelmaier's arguments.

We need some Lemmata before we can give the proof.

Lemma 1.7.2. (Weakening Lemma) *If* $\vdash^{\alpha}_{m} \Gamma \Rightarrow \varphi$, *then* $\vdash^{\alpha}_{m} \Gamma, \Delta \Rightarrow \varphi$. $\square$

Lemma 1.7.3. (Inversion)

i. If $\vdash^{\alpha}_{m} \Gamma \Rightarrow \varphi \to \psi$, *then* $\vdash^{\alpha}_{m} \Gamma, \varphi \Rightarrow \psi$.
ii. If $\vdash^{\alpha}_{m} \Gamma, \varphi \to \psi \Rightarrow \chi$, *then* $\vdash^{\alpha}_{m} \Gamma, \psi \Rightarrow \chi$.
iii. If $\vdash^{\alpha}_{m} \Gamma, (\varphi \to \psi) \to \chi \Rightarrow \vartheta$, *then* $\vdash^{\alpha}_{m} \Gamma, \varphi, \psi \to \chi \Rightarrow \vartheta$.

Proof. By induction on α. We only treat one case of iii. Assume that

$$\vdash^{\alpha}_{m} \Gamma, (\varphi \to \psi) \to \chi \Rightarrow \vartheta$$

was inferred from

$$\vdash^{\alpha-1}_{m} \Gamma, (\varphi \to \psi) \to \chi \Rightarrow \varphi \to \psi \quad \text{and} \quad \vdash^{\alpha-1}_{m} \Gamma, (\varphi \to \psi) \to \chi, \chi \Rightarrow \vartheta.$$

By induction hypothesis we get

$$\vdash^{\alpha-1}_{m} \Gamma, \varphi, \psi \to \chi \Rightarrow \varphi \to \psi \quad \text{and} \quad \vdash^{\alpha-1}_{m} \Gamma, \varphi, \psi \to \chi, \chi \Rightarrow \vartheta.$$

Hence by i

$$\vdash^{\alpha-1}_{m} \Gamma, \varphi, \psi \to \chi \Rightarrow \psi.$$

Now $\to$–left yields $\vdash^{\alpha}_{m} \Gamma, \varphi, \psi \to \chi \Rightarrow \vartheta$. $\square$

Lemma 1.7.4. (Cut Elimination Lemma)

i. If $\vdash^{\alpha}_{0} \Gamma \Rightarrow \varphi$ *and* $\vdash^{\beta}_{0} \Gamma, \varphi \Rightarrow \psi$ *and* φ *is atomic, then* $\vdash^{\alpha+\beta}_{0} \Gamma \Rightarrow \psi$.
ii. If $\vdash^{\alpha}_{0} \Gamma \Rightarrow \varphi \to \psi$ *and* $\vdash^{\beta}_{0} \Gamma, \varphi \to \psi \Rightarrow \chi$ *and* φ *is atomic and* $\beta \leq \alpha$, *then* $\vdash^{\alpha+\beta}_{m} \Gamma \Rightarrow \chi$ *with* $m = \deg(\psi)$.
iii. If $\vdash^{\alpha}_{0} \Gamma \Rightarrow (\varphi \to \psi) \to \chi$ *and* $\vdash^{\beta}_{0} \Gamma, (\varphi \to \psi) \to \chi \Rightarrow \vartheta$, *then* $\vdash^{\alpha+\beta+2}_{m} \Gamma \Rightarrow \vartheta$, *with* $m = \deg(\varphi \to \psi) + \deg(\psi \to \chi)$.

Proof. i. By induction on β. ii. Consider also

ii'. If $\vdash^{\alpha}_{0} \Gamma \Rightarrow \varphi \to \psi$ and $\vdash^{\alpha}_{0} \Gamma, \psi \Rightarrow \chi$ and $\vdash^{\beta}_{0} \Gamma, \varphi \to \psi \Rightarrow \varphi$ and φ is atomic and $\beta \leq \alpha$, then $\vdash^{\alpha+\beta+1}_{m} \Gamma \Rightarrow \chi$, with $m = \deg(\psi)$.

We prove ii and ii' simultaneously by induction on β.

ii. Assume

$$\vdash^{\beta-1}_{0} \Gamma, \varphi \to \psi \Rightarrow \varphi \text{ and } \vdash^{\beta-1}_{0} \Gamma, \varphi \to \psi, \psi \Rightarrow \chi.$$

Then we have $\vdash^{\alpha}_{0} \Gamma, \psi \Rightarrow \chi$ (since $\beta \leq \alpha$). Hence induction hypothesis ii' yields

$$\vdash^{\alpha+\beta}_{\deg(\psi)} \Gamma \Rightarrow \chi.$$

ii'. *Case* $\beta = 0$. Then $\varphi \in \Gamma$, hence

$$\vdash^{\alpha+1}_{\deg(\psi)} \Gamma \Rightarrow \chi.$$

Case $\vdash^{\beta-1}_{0} \Gamma, \varphi \to \psi \Rightarrow \varphi$ *and* $\vdash^{\beta-1}_{0} \Gamma, \varphi \to \psi, \psi \Rightarrow \varphi$. Then the claim follows immediately from the induction hypothesis.

Case $\vdash_0^{\beta-1} \Gamma, \varphi \to \psi \Rightarrow \vartheta$ *and* $\vdash_0^{\beta-1} \Gamma, \varphi \to \psi, \eta \Rightarrow \varphi$ *with* $\vartheta \to \eta \in \Gamma$ *and* $\vartheta \to \eta \neq \varphi \to \psi$. Then induction hypothesis ii yields

$$\vdash_{\deg(\psi)}^{\alpha+\beta-1} \Gamma \Rightarrow \vartheta,$$

and induction hypothesis ii' yields

$$\vdash_{\deg(\psi)}^{\alpha+\beta} \Gamma, \eta \Rightarrow \chi.$$

Now $\to$–left gives

$$\vdash_{\deg(\psi)}^{\alpha+\beta+1} \Gamma \Rightarrow \chi.$$

iii. By induction on β. Assume

$$\vdash_0^{\beta-1} \Gamma, (\varphi \to \psi) \to \chi \Rightarrow \varphi \to \psi \text{ and } \vdash_0^{\beta-1} \Gamma, (\varphi \to \psi) \to \chi, \chi \Rightarrow \vartheta.$$

From $\vdash_0^{\alpha} \Gamma \Rightarrow (\varphi \to \psi) \to \chi$ we get $\vdash_0^{\alpha} \Gamma, \varphi \to \psi \Rightarrow \chi$, hence $\vdash_0^{\alpha} \Gamma, \psi \Rightarrow \chi$, hence $\vdash_0^{\alpha+1} \Gamma \Rightarrow \psi \to \chi$. On the other hand, the Inversion Lemma, Parts i and iii, yields $\vdash_0^{\beta-1} \Gamma, \varphi, \psi \to \chi \Rightarrow \psi$, hence $\vdash_0^{\beta} \Gamma, \psi \to \chi \Rightarrow \varphi \to \psi$. So

$$\vdash_{\deg(\psi \to \chi)}^{\alpha+\beta+1} \Gamma \Rightarrow \varphi \to \psi.$$

Furthermore, from $\vdash_0^{\beta-1} \Gamma, (\varphi \to \psi) \to \chi, \chi \Rightarrow \vartheta$ we get $\vdash_0^{\beta-1} \Gamma, \chi \Rightarrow \vartheta$. Since $\vdash_0^{\alpha} \Gamma, \varphi \to \psi \Rightarrow \chi$, a cut yields

$$\vdash_{\deg(\chi)}^{\alpha+\beta} \Gamma, \varphi \to \psi \Rightarrow \vartheta.$$

Now one further cut gives $\vdash_m^{\alpha+\beta+2} \Gamma \Rightarrow \vartheta$. $\square$

We now prove the Cut Elimination Theorem, by induction on α. Assume

$$\vdash_k^{\alpha-1} \Gamma \Rightarrow \chi \text{ and } \vdash_k^{\alpha-1} \Gamma, \chi \Rightarrow \varphi$$

with $m+1 = k + \deg(\chi)$. If $k \neq 0$, then by induction hypothesis

$$\vdash_{k-1}^{2(\alpha-1)} \Gamma \Rightarrow \chi \text{ and } \vdash_{k-1}^{2(\alpha-1)} \Gamma, \chi \Rightarrow \varphi,$$

hence by cut $\vdash_m^{2\alpha} \Gamma \Rightarrow \varphi$, since $k - 1 + \deg(\chi) = m$. If $k = 0$, then the claim follows from the Cut Elimination Lemma together with (1). $\square$

An interesting consequence of the fact that we have a (Kalmar) elementary bound on the length of the cut free derivation given by our algorithm in terms of the original derivation is the following: The cut elimination algorithm $d \mapsto d^{\mathrm{cf}}$ just described is essentially different from normalization $d \mapsto d^{\mathrm{nf}}$, in the sense that there cannot exist elementary translations $d \mapsto d^{\mathrm{seq}}$, $d \mapsto d^{\mathrm{nat}}$ from derivations in natural deductions to derivations in the sequent calculus and back, such that $d^{\mathrm{nf}} = ((d^{\mathrm{seq}})^{\mathrm{cf}})^{\mathrm{nat}}$. For then $d \mapsto d^{\mathrm{nf}}$ would be elementary, which it isn't by the counterexample in Section 1.4.

1.8. A decision algorithm for implicational logic

It is clearly decidable whether a given formula φ is derivable in minimal implicational logic: just search systematically for a normal derivation of φ. This search must terminate, since by the subformula property there are only finitely many such normal derivations. However it does not seem to be a good idea to try to implement this algorithm

Here we present another decision algorithm which is easy to implement and also seems to be rather efficient in cases of interest. It also amounts to searching for a "normal" proof, but now in a special calculus LH due to (Hudelmaier 1989), designed with the intention that most rules should be invertible. Again our formulation of LH and most proofs are taken from (Buchholz 1989).

The calculus LH is again a sequent calculus. To distinguish it from Gentzen's sequent calculus discussed in Section 1.7 we write

$$\vdash^{\alpha}_{LH} \Gamma \Rightarrow \varphi$$

if the sequent $\Gamma \Rightarrow \varphi$ is derivable with height $\leq \alpha$ in LH. The rules of LH are the following; again we write Γ, φ for $\Gamma \cup \{\varphi\}$

Axiom. $\vdash^{\alpha}_{LH} \Gamma, \varphi \Rightarrow \varphi$ for φ atomic.

H→–*right.* If $\vdash^{\alpha}_{LH} \Gamma, \varphi \Rightarrow \varphi$, then $\vdash^{\alpha+1}_{LH} \Gamma \Rightarrow \varphi \to \psi$.

