GOODSTEIN'S THEOREM, ϵ_0, AND UNPROVABILITY

HENRY TOWSNER

Part 1. A Proof of Goodstein's Theorem

1. The Statement of Goodstein's Theorem

In order to even state Goodstein's Theorem, we need to define the *hereditary base b notation*.

First, recall how base b notation works (in some base $b \geq 2$): we write a number n in the form

$$n = b^k c_k + b^{k-1} c_{k-1} + \cdots + b c_1 + c_0 = \sum_{i=0}^{k} b^i c_i$$

where each coefficient $c_i \in [0, b)$.

In the heritary base b notation, we take this a step further by insisting that the exponents are also written in hereditary base b notation. For example, to write 26 in ordinary base 2 notation, we would write

$$26 = 2^4 + 2^3 + 2$$

(that is, $k = 4$, $c_4 = c_3 = c_1 = 1$ while $c_2 = c_0 = 0$).

In hereditary base 2 notation, we keep going: the exponent 3 becomes $2 + 1$, which becomes $2^1 + 2^0$, which becomes $2^{2^0} + 2^0$. So in full hereditary base 2 notation,

$$26 = 2^{2^{2^{2^0}}} + 2^{2^{2^0}+2^0} + 2^{2^0}.$$

Since, in any base, $b^0 = 1$, we'll usually omit the very last step, and write

$$26 = 2^{2^{2^1}} + 2^{2^1+1} + 2^1.$$

Definition 1.1. The *hereditary base change function* $B_{b,d}(n)$ is a function from positive natural numbers to positive natural numbers which can be calculated by:

- writing n in hereditary base b notation, and then
- changing every b to a d.

For instance,

$$B_{2,3}(26) = 3^{3^{3^1}} + 3^{3^1+1} + 3^1 = 7625597485071.$$

As this example indicates, when $b < d$, the hereditary base change tends to make numbers grow by a lot.

Date: March 20, 2020.

Definition 1.2. The *Goodstein sequence starting with n* is the sequence given recursively by:

- $g_1(n) = n$,
- if $g_i(n) = 0$, the sequence terminates,
- if $g_i(n) > 0$ then $g_{i+1}(n) = B_{i+1,i+2}(g_i(n)) - 1$.

That is, the sequence begins with n, which we think of as being written in hereditary base 2 notation. At each step we first increment the base by 1, and then subtract 1.

For example, the Goodstein sequence starting with 2 is

$$2, 2, 1, 0 :$$

initially we have $g_1 = 2 = 2^1$. Then $g_2 = B_{2,3}(2) - 1 = 3^1 - 1 = 2$. After that, the base change operation doesn't do anything, so the sequence just decreases to 0.

The Goodstein sequence starting with 3 is

$$3, 3, 3, 2, 1, 0 :$$

initially we have $g_1 = 3 = 2^1 + 1$, so $g_2 = B_{2,3}(3) - 1 = 3^1 + 1 - 1 = 3$, then $g_3 = B_{3,4}(3) - 1 = 4^1 - 1 = 3$, and then the sequence just decreases.

The Goodstein sequence starting with 4 is more dramatic:

$$4, 26, 41, 60, 83, 109, 139, 173, 211, 253, \ldots.$$

initially we have $g_1 = 4 = 2^2$, so $g_2 = B_{2,3}(4) = 3^3 - 1 = 26 = 3^2 \cdot 2 + 3 \cdot 2 + 2$, $g_3 = B_{3,4}(26) = 4^2 \cdot 2 + 4 \cdot 2 + 2 - 1 = 41$, and so on. In fact, this sequence does eventually terminate, but not quickly: it takes more than 10^{10^8} steps.

As n gets bigger, the Goodstein sequences take even longer to terminate.

This inspires the theorem we are interested in:

Theorem 1.3 (Goodstein's Theorem)**.** *For every n, the Goodstein sequence starting with n terminates.*

We will be able to prove this, but the proof is more complicated than the initial set-up suggests. We'll also be able to show that this result is, in a real sense, intrinsically difficult to prove: this theorem isn't provable in Peano Arithmetic. Since the axioms of Peano Arithmetic capture most "ordinary mathematical reasoning", this shows that the proof really must use something beyond common methods.

Along the way we'll show a related fact: that the number of steps it takes a Goodstein sequence to terminate really is very, very large. Let $\mathcal{G}(n)$ be the number of steps it takes the Goodstein sequence starting with n to terminate. Then the function $\mathcal{G}(n)$ grows extremely fast. In order to make this precise, we'll need to describe a family of functions, the *fast-growing functions*, and show that $\mathcal{G}(n)$ shows up very high in this hierarchy.

2. Ordinals as Timers

We will define a "timer": a value $o(g_i(n))$ with the property that, even though $g_{i+1}(n)$ could be much larger than $g_i(n)$, we will still have $o(g_{i+1}(n)) < o(g_i(n))$. Somehow, the timer function will detect a way that the sequence is always decreasing, even though it's numeric value is increasing. We can think of the timer as "counting down" the time until the sequence must terminate.

Of course, it would be most convenient if we knew in advance how many steps it takes the Goodstein sequence to terminate: then we could take $o(g_1(n))$ to be this number, and each $o(g_i(n))$ could be the natural number $o(g_1(n)) - (i-1)$; then this sequence would decrease at each step, and we would know the sequence terminated exactly when this counter hit 0.

The problem is that figuring out what natural number to assign is as hard as proving the theorem in the first place, so that won't help us. Instead, our timer $o(g_i(n))$ will have to be from a bigger set than the natural numbers—it will be an *ordinal*.

The crucial property we would want our "timer" to have is that we can't count down forever.

Definition 2.1. A set $\mathcal{S}$ with an ordering $\prec$ is *well-founded* if any sequence

$$s_1 \succ s_2 \succ \cdots$$

is finite.

Of course we need the value of our timer to be well-founded. For example, the integers are not well-founded, and would be useless timers: we could have $o(g_1(n)) = 0$, $o(g_2(n)) = -1$, $o(g_3(n)) = -2$, and so on; then the timer value is getting smaller and smaller, but there's no reason it can't go on forever.

Well-foundedness prevents this: every decreasing sequence stops after finitely many steps. For example, the first ordinal past the natural numbers is ω. ω is larger than n for any natural number n. Even though ω is "infinite", the only things smaller than ω are natural numbers, so any decreasing sequence beginning with ω stops in finitely many steps.

If we assign ω as the timer at some stage, we're saying that we don't know how many steps are left—but that after the next step, we will. If $o(g_1(n)) = \omega$ then $o(g_2(n))$ must be a natural number, and then we can only have $o(g_2(n))$ additional steps before hitting 0.

We can imagine an ordinal bigger than ω, which we call $\omega\#1$. (# is a version of addition.) $\omega\#1 > \omega$, so we could have $o(g_1(n)) = \omega\#1$, $o(g_2(n)) = \omega$, but then $o(g_3(n))$ is a natural number, so there are at most $o(g_3(n))$ many steps left.

Continuing in this way, we can imagine a larger ordinal $\omega\#\omega$: $\omega\#\omega$ is larger than $\omega\#n$ for every n, so if we count down from $\omega\#\omega$, it looks something like

$$\omega\#\omega > \omega\#n > \omega\#n{-}1 > \omega\#n{-}2 > \cdots > \omega\#1 > \omega > m > m{-}1 > \cdots > 2 > 1 > 0.$$

We could think of $\omega\#\omega$ as an instance of multiplication—$\omega \cdot 2$—which suggests $\omega \cdot 3$ and so on, and more generally the ordinals $\omega \cdot m\#n$ for any natural numbers m and n, ordered so that

$$\omega \cdot m\#n > \omega \cdot m\#(n-1) > \cdots > \omega \cdot m > \omega \cdot (m-1)\#n'$$

for any n', and continuing down in this fashion.

3. The Ordinals Below ϵ_0

We need the ordinals less than a particularly important ordinal, ϵ_0.

Definition 3.1. We define the *ordinals below* ϵ_0 inductively as follows:

- 0 is the only ordinal of height 0,
- if $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_k$ is a finite, nonincreasing sequence of ordinals of height $\leq n$ then
$$\omega^{\alpha_1}\#\omega^{\alpha_2}\#\cdots\omega^{\alpha_k}$$
is an ordinal of height $n+1$,
- $0 < \omega^{\alpha_1}\#\omega^{\alpha_2}\#\cdots\omega^{\alpha_k}$,
- if $\omega^{\alpha_1}\#\omega^{\alpha_2}\#\cdots\omega^{\alpha_k}$ and $\omega^{\beta_1}\#\omega^{\beta_2}\#\cdots\omega^{\beta_r}$ are ordinals, we say
$$\omega^{\alpha_1}\#\omega^{\alpha_2}\#\cdots\omega^{\alpha_k} < \omega^{\beta_1}\#\omega^{\beta_2}\#\cdots\omega^{\beta_r}$$
if either:
  - when $i \leq k$ is least so that $\alpha_i \neq \beta_i$, $\alpha_i < \beta_i$, or
  - for every $i \leq k$, $\alpha_i = \beta_i$ and $k < r$.

The *ordinals below* ϵ_0 are the ordinals of height n for some natural number n.

So the ordinals of height 1 have the form

$$\underbrace{\omega^0\#\omega^0\#\cdots\omega^0}_{k \text{ times}}.$$

Recognizing that $\omega^0 = 1$, this represents the natural number k. So the ordinals of height 1 are exactly the natural numbers ordered as usual.