H→–*left–atomic.* If $\vdash^{\alpha}_{LH} \Gamma, \varphi, \psi \Rightarrow \chi$ and φ is atomic, then $\vdash^{\alpha+1}_{LH} \Gamma, \varphi, \varphi \to \psi \Rightarrow \chi$.

H→–*left*–→. If $\vdash^{\alpha}_{LH} \Gamma, \varphi, \psi \to \chi \Rightarrow \psi$ and $\vdash^{\alpha}_{LH} \Gamma, \chi \Rightarrow \vartheta$, then $\vdash^{\alpha+1}_{LH} \Gamma, (\varphi \to \psi) \to \chi \Rightarrow \vartheta$.

Hudelmaier has observed — and we will prove it below — that this calculus is equivalent to minimal implicational logic. Now the point in these strange rules is that they are all invertible, with the sole exception of the last one which is only "half invertible":

Inversion Lemma 1.8.1.
i. If $\vdash^{\alpha}_{LH} \Gamma \Rightarrow \varphi \to \psi$, *then* $\vdash^{\alpha}_{LH} \Gamma, \varphi \Rightarrow \psi$.
ii. If $\vdash^{\alpha}_{LH} \Gamma, \varphi, \varphi \to \psi \Rightarrow \chi$ *and* φ *is atomic, then* $\vdash^{\alpha}_{LH} \Gamma, \varphi, \psi \Rightarrow \chi$.
iii. If $\vdash^{\alpha}_{LH} \Gamma, (\varphi \to \psi) \to \chi \Rightarrow \vartheta$, *then* $\vdash^{\alpha}_{LH} \Gamma, \chi \Rightarrow \vartheta$. □

Clearly the last rule H→–left–→ cannot be fully invertible. A counterexample is

$$P, ((Q \to \bot) \to \bot) \to Q \Rightarrow P,$$

which is clearly derivable, whereas

$$P, Q \to \bot, \bot \to Q \Rightarrow \bot$$

is not. Now the decision algorithm derived from the Inversion Lemma runs as follows. Given a sequent $\Gamma \Rightarrow \varphi$, first apply Parts i and ii of the Inversion Lemma as long as possible. If you end up with a sequent which does not contain left–iterated implications $(\varphi \to \psi) \to \chi$, then by the form of the LH–rules it is derivable if and only if it is an axiom. Now assume there are some left–iterated implications $(\varphi \to \psi) \to \chi$ among the premiss–formulas Γ. Choose one of them (this step may lead to backtracking!), form its premisses according to the rule H→–left–→, and continue with both sequents.

An example for the necessity of backtracking is the sequent

$$\neg\neg Q \to Q, \neg\neg(P \to Q), P \Rightarrow Q.$$

If the second occurence of a left–iterated implication is choosen, i.e. $\neg\neg(P \to Q)$, we obtain by H$\to$–left–$\to$

$$\neg\neg Q \to Q, P \to Q, \bot \to \bot, P \Rightarrow \bot$$

and

$$\neg\neg Q \to Q, \bot, P \Rightarrow Q.$$

Now the first of these sequents is clearly underivable, hence this path in the search–tree fails, and we have to backtrack and choose the other left–iterated implication $\neg\neg Q \to Q$ instead.

It remains to be shown that the calculus LH is equivalent to minimal implicational logic. One direction is easy, namely that any sequent derivable in LH is also derivable in minimal implicational logic. We only consider the rule H$\to$–left–$\to$, and argue informally. So assume Γ and $(\varphi \to \psi) \to \chi$. Then clearly $\psi \to \chi$ (for if we assume ψ, we certainly have $\varphi \to \psi$ hence χ). So by the first premiss $\varphi \to \psi$, hence χ by our assumption, hence ϑ by the second premiss.

For the other direction we need a Lemma.

Lemma 1.8.2. *If* $\vdash^{\alpha}_{LH} \Gamma, \varphi \to \psi \Rightarrow \varphi$ *and* $\vdash_{LH} \Gamma, \psi \Rightarrow \vartheta$ *and* φ *is atomic, then* $\vdash_{LH} \Gamma, \varphi \to \psi \Rightarrow \vartheta$.

The proof is by induction on α.

1. Assume $\varphi \in \Gamma$. Then from $\vdash_{LH} \Gamma, \psi \to \vartheta$ we get $\vdash_{LH} \Gamma, \varphi \to \psi \Rightarrow \vartheta$ by H$\to$–left–atomic.

2. Let $\Gamma = \Delta, (\varphi_1 \to \psi_1) \to \chi_1$ and assume that

$$\vdash^{\alpha}_{LH} \Delta, (\varphi_1 \to \psi_1) \to \chi_1, \varphi \to \psi \Rightarrow \varphi$$

was inferred from

$$\vdash^{\alpha-1}_{LH} \Delta, \varphi_1, \psi_1 \to \chi_1, \varphi \to \psi \Rightarrow \psi_1 \tag{1.2}$$

and

$$\vdash^{\alpha-1}_{LH} \Delta, \chi_1, \varphi \to \psi \Rightarrow \varphi \tag{1.3}$$

by H$\to$–left–$\to$. First note that from the second premiss of the Lemma we get

$$\vdash_{LH} \Delta, \chi_1, \psi \Rightarrow \vartheta \tag{1.4}$$

by the Inversion Lemma, Part iii. Now from 1.3 and 1.4 we obtain by the induction hypothesis

$$\vdash_{LH} \Delta, \chi_1, \varphi \to \psi \Rightarrow \vartheta. \tag{1.5}$$

The rule H$\to$–left–$\to$ yields form 1.2 and 1.5

$$\vdash_{LH} \Delta, (\varphi_1 \to \psi_1) \to \chi_1, \varphi \to \psi \Rightarrow \vartheta.$$

3. Let $\Gamma = \Delta, \varphi_1, \varphi_1 \to \psi_1$ with φ_1 atomic and different from φ, and assume that

$$\vdash^{\alpha}_{LH} \Delta, \varphi_1, \varphi_1 \to \psi_1, \varphi \to \psi \Rightarrow \varphi$$

was inferred from

$$\vdash^{\alpha-1}_{LH} \Delta, \varphi_1, \psi_1, \varphi \to \psi \Rightarrow \varphi \tag{1.6}$$

by H$\to$–left–atomic. First note that from the second premiss of the Lemma we get

$$\vdash_{LH} \Delta, \varphi_1, \psi_1, \psi \Rightarrow \vartheta \tag{1.7}$$

by the Inversion Lemma, Part ii. Now from 1.6 and 1.7 we obtain by the induction hypothesis

$$\vdash_{LH} \Delta, \varphi_1, \psi_1, \varphi \to \psi \Rightarrow \vartheta. \tag{1.8}$$

An application of H$\to$–left–atomic to 1.8 yields

$$\vdash_{LH} \Delta, \varphi_1, \varphi_1 \to \psi_1, \varphi \to \psi \Rightarrow \vartheta.$$

□

Theorem 1.8.3. *If* $\vdash \Gamma \Rightarrow \vartheta$, *then* $\vdash_{LH} \Gamma \Rightarrow \vartheta$.

Proof. By the Cut Elimination Theorem in Section 1.7 it suffices to prove that $\vdash^{\alpha}_{0} \Gamma \Rightarrow \vartheta$ implies $\vdash_{LH} \Gamma \Rightarrow \vartheta$. This is done by induction on α. Since the claim is obvious for the rules *Axiom* and $\to$–*right* and since *Cut* cannot occur, we only treat the rule $\to$–*left*.

Case 1. Let $\Gamma = \Delta, \varphi \to \psi$ with φ atomic and assume that

$$\vdash^{\alpha}_{0} \Delta, \varphi \to \psi \Rightarrow \vartheta$$

was inferred from

$$\vdash^{\alpha-1}_{0} \Delta, \varphi \to \psi \Rightarrow \vartheta \tag{1.9}$$

and

$$\vdash^{\alpha-1}_{0} \Delta, \varphi \to \psi, \psi \Rightarrow \vartheta. \tag{1.10}$$

From 1.10 we get by the Inversion Lemma 1.7.3, Part ii

$$\vdash^{\alpha-1}_{0} \Delta, \psi \Rightarrow \vartheta. \tag{1.11}$$

By induction hypothesis we can replace $\vdash^{\alpha-1}_{0}$ in 1.9 and 1.11 by $\vdash_{LH}$. Now the Lemma yields

$$\vdash_{LH} \Delta, \varphi \to \psi \Rightarrow \vartheta.$$

Case 2. Let $\Gamma = \Delta, (\varphi \to \psi) \to \chi$ and assume that

$$\vdash^{\alpha}_{0} \Delta, (\varphi \to \psi) \to \chi \Rightarrow \vartheta$$

was inferred from

$$\vdash^{\alpha-1}_{0} \Delta, (\varphi \to \psi) \to \chi \Rightarrow \varphi \to \psi \tag{1.12}$$

and

$$\vdash^{\alpha-1}_{0} \Delta, (\varphi \to \psi) \to \chi, \chi \Rightarrow \vartheta. \tag{1.13}$$

From 1.12 we get by the Inversion Lemma 1.7.3, Part iii

$$\vdash_0^{\alpha-1} \Delta, \varphi, \psi \to \chi \Rightarrow \varphi \to \psi$$

and hence

$$\vdash_0^{\alpha-1} \Delta, \varphi, \psi \to \chi \Rightarrow \psi. \tag{1.14}$$

From 1.13 we get by the Inversion Lemma 1.7.3, Part ii

$$\vdash_0^{\alpha-1} \Delta, \chi \Rightarrow \vartheta. \tag{1.15}$$

By induction hypothesis we can replace $\vdash_0^{\alpha-1}$ in 1.14 and 1.15 by $\vdash_{LH}$. Now H$\to$–left–$\to$ yields

$$\vdash_{LH} \Delta, (\varphi \to \psi) \to \chi \Rightarrow \vartheta. \Box$$

It is also possible to prove the Theorem directly for natural deductions in minimal implicational logic. We sketch the proof. So let a normal derivation of ϑ from assumptions Γ be given. We may assume that in any branch (see Section 2.4) of this normal derivation the minimal formula (see Section 2.4) is atomic, and use induction on the length of this derivation.

Case 1. $\Gamma = \Delta, \varphi \to \psi$ with φ atomic. In

$$\dfrac{\varphi \to \psi \qquad \begin{matrix} \vert \\ \varphi \end{matrix}}{\begin{matrix} \psi \\ \vert \\ \vartheta \end{matrix}}$$

we can apply the induction hypothesis to the subderivations of φ from $\Delta, \varphi \to \psi$ and of ϑ from Δ, ψ (any assumption $\varphi \to \psi$ here can be cancelled, since we already have assumed ψ). So we get

$$\vdash_{LH} \Delta, \varphi \to \psi \Rightarrow \varphi \quad \text{and} \quad \vdash_{LH} \Delta, \psi \to \vartheta,$$

and the claim follows by the Lemma.