The ordinals of level 2 then have forms like

$$\omega^7\#\omega^7\#\omega^7\#\omega^5\#\omega^2\#\omega^2\#\omega^0\#\omega^0.$$

We don't use multiplication—it is more natural for us to simply add repeatedly.

We use the usual abbreviations: k for $\underbrace{\omega^0\#\omega^0\#\cdots\omega^0}_{k \text{ times}}$ and ω for ω^1 (which is really ω^{ω^0}).

The ordering on ordinals is "lexicographic in their exponent". To compare $\omega^7\#\omega^0$ and $\omega^6\#\omega^6\#\omega^5$, we look at the lists of ordinals that define them: the first is defined by $7 \geq 0$ and the second by $6 \geq 6 \geq 5$. Since $7 > 6$, $\omega^7\#\omega^0$ is larger.

To compare $\omega^{\omega^2}\#\omega^\omega$ to $\omega^{\omega^2}\#\omega^8$, we compare the list of exponents: the list $\omega^2 \geq \omega^1$ from the first ordinal with the list $\omega^2 \geq 8$ from the second ordinal; since the first component of the lists are the same, we go to the second component. Since $\omega^1 \geq 8$, $\omega^{\omega^2}\#\omega^\omega$ is larger.

For one more example, compare $\omega^{\omega^2}\#\omega^\omega\#\omega^8$ to $\omega^{\omega^2}\#\omega^\omega$. The lists of exponents are $\omega^2 \geq \omega^1 \geq 8$ and $\omega^2 \geq \omega^1$. The first components of both lists are ω^2, so this doesn't distinguish the ordinals and we go to the second component. Both lists have ω^1 as the second component, so this doesn't distinguish them either. At this point the second list has run out, so the longer list is the larger ordinal: $\omega^{\omega^2}\#\omega^\omega\#\omega^8$ is larger.

We call these the ordinals below ϵ_0 because ϵ_0 is the name of the smallest ordinal which is larger than all ordinals of height n for any n.

4. Well-foundedness

We first note that this construction adds all new ordinals "at the top": the ordinals of height 0 are smaller than the new ordinals of height 1 are smaller than the new ordinals of height 2 and so on.

Lemma 4.1. *If α is an ordinal of height n and β is an ordinal of height $m > n$ then $\alpha < \beta$.*

Proof. By induction on n. When $n = 0$ this is immediate from the definition, since the only ordinal of height 0 is 0 itself, which is smaller than all other ordinals.

Suppose this is true for ordinals of height $n-1$ and that α is an ordinal of height n, and therefore not 0, so

$$\alpha = \omega^{\alpha_1}\# \cdots \#\omega^{\alpha_k}.$$

Since β has height $m > n$, β is also not 0, so

$$\beta = \omega^{\beta_1}\# \cdots \#\omega^{\beta_r}.$$

If β_1 is an ordinal of height $n-1$ then, by the inductive hypothesis, also every β_i has height $n-1$, which would mean that β is an ordinal of height n. Since this is not the case, β_1 must have height $\geq n$.

On the other hand, since α does have height n, in particular α_1 has height $\leq n-1$. Therefore, by the inductive hypothesis, $\alpha_1 < \beta_1$, and therefore $\alpha < \beta$. □

Theorem 4.2. *The ordinals below ϵ_0 are well-founded.*

Proof. We need to show that there is no infinite decreasing sequence of ordinals

$$\gamma_0 > \gamma_1 > \cdots .$$

Suppose, for a contradiction, that there were such an infinite decreasing sequence. The heights would be nonincreasing, so in particular every ordinal in the list would have height $\leq$ the height of γ_0. But γ_0 has some finite

height n, so one of the finitely many heights $\leq n$ contains infinitely many of the γ_i.

So it suffices to show that, for each n, the ordinals of height n are well-founded. We proceed by induction on n. Certainly the ordinals of height 0 are well-founded, because there's only one of them, 0.

So suppose the ordinals of height $n-1$ are well-founded, but that there is an infinite decreasing sequence of ordinals

$$\gamma_0 > \gamma_1 > \cdots$$

of height n. Each of these ordinals can be written

$$\gamma_i = \omega^{\alpha_{i,1}} \# \omega^{\alpha_{i,2}} \# \cdots \# \omega^{\alpha_{i,k_i}}.$$

Consider the first components, $\alpha_{i,1}$. By the comparison rule for ordinals, since $\gamma_{i+1} < \gamma_i$, we must have $\alpha_{i+1,1} \leq \alpha_{i,1}$. These are all ordinals of height $n-1$, so they can't strictly decrease infinitely often: there must be some α_1 so that $\alpha_{i,1} = \alpha_1$ for all but finitely many i.

Consider these ordinals

$$\gamma_i = \omega^{\alpha_1} \# \omega^{\alpha_{i,2}} \# \cdots \# \omega^{\alpha_{i,k_i}}.$$

Since $\gamma_i > \gamma_{i+1}$, we must have $\alpha_{i,2} \geq \alpha_{i+1,2}$. Again these are ordinals of height $\leq n-1$, so they can't strictly decrease infinitely often, so there must be some α_2 so that $\alpha_{i,2} = \alpha_2$ for all but finitely many i.

Continuing in this way, we obtain an infinite sequence of ordinals $\alpha_1, \alpha_2, \ldots, \alpha_k, \ldots$ such that, for each k, for all but finitely many i, γ_i has the form

$$\gamma_i = \omega^{\alpha_1} \# \omega^{\alpha_2} \# \cdots \# \omega^{\alpha_k} \# \cdots .$$

Since the α_k have height $\leq n-1$, they cannot decrease infinitely often, so there must be some single value α^* so that $\alpha_k = \alpha^*$ for all but finitely many k. That is, the sequence goes $\alpha_1, \alpha_2, \ldots, \alpha_{k-1}, \alpha^*, \alpha^*, \ldots, \alpha^*, \ldots$.

Consider any i which reaches this value α^*: some γ_{i_0} of the form

$$\gamma_{i_0} = \omega^{\alpha_1} \# \cdots \# \omega^{\alpha_{k-1}} \# \omega^{\alpha^*} \# \cdots .$$

Since γ_{i_0} is a finite sum, the term ω^{α^*} can only appear finitely many, say m times, in γ_{i_0}. This means γ_{i_0} has either the form

$$\gamma_{i_0} = \omega^{\alpha_1} \# \cdots \# \underbrace{\omega^{\alpha_{k-1}} \# \omega^{\alpha^*} \# \cdots \omega^{\alpha^*}}_{m \text{ times}}$$

or

$$\gamma_{i_0} = \omega^{\alpha_1} \# \cdots \# \underbrace{\omega^{\alpha_{k-1}} \# \omega^{\alpha^*} \# \cdots \omega^{\alpha^*}}_{m \text{ times}} \# \omega^{\gamma} \# \cdots$$

with $\gamma < \alpha^*$.

However since α^* appears infinitely many times in the sequence of α's, for all but finitely many i,

$$\gamma_i = \omega^{\alpha_1} \# \cdots \# \omega^{\alpha_{k-1}} \# \underbrace{\omega^{\alpha^*} \# \cdots \# \omega^{\alpha^*}}_{m+1 \text{ times}} \# \cdots .$$

In particular, there is some $i_1 > i_0$ of this form. But, comparing the ordinals using the rule for comparing sums of exponential terms gives $\gamma_{i_0} < \gamma_{i_1}$. This is a contradiction. □

5. Proving Goodstein's Theorem

We add an additional case to the hereditary base change function: we define $B_{b,\omega}(n)$ to be the ordinal given by:

- writing n in hereditary base b notation,
- changing every b to an ω, and
- replacing $+$ with $\#$.

For example,

$$B_{2,\omega}(26) = \omega^{\omega^\omega} \# \omega^{\omega+1} \# 2.$$

Then for every b, n, $B_{b,\omega}(n)$ is an ordinal below ϵ_0.

Lemma 5.1. *If $x < y$ then $B_{b,\omega}(x) < B_{b,\omega}(y)$.*

Proof. It suffices to show by induction on x that $B_{b,\omega}(x) < B_{b,\omega}(x+1)$.

Consider what happens when we begin writing $x+1$ in base b notation; we are mostly interested in the final term, so let us write

$$x + 1 = a + b^d$$

for some a and some d (which are themselves written in hereditary base b notation). Therefore $B_{b,\omega}(x+1) = \alpha \# \omega^\delta$ for corresponding ordinals α and δ.

Compare this to x. If $d = 0$ then $x = a$ and $B_{b,\omega}(x) = \alpha < B_{b,\omega}(x+1)$. If $d > 0$ then x is equal to a sum

$$x = a + b^{d-1} \cdot (b-1) + b^{d-2} \cdot (b-1) + \cdots + b^0 \cdot (b-1),$$

and therefore

$$B_{b,\omega}(x) = \alpha \# \omega^{B_{b,\omega}(d-1)} \cdots (b-1) \# \cdots \omega^0 (b-1).$$

By the inductive hypothesis, $B_{b,\omega}(d-i) < B_{b,\omega}(d) = \delta$ for all $i > 0$, so in particular $B_{b,\omega}(x) > B_{b,\omega}(x+1)$.