Case 2. $\Gamma = \Delta, (\varphi \to \psi) \to \chi$. Replace an uppermost occurrence of the assumption $(\varphi \to \psi) \to \chi$

$$\dfrac{(\varphi \to \psi) \to \chi \qquad \dfrac{\dfrac{\begin{matrix} \vert \\ \varphi \to \psi \end{matrix} \qquad \varphi}{\psi}}{\varphi \to \psi}}{\begin{matrix} \chi \\ \vert \\ \vartheta \end{matrix}}$$

by

$$\dfrac{\varphi \to \psi \qquad \dfrac{\begin{matrix} | \\ \varphi \to \psi \end{matrix} \qquad \varphi}{\psi}}{\begin{matrix} \chi \\ | \\ \vartheta \end{matrix}}$$

Apply the induction hypothesis to the subderivation of ψ from $\Delta, \varphi, \psi \to \chi$, and of ϑ from Δ, χ (any assumption $(\varphi \to \psi) \to \chi$ here can be cancelled, since we already have assumed χ). So we get

$$\vdash_{LH} \Delta, \varphi, \psi \to \chi \Rightarrow \psi \quad \text{and} \quad \vdash_{LH} \Delta, \chi \Rightarrow \vartheta,$$

and the claim follows by an application of H$\to$–left–$\to$.

2. Normalization for first–order logic

We restrict our attention to the $\to\forall$–fragment of first–order logic with just introduction and elimination rules for both symbols, i.e. with minimal logic formulated in natural deduction style. This restriction does not mean a loss in generality, since it is well known that full classical first–order can be embedded in this system; the argument for that fact is sketched in Section 2.1. Equality is not treated as a logical symbol, but can be added via suitable equality axioms.

We extend our results and estimates on normalization to first–order logic by the method of collapsing types. Applications include the subformula property, Herbrand's theorem and the interpolation theorem.

2.1. The $\to\forall$–fragment as a typed λ–calculus

Assume that a fixed (at most countable) supply of function variables $f, g, h, \ldots$ and predicate variables $P, Q, \ldots$ is given, each with an arity ≥ 0. *Terms* are built up from object variables x, y, z by means of $fr_1 \ldots r_m$. *Formulas* are built up from prime formulas $Pr_1 \ldots r_m$ by means of $(\varphi \to \psi)$ and $\forall x \varphi$. *Derivations* are built up from assumption variables x^φ, y^φ by means of the rule $\to^+$ of implication introduction

$$(\lambda x^\varphi r^\psi)^{\varphi \to \psi},$$

the rule $\to^-$ of implication elimination

$$(t^{\varphi \to \psi} s^\varphi)^\psi,$$

the rule $\forall^+$ of $\forall$–introduction

$$(\lambda x r^\varphi)^{\forall x \varphi},$$

provided that no assumption variable y^ψ free in r^φ has x free in its type ψ, and finally the rule $\forall^-$ of $\forall$–elimination

$$(t^{\forall x \varphi} s)^{\varphi_x[s]}.$$

Each of the rules $\to^+, \forall^+$ and $\forall^-$ has a uniquely determined derivation as its *premiss*, whereas $\to^-$ has the two derivations $t^{\varphi\to\psi}$ and s^φ as premisses. Here $t^{\varphi\to\psi}$ is called the *main premiss* and s^φ is called the *side premiss*.

As an example we give a derivation of

$$\forall x(Px \to Qx) \to (\forall x Px \to \forall x Qx).$$

Such a derivation is

$$\lambda u^{\forall x(Px \to Qx)} \lambda v^{\forall x Px} \lambda x((ux)(vx)).$$

Derivations can be easily written in the more usual tree form. We will continue to use the word term for derivations (as long as this does not lead to confusion with the notion of (object) term inherent in first–order logic), and type for formula.

Note that our ($\to\forall$–fragment of) minimal logic contains full classical first–order logic. This can be seen as follows:

1. Choose a particular propositional variable and denote it by $\bot$ (falsity). Associate with any formula φ in the language of classical first–order logic a finite list $\vec{\varphi}$ of formulas in our $\to\forall$–fragment, by induction on φ:

$$\begin{aligned}
P\vec{r} &\mapsto P\vec{r} \\
\neg\varphi &\mapsto \vec{\varphi} \to \bot \\
\varphi \to \psi &\mapsto \vec{\varphi} \to \psi_1, \ldots, \vec{\varphi} \to \psi_n \\
\varphi \land \psi &\mapsto \vec{\varphi}, \vec{\psi} \\
\varphi \lor \psi &\mapsto (\vec{\varphi} \to \bot), (\vec{\psi} \to \bot) \to \bot \\
\forall x \varphi &\mapsto \forall x \varphi_1, \ldots, \forall x \varphi_m \\
\exists x \varphi &\mapsto \forall x(\vec{\varphi} \to \bot) \to \bot
\end{aligned}$$

2. In any model $\mathcal{M}$ where $\bot$ is interpreted by falsity, we clearly have that a formula φ in the language of full first–order logic holds under an assignment α iff all formulas in the assigned sequence $\vec{\varphi}$ hold under α.

3. Our derivation calculus for the $\to\forall$–fragment is complete in the following sense: A formula φ is derivable from stability assumptions

$$\forall\vec{x}(\neg\neg P\vec{x} \to P\vec{x})$$

for all predicate symbols P in φ iff φ is valid in any model under any assignment.

2.2. Strong normalization

Here we use the method of collapsing types (cf.(Troelstra and van Dalen 1988,p.560)) to transfer our results and estimates concerning strong normalization from implicational logic to first–order logic.

The notions concerning conversion introduced in Section 1.2 can be easily extended to first–order logic. In particular, we have

$$(\lambda\vec{x}x.r)\vec{s}s \quad \text{converts into} \quad (\lambda\vec{x}.r_x[s])\vec{s},$$

where the variables $\vec{x}, x$ now can be either assumption variables or else object variables. The rules generating the relation $r \to r'$ are extended by requiring $r \to r$ for object terms r of our first–order logic. Again a derivation is said to be in *normal form* if it does not contain a convertible subderivation.

For any formula φ of first–order logic we define its *collapse* φ^c by

$$\begin{aligned} (P\vec{r})^c &\equiv P \\ (\varphi \to \psi)^c &\equiv \varphi^c \to \psi^c \\ (\forall x\varphi)^c &\equiv \top \to \varphi^c \end{aligned}$$

where $\top :\equiv \bot \to \bot$ with $\bot$ a fixed propositional variable (i.e. $\top$ means truth). The *level* of a formula φ of first–order logic is defined to be the level of its collapse φ^c. For any derivation r^ψ in first–order logic we can now define its *collapse* $(r^\psi)^c$. It is plain from this definition that for any derivation r^ψ in first–order logic with free assumption variables $x_1^{\varphi_1}, \dots, x_m^{\varphi_m}$ the collapse $(r^\psi)^c$ is a derivation $(r^c)^{\psi^c}$ in implicational logic with free assumption variables $x_1^{\varphi_1^c}, \dots, x_m^{\varphi_m^c}$.

$$\begin{aligned} (x^\varphi)^c &\equiv x^{\varphi^c} \\ (\lambda x^\varphi r)^c &\equiv \lambda x^{\varphi^c} r^c \\ (t^{\varphi\to\psi} s)^c &\equiv t^c s^c \\ (\lambda x r)^c &\equiv \lambda x^\top r^c \\ (t^{\forall x\varphi} s)^c &\equiv t^c(\lambda z^\bot z^\bot)^\top \end{aligned}$$

Note that for any derivation r^ψ, assumption variable x^φ and derivation s^φ we have that $r^c[s^c]$ is a derivation in implicational logic (where the substitution of s^c is done for the assumption variable x^{φ^c}), which is the collapse of $r[s]$. Also for any derivation r^ψ, object variable x and term s we have that $r_x[s]$ is a derivation of $\psi_x[s]$ with collapse $(r_x[s])^c \equiv r^c$.

Lemma 2.2.1. *If $r \to_1 r'$ in first–order logic, then $r^c \to_1 (r')^c$ in implicational logic.*

The proof is by induction on the generation of $r \to_1 r'$. We only treat the case

$$(\lambda x r)s \to_1 r_x[s].$$

If x is an assumption variable, then

$$\begin{aligned} ((\lambda x^{\varphi} r^{\psi}) s^{\varphi})^c &\equiv (\lambda x^{\varphi^c} r^c) s^c \\ &\to_1 r^c[s^c] \\ &\equiv (r[s])^c), \end{aligned}$$

by the note above. If x is an object variable, then

$$\begin{aligned} ((\lambda x^{\varphi}) s)^c &\equiv (\lambda x^{\top} r^c)(\lambda z^{\perp} z^{\perp})^{\top} \\ &\to_1 r^c \\ &\equiv (r[s])^c), \end{aligned}$$

again by the note above. $\square$

Hence from Theorem 1.5.1 we can conclude

Theorem 2.2.2. *Any derivation r in first–order logic is strongly normalizable.* $\square$

Also we can apply Theorem 1.6.1 to obtain an upper bound for the length of arbitrary reduction sequences.

Theorem 2.2.3. *Let r be a derivation in first–order logic of a formula of level 0, i.e. a prime formula. Let r^c be the collapse of r into implicational logic. Let m be a bound for the levels of subterms of r^c and $k \geq 2$ be a bound for the arities of subterms of r^c. Then the length of an arbitrary reduction sequence for r with respect to $\to_1$ is bounded by*

$$k^{2_m(m+2\cdot\mathrm{height}(r^c)+2k+2)}.$$

2.3. Uniqueness

The Church–Rosser Theorem and hence the uniqueness of the normal form for derivations in first–order logic can be proved exactly as in Section 1.3. We do not repeat this here.

2.4. Applications

Here we want to draw some conclusions from the fact that any derivation in first–order logic can be transformed into normal form. The arguments in this section are based on Prawitz' book (Prawitz 1965). We begin with an analysis of the form of normal derivations.

Let a derivation r be given. A sequence $r_1, \dots, r_m$ of subderivations of r is a *branch* if

1. r_1 is an assumption variable,
2. r_i is the main premiss of r_{i+1}, and
3. r_m is either the whole derivation r or else the side premiss of an instance of the rule $\to^-$ within r.

It is obvious that any subderivation of r belongs to exactly one branch. The *order* of the branch ending with the whole derivation r is defined to be 0, and if the order of the branch through the main premise t of some instance $t^{\varphi\to\psi}s^{\varphi}$ of the rule $\to^-$ in r is k, then the order of the branch ending with that s^{φ} is defined to be $k+1$.

The relation "φ is a *subformula* of ψ" is defined to be the transitive and reflexive closure of the relation "immediate subformula", defined by

1. φ and ψ are immediate subformulas of $\varphi \to \psi$,
2. $\varphi_x[r]$ is an immediate subformula of $\forall x\varphi$.