□

We associate a sequence of ordinals to the Goodstein sequence:

$$\gamma_i(n) = B_{i+1,\omega}(g_i(n)).$$

Lemma 5.2. *For every i, n, $\gamma_{i+1}(n) < \gamma_i(n)$.*

Proof.

$$\gamma_{i+1}(n) = B_{i+2,\omega}(g_{i+1}(n)) = B_{i+2,\omega}(B_{i+1,i+2}(g_i(n)) - 1).$$

On the other hand, $\gamma_i(n) = B_{i+1,\omega}(g_i(n)) = B_{i+2,\omega}(B_{i+1,i+2}(g_i(b)) - 1)$. So, taking $x = B_{i+1,i+2}(g_i(b))$, $\gamma_i(n) = B_{i+2,\omega}(x+1)$ while $\gamma_{i+1}(n) = B_{i+2,\omega}(x)$. Therefore by the previous lemma, $\gamma_{i+1}(n) < \gamma_i(n)$.

□

Theorem 5.3 (Goodstein's Theorem)**.** *For every n, there is an i so that $g_i(n) = 0$.*

Proof. The sequence $\gamma_i(n)$ is a strictly decreasing sequence of ordinals below ϵ_0, and therefore must be finite, so the sequence $g_i(n)$ must also be finite. □

Part 2. The Goodstein Function Grows Quickly

6. Fast-Growing Functions

Definition 6.1. We define the *fundamental sequence* of a non-zero ordinal below ϵ_0:

- $(\gamma\#1)[n] = \gamma$,
- when $\alpha > 0$,

$$(\gamma\#\omega^\alpha)[n] = \gamma\#\underbrace{\omega^{\alpha[n]}\#\cdots\#\omega^{\alpha[n]}}_{n \text{ times}}.$$

So $m[n] = m - 1$, $\omega[n] = n$, $\omega^2[n] = \omega \cdot n$, $\omega^\omega = \omega^n \cdot n$, and so on.

Lemma 6.2. *For any n and any $\alpha > 0$, $\alpha[n] < \alpha$.*

Proof. By induction on α. We have $\alpha = \alpha'\#\omega^\beta$. When $\beta = 0$, $\alpha[n] = \alpha' < \alpha$. When $\beta > 0$, $\alpha[n] = \alpha'\#\omega^{\beta[n]}n$. By the inductive hypothesis, $\beta[n] < \beta$, so $\alpha'\#\omega^{\beta[n]}n < \alpha'\#\omega^\beta = \alpha$. □

Definition 6.3. We define a family of functions $h_\alpha : \mathbb{N} \to \mathbb{N}$:

- $h_0(n) = n$,
- if $\alpha > 0$, $h_\alpha(n) = h_{\alpha[n]}(n+1)$.

So $h_m(n) = n + m$, $h_\omega(n) = h_n(n+1) = 2n + 1$, $h_{\omega\#\omega}(n) = h_{\omega\#n}(n + 1) = h_\omega(2n + 1) = 4n + 3$. More generally, $h_{\omega\cdot k}(n)$ is roughly $2^k n$, so $h_{\omega^2}(n) = h_{\omega\cdot n}(n+1) \geq 2^{n+1}n$.

The function $h_{\omega^2\cdot k}(n)$ corresponds to a tower of k powers of 2:

$$h_{\omega^2\cdot k}(n) \geq \underbrace{2^{2^{\cdots^{2^{n+1}}}}}_{k \text{ times}}.$$

That means that $h_{\omega^3}(n)$ is a tower of exponents whose height depends on the value n:

$$h_{\omega^3}(n) \geq \underbrace{2^{2^{\cdots^{2^{n+1}}}}}_{n \text{ times}}.$$

In general, we can think of each new power of ω as "iterating" the previous one: h_{ω^0} is "add one", so h_{ω^1} is "add one n times"—that is, double the input. Then h_{ω^2} is "double n times", which is roughly raising to a power, and h_{ω^3} is "raise to a power n times". h_{ω^4} means to iterate that operation n times; this is sometimes called a "Wowzer" function.

h_{ω^ω} is larger than any of these iterations—it's somehow diagonalizing over the concept of iterating, so that $h_{\omega^\omega}(n)$ means "use n levels of iteration". It's

roughly the same order of magnitude as the function called the Ackermann function.

But we're barely starting: $h_{\omega^{\omega+1}}$ is the function which iterates the Ackermann function n times. And this hierarchy keeps going, to functions like $h_{\omega^{\omega^2}}$ and $h_{\omega^{\omega^{\omega^\omega}}}$ and so on.

There's an essential tool for thinking about these functions. Suppose we start with an ordinal α and look at the sequence $\alpha > \alpha[n] > \alpha[n][n+1] > \cdots$. This is a decreasing sequence of ordinals, so it must be finite. It turns out that $h_\alpha(n)$ is the length of this sequence.

Lemma 6.4. *For $\alpha > 0$, $h_\alpha(n)$ is the smallest value $n + k$ such that*

$$\alpha[n][n+1][n+2]\cdots[n+k] = 0.$$

Proof. By induction on $\alpha > 0$. When $\alpha = \alpha' \# \omega^0$, $h_\alpha(n) = h_{\alpha'}(n+1) = n + 1 + k$ where, by the inductive hypothesis,

$$\alpha'[n+1][n+1+1][n+1+2]\cdots[n+1+k] = 0.$$

Therefore

$$\alpha[n][n+1][n+2][n+3]\cdots[n+k+1] = \alpha'[n+1][n+1+1][n+1+2]\cdots[n+1+k] = 0.$$

When $\alpha = \alpha' \# \omega^\beta$ with $\beta > 0$, $h_\alpha(n) = h_{\alpha' \# \omega^{\beta[n]} \cdot n}(n+1) = n + 1 + k$ where, by the inductive hypothesis,

$$\alpha' \# \omega^{\beta[n]} \cdot n[n+1][n+1+1]\cdots[n+1+k] = 0,$$

so

$$\alpha[n][n+1]\cdots[n+k+1] = 0$$

as needed. □

7. The Goodstein Function Grows Quickly

Definition 7.1. The *Goodstein function* is defined by setting $\mathcal{G}(n)$ to be the value of i so that $g_i(n) = 0$.

Theorem 7.2. *For every $\alpha < \epsilon_0$, there is some n so that, for all $m \geq n$, $\mathcal{G}(m) > h_\alpha(m)$.*

We can outline the proof with what we have, though working out the details is a bit more technical.

Proof Sketch. Naturally, we're going to use the ordinals $\gamma_i(n)$ to compare the Goodstein sequences to fast-growing functions.

The key observation is that we expect the $\gamma_i(n)$ sequences to decrease "less quickly" than sequences $\alpha > \alpha[n] > \cdots$. More precisely, we expect that if $\gamma_i(n) \geq \alpha$ then $\gamma_{i+1}(n) \geq \alpha[i+1]$. This isn't quite true in general—it's not even true that $\beta \geq \alpha$ implies $\beta[i] \geq \alpha[i]$ (the counterexamples involve cases where α is something like $\omega \cdot (i + 100) + 1$). However this will be true once we promise that α doesn't have any "coefficients" which are too big.

So suppose that $\gamma_m(m) \geq \alpha$. Then we claim that $\mathcal{G}(m) \geq h_\alpha(m)$. The idea is to compare the sequence of ordinals

$$\gamma_m(m) > \gamma_{m+1}(m) > \gamma_{m+2}(m) > \cdots$$

to the sequence

$$\alpha > \alpha[m+1] > \alpha[m+1][m+2] > \cdots .$$

Inductively, we can show that the top sequence is always larger: once we assume that $\gamma_m(m) \geq \alpha$, we should have $\gamma_{m+1}(m) \geq \gamma_m(m)[m] \geq \alpha[m]$, and so on. Since $h_\alpha(m)$ is the value where α reaches 0, that means $\mathcal{G}(m) \geq h_\alpha(m)$.

It's not enough to find one m where all this works: we need to show that it happens for all $m \geq n$ for some n. So we also need to show that when $m \geq n$, $\gamma_m(m) \geq \gamma_n(n)$, so that once we find one n with $\gamma_n(n)$, the argument above will work for all $m \geq n$.

Finally, we need to show that there actually is some n such that $\gamma_n(n) \geq \alpha$. It's not hard to find an n so that $\gamma_1(n) \geq \alpha$—let n be the result of taking the ordinal α and replacing every ω with a 2. The result might give $\gamma_1(n) > \alpha$ (for instance, if $\alpha = \omega \cdot 3$ then $n = 2 \cdot 3 = 6$ and $\gamma_1(6) = \omega^\omega + \omega$), but at least $\gamma_1(n) \geq \alpha$.

It turns out that once $n \geq 4$, we will then have $\gamma_{2n}(2n) \geq \alpha$; the reason is that $\gamma_i(2n) \geq \gamma_1(n) \# \gamma_i(n)$, and since $n \geq 4$ ensures that $\gamma_n(n) \geq \omega$, in particular we have $\gamma_n(2n) \geq \gamma_1(n) \# \gamma_n(n) \geq \alpha \# \omega$. Then we can compare

$$\gamma_n(2n) > \gamma_{n+1}(2n) > \gamma_{n+2}(2n) > \cdots$$

to

$$\alpha \# \omega > \alpha \# \omega[n+1] > \alpha \# \omega[n+1][n+2] > \cdots .$$

For the same reasons as above, we should have $\gamma_{n+i}(2n) \geq \alpha \# \omega[n+1][n+2] \cdots [n+i]$. But the latter sequence is

$$\alpha \# \omega > \alpha \# (n+1) > \alpha \# n > \cdots > \alpha,$$

so $\gamma_{2n}(2n) \geq \alpha$. □

8. Lemmas about Complexity of Ordinals

In order to fill in the gaps in the sketchy proof of Theorem 7.2, we need to work out some technical properties about how fundamental sequences work. The recurring difficulty while trying to work with fundamental sequences is that they're not quite order-preserving in the way we'd like them to be. For instance, even though $\omega \cdot 100 < \omega^2$, $\omega^2[3] = \omega \cdot 3$ while $\omega \cdot 100[n] = \omega \cdot 99 + n > \omega \cdot 3$.

This problem goes away if we pay attention to the coefficient 100: once $n \geq 100$, we do have $\omega^2[n] > \omega \cdot 100[n]$ as we'd expect.

Definition 8.1. We define $\tau(\alpha)$ by:

- $\tau(0) = 0$,

- When $\alpha > 0$, write $\alpha = \omega^{\alpha_1} c_1 \# \cdots \# \omega^{\alpha_k} c_k$ where $\alpha_1 > \cdots > \alpha_k$; then
$$\tau(\alpha) = \max\{c_1, \ldots, c_k, \tau(\alpha_1), \ldots, \tau(\alpha_k)\}.$$

That is, τ is the largest coefficient appearing in α (including recursively in the exponents in α).

Definition 8.2. When α is an ordinal below ϵ_0 and b is a natural number, $B_{\omega,b}(\alpha)$ is defined by induction on α:

- $B_{\omega,b}(0) = 0$,
- $B_{\omega,b}(\omega^{\alpha_1} \# \cdots \# \omega^{\alpha_k}) = b^{B_{\omega,b}(\alpha_1)} + \cdots + b^{B_{\omega,b}(\alpha_k)}$.

For instance, $B_{\omega,2}(\omega^\omega \# \omega) = 2^2 + 2 = 6$. Note that $B_{\omega,b}(\alpha)$ need not be in hereditary base b notation; for instance, $B_{\omega,2}(\omega \# \omega) = 2 + 2 = 4 = 2^2$.

Lemma 8.3. $\tau(B_{b,\omega}(n)) < b$.

Proof. When n is written in hereditary base b, the coefficients are digits in base b, which must be $< b$, so each term in $B_{b,\omega}(n)$ is repeated at most $b-1$ times, so $\tau(B_{b,\omega}(n)) < b$. □

Lemma 8.4. *$B_{\omega,b}$ and $B_{b,\omega}$ give a bijection between $\mathbb{N}$ and the ordinals α below ϵ_0 with $\tau(\alpha) < b$.*

Proof. By the previous lemma, $B_{b,\omega}$ is a function from $\mathbb{N}$ to ordinals below ϵ_0 with $\tau(\alpha) < b$. Moreover, by the definition, $B_{\omega,b}(B_{b,\omega}(n)) = n$, since we obtain this number by replacing b with ω and then immediately replacing each ω with b again.

Convsersely, if α is in hereditary base b notation then $B_{b,\omega}(B_{\omega,b}(\alpha)) = \alpha$, by the same argument: $B_{b,\omega}(B_{\omega,b}(\alpha))$ is obtained by replacing the ω's in α with b's, and then replacing the b's with ω's.

So $B_{b,\omega}$ and $B_{\omega,b}$ are two-sided inverses, and therefore give a bijection. □

Lemma 8.5. *If $\gamma_i(n) = \alpha \# \omega^\beta$ where $\beta > 0$ then*
$$\gamma_{i+1}(n) = \alpha \# \sum_{\beta' < \beta, \tau(\beta') \le i+1} \omega^{\beta'}(i+1).$$

Proof. Consider the number $g_{i+1}(n) + 1$ written in hereditary base $i+2$ notation, $a + (i+2)^b$ where $B_{i+2,\omega}(a) = \alpha$ and $B_{i+2,\omega}(b) = \beta$. When we subtract 1, we get $a + \sum_{b' < b} (i+2)^{b'}(i+1)$. Since $B_{i+2,\omega}$ is a bijection between the $b' < b$ and the $\beta' < \beta$ with $\tau(\beta') \le i+1$,
$$B_{i+2,\omega}(a + \sum_{b' < b} (i+2)^{b'}(i+1)) = \alpha \# \sum_{\beta' < \beta, \tau(\beta') \le i+1} \omega^{\beta'}(i+1).$$

□

Lemma 8.6. *If $\beta < \gamma_i(n)$ and $\tau(\beta) \le i+1$ then $\beta \le \gamma_{i+1}(n)$.*

Proof. Write $\gamma_i(n) = \alpha \# \omega^{\alpha'}$. If $\alpha' = 0$ then $\beta \le \alpha = \gamma_{i+1}(n)$.

So suppose $\alpha' > 0$. $\gamma_{i+1}(n) \ge \alpha$, so if $\beta < \alpha$ we are done. So suppose $\beta = \alpha \# \beta'$. Since $\tau(\beta) \le i+1$, in particular $\tau(\beta') \le i+1$ and $\beta < \alpha'$, so $\beta' \le \sum_{\alpha'' < \alpha', \tau(\alpha') \le i+1} \omega^{\alpha''}(i+1)$, so $\beta \le \gamma_{i+1}(n)$. □

9. Completing the Proof

Lemma 9.1. *For all $m > n$ and all i, $\gamma_i(m) > \gamma_i(n)$.*

Proof. We will show, by induction on i, that $\gamma_i(n) \leq \gamma_{i+1}(m)$; since $\gamma_{i+1}(m) < \gamma_i(m)$, the claim follows.

When $i = 1$, $\gamma_1(n) = B_{2,\omega}(n) < B_{2,\omega}(m) = \gamma_1(m)$, and since $\tau(\gamma_1(n)) \leq 1 \leq 2$, $\gamma_1(n) \leq \gamma_2(m)$. Suppose $\gamma_i(n) \leq \gamma_{i+1}(m)$; then $\gamma_{i+1}(n) < \gamma_{i+1}(m)$ and $\tau(\gamma_{i+1}(n)) \leq i+1$, so $\gamma_{i+1}(n) \leq \gamma_{i+2}(m)$. □

Lemma 9.2. *For every $\alpha < \epsilon_0$, there is an n so that, for all $m \geq n$, $\gamma_1(m) \geq \alpha$.*

Proof. Observe that $B_{2,\omega}(B_{\omega,2}(\alpha)) \geq \alpha$—when these are not equal, it is because $\tau(\alpha) \geq 2$, so $B_{\omega,2}(\alpha)$, written in hereditary base 2, has terms which combine to terms of higher exponent. But this only makes the ordinal $B_{2,\omega}(B_{\omega,2}(\alpha))$ larger.

So when $m \geq B_{\omega,2}(\alpha)$, we have

$$\gamma_1(m) \geq \gamma_1(B_{\omega,2}(\alpha)) = B_{2,\omega}(B_{\omega,2}(\alpha)) \geq \alpha.$$

□

Lemma 9.3. *For all $i \leq \mathcal{G}(n)$, $\gamma_i(2n) > \beta\#\gamma_i(n)$.*

Proof. Let $\beta = \gamma_1(n)$. We proceed by induction on i.

For $i = 1$, write n in binary as $n = 2^{c_1} + \cdots + 2^{c_k}$, so

$$\gamma_1(n) = \omega^{B_{2,\omega}(c_1)}\# \cdots \#\omega^{B_{2,\omega}(c_k)}.$$

Then

$$\begin{aligned}\gamma_1(2n) &= \omega^{B_{2,\omega}(c_1+1)}\# \cdots \#\omega^{B_{2,\omega}(c_k+1)} \\ &\geq \omega^{B_{2,\omega}(c_1)+1}\# \cdots \#\omega^{B_{2,\omega}(c_k)+1} \\ &> \omega^{B_{2,\omega}(c_1)}2\# \cdots \#\omega^{B_{2,\omega}(c_k)}2 \\ &= \gamma_1(n)\#\gamma_1(n).\end{aligned}$$

Suppose we have $\gamma_i(2n) > \beta\#\gamma_i(n)$. Since $\tau(\beta) \leq 1$ and $\tau(\gamma_i(n)) \leq i$, $\tau(\beta\#\gamma_i(n)) \leq i+1$, so $\gamma_{i+1}(2n) \geq \beta\#\gamma_i(n)) > \beta\#\gamma_{i+1}(n)$. □

Lemma 9.4. *For all $n \geq 4$, $\gamma_n(n) \geq \omega$.*

Proof. By induction on n. When $n = 4$, this is a calculation: $g_4(4) = 60$, so $\gamma_4(4) = B_{5,\omega}(60) = \omega^2 2\#\omega \cdot 2$.

Suppose $\gamma_n(n) \geq \omega$. Then $\gamma_n(n+1) > \gamma_n(n) \geq \omega$, and since $\tau(\omega) = 1 < n+1$, also $\gamma_{n+1}(n+1) \geq \omega$. □

Lemma 9.5. *For all $\alpha < \epsilon_0$ there is an n so that, for all $m \geq n$, $\gamma_m(m) \geq \alpha$.*

Proof. For convenience, we "round α up" by replacing any finite terms with an extra ω—that is, replace $\alpha'\#k$ with $\alpha'\#\omega \geq \alpha'\#k$.