We will also need the notion "φ is a *strictly positive subformula* of ψ", which is defined to be the transitive and reflexive closure of the relation "immediate strictly positive subformula", defined by

1. ψ is an immediate strictly positive subformula of $\varphi \to \psi$,
2. $\varphi_x[r]$ is an immediate strictly positive subformula of $\forall x\varphi$.

In a normal derivation r any branch $r_1^{\varphi_1}, \dots, r_m^{\varphi_m}$ has a rather perspicious form: all elimination rules must come before all introduction rules. Hence, if i is maximal such that $r_i^{\varphi_i}$ ends with an elimination rule, then φ_i must be a strictly positive subformula of all φ_j for $j \neq i$. This φ_i is called the *minimal formula* of the branch. Also, any φ_j with $j \leq i$ is a strictly positive subformula of φ_1, and any φ_j with $j \geq i$ is a strictly positive subformula of φ_m.

Theorem 2.4.1. (Subformula property) *If r^{φ} is a normal derivation with free assumption variables among $x_1^{\varphi_1}, \dots, x_m^{\varphi_m}$ and s^{ψ} is a subderivation of r^{φ}, then ψ is a subformula of φ or of some φ_i.*

The proof is by induction on the order of branches in r, using the property of branches in normal derivations mentioned above. $\square$

We write $\varphi_1, \dots, \varphi_m \vdash \varphi$ to mean that there is a derivation r^{φ} with free assumption variables among $x_1^{\varphi_1}, \dots, x_m^{\varphi_m}$.

Theorem 2.4.2. (Herbrand) *Assume that $\forall \vec{x}_1\varphi_1, \dots, \forall \vec{x}_m\varphi_m \vdash \psi$ with quantifier–free φ_1, ..., φ_m, ψ. Then we can find $\vec{r}_{11}, \dots, \vec{r}_{1n_1}, \dots, \vec{r}_{m1}, \dots, \vec{r}_{mn_m}$ such that*

$$\varphi_1[\vec{r}_{11}], \dots, \varphi_1[\vec{r}_{1n_1}], \dots, \varphi_m[\vec{r}_{m1}], \dots, \varphi_m[\vec{r}_{mn_m}] \vdash \psi$$

Proof. To simplify notation let us assume $\forall x\varphi \vdash \psi$ with quantifier–free φ, ψ. By Section 2.2 we can construct from the given derivation a normal derivation r^ψ with free assumption variables among $x^{\forall x\varphi}$. By induction on the order of branches it is easy to see that any branch must end with the derivation of a quantifier–free formula and must begin with the rule $\forall^-$, i. e. with $x^{\forall x\varphi} r_i$. Now replace any such subderivation by $y_i^{\varphi[r_i]}$, with new assumption variables y_i. $\square$

Our next application is the Craig interpolation theorem. We shall use the notation $\varphi_1, \dots, \varphi_m \vdash^c \varphi$ (c for classical) to mean that there is a derivation r^φ with free assumption variables among $x_1^{\varphi_1}, \dots, x_m^{\varphi_m}$ and some stability assumptions $y^{\forall \vec{x}(\neg\neg P\vec{x} \to P\vec{x})}$ for P predicate variable in $\vec{\varphi}, \varphi$, where again $\neg\psi$ denotes $\psi \to \bot$ with a fixed propositional variable $\bot$.

Theorem 2.4.3. **(Interpolation)** *Assume $\Gamma, \Delta \vdash \varphi$. Then we can find a finite list $\vec{\gamma}$ of formulas such that*

$$\Gamma \vdash^c \vec{\gamma} \text{ and } \vec{\gamma}, \Delta \vdash^c \varphi$$

(where $\Gamma \vdash^c \vec{\gamma}$ means $\Gamma \vdash^c \gamma_i$ for each γ_i in $\vec{\gamma}$), and any object or predicate variable free in $\vec{\gamma}$ occurs free both in Γ and in Δ, φ.

For the proof we shall use a somewhat more explicit formulation of the theorem: Let r^φ be a derivation with free assumption variables among $\vec{u}^\Gamma, \vec{v}^\Delta$. Then we can find a finite list $\vec{r_1}^{\vec{\gamma}}$ of derivations with free assumption variables among $\vec{u}^\Gamma$ and stability assumptions and a derivation r_2^φ with free assumption variables among $\vec{y}^{\vec{\gamma}}, \vec{v}^\Delta$ and stability assumptions, such that any object or predicate variable free in $\vec{\gamma}$ occurs free both in Γ and in Δ, φ.

For brevity we shall not mention stability assumptions any more (they will only be used in Case $2b(ii)$ below), and write "r^φ with $\vec{u}^\Gamma$", to mean the derivation r^φ with free assumption variables among $\vec{u}^\Gamma$.

The proof is by induction on the height of the given derivation, which by Section 2.2 we can assume to be normal. We distinguish two cases according to whether it ends with an introduction rule (i. e. $\to^+$ or $\forall^+$) or with an elimination rule.

Case 1a. $(\lambda x^\varphi r^\psi)^{\varphi \to \psi}$ with $\vec{u}^\Gamma, \vec{v}^\Delta$. By induction hypothesis for r^ψ with $x^\varphi, \vec{u}^\Gamma, \vec{v}^\Delta$ we have $\vec{r}_1^{\vec{\gamma}}$ with $\vec{u}^\Gamma$ and r_2^ψ with $\vec{y}^{\vec{\gamma}}, x^\varphi, \vec{v}^\Delta$. An application of $\to^+$ to the latter derivation yields $(\lambda x^\varphi r_2^\psi)^{\varphi \to \psi}$ with $\vec{y}^{\vec{\gamma}}, \vec{v}^\Delta$.

Case 1b. $(\lambda x r^A)^{\forall x\varphi}$ with $\vec{u}^\Gamma, \vec{v}^\Delta$, where x is not free in Γ, Δ. By induction hypothesis for r^φ with $\vec{u}^\Gamma, \vec{v}^\Delta$ we have $\vec{r}_1^{\vec{\gamma}}$ with $\vec{u}^\Gamma$ and r_2^φ with $\vec{y}^{\vec{\gamma}}, \vec{v}^\Delta$. Since x is not free in Γ, we know that x is not free in $\vec{\gamma}$. An application of $\forall^+$ to the latter derivation yields $(\lambda x r_2^\varphi)^{\forall x\varphi}$ with $\vec{y}^{\vec{\gamma}}, \vec{v}^\Delta$.

Case 2a. $(w^{\chi \to \vartheta} s^\chi \vec{t})^A$ with $\vec{u}^\Gamma, \vec{v}^\Delta$.

Subcase i. $w^{\chi \to \vartheta}$ is among $\vec{u}^\Gamma$. By induction hypothesis for s^χ with $\vec{u}^\Gamma, \vec{v}^\Delta$ we have $\vec{s}_1^{\vec{\gamma}}$ with $\vec{v}^\Delta$ and s_2^χ with $\vec{y}^{\vec{\gamma}}, \vec{u}^\Gamma$. By induction hypothesis for $(u^\vartheta \vec{t})^\varphi$ with $u^\vartheta, \vec{u}^\Gamma, \vec{v}^\Delta$ we have $\vec{t}^{\vec{\delta}}$ with $u^\vartheta, \vec{u}^\Gamma$ and t_2^φ with $\vec{z}^{\vec{\delta}}, \vec{v}^\Delta$. From these derivations we obtain

$$(\lambda \vec{y}^{\vec{\gamma}} (\vec{t}_1^{\vec{\delta}})_u [w^{\chi \to \vartheta} s_2^\chi])^{\vec{\gamma} \to \vec{\delta}} \quad \text{with} \quad \vec{u}^\Gamma$$

and

$$(t_2^\varphi)_{\vec{z}} [\vec{x}^{\vec{\gamma} \to \vec{\delta}} \vec{s}_1^{\vec{\gamma}}] \quad \text{with} \quad \vec{x}^{\vec{\gamma} \to \vec{\delta}}, \vec{v}^\Delta,$$

where $\vec{\gamma} \to \vec{\delta}$ means $\vec{\gamma} \to \delta_1, \dots, \vec{\gamma} \to \delta_n$.

Subcase ii. $w^{\chi\to\vartheta}$ is among $\vec{v}^\Delta$. By induction hypothesis for s^χ with $\vec{u}^\Gamma, \vec{v}^\Delta$ we have $\vec{s}_1^{\vec{\gamma}}$ with $\vec{u}^\Gamma$ and s_2^χ with $\vec{y}^{\vec{\gamma}}, \vec{v}^\Delta$. By induction hypothesis for $(u^\vartheta \vec{t})^\varphi$ with $u^\vartheta, \vec{u}^\Gamma, \vec{v}^\Delta$ we have $\vec{t}_1^{\vec{\delta}}$ with $\vec{u}^\Gamma$ and t_2^φ with $\vec{z}^{\vec{\Gamma}}, u^\vartheta, \vec{v}^\Delta$. From these derivations we obtain

$$(\vec{s}_1, \vec{t}_1)^{\vec{\gamma},\vec{\delta}} \quad \text{with} \quad \vec{u}^\Gamma$$

and

$$(t_2^\varphi)_u[w^{\chi\to\vartheta} s_2^\chi] \quad \text{with} \quad \vec{y}^{\vec{\gamma}}, \vec{z}^{\vec{\delta}}, \vec{v}^\Delta.$$

Case 2b. $w^{\forall x\chi} s\vec{t}$ with $\vec{u}^\Gamma, \vec{v}^\Delta$.

Subcase i. $w^{\forall x\chi}$ is among $\vec{u}^\Gamma$. By induction hypothesis for $(u^{\chi[s]}\vec{t})^\varphi$ with $u^{\chi[s]}$, $\vec{u}^\Gamma$, $\vec{v}^\Delta$ we have $\vec{t}_1^{\vec{\gamma}}$ with $u^{\chi[s]}, \vec{u}^\Gamma$ and t_2^φ with $\vec{y}^{\vec{\gamma}}, \vec{v}^\Delta$. Let $\vec{z}$ be all variables free in $\vec{\gamma}$ that are in s, but not free in Γ. We now construct derivations

$$(\lambda\vec{z}(\vec{t}_1^{\vec{\gamma}})_u[w^{\forall x\chi} s])^{\forall\vec{z}\vec{\gamma}} \quad \text{with} \quad \vec{u}^\Gamma$$

and

$$(t_2^\varphi)_{\vec{y}}[\vec{z}^{\forall\vec{z}\vec{\gamma}}\vec{z}] \quad \text{with} \quad \vec{x}^{\forall\vec{z}\vec{\gamma}}, \vec{v}^\Delta,$$

where $\forall\vec{z}\vec{\gamma}$ means $\forall\vec{z}\gamma_1, \dots, \forall\vec{z}\gamma_m$. Note that any object or predicate variable free in $\forall\vec{z}\vec{\gamma}$ is both free in Δ, φ and free in Γ.