We let $n = \max\{B_{\omega,2}(\alpha), 4, \tau(\alpha)\}$. For any $m \geq n$, we have $\gamma_m(2m) > \gamma_1(m)\#\gamma_m(m) \geq \alpha\#\omega$. Therefore $\gamma_{m+1}(2m) \geq \alpha\#m\#1$, $\gamma_{m+2}(2m) \geq \alpha\#m$,

and in general, for $k < m+2$, $\gamma_{m+k}(2m) \geq \alpha\#(m+2-k)$. In particular, $\gamma_{2m}(2m) \geq \alpha\#1 > \alpha$.

Also $\gamma_{2m}(2m+1) \geq \gamma_{2m}(2m) \geq \alpha\#1$, so $\gamma_{2m+1}(2m+1) \geq \alpha$. So for all $m \geq 2n$, $\gamma_m(m) \geq \alpha$. □

Lemma 9.6. *When $\alpha > 0$, $\tau(\alpha[i]) \leq \max\{\tau(\alpha), i\}$.*

Proof. By induction on $\alpha = \alpha'\#\omega^\beta$. If $\beta = 0$ then $\alpha[i] = \alpha'$, so $\tau(\alpha[i]) \leq \tau(\alpha)$.

If $\beta > 0$ then $\tau(\alpha[i]) = \tau(\alpha'\#\omega^{\beta[i]}i) \leq \max\{\tau(\alpha'), \tau(\beta[i]), i\} \leq \max\{\tau(\alpha), i\}$. □

Lemma 9.7. *If $\gamma_i(n) \geq \alpha > 0$ and $\tau(\alpha) \leq i$ then $\gamma_{i+1}(n) \geq \alpha[i]$.*

Proof. Since $\alpha[i] < \alpha \leq \gamma_i(n)$ and $\tau(\alpha) \leq i$, we have $\gamma_{i+1}(n) \geq \alpha[i]$. □

Theorem 9.8. *For every $\alpha < \epsilon_0$, there is some n so that, for all $m \geq n$, $\mathcal{G}(m) \geq h_\alpha(m)$.*

Proof. Let α be given and choose $n = 2\max\{B_{\omega,2}(\alpha), 4, \tau(\alpha)\}$. Consider any $m \geq n$, so $\gamma_m(m) \geq \alpha$. We claim, by induction on $i \leq h_\alpha(m) - m$, that

$$\gamma_{m+i}(m) \geq \alpha[m+1][m+2]\cdots[m+i].$$

For $i = 0$, this is given by our choice of m.

Suppose $\gamma_{m+i}(m) \geq \alpha[m+1][m+2]\cdots[m+i]$. Let $\alpha' = \alpha[m+1][m+2]\cdots[m+i]$. Then $\alpha'[m+i+1] < \alpha'$ and $\tau(\alpha'[m+i+1]) \leq \max\{\tau(\alpha'), m+i+1\} = m+i+1$, so $\alpha[m+i+1] \leq \gamma_{m+i+1}(m)$.

In particular, for $i < h_\alpha(m) - m$, $0 < \alpha[m+1]\cdots[m+i] \leq \gamma_{m+i}(m)$, so $\mathcal{G}(m) > m+i$. Therefore $\mathcal{G}(m) - m \geq h_\alpha(m) - m$, and therefore $\mathcal{G}(m) \geq h_\alpha(m)$. □

Part 3. Unprovability

10. Introduction to Unprovability

What remains is showing that Goodstein's Theorem is not provable in Peano arithmetic. We can't work formally with things in an informal language like English. Instead, Peano arithmetic works with *formulas* in the *language of arithmetic*.

Typically the language of arithmetic considers *terms* which are built from zero, successor, addition, and multiplication. However it turns out not to change anything for our purposes to include additional symbols, so we will freely use any natural numbers, exponentiation, the base change function, and the function $g_i(n)$ evaluating the i-th term of the Goodstein sequence starting with n. (Once we finish the proof, one can go back and see why these additional functions were harmless—roughly speaking, it is because these functions are bounded by functions h_α with $\alpha < \epsilon_0$.) On the other hand, we would not want to add the function $\mathcal{G}(n)$: if we already know that, for each n, the value $\mathcal{G}(n)$ exists then there is nothing left to prove!

All formulas are built from *atomic formulas*. The atomic formulas allow us to compare terms using $=$ and $<$—our atomic formulas are things like $1+1=2$, $B_{4,5}(15)>6$, or $1\cdot 3=0$ (a formula need not be true!). We also allow terms with variables, like $1+x=3$ or $2\cdot x^y<z$.

Following the convention in logic, we will combine formulas using $\wedge$ to represent "and", $\vee$ to represent "or", $\neg$ to represent "not", $\forall$ to represent "for every", and $\exists$ to represent "there exists".

Unlike some systems, we do not have an implication $\rightarrow$; instead we think of $\phi\rightarrow\psi$ as an abbreviation for $\neg\phi\vee\psi$.

In particular, Goodstein's Theorem can be expressed by

$$\forall x\ \exists y\ g_y(x)=0.$$

We will use lower case Greek letters, especially ϕ and ψ, to refer to general formulas.

When ϕ is a formula, it might have *free variables*. For instance, $2^x<17$ is a formula, but x is free—we wouldn't say $2^x<17$ is true or false without some context explaining what x is. On the other hand, in the formula $\exists x\, 2^x<17$, x is *bound*: we would say that $\exists x\, 2^x<17$ is true, because there is such an x: since $2^4<17$ is true, also $\exists x\, 2^x<17$ is true.

When we want to say this with a general formula, we'll write $\phi[x\mapsto n]$ for the formula where we replace the variable x with the number n, so when ϕ is $2^x<17$, $\phi[x\mapsto 4]$ is $2^4<17$.

Definition 10.1. A formula is a *sentence* if it has no *free variables*: that is, if every variable is bound by a quantifier.

So "$\forall x\ \exists y\ g_y(x)=0$" is a sentence but "$\exists y\ g_y(x)=0$" is not, because x is free in the latter. The sentences are the only formulas we can reasonably expect to have truth values: it could be that $\exists y\ g_y(x)=0$ is sometimes true and sometimes false, depending on what the value of x is.

Definition 10.2. PA^+ is the collection of axioms consisting of:

- All sentences of the form
$$\forall x_1\cdots\forall x_m\,\phi$$
where ϕ has no quantifiers and which are true (in the standard natural numbers). (This includes the case where $m=0$. Such a sentence is called *universal*.)
- For every formula ϕ, the induction axiom
$$\neg\phi[x\mapsto 0]\vee\exists x\,(\phi\wedge\neg\phi[x\mapsto x+1])\vee\forall x\,\phi.$$

Remark 10.3. One might object that PA^+ is not even a proper axiom system. For instance, we have taken

$$\forall x\forall y\forall z\forall c(c<3\vee\neg(x^c+y^c=z^c))$$

to be an axiom. But this is Fermat's Last Theorem, and we only know it is an axiom because of a very difficult proof. There are other universal

sentences where we do not even know whether or not they are axioms in PA^+!

The choice of both the formulation of the language of arithmetic (with all the extra function symbols) and the definition of PA^+ (with its many axioms) here is motivated by the following trade-off. The conventional definitions are much more parsimonious, and (most of) the definition here is recovered by some painstaking work showing that more complicated notations can be encoded using the basic definitions, and that one has included, not all true universal sentences, but a sufficiently large subset which we can still actually list.

The definitions here are suited to our purpose for two reasons. First, they let us skip over issues about how precisely one encodes the Goodstein sequence using only addition and multiplication. Second, it lets us highlight the structure of the proof: we'll see exactly where in our proof these function symbols but not others, and universal axioms but not existential ones, can be used.

Therefore, for a first presentation of this unprovability result, I think this is a convenient approach, focusing on the proof of unprovability rather than on the details of properly establishing Peano Arithmetic. If you aren't already familiar with a more rigorous presentation of Peano Arithmetic, I hope these notes might inspire you to learn about it elsewhere and then revisit this proof in light of that.

11. An Outline

Since the proof of unprovability is fairly complicated, we briefly outline what needs to be done.

First, in Section 12, we define a special form a formula can have, *negation normal form*, and prove that every formula is equivalent to one in negation normal form. This form will let us exploit some of the "duality" between $\wedge$ and $\vee$ and between $\forall$ and $\exists$.

In Section 13, we introduce the formal proof system Z_∞ we will work with. However we will see in Section 15 that this system is actually true powerful—Theorem 14.2 says that Z_∞ actually proves every true statement about the natural numbers—so we'll define a restriction of Z_∞ by adding ordinal bounds.

In Section 16, we will verify that this restricted system is still sufficiently strong, proving in Theorem 16.7 that anything we can prove from the aioms of PA^+, we can also prove in Z_∞ with a bounded deduction.

In Section 17, we'll give an even more restricted part of Z_∞ by considering *cut-free* deductions. In particular, in Theorem 17.1, we'll prove that there can't be a bounded, cut-free deduction of Goodstein's Theorem in PA^+.

Section 19 will be the culmination of this work, where we prove the most difficult theorem we have encountered, Theorem 19.9, which shows that we can transform deductions in Z_∞ into bounded deductions.

Combining Theorem 16.7, Theorem 19.9, and Theorem 17.1, we see that PA^+ cannot prove Goodstein's Theorem: we could take a proof of Goodstein's Theorem in PA^+ and translate it first into a bounded deduction, and then a cut-free bounded deduction, which does not exist.

12. Negation Normal Form

It is convenient to put all formulas into a normal form.

Definition 12.1. We say a formula ϕ is in *negation normal form* if the only negations in the formula are negations of atomic formulas.

This is harmless, because every formula is equivalent to a formula in negation normal form. The main step in proving this is noticing that we can negate a formula in negation normal form by "pushing in" the negations.

Lemma 12.2. *If ϕ is in negation normal form then there is a formula $\sim\phi$ in negation normal form equivalent to $\neg\phi$.*

Proof. By induction on ϕ. If ϕ is atomic, $\neg\phi$ itself is in negation normal form, so $\sim\phi$ is $\neg\phi$.

If ϕ is $\neg\phi'$ then, since ϕ is in negation normal form, ϕ' must be atomic, so $\sim\phi$ is ϕ', which is equivalent to $\neg\neg\phi'$.

If ϕ is $\phi_0 \wedge \phi_1$ then the inductive hypothesis gives us $\sim\phi_0$ and $\sim\phi_1$ in negation normal form, and since $\neg(\phi_0 \wedge \phi_1)$ is equivalent to $(\neg\phi_0) \vee (\neg\phi_1)$, we may take $\sim\phi$ to be $(\sim\phi_0) \vee (\sim\phi_1)$.

Similarly, if ϕ is $\phi_0 \vee \phi_1$ then the inductive hypothesis gives us $\sim\phi_0$ and $\sim\phi_1$ in negation normal form, and since $\neg(\phi_0 \vee \phi_1)$ is equivalent to $(\neg\phi_0) \wedge (\neg\phi_1)$, we may take $\sim\phi$ to be $(\sim\phi_0) \wedge (\sim\phi_1)$.

If ϕ is $\forall x\psi$ then the inductive hypothesis gives us $\sim\psi$ in negation normal form, and since $\neg\forall x\psi$ is equvialent to $\exists x\neg\psi$, we may take $\sim\phi$ to be $\exists x\sim\psi$.

Similarly, if ϕ is $\exists x\psi$ then the inductive hypothesis gives us $\sim\psi$ in negation normal form, and since $\neg\exists x\psi$ is equvialent to $\forall x\neg\psi$, we may take $\sim\phi$ to be $\forall x\sim\psi$. □

Theorem 12.3. *For every ϕ, there is an equivalent formula ϕ' in negation normal form.*

Proof. By induction on ϕ. If ϕ is atomic then it is automatically in negation normal form.

If ϕ is $\psi_0 \wedge \psi_1$ or $\psi_0 \vee \psi_1$ then, by the inductive hypothesis, there are ψ_0' and ψ_1' in negation normal form equivalent to ψ_0 and ψ_1, and $\psi_0' \wedge \psi_1'$ or $\psi_0' \vee \psi_1'$ is in negation normal form and equivalent to ϕ.

If ϕ is $\forall x\psi$ or $\exists x\psi$ then, by the inductive hypothesis, there is a ψ' in negation normal form equivalent to ψ, so $\forall x\psi'$ and $\exists x\psi'$ are in negation normal form.

Finally, if ϕ is $\neg\psi$ then, by the inductive hypothesis, we have ψ' in negation normal form equivalent to ψ, so $\sim\psi'$ is in negation normal form and is equivalent to ϕ. □

13. Z_∞-DEDUCTIONS

To avoid confusion between ordinary mathematical proof and the notion of formal provability in a logical system we are about to describe, we will call these formal constructions *deductions*.

The main notion we will work with is a system we call Z_∞-deductions. As we will see, it will be convenient to deduce, not just individual sentences, but finite *multisets* of sentences. A multiset is like a set except that we're allowed to have multiple copies of the same thing in it. We will denote these finite multisets by capital Greek letters like Γ, Δ, and Σ.

When we say we have deduced a set Γ, what we mean is that we have concluded that *at least one* of the sentences in Γ must be true, but we may not know which one. (For instance, we will always be able to deduce $\{\phi, \sim\phi\}$—one of ϕ and the negation of ϕ must be true.) It is conventional to abbreviate unions with commas in this context, so when we write Γ, ϕ, we mean the multiset $\Gamma \cup \{\phi\}$ which contains everything in Γ and also ϕ (and, if Γ already contained ϕ then Γ, ϕ contains one more copy of ϕ than Γ did).

The basic idea in Z_∞ is that we will be able to deduce sentences—and, more generally, finite multsets of sentences—according to certain rules which will tell us that we can deduce a multiset from other multisets. These rules should be *sound*: that is, we should agree that if we can deduce the premises then it is reasonable to deduce the conclusion.

Our deductions can't go on forever: there must be some things we can simply deduce on their own. For us, these will come from the atomic sentences. Note that for *atomic* sentences, we can always determine whether or not the sentence is true: the only atomic formulas are equalities like $s = t$ or inequalities like $s < t$, where s and t are terms built from numbers and our allowed functions (addition, multiplication, base change, and so on). Since this is a sentence, there are no variables in s or t, so we can can always just calculate both sides and determine whether the sentence is true.

So if ϕ is a true atomic sentence and $\phi \in \Gamma$, we may simple deduce ϕ:

- If ϕ is a true atomic sentence and $\phi \in \Gamma$,

$$Z_\infty \vdash \Gamma.$$

We call this rule True—our deduction is justified simply by the observation that ϕ is true (and the fact that, as an atomic sentence, ϕ is simple enough to make this observation).

Conversely, if ϕ is a *false* atomic sentence then we can observe that $\neg\phi$ must be true:

- If ϕ is a false atomic sentence and $\neg\phi \in \Gamma$,

$$Z_\infty \vdash \Gamma.$$

We call this rule True as well.

Atomic and negated atomic sentence are the only sentence we consider simple enough to justify by simply observing their truth. All other rules

will tell us how to justify sentences (or finite multisets of sentences) only by looking at other sentences we have already justified.

If we want to deduce $\phi \wedge \psi$, we would typically do this by separately deducing ϕ and ψ:

- If $Z_\infty \models \Gamma, \phi$ and $Z_\infty \models \Gamma, \psi$ then $Z_\infty \models \Gamma, \phi \wedge \psi$.

We call this rule $\mathsf{I}\bigwedge$; the I stands for "introduction": in the conclusion $\Gamma, \phi \wedge \psi$, we have *introduced* a new formula, $\phi \wedge \psi$. In this case we call $\phi \wedge \psi$ the *main formula* of this rule.

If we want to deduce $\phi \vee \psi$, we only need to deduce one of ϕ and ψ:

- If $Z_\infty \models \Gamma, \phi$ then $Z_\infty \models \Gamma, \phi \vee \psi$.
- If $Z_\infty \models \Gamma, \psi$ then $Z_\infty \models \Gamma, \phi \vee \psi$.

We call both versions of this rule $\mathsf{I}\bigvee$, and $\phi \vee \psi$ is the main formula of this rule.

To justify $\forall x\, \phi$, we would want ϕ to be true for each possible value of x—that is, for each natural number n, we would want $\phi[x \mapsto n]$ to be true.

- If, for every n, $Z_\infty \models \Gamma, \phi[x \mapsto n]$ then $Z_\infty \models \Gamma, \forall x\, \phi$.

We call this rule $\mathsf{I}\forall$, and $\forall x\, \phi$ is the main formula.

Similarly, to justify $\exists x\, \phi$, we only need to know that ϕ is true for some specific value of x.

- If there is a closed term t so that $Z_\infty \models \Gamma, \phi[x \mapsto t]$ then $Z_\infty \models \Gamma, \exists x\, \phi$.

(A closed term is a term with no variables in it.) We call this rule $\mathsf{I}\exists$, and $\exists x\, \phi$ is the main formula of this rule.

There are three more rules we need to work with our system. The first two are what we call *structural rules*: they doesn't contain important logical content, it's just needed to make our deductions function. One way we could justify Γ is if we had already justified some subset $\Delta \subseteq \Gamma$: we only need one sentence in our set to be true, so if we know that one of the sentences in Δ is true, we certainly know that one of the sentences in Γ is true as well.[1]

- If $Z_\infty \models \Delta$ and $\Delta \subseteq \Gamma$ then $Z_\infty \models \Gamma$.

We call this rule W (for "weakening"). There is no main formula for this rule.

The other structural rule is that if we have multiple copies of a formula, we could combine them:

- If $Z_\infty \models \Gamma, \phi, \phi$ then $Z_\infty \models \Gamma, \phi$.

We call this rule C (for "contraction"). ϕ is the main formula.

The last rule records a common proof technique that the rules above don't directly cover: breaking an argument into cases. Another way we could

[1]We won't usually worry to much about the technicalities of multisets, but if we're being careful, we should note that saying $\Delta \subseteq \Gamma$ for multisets means that whenever ϕ is in Δ with multiplicity m, ϕ is in Γ with multiplicity $n \geq m$.

justify Γ is to consider two separate situations, one where some formula ϕ is true, and another where ϕ is not true.

- If both $Z_\infty \vdash \Gamma, \phi$ and $Z_\infty \vdash \Gamma, \sim\phi$ then $Z_\infty \vdash \Gamma$.

We call this the Cut rule. We call ϕ the main formula, but note that ϕ behaves differently than in all other rules, since ϕ does *not* appear in the conclusion.

We need to establish that it is possible to deduce anything interesting. As a starting point—and as a little practice using these rules—basically any interesting deduction is going to need the fact that $Z_\infty \vdash \phi, \sim\phi$ for every formula ϕ, and the proof illustrates nicely how the rules work and pair up.