Subcase ii. $w^{\forall x\chi}$ is among $\vec{v}^\Delta$. By induction hypothesis for $(u^{\chi[s]}\vec{t})^\varphi$ with $u^{\chi[s]}$, $\vec{u}^\Gamma$, $\vec{v}^\Delta$ we have $\vec{t}_1^{\vec{\delta}}$ with $\vec{u}^\Gamma$ and t_2^φ with $\vec{y}^{\vec{\delta}}, u^{C[s]}, \vec{v}^\Delta$. Let $\vec{z}$ be all variables free in $\vec{\delta}$ that are in s, but not free in Δ, φ. We now construct derivations

$$(\lambda v^{\forall\vec{z}\neg\vec{\delta}}(v\vec{z}\vec{t}_1))^{\neg\forall\vec{z}\neg\vec{\delta}} \quad \text{with} \quad \vec{u}^\Gamma$$

and

$$(t^{\neg\neg\varphi\to\varphi}\lambda v^{\neg\varphi}.x^{\neg\forall\vec{z}\neg\vec{\gamma}}\lambda\vec{z}\lambda\vec{y}(v(t_2^\varphi)_u[w^{\forall x\chi}s]))^\varphi \quad \text{with} \quad x^{\neg\forall\vec{z}\neg\vec{\delta}}, \vec{v}^\Delta$$

and stability assumptions (which are used to build $t^{\neg\neg\varphi\to\varphi}$). Note again that any object or predicate variable free in $\neg\forall\vec{z}\neg\vec{\delta}$ is both free in Γ and free in Δ, φ. $\square$

3. Normalization for arithmetic

3.1. Ordinal notations

We want to discuss the derivability and underivability of initial cases of transfinite induction in arithmetical systems. In order to do that we shall need some knowledge and notations for ordinals. Now we do not want to assume set theory here; hence we introduce a certain initial segment of the ordinal (the ordinals $< \varepsilon_0$) in a formal, combinatorial way, i.e. via ordinal notations. Our treatment is based on the Cantor normal form for ordinals; cf. (Bachmann 1955). We also introduce some elementary relations and operations for such ordinal notations, which will be used later.

We define the two notions

- α is an ordinal notation
- $\alpha < \beta$ for ordinal notations α, β

simultaneously by induction:

1. If $\alpha_m, \dots, \alpha_0$ are ordinal notations and $\alpha_m \geq \dots \geq \alpha_0$ (where $\alpha \geq \beta$ means $\alpha > \beta$ or $\alpha = \beta$), then
$$\omega^{\alpha_m} + \dots + \omega^{\alpha_0}$$
is an ordinal notation. Note that the empty sum denoted by 0 is allowed here.
2. If $\omega^{\alpha_m} + \dots + \omega^{\alpha_0}$ and $\omega^{\beta_n} + \dots + \omega^{\beta_0}$ are ordinal notations, then
$$\omega^{\alpha_m} + \dots + \omega^{\alpha_0} < \omega^{\beta_n} + \dots + \omega^{\beta_0}$$
iff there is an $i \geq 0$ such that $\alpha_{m-i} < \beta_{n-i}$, $\alpha_{m-i+1} = \beta_{n-i+1}, \dots, \alpha_m = \beta_n$, or else $m < n$ and $\alpha_m = \beta_n, \dots, \alpha_0 = \beta_{n-m}$

It is easy to see (by induction on the levels in the inductive definition) that $<$ is a linear order with 0 being the smallest element.

We shall use the notation 1 for ω^0, a for $\omega^0 + \dots + \omega^0$ with a copies of ω^0 and $\omega^\alpha a$ for $\omega^\alpha + \dots + \omega^\alpha$ again with a copies of ω^α.

We now define addition for ordinal notations:

$$\omega^{\alpha_m} + \dots + \omega^{\alpha_0} + \omega^{\beta_n} + \dots + \omega^{\beta_0} := \omega^{\alpha_m} + \dots + \omega^{\alpha_i} + \omega^{\beta_n} + \dots + \omega^{\beta_0}$$

where i is minimal such that $\alpha_i \geq \beta_n$.

It is easy to see that $+$ is an associative operation which is strictly monotonic in the second argument and weakly monotonic in the first argument. Note that $+$ is not commutative: $1 + \omega = \omega \neq \omega + 1$.

The *natural* (or Hessenberg) sum of two ordinal notations is defined by

$$(\omega^{\alpha_m} + \dots + \omega^{\alpha_0}) \# (\omega^{\beta_n} + \dots + \omega^{\beta_0}) := \omega^{\gamma_{m+n}} + \dots + \omega^{\gamma_0},$$

where $\gamma_{m+n}, \dots, \gamma_0$ is a decreasing permutation of $\alpha_m, \dots, \alpha_0, \beta_n, \dots, \beta_0$.

Again it is easy to see that # is associative, commutative and strictly monotonic in both arguments.

We will also need to know how ordinal notations of the form $\beta + \omega^\alpha$ can be approximated from below. First note that

$$\delta < \alpha \to \beta + \omega^\delta a < \beta + \omega^\alpha.$$

Furthermore, for any $\gamma < \beta + \omega^\alpha$ we can find a $\delta < \alpha$ and an a such that

$$\gamma < \beta + \omega^\delta a.$$

We now define 2^α for ordinal notations α. Let $\alpha_m \geq \cdots \alpha_0 \geq \omega > k_n \geq \cdots \geq k_1 > 0$. Then

$$2^{\omega^{\alpha_m}+\cdots+\omega^{\alpha_0}+\omega^{k_n}+\cdots+\omega^{k_1}+\omega^0 a} := \omega^{\omega^{\alpha_m}+\cdots+\omega^{\alpha_0}+\omega^{k_n-1}+\cdots+\omega^{k_1-1}} 2^a.$$

It is easy to see that $2^{\alpha+1} = 2^\alpha + 2^\alpha$ and that 2^α is strictly monotonic in α.

In order to work with ordinal notations in a purely arithmetical system we set up a bijection between ordinal notations and nonnegative integers (i.e., a Gödel numbering). For its definition it is useful to refer to ordinal notations in the form

$$\omega^{\alpha_m} a_m + \cdots + \omega^{\alpha_0} a_0 \quad \text{with} \quad \alpha_m > \cdots > \alpha_0.$$

For any ordinal notation α we define its Gödel number $|\alpha|$ inductively by

$$|0| := 0,$$

$$|\omega^{\alpha_m} a_m + \cdots + \omega^{\alpha_0} a_0| := (\prod_{i \leq m} p_{|\alpha_i|}^{a_i}) - 1.$$

For any nonnegative integer x we define its corresponding ordinal notation $\mathrm{o}(x)$ inductively by

$$\mathrm{o}(0) = 0$$

$$\mathrm{o}((\prod_{i \leq m} p_i^{a_i}) - 1) = \sum_{i \leq m} \omega^{\mathrm{o}(i)} a_i$$

where the sum is to be understood as the natural sum.

Lemma 3.1.1.

1. $\mathrm{o}(|\alpha|) = \alpha$,
2. $|\mathrm{o}(x)| = x$.

This can be proved easily by induction. $\square$

Hence we have a bijection between ordinal notations and nonnegative integers. Using this bijection we can transfer our relations and operations on ordinal notations to computable relations and operations on nonnegative integers. We will use the notations

$$\begin{aligned} x \prec y \quad & \text{for} \quad o(x) < o(y), \\ \omega^x \quad & \text{for} \quad |\omega^{o(x)}|, \\ x \oplus y \quad & \text{for} \quad |o(x) + o(y)|. \end{aligned}$$

3.2. Provability of initial cases of transfinite induction

We now set up some formal systems of arithmetic and derive initial cases of the principle of transfinite induction in them, i.e. of

$$\forall x(\forall y \prec x : Py \to Px) \to \forall x \prec a : Px$$

for some numeral a and a predicate variable P. In Section 3.4 we will see that our results here are optimal in the sense that for larger segments of the ordinals transfinite induction is underivable. All these results are due to (Gentzen 1943).

Our arithmetical systems are based on a fixed (possibly countably infinite) supply of function constants and predicate constants which are assumed to denote fixed functions and predicates on the nonnegative integers for which a computation procedure is known. Among the function constants there must be a constant S for the successor function and 0 for (the 0–place function) zero. Among the predicate constants there must be a constant $=$ for equality and $\perp$ for (the 0–place predicate) falsity. In order to formulate the general principle of transfinite induction we also assume that predicate variables $P, Q, \dots$ are present.

Terms are built up from object variables x, y, z by means of $fr_1 \dots r_m$, where f is a function constant. We identify closed terms which have the same value; this is a convenient way to express in our formal systems the assumption that for each function constant a computation procedure is known. Terms of the form $SS \dots S0$ are called *numerals*. We use the notation S^i0 or even i for them. *Formulas* are built up from prime formulas $Pr_1 \dots r_m$ with P a predicate constant or a predicate variable by means of $(\varphi \to \psi)$ and $\forall x\varphi$. As usual we abbreviate $\varphi \to \perp$ by $\neg\varphi$.

The *axioms* of our arithmetical systems will always include the Peano–axioms

$$\forall xy(Sx = Sy \to x = y),$$

$$\forall x(Sx \neq 0).$$

Any instance of the induction scheme

$$\varphi[0], \forall x(\varphi[x] \to \varphi[Sx]) \to \forall x\varphi[x]$$

with φ an arbitrary formula is an axiom of full arithmetic Z. We will also consider subsystems Z_k of Z where the formulas φ in the induction scheme are restricted to Π^0_k–formulas; the latter notion is defined inductively, as follows.

1. Any prime formula $P\vec{r}$ is a Π^0_k–formula, for any $k \geq 1$.
2. If φ is quantifier–free and ψ is a Π^0_k–formula, then $\varphi \to \psi$ is a Π^0_k–formula.
3. If φ is a Π^0_k–formula and ψ is a Π^0_l–formula, then $\varphi \to \psi$ is a Π^0_p–formula with $p = \max(k + 1, l)$.
4. If φ is a Π^0_k–formula, then so is $\forall x\varphi$.

Note that a formula is a Π^0_k–formula iff it is logically equivalent to a formula with a prefix of k alternating quantifiers beginning with $\forall$ and a quantifier–free kernel. For example, $\forall x\exists y\forall zPxyz$ is a Π^0_3–formula. In addition, in any arithmetical system we have the equality axioms

$$\forall x(x = x),$$

$$\forall \vec{x}\vec{y}(x_1 = y_1, \ldots, x_m = y_m \to f\vec{x} = f\vec{y}),$$
$$\forall \vec{x}\vec{y}(x_1 = y_1, \ldots, x_m = y_m, P\vec{x} \to P\vec{y})$$

for any function constant f and predicate constant or predicate variable P. We also require for any such P the stability axioms

$$\forall \vec{x}(\neg\neg P\vec{x} \to P\vec{x}).$$

We express our assumption that for any predicate constant a decision procedure is known by adding the axiom

$$P(S^{i_1}0)\ldots(S^{i_m}0)$$

whenever $P\vec{i}$ is true, and

$$\neg P(S^{i_1}0)\ldots(S^{i_m}0)$$

whenever $P\vec{i}$ is false.

We finally allow in any of our arithmetical systems an arbitrary supply of true Π^0_1–formulas as axioms. Our (positive and negative) results concerning initial cases of transfinite recursion will not depend on which of those axioms we have chosen, except that for the positive results we always assume

$$\forall x(x \not\prec 0) \tag{3.1}$$

$$\forall yz(z \prec y \oplus \omega^0, z \not\prec y, z \neq y \to \bot) \tag{3.2}$$

$$\forall x(x \oplus 0 = x) \tag{3.3}$$

$$\forall xyz(x \oplus (y \oplus z) = (x \oplus y) \oplus z) \tag{3.4}$$

$$\forall x(0 \oplus x = x) \tag{3.5}$$

$$\forall x(\omega^x 0 = 0) \tag{3.6}$$

$$\forall xy(\omega^x(Sy) = \omega^x y \oplus \omega^x) \tag{3.7}$$

$$\forall xyz(z \prec y \oplus \omega^x, x \neq 0 \to z \prec y \oplus \omega^{fxyz}(gxyz)) \tag{3.8}$$

$$\forall xyz(z \prec y \oplus \omega^x, x \neq 0 \to fxyz \prec x) \tag{3.9}$$

where in 3.9 f and g are function constants.

Theorem 3.2.1. (Gentzen) *Transfinite induction up to ω_n (with $\omega_1 := \omega, \omega_{n+1} := \omega^{\omega_n}$) i.e. the formula*

$$\forall x(\forall y \prec x : \varphi[y] \to \varphi[x]) \to \forall x \prec \omega_n : \varphi[x]$$

is derivable in Z.

Proof: To any formula φ we assign a formula φ^+ (with respect to a fixed variable x) by

$$\varphi^+ :\equiv \forall y(\forall z \prec y : \varphi_x[z] \to \forall z \prec y \oplus \omega^x : \varphi_x[z]).$$

We first show

$$\varphi \text{ is progressive } \to \varphi^+ \text{ is progressive},$$

where "ψ is *progressive*" means $\forall x(\forall y \prec x : \psi[y] \to \psi[x])$. So assume that φ is progressive and

$$\forall y \prec x : \varphi^+[y]. \tag{3.10}$$

We have to show $\varphi^+[x]$. So assume further

$$\forall z \prec y : \varphi[z] \tag{3.11}$$

and $z \prec y \oplus \omega^x$. We have to show $\varphi[z]$. Case $x = 0$. From $z \prec y \oplus \omega^0$ we have by 3.2 $z \prec y \vee z = y$. If $z \prec y$, then $\varphi[z]$ follows from 3.11, and if $z = y$, then $\varphi[z]$ follows from 3.11 and the progressiveness of φ. Case $x \neq 0$. From $z \prec y \oplus \omega^x$ we obtain $z \prec y \oplus \omega^{fxyz} gxyz$ by 3.8 and $fxyz \prec x$ by 3.9. From 3.10 we obtain $\varphi^+[fxyz]$. By the definition of φ^+ we get

$$\forall u \prec y \oplus \omega^{fxyz} v : \varphi[u] \to \forall u \prec (y \oplus \omega^{fxyz} v) \oplus \omega^{fxyz} : \varphi[u]$$

and hence, using 3.4 and 3.7

$$\forall u \prec y \oplus \omega^{fxyz} v : \varphi[u] \to \forall u \prec y \oplus \omega^{fxyz}(Sv) : \varphi[u].$$

Also from 3.11 and 3.6, 3.3 we obtain

$$\forall u \prec y \oplus \omega^{fxyz} 0 : \varphi[u].$$

Using an appropriate instance of the induction scheme we can conclude

$$\forall u \prec y \oplus \omega^{fxyz} gxyz : \varphi[u]$$

and hence $\varphi[z]$.

We now show, by induction on n, how to obtain a derivation of

$$\forall x(\forall y \prec x : \varphi[y] \to \varphi[x]) \to \forall x \prec \omega_n : \varphi[x].$$

So assume the left–hand side, i.e. assume that φ is progressive. Case 0. From $x \prec \omega_0$ we get $x = 0$ by 3.5, 3.2 and 3.1, and $\varphi[0]$ follows from the progressiveness of φ by 3.1. Case $n+1$. Since φ is progressive, by what we have shown above also φ^+ is progressive. Applying the induction hypothesis to φ^+ yields $\forall x \prec \omega_n : \varphi^+[x]$, and hence $\varphi^+[\omega_n]$ by the progressiveness of $\varphi^+[x]$. Now the definition of φ^+ (together with 3.1 and 3.5) yields $\forall z \prec \omega^{\omega_n} : \varphi[z]$. $\square$

Note that in these derivations the induction scheme was used for formulas of unbounded complexity.

We now want to refine Theorem 3.2.1 to a corresponding result for the subsystems Z_k of Z. Note first that if φ is a Π^0_k–formula, then the formula φ^+ constructed in the proof of Theorem 3.2.1 is a Π^0_{k+1}–formula, and for the proof of

$$\varphi \text{ is progressive } \to \varphi^+ \text{ is progressive}$$

we have used induction with a Π^0_k induction formula.

Now let φ be a Π^0_1–formula, and let $\varphi^0 :\equiv \varphi, \varphi^{i+1} :\equiv (\varphi^i)^+$. Then φ^k is a Π^0_{k+1}–formula, and hence in Z_k we can derive that if φ is progressive, then also $\varphi^1, \varphi^2, \dots \varphi^k$ are all progressive. Let $\omega_1[m] := m, \omega_{i+1}[m] = \omega^{\omega_i[m]}$. Since in Z_k we can derive that φ^k is progressive, we can also derive $\varphi^k[0], \varphi^k[1], \varphi^k[2]$ and generally $\varphi^k[m]$ for any m, i.e. $\varphi^k[\omega_1[m]]$. But since

$$\varphi^k \equiv (\varphi^{k-1})^+ \equiv \forall y(\forall z \prec y : \varphi^{k-1}[z] \to \forall z \prec y \oplus \omega^x : \varphi^{k-1}[z]),$$

we first get (with $y = 0$) $\forall z \prec \omega_2[m] : \varphi^{k-1}[z]$ and then $\varphi^{k-1}[\omega_2[m]]$ by the progressiveness of φ^{k-1}. Repeating this argument we finally obtain $\varphi^0[\omega_{k+1}[m]]$. Hence we have

Theorem 3.2.2. *Let φ be a Π^0_1–formula. Then in Z_k we can derive transfinite induction for φ up to $\omega_{k+1}[m]$ for any m, i.e.*

$$Z_k \vdash \forall x(\forall y \prec x : \varphi[y] \to \varphi[x]) \to \forall x \prec \omega_{k+1}[m] : \varphi[x]. \square$$

If more generally we start out with a Π^0_l–formula φ instead, where $1 \leq l \leq k$, then a similar argument yields the following result of (Parsons 1973)

Theorem 3.2.3. *Let φ be a Π^0_l–formula, $1 \leq l \leq k$. Then in Z_k we can derive transfinite induction for φ up to $\omega_{k+2-l}[m]$ for any m, i.e.*

$$Z_k \vdash \forall x(\forall y \prec x : \varphi[y] \to \varphi[x]) \to \forall x \prec \omega_{k+2-l}[m] : \varphi[x]. \square$$

Our next aim is to prove that these bounds are sharp. More precisely, we will show that in Z (no matter how many true Π^0_1–formulas we have added as axioms) one cannot derive transfinite induction up to ε_0, i.e. the formula

$$\forall x(\forall y \prec x : Py \to Px) \to \forall x Px$$

with a free predicate variable P, and that in Z_k one cannot derive transfinite induction up to ω_{k+1}, i.e. the formula

$$\forall x(\forall y \prec x : Py \to Px) \to \forall x \prec \omega_{k+1} : Px.$$

This will follow from the method of normalization applied to arithmetical systems, which we have to develop first.

3.3. Normalization for arithmetic with the ω–rule

We will show in Section 3.5 that a normalization theorem does not hold for a system of arithmetic like Z in Section 3.2, in the sense that for any formula φ derivable in Z there is a derivation of the same formula φ in Z which only uses formulas of a level bounded by the level of φ. The reason for this failure is the presence of the induction axioms, which can be of arbitrary level.

Here we remove that obstacle against normalization and replace the induction axioms by a rule with infinitely many premisses, the so–called ω–rule (suggested by Hilbert and studied by Lorenzen, Novikov and Schütte), which allows to conclude $\forall x\varphi[x]$ from $\varphi[0]$, $\varphi[1]$, $\varphi[2]$, ….

Clearly this ω–rule can also be used to replace the rule $\forall^+$. As a consequence we do not need to consider free object variables.

So we introduce the system Z^∞ of ω–arithmetic as follows. Z^∞ has the same language and — apart from the induction axioms — the same axioms as Z. Derivations in Z^∞ are infinite objects; they are built up from assumption variables x^φ, y^ψ and constants ax^φ for any axiom φ of Z other than an induction axiom by means of the rules

$$(\lambda x^\varphi r^\psi)^{\varphi \to \psi}$$

$$(t^{\varphi\to\psi} s^\varphi)^\psi$$

$$\langle r_i^{\varphi[i]} \rangle_{i<\omega}^{\forall x \varphi}$$

$$(t^{\forall x \varphi} i)^{\varphi[i]}$$

denoted by $\to^+, \to^-, \omega$ and $\forall^-$, respectively.

More precisely, we define the notion of an $\vec{x}$–derivation (i. e. a derivation in Z^∞ with free assumption variables among $\vec{x}$) of height $\leq \alpha$ and degree $\leq k$ inductively, as below.

Note that derivations are infinite objects now. They may be viewed as mappings from finite sequences of natural numbers (= nodes in the derivation tree) to lists of data including the formula appearing at that node, the rule applied last, a list of assumption variables including all those free in the subderivation (starting at that node), a bound on the height of the subderivation, and a bound on the degree of the subderivation.

Intuitively, the degree of a derivation is the least number $\geq$ the level of any subderivation $\lambda x r$ in a context $(\lambda x r)s$ or $\langle r_i \rangle_{i<\omega}$ in a context $\langle r_i \rangle_{i<\omega} j$, where the level of a derivation is the level of its type, i.e. the formula it derives. This notion of a degree is needed for the normalization proof we give below.