14. Examples in Z_∞

Theorem 14.1. *For every ϕ, $Z_\infty \vdash \phi, \sim\phi$.*

Proof. The proof is by induction on the formula ϕ.

When ϕ either an atomic formula or a negated atomic formula, either ϕ or $\sim\phi$ must be true, so either way the True axiom gives us $Z_\infty \vdash \phi, \neg\phi$.

Suppose ϕ is a conjunction—say, ϕ is $\psi_0 \wedge \psi_1$, so $\sim\phi$ is $(\sim\psi_0) \vee (\sim\psi_1)$. It is typical when looking for deductions to work backwards: to start with the conclusion and identify what rule we could have used to deduce it.

We wish to conclude

$$\psi_0 \wedge \psi_1, (\sim\psi_0) \vee (\sim\psi_1).$$

The most natural ways to deduce this would be using either an $\mathsf{I}\wedge$ rule or an $\mathsf{I}\vee$ rule. $\mathsf{I}\vee$ requires us to make a choice—we would need to choose either $\psi_0 \wedge \psi_1, \sim\psi_0$ or $\psi_0 \wedge \psi_1, \sim\psi_1$—and since we would rather put off making choices, we try deducing this using $\mathsf{I}\wedge$ from two simpler sets

$$\psi_0, (\sim\psi_0) \vee (\sim\psi_1) \text{ and } \psi_1, (\sim\psi_0) \vee (\sim\psi_1).$$

We need to deduce each of these, and now $\mathsf{I}\vee$ looks more tempting: we can deduce $\psi_0, (\sim\psi_0) \vee (\sim\psi_1)$ from $\psi_0, \sim\psi_0$ and $\psi_1, \sim\psi_1$ respectively. And, by the inductive hypothesis, we can assume we already know how to deduce these.

It's customary to write these deductions as trees, so we can gather up the deduction we just described as follows:

$$\dfrac{\dfrac{\dfrac{\vdots}{\psi_0, \sim\psi_0}}{\psi_0, (\sim\psi_0) \vee (\sim\psi_1)}\,\mathsf{I}\vee \qquad \dfrac{\dfrac{\vdots}{\psi_1, \sim\psi_1}}{\psi_1, (\sim\psi_0) \vee (\sim\psi_1)}\,\mathsf{I}\vee}{\psi_0 \wedge \psi_1, (\sim\psi_0) \vee (\sim\psi_1)}\,\mathsf{I}\wedge$$

Note that even though we figure out what this deduction should look like by starting at the bottom, we write it like a proof, starting at the top with things we already know we can deduce and working down towards the bottom, justifying each step based on steps above it.

This notation is a convenient shorthand, which makes it much easier to quickly describe deductions. We can consider the other cases for what ϕ might be.

When ϕ is a disjunction, the argument is entirely symmetric.

When ϕ is $\forall x\, \psi$ (or, symmetrically, when ϕ starts with an $\exists$ quantifier), we can use the deduction

$$\dfrac{\dfrac{\dfrac{\vdots}{\psi[x \mapsto 0], \sim\psi[x \mapsto 0]}}{\psi[x \mapsto 0], \exists x \sim\psi}\,\mathsf{I}\exists \quad \cdots \quad \dfrac{\dfrac{\vdots}{\psi[x \mapsto n], \sim\psi[x \mapsto n]}}{\psi[x \mapsto n], \exists x \sim\psi}\,\mathsf{I}\exists \quad \cdots}{\forall x\, \psi, \exists x \sim\psi}\,\mathsf{I}\forall$$

To discover this deduction, we can work the same way as for the conjunction: we know that we want to use either $\mathsf{I}\forall$ or $\mathsf{I}\exists$, and we prefer $\mathsf{I}\forall$ because we want to postpone making the choice that $\mathsf{I}\forall$ requires. So we then need to deduce $\psi[x \mapsto n], \exists x \sim\psi$ for every natural number n, which we can do in a uniform way by using $\mathsf{I}\exists$. To use $\mathsf{I}\exists$, we do need to make a choice—which number we will use to replace x—but of course, now we have the number n to use as our guess. □

As this proof illustrates, when $Z_\infty \vdash \Gamma$, we can think of this as corresponding to a tree: some rule justified Γ from some other finite sets, which were in turn justified by other finite sets, and so on, reaching back until we ultimately encounter leaves justified by True.

It's actually much too easy to deduce things in Z_∞; in fact, we don't really gain anything over ordinary mathematical proof. For instance, certainly $Z_\infty \vdash \forall x\, \exists y\, g_y(x) = 0$: we know that Goodstein's Theorem is true, so for each n we know that $g_{\mathcal{G}(n)}(n) = 0$ is true. But then, by True, taking $m = \mathcal{G}(n)$ we have $Z_\infty \vdash g_m(n) = 0$, and so for each n we have $Z_\infty \vdash \exists y\, g_y(n) = 0$ by $\mathsf{I}\exists$, and so by $\mathsf{I}\forall$ we have $Z_\infty \vdash \forall x\, \exists y\, g_y(x) = 0$.

In fact, if ϕ is *any* sentence true of the natural numbers, we have $Z_\infty \vdash \phi$, by a similar argument.

Theorem 14.2. *If ϕ is a true sentence, $Z_\infty \vdash \phi$.*

Proof. By induction on ϕ. If ϕ is atomic then, since ϕ is true by assumption, $Z_\infty \vdash \phi$ using the rule True.

If ϕ is $\psi_0 \wedge \psi_1$ then both ψ_0 and ψ_1 must be true. By the inductive hypothesis, $Z_\infty \vdash \psi_0$ and $Z_\infty \vdash \psi_1$, so using the $\mathsf{I}\wedge$ rule, $Z_\infty \vdash \psi_0 \wedge \psi_1$.

If ϕ is $\psi_0 \vee \psi_1$ then one of ψ_0 and ψ_1 must be true. Say ψ_0 is true (the ψ_1 case is symmetric). Then, by the inductive hypothesis, $Z_\infty \vdash \psi_0$, so by $\mathsf{I}\vee$, $Z_\infty \vdash \psi_0 \vee \psi_1$.

If ϕ is $\forall x\, \psi$ then, for each natural number n, $\psi[x \mapsto n]$ must be true, so by the inductive hypothesis $Z_\infty \vdash \psi[x \mapsto n]$, and then using the $\mathsf{I}\forall$ rule, $Z_\infty \vdash \forall x\, \psi$.

Finally, if ϕ is $\exists x\, \psi$ then there must be some natural number n so that $\psi[x \mapsto n]$ is true, so by the inductive hypothesis $Z_\infty \vdash \psi[x \mapsto n]$, and then by the I$\exists$ rule, $Z_\infty \vdash \exists x\, \pi$. □

So, as defined, Z_∞ isn't useful: knowing that something has a deduction gives us no more information than just knowing that we proved it somehow.

15. Ordinal Bounds on Deductions

A deduction system is supposed to be a restricted form of argument. We shouldn't be able to just import everything we know about the natural numbers into it directly, because if we can do that, these deductions aren't doing anything for us. The point of a formal deduction system is that it asks us to decide in advance what the valid kinds of reasoning are.

In particular, we want to learn something about deductions of Goodstein's Theorem: deductions of $\forall x\, \exists y\, g_y(x) = 0$. But the deduction of $\forall x\, \exists y\, g_y(x) = 0$ described by Theorem 14.2 is as complicated as the proof—it requires already knowing the exact value of $\mathcal{G}(n)$ for each n.

The more common way to do this would be to restrict the I$\forall$ rule, by replacing it with some specific, finitary ways we might have obtained proofs of $\Gamma, \phi[x \mapsto n]$ for each n—one common choice would be to have a "generalization" rule, where we demand a single deduction of Γ, ϕ (with the variable x free in ϕ). This is like demanding that the deduction of $\Gamma, \phi[x \mapsto n]$ not really depend on n—that, no matter what n is, the deduction of $\Gamma, \phi[x \mapsto n]$ consists of the same rules in the same order.

It's more convenient for us to take a different approach: we're going to do is restrict the I$\exists$ rule. When we want to deduce $\Gamma, \exists x\, \phi$ from $\Gamma, \phi[x \mapsto n]$, we going to put a bound on how big n can be, so that we can't always just guess the correct value of n. If n is a small number, that's fine, but if n is really big, we won't be allowed to just use I$\exists$: our restricted deductions will force us to find some more abstract way to prove that this number exists.

(Of course, as soon as we've finished defining our restricted system of deductions, our next obligation will be to prove that they're not *too* restricted—that a lot of arguments we would like to make are still valid.)

We have to decide what too big is. Presumably, it's going to depend on the numbers in the formula ϕ. The number $10^{10^{10}} + 1$ is pretty big, but if we want to prove $\exists x\, 10^{10^{10}} < x$, it's not so unreasonable to use I$\exists$ and say that since $10^{10^{10}} < 10^{10^{10}} + 1$ is true, so is $\exists x\, 10^{10^{10}} < x$. After all, while $10^{10^{10}} + 1$ is a large number, it's not really that much bigger than $10^{10^{10}}$ which is already in the formula.