- $*$ Any assumption variable x^φ and any axiom ax^φ is an $\vec{x}$–derivation of height $\leq \alpha$ and degree $\leq k$, for any list $\vec{x}$ of assumption variables (containing x in the first case), ordinal α and number k.
- $\to^+$ If r^ψ is an $\vec{x}, x, \vec{y}$–derivation of height $\leq \alpha_0 < \alpha$ and degree $\leq k$, then $(\lambda x^\varphi r^\psi)^{\varphi \to \psi}$ is an $\vec{x}, \vec{y}$–derivation of height $\leq \alpha$ and degree $\leq k$.
- $\to^-$ If $t^{\varphi \to \psi}$ and s^φ are $\vec{x}$–derivations of heights $\leq \alpha_i < \alpha$ and degrees $\leq k_i \leq k$ $(i = 1, 2)$, then $(t^{\varphi \to \psi} s^\varphi)^\psi$ is an $\vec{x}$–derivation of height $\leq \alpha$ and degree $\leq m$ with $m = \max(k, \mathrm{lev}(\varphi \to \psi))$, if $t^{\varphi \to \psi}$ is generated by the rule $\to^+$, or of degree $\leq k$ otherwise.
- ω If $r_i^{\varphi[i]}$ are $\vec{x}$–derivations of heights $\leq \alpha_i < \alpha$ and degrees $\leq k_i \leq k$ $(i < \omega)$, then $\langle r_i^{\varphi[i]} \rangle_{i<\omega}^{\forall x \varphi}$ is an $\vec{x}$–derivation of height $\leq \alpha$ and degree $\leq k$.
- $\forall^-$ If $t^{\forall x \varphi}$ is an $\vec{x}$–derivation of height $\leq \alpha_0 < \alpha$ and degree $\leq k$, then $(t^{\forall x \varphi} i)^{\varphi[i]}$ is an $\vec{x}$–derivation of height $\leq \alpha$ and degree $\leq m$ with $m = \max(k, \mathrm{level}\ \forall x \varphi)$, if $t^{\forall x \varphi}$ is generated by the rule ω, or of degree $\leq k$ otherwise.

We now embed our systems Z_k (i. e. arithmetic with induction restricted to Π_k^0–formulas) and hence Z into Z^∞.

Lemma 3.3.1. *Let r^ψ be a derivation in Z_k with free assumption variables among $\vec{x}^{\vec{\varphi}}$ which contains $\leq m$ instances of the induction scheme all with induction formulas of level $\leq k$. Let σ be a substitution of numerals for object variables such that $\vec{\varphi}\sigma, \psi\sigma$ do not contain free object variables. Then we can find an $\vec{x}^{\vec{\varphi}\sigma}$–derivation $(r_\infty)^{\psi\sigma}$ in Z^∞ of height $\leq \omega^m + h$ for some $h < \omega$ and degree $\leq k$.*

Proof. First note that from any normal derivation in first–order logic we can construct a normal derivation r_0^ψ with the same free assumption variable $\vec{x}^{\vec{\varphi}}$, such that in r_0^ψ any branch has a prime formula as its minimal formula (cf. Section 2.4). For if φ is a minimal formula which is not prime we can first apply elimination rules until a prime formula is reached and later build a up again by the corresponding introduction rules.

The lemma is proved by induction on the height of the given derivation r. By the Normalization Theorem 2.2.3 and the note above we can assume that r is normal with prime minimal formulas. The only case which requires some argument is when r consists of two applications of $\to^-$ to an instance of the induction scheme. Then r must have the form

$$\mathrm{ax}^{\varphi[0],\forall x(\varphi[x]\to\varphi[Sx])\to\forall x\varphi[x]} s^{\varphi[0]}(\lambda x\lambda y^{\varphi[x]} t^{\varphi[Sx]}).$$

By induction hypothesis we obtain derivations

$$\begin{array}{rll} s_\infty^{\varphi[0]} & \text{of height} & \leq \omega^{m-1}+h_0 \\ t_\infty^{\varphi[1]}[s_\infty^{\varphi[0]}] & \text{of height} & \leq \omega^{m-1}\cdot 2+h_1, \\ t_\infty^{\varphi[2]}[t_\infty^{\varphi[1]}[s_\infty^{\varphi[0]}]] & \text{of height} & \leq \omega^{m-1}\cdot 3+h_2 \end{array}$$

and so on, all of degree $\leq k$. Combining all these derivations of $\varphi[i]$ as premisses of the ω–rule yields a derivation t_∞ of $\forall x\varphi[x]$ of height $\leq \omega^m$ and degree $\leq k$. □

A derivation is called *convertible* if it is of the form $(\lambda xr)s$ or else $\langle r_i\rangle_{i<\omega}j$, which can be converted into $r_x[s]$ or r_j, respectively. Here $r_x[s]$ is obtained from r by substituting s for all free occurences of x in r. A derivation is called *normal* if it does not contain a convertible subderivation. Note that a derivation of degree 0 must be normal.

We want to define an operation which by repeated conversions transforms a given derivation into a normal one with the same end formula and no more assumption variables. The methods employed in Sections 1 and 2 to achieve such a task have to be adapted properly in order to deal with the new situation of infinitary derivations. Here we give a particularly simple argument due to (Tait 1965).

Lemma 3.3.2. *If r is an $\vec{x}, x^\varphi, \vec{y}$–derivation of height $\leq \alpha$ and degree $\leq k$ and s^φ is an $\vec{x}, \vec{y}$–derivation of height $\leq \beta$ and degree $\leq l$, then $r_x[s]$ is an $\vec{x}, \vec{y}$–derivation of height $\leq \beta+\alpha$ and degree $\leq \max(k, l, \mathit{level}\, s)$.*

This is proved by a straightforward induction on the height of r. □

Lemma 3.3.3. *For any $\vec{x}$–derivation r^φ of height $\leq \alpha$ and degree $\leq k+1$ we can find an $\vec{x}$–derivation $(r^k)^\varphi$ of height $\leq 2^\alpha$ and degree $\leq k$.*

The proof is by induction on α. The only case which requires some argument is when r is of the form ts with t of height $\leq \alpha_1 < \alpha$ and s of height $\leq \alpha_2 < \alpha$. We first consider the subcase where $t^k = \lambda x t_1$ and $\mathrm{lev}(t) = k+1$. Then $\mathrm{lev}(s) \leq k$ by the definition of level, and hence $(t_1)_x[s^k]$ has degree $\leq k$ by Lemma 3.3.2. Furthermore, also by Lemma 3.3.2, $(t_1)_x[s^k]$ has height $\leq 2^{\alpha_2}+2^{\alpha_1} \leq 2^{\max(\alpha_2,\alpha_1)+1} \leq 2^\alpha$. Hence we can take $(ts)^k$ to be $(t_1)_x[s^k]$. If we are not in the above subcase, we can simply take $(ts)^k$ to be $t^k s^k$. This derivation clearly has height $\leq 2^\alpha$. Also it has degree $\leq k$, which can be seen as follows. If $\mathrm{lev}(t) \leq k$ we are done. If however $\mathrm{lev}(t) \geq k+2$, then t must be of the form $t_0 t_1 \ldots t_m$ for some assumption variable or axiom t_0 (since r has degree $\leq k+1$). But then t^k has the form $t_0 t_1^k \ldots t_m^k$ and we are done again. (To be completely precise, this last statement has to be added to the formulation of the Lemma above and proved simultaneously with it). □

As an immediate consequence we obtain

Theorem 3.3.4. (Normalization for Z^∞) *For any $\vec{x}$–derivation r^φ of height $\leq \alpha$ and degree $\leq k$ we can find a normal $\vec{x}$–derivation $(r^*)^\varphi$ of height $\leq 2_k\alpha$ (where $2_0\alpha = \alpha, 2_{m+1}\alpha = 2^{2_m^\alpha}$).*

3.4. Unprovable initial cases of transfinite induction

We now apply the technique of normalization for arithmetic with the ω–rule for a proof that transfinite induction up to ε_0 is underivable in Z, i.e. of

$$Z \not\vdash \forall x(\forall y \prec x : Py \to Px) \to \forall x Px$$

with a predicate variable P, and that transfinite induction up to ω_{k+1} is underivable in Z_k, i.e. of

$$Z_k \not\vdash \forall x(\forall y \prec x : Py \to Px) \to \forall x \prec \omega_{k+1} : Px.$$

Our proof is based on an idea of Schütte, which consists in adding a so–called *progression rule* to the infinitary systems. This rule allows to conclude Pj (where j is any numeral) from all Pi for $i \prec j$.

More precisely, we define the notion of an $\vec{x}$–derivation in $Z^\infty +$ Prog(P) of height $\leq \alpha$ and degree $\leq k$ by the inductive clauses of Section 3.2 and the additional clause Prog(P):

If r_i^{Pi} are $\vec{x}$–derivations of heights $\leq \alpha_i < \alpha$ and degrees $\leq k_i \leq k$ $(i \prec j)$, then $\langle r_i^{Pi} \rangle_{i \prec j}^{Pj}$ is an $\vec{x}$–derivation of height $\leq \alpha$ and degree $\leq k$.

Since this progression rule only deals with derivations of prime formulas it does not affect the degrees of derivations. Hence the proof of normalization for Z^∞ carries over unchanged to $Z^\infty +$ Prog(P). In particular we have

Lemma 3.4.1. *For any $\vec{x}$–derivation r^φ in $Z^\infty +$ Prog(P) of height $\leq \alpha$ and degree $\leq k+1$ we can find an $\vec{x}$–derivation $(r^k)^\varphi$ in $Z^\infty +$ Prog(P) of height $\leq 2^\alpha$ and degree $\leq k$.*

We now show that from the progression rule for P we can easily derive the progressiveness of P.

Lemma 3.4.2. *We have a normal derivation of $\forall x(\forall y \prec x : Py \to Px)$ in $Z^\infty +$ Prog(P) with height ≤ 5.*

Proof. By the ω–rule it suffices to derive $\forall y \prec j : Py \to Pj$ for any j with height ≤ 4. We argue informally. Assume $\forall y \prec j : Py$. By $\forall^-$ we have $i \prec j \to Pi$ for any i. Now for any $i \prec j$ we have $i \prec j$ as an axiom; hence Pi for any such i. An application of the progression rule yields Pj, with a derivation of height ≤ 3. Now by $\to^+$ and ω the claim follows. □

The crucial observation now is that a normal derivation of $P|\beta|$ must essentially have a height of at least β. However, to obtain the right estimates for our subsystems Z_k we cannot apply Lemma 3.4.1 down to degree 0 (i.e. to the normal form) but must stop already at degree 1. Such derivations, i.e. those of degree ≤ 1, will be called *almost normal*; they can also be analyzed easily. An almost normal derivation r in $Z^\infty +$ Prog(P) is called a $P|\vec{\alpha}|, \neg P|\vec{\beta}|$*–refutation* if r derives a formula $\vec{\varphi} \to \psi$ with $\vec{\varphi}$ and the free assumptions in r among $P|\vec{\alpha}| :\equiv P|\alpha_1|, \ldots, P|\alpha_m|$ and $\neg P|\vec{\beta}| :\equiv \neg P|\beta_1|, \ldots, \neg P|\beta_n|$ and true prime formulas, and ψ a false prime formula or else among $\neg P|\vec{\beta}|$.

Lemma 3.4.3. *Let r be an almost normal $P|\vec{\alpha}|, \neg P|\vec{\beta}|$–refutation of height $\leq |r|$ with $\vec{\alpha}$ and $\vec{\beta}$ disjoint. Then*

$$\min \vec{\beta} \leq |r| + \#\vec{\alpha},$$

where $\#\vec{\alpha}$ denotes the number of ordinals in $\vec{\alpha}$.

Proof. By induction on $|r|$. Note that we may assume that r does not contain either ω or else $\forall^-$. Note also that r cannot be an equality axiom $\mathrm{ax}^{P|\gamma|,|\gamma|=|\delta|\to P|\delta|}$ with $\gamma = \delta$ true, since we have assumed that $\vec{\alpha}$ and $\vec{\beta}$ are disjoint. We distinguish cases according to the last rule in r.

Case $\to^+$. By our definition of refutations the claim follows immediately from the induction hypothesis.

Case $\to^-$. Then $r \equiv t^{\varphi\to(\vec{\varphi}\to\psi)}s^{\varphi}$. If φ is a true prime formula, the claim follows from the induction hypothesis for t. If φ is a false prime formula, the claim follows from the induction hypothesis for s. If φ is $\neg\neg P|\gamma|$ (and hence $t \equiv \mathrm{ax}^{\forall x(\neg\neg Px\to Px)}|\gamma|$), then since the level of $\neg\neg P|\gamma|$ is 2 the derivation $s^{\neg\neg P|\gamma|}$ must end with an introduction rule, i. e. $s \equiv \lambda x^{\neg P|\gamma|}s_0^{\perp}$ (for otherwise, since no axiom contains some $\neg\neg Pr_0$ as a strictly positive subformula, we would get a contradiction against the assumption that r has degree ≤ 1). The claim now follows from the induction hypothesis for s_0. The only remaining case is when φ is $P|\gamma|$. Then t is an almost normal $P|\gamma|, P|\vec{\alpha}|, \neg P|\vec{\beta}|$ –refutation and s is an almost normal $P|\vec{\alpha}|, \neg P|\vec{\beta}|, \neg P|\gamma|$ –refutation. We may assume that γ is not among $\vec{\alpha}$, since otherwise the claim follows immediately from the induction hypothesis for t. Hence we have by the induction hypothesis for t

$$\min \vec{\beta} \leq |t| + \#\vec{\alpha} + 1 \leq |r| + \#\vec{\alpha}.$$

Case Prog(P). Then $r \equiv \langle r_\delta^{P|\delta|}\rangle_{\delta<\gamma}^{P|\gamma|}$. By induction hypothesis, since r_δ is a $P|\vec{\alpha}|$, $\neg P|\vec{\beta}|$, $\neg P|\delta|$ –refutation, we have for all $\delta < \gamma$

$$\min(\vec{\beta}, \delta) \leq |r_\delta| + \#\vec{\alpha} < |r| + \#\vec{\alpha}$$

and hence

$$\min(\vec{\beta}, \gamma) \leq |r| + \#\vec{\alpha}.$$

□

Now we can show the following result of (Mints 1971) and (Parsons 1973)

Theorem 3.4.4. *Transfinite induction up to ε_0 is underivable in Z, i.e.*

$$Z \nvdash \forall x(\forall y \prec x : Py \to Px) \to \forall x Px$$

with a predicate variable P, and transfinite induction up to ω_{k+1} is underivable in Z_k, i.e.

$$Z_k \nvdash \forall x(\forall x \prec x : Py \to Px) \to \forall x \prec \omega_{k+1} : Px.$$

Proof. We restrict ourselves to the second part. So assume that transfinite induction up to ω_{k+1} is derivable in Z_k. Then by the embedding of Z_k into Z^∞ (Lemma 3.3.1) and the normal derivability of the progressiveness of P in $Z^\infty + \mathrm{Prog}(P)$ with finite height (Lemma 3.4.2) we can conclude that $\forall x \prec \omega_{k+1} : Px$ is derivable in $Z^\infty + \mathrm{Prog}(P)$ with height $< \omega^m + h$ for some $m, h < \omega$ and degree $\leq k$. Now $k-1$ applications of Lemma 3.4.1 yield a derivation of the same formula $\forall x \prec \omega_{k+1} : Px$ in $Z^\infty + \mathrm{Prog}(P)$ with height $\leq \gamma \leq 2_{k-1}(\omega^m + h) < \omega_{k+1}$ and degree ≤ 1, hence also a derivation of $P|\gamma + 1|$ in $Z^\infty + \mathrm{Prog}(P)$ with height $\leq \gamma$ and degree ≤ 1. But this contradicts Lemma 3.4.3. □

3.5. Normalization for arithmetic is impossible

The normalization theorem for first–order logic applied to arithmetic Z is not particularly useful since we may have used in our derivation induction axioms of arbitrary complexity. Hence it is tempting to first eliminate the induction scheme in favour of an induction rule allowing to conclude $\forall x\varphi[x]$ from a derivation of $\varphi[0]$ and a derivation of $\varphi[Sx]$ with an additional assumption $\varphi[x]$ to be cancelled at this point (note that this rule is equivalent to the induction scheme), and then to try to normalize the resulting derivation in the new system Z with the induction rule. We will apply our results from Section 3.4 to show that even a very weak form of the normalization theorem cannot hold in Z with the induction rule.

Theorem 3.5.1. *The following weak form of a normalization theorem for Z with the induction rule is false: For any $\vec{x}^{\vec{\varphi}}$–derivation r^ψ with $\vec{\varphi}, \psi$ Π^0_l–formulas there is an $\vec{x}^{\vec{\varphi}}$–derivation $(r^*)^\psi$ containing only Π^0_k–formulas, with k depending only on l.*

Proof. Assume that such a normalization theorem would hold. Consider the Π^0_3–formula

$$\forall x(\forall y \prec x : Py \to Px) \to \forall x \prec \omega_{n+1} : Px$$

expressing transfinite induction up to ω_{n+1}. By Theorem 3.2.1 it is derivable in Z. Hence there exists a derivation of the same formula containing only Π^0_k–formulas, for some k independent of n. Hence Z_k derives transfinite induction up to ω_{n+1} for any n. But this clearly contradicts Theorem 3.4.1. □

Bibliography

Bachmann, H.: Transfinite Zahlen. Berlin: Springer 1955

Barendregt, H.P.: The lambda calculus. Amsterdam: North–Holland 1984

Buchholz, W.: Bemerkungen zur Dissertation Hudelmaier. Unveröffentlichtes Manuskript, München: 1989

Church, A.: The calculi of lambda–conversion. Annals of Math. Studies No.6, Princeton: 1941

Diller, J.: Zur Berechenbarkeit primitiv–rekursiver Funktionale endlicher Typen. In: Contributions to mathematical logic (K. Schütte, ed.), pp. 109–120. Amsterdam: North–Holland 1968

Gandy, R.O.: An early proof of normalization. In: To H.B. Curry: Essays on combinatory logic, lambda calculus and formalism (J.P. Seldin and J.R. Hindley, eds.), pp. 453–455. Academic Press 1980a

Gandy, R.O.: Proofs of strong of normalization. In: To H.B. Curry: Essays on combinatory logic, lambda calculus and formalism (J.P. Seldin, J.R. Hindley, eds.) pp. 457–477. Academic Press 1980b

Gentzen, G.: Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie. Mathematische Annalen *119*, 140–161 (1943)

Girard, J.Y.: Proof theory and logical complexity. Napoli: Bibliopolis 1987

Howard, W.A.: The formulae–as–types notion of construction. In: To H.B. Curry: Essays on combinatory logic, lambda calculus and formalism (J.P. Seldin and J.R. Hindley, eds.), pp. 479–490. Academic Press 1980a

Howard, W.A.: Ordinal analysis of terms of finite type. J. Symbolic Logic *45*, 493–504 (1980b)

Hudelmaier, J.: Bounds for cut–elimination in intuitionistic propositional logic. Dissertation, Tübingen: 1989

Mints, G.E.: Exact estimates of the provability of transfinite induction in the initial segmants of arithmetic. Zapiski Nauch. Sem Leningrad *20*, 134–144 (1971). Translated in: J. Soviet Math. *1*, 85–91 (1973)

Parsons, C.: Transfinite induction in subsystems of number theory (abstract) J. Symbolic Logic *36*, 544–545 (1973)

Prawitz, D.: Natural deduction. Stockholm: Almqvist & Wiksell 1965

Sanchis, L.E.: Functionals defined by recursion. Notre Dame J. Formal Logic *8*, 161–174 (1967)

Schütte, K.: Proof theory. Berlin: Springer 1977

Schwichtenberg, H.: Proof theory: Some applications of cut–elimination. In: Handbook of mathematical logic (J. Barwise, ed.), pp. 867–895. Amsterdam: North–Holland 1977

Schwichtenberg, H.: Complexity of normalization in the pure typed λ–calculus. In: The L.E.J. Brouwer centenary symposium (A.S. Troelstra and D. van Dalen, eds.), pp. 453–458. Amsterdam: North–Holland 1982

Schwichtenberg, H.: A normal form for natural deductions in a type theory with realizing terms. In: Atti del congresso logica e filosofia della scienza, oggi. Vol.1–Logica. (V.M. Abrusci and E.Casari, eds.), pp. 95–138. Bologna: CLUEB 1986

Schwichtenberg, H.: An upper bound for reduction sequences in the typed λ–calculus. Archive Math. Logic, to appear 1990

Statman, R.: The typed λ–calculus is not elementary recursive. Theoretical Computer Science *9*, 73–81 (1979)

Tait, W.W.: Infinitely long terms of transfinite type. In: Formal systems and recursive functions (J. Crossley and M. Dummett, eds.), pp. 176–185. Amsterdam: North–Holland 1965

Takeuti, G.: Proof theory. Second edition. Amsterdam: North–Holland 1987

Troelstra, A., editor. Metamathematical investigations of intuitionistic arithmetic and analysis. Lecture Notes in Mathematics, Vol. 344, Berlin: Springer 1973

Troelstra, A. and van Dalen, D.: Constructivism in mathematics. An introduction. Studies in Logic and the Foundations of Mathematics, Vol. 121, 123 Amsterdam: North–Holland 1988