Even if we put this restriction on the witnesses we get in the I$\exists$ rule, there's another way we could smuggle large numbers into our proofs: by making the proofs very tall. We could deduce $\exists y\, g_{500}(y) = 0$ from the finite set $\{\exists y\, g_{500}(y) = 0,\ g_{500}(501) = 0\}$—remember that this means we only promise that one of the two statements in the set is true. Of course, $g_{500}(501) = 0$

happens to be false, but it gives us license to derive this, using $\mathsf{I}\exists$ again, from

$$\{\exists y\, g_{500}(y) = 0,\ g_{500}(501) = 0, g_{500}(502) = 0\}$$

which we can derive using $\mathsf{I}\exists$ from

$$\{\exists y\, g_{500}(y) = 0,\ g_{500}(501) = 0, g_{500}(502) = 0, g_{500}(503) = 0\},$$

and so on. By making this deduction very, very tall, we'll eventually reach the correct value of $\mathcal{G}(500)$, giving us a valid Z_∞ deduction.

This, also, seems like cheating—again, having a formal deduction isn't telling us any more than an ordinary mathematical proof. So when we define a restricted deduction, we need to bound two things: how much the numbers can grow from one step to the one above it, and how many steps there can be.

Our bounds will have two parts, an ordinal α and a number k, and we will write $Z_\infty \left|\frac{\alpha,k}{}\right. \Gamma$ to indicate a deduction bounded by α and k. α will be a bound on how "tall" our deduction can be: we will require that all premises be bounded by ordinals $< \alpha$. In particular, this enforces that our deduction be finite in a specific way: if we deduce something with ordinal bound α, the previous step must have bound $\alpha' < \alpha$, and the step before that must have bound $\alpha'' < \alpha' < \alpha$, and since the ordinals are well-founded, this process must eventually stop.

k will help restrict how "complicated" the steps can be: it will tell us that the numbers we work with should be less than k, or at least not too much bigger than k.

In particular, when the previous step has bound $\beta < \alpha$, we shouldn't let β be too complicated. In fact, what we really want is that $\beta = \alpha[k]$, where k is the numeric bound. That would make sure that $\beta < \alpha$, while also ensuring that β isn't smuggling in a lot of extra complexity. It turns out to be a little inflexible to require that β be exactly equal to $\alpha[k]$; what's convenient is to require that the ordinal bound always have complexity $< k$.

Formally, we defined the restricted rules of deductions

$$Z_\infty \left|\frac{\alpha,k}{}\right. \Delta$$

inductively by:

- If ϕ is a true atomic formula, $\phi \in \Gamma$, and $\tau(\alpha) < k$ then $Z_\infty \left|\frac{\alpha,k}{}\right. \Gamma$,
- If ϕ is a false atomic formula, $\neg\phi \in \Gamma$, and $\tau(\alpha) < k$ then $Z_\infty \left|\frac{\alpha,k}{}\right. \Gamma$,
- If there are $\beta, \gamma < \alpha$ with both $\tau(\beta), \tau(\gamma) < k$ so that $Z_\infty \left|\frac{\beta,k}{}\right. \Gamma, \phi$ and $Z_\infty \left|\frac{\gamma,k}{}\right. \Gamma, \psi$ then $Z_\infty \left|\frac{\alpha,k}{}\right. \Gamma, \phi \wedge \psi$,
- If there is a $\beta < \alpha$ with $\tau(\beta) < k$ so that $Z_\infty \left|\frac{\beta,k}{}\right. \Gamma, \phi$ then $Z_\infty \left|\frac{\alpha,k}{}\right. \Gamma, \phi \vee \psi$,
- If there is a $\beta < \alpha$ with $\tau(\beta) < k$ so that $Z_\infty \left|\frac{\beta,k}{}\right. \Gamma, \psi$ then $Z_\infty \left|\frac{\alpha,k}{}\right. \Gamma, \phi \vee \psi$,

- If, for every n, there is a $\beta_n < \alpha$ with $\tau(\beta_n) < \max\{k, n\}$ so that $Z_\infty \vdash^{\beta_n, \max\{k,n\}} \Gamma, \phi[n]$ then $Z_\infty \vdash^{\alpha,k} \Gamma, \forall x\, \phi$,
- If there is some closed term t such that the value of t is $\leq h_\alpha(k)$ and some $\beta < \alpha$ with $\tau(\beta) < k$ such that $Z_\infty \vdash^{\beta,k} \Gamma, \phi[x \mapsto t]$ then $Z_\infty \vdash^{\alpha,k} \Gamma, \exists x\, \phi$,
- If there is a $\beta < \alpha$ with $\tau(\beta) < k$ so that $Z_\infty \vdash^{\beta,k} \Delta$ and $\Delta \subseteq \Gamma$ then $Z_\infty \vdash^{\alpha,k} \Gamma$,
- If there is a $\beta < \alpha$ with $\tau(\beta) < k$ so that $Z_\infty \vdash^{\beta,k} \Gamma, \phi, \phi$ then $Z_\infty \vdash^{\alpha,k} \Gamma, \phi$,
- If there are $\beta, \gamma < \alpha$ with $\tau(\beta), \tau(\gamma) < k$ so that both $Z_\infty \vdash^{\beta,k} \Gamma, \phi$ and $Z_\infty \vdash^{\gamma,k} \Gamma, \sim\phi$ then $Z_\infty \vdash^{\alpha,k} \Gamma$.

We are left with establishing two facts. First, that many interesting things can be deduced in Z_∞ using restricted deductions. To demonstrate this, we will show that the restricted version of Z_∞ includes Peano arithmetic: we will show that we can prove each of the axioms of Peano arithmetic using short deductions (that is, deductions with a small ordinal bound).

Next, we need to establish that restricted deductions are restricted in an interesting way. This will split into two pieces. We will show that deductions with an additional restriction—that they not use the Cut rule—are limited. In particular, we will show that it is not possible to prove Goodstein's Theorem using a short (ordinal bound $< \epsilon_0$) deduction without the cut rule. Second (and this will be quite difficult), we will prove that we can eliminate the Cut rule: if we have a deduction which does use the Cut rule, we can transform the deduction in a systematic way to get rid of all uses of the Cut rule. The price of this elimination will be to increase the ordinal bound, but in a controlled way—in particular, a way that keeps us below the bound ϵ_0.

16. From PA^+ to Z_∞

We want to show that anything which can be proven from the axioms of PA^+ can be deduced in Z_∞ in a sufficiently nice way.

We need to verify that the axioms of Peano arithmetic can all be derived using bounded Z_∞ deductions. Peano arithmetic has two groups of axioms. The first group consists of a number of basic facts about the symbols, namely:

- $\forall x \forall y\, \neg(x + 1 = y + 1) \vee (x = y)$,
- $\forall x\, \neg(x + 1 = 0)$,
- $\forall x\, x + 0 = x$
- $\forall x \forall y\, x + (y + 1) = (x + y) + 1$,
- $\forall x\, x \cdot 0 = 0$
- $\forall x \forall y\, x \cdot (y + 1) = x \cdot y + x$.

When we include extra symbols, like exponentiation or $g_i(n)$, we need to include additional corresponding "defining axioms". The requirement is that

these should be *universal*: that is, they must have the form

$$\forall x_1 \forall x_2 \cdots \forall x_k \, \psi$$

where ψ does not contain any quantifiers.

Lemma 16.1. *If ψ is a true universal sentence with k variables then $Z_\infty \left|\frac{}{}^{k,k+1}\right. \psi$.*

Proof. Consider a sentence with $k = 2$, like $\forall x \forall y \, x + (y+1) = (x+y) + 1$. We have the deduction

$$\dfrac{\cdots \quad \dfrac{\dfrac{}{n + (m+1) = (n+m) + 1}\,\mathsf{True} \quad \cdots}{\forall y \, n + (y+1) = (n+y) + 1}\,\mathsf{I}\forall \quad \cdots}{\forall x \forall y \, x + (y+1) = (x+y) + 1}\,\mathsf{I}\forall$$

□

The remaining axioms of Peano arithmetic are the induction axioms: for every formula ϕ with the free variable x, we have

$$\sim\phi[x \mapsto 0], \exists x \left[\phi \wedge (\sim\phi[x \mapsto x+1])\right], \forall x \, \phi.$$

It will be useful to have a measure of complexity for formulas, to help guide our inductive arguments. The simplest notion we could use is "rank", which simply assigns a number to a formula based on how many "steps" it takes to build.

Definition 16.2. The *rank* of a formula $rk(\phi)$ is defined inductively by:

- the rank of an atomic formula is 0,
- the rank of a negated atomic formula is 0,
- the ranks of $\phi \wedge \psi$ and $\phi \vee \psi$ are $\max\{rk(\phi), rk(\psi)\} + 1$,
- the ranks of $\forall x \, \phi$ and $\exists x \, \phi$ are $rk(\phi) + 1$.

Lemma 16.3. *For every ϕ with $rk(\phi) = n$ then, for any $m > 2n$, $Z_\infty \left|\frac{}{}^{2n,m}\right. \phi, \sim\phi$.*

Proof. This follows just be counting the number of steps in our argument above that $Z_\infty \vdash \phi, \sim\phi$.

□

We will need to observe that we can always make the numeric part of the bound bigger:

Lemma 16.4. *If $Z_\infty \left|\frac{}{}^{\alpha,k}\right. \Gamma$ and $k < m$ then $Z_\infty \left|\frac{}{}^{\alpha,m}\right. \Gamma$.*

Proof. By induction on the deduction: we may simply replace the bound k with m at every step of the deduction. □

Lemma 16.5. *For any ϕ and any $m > \max\{4, 2rk(\phi)\}$,*

$$Z_\infty \left|\frac{}{}^{\omega \cdot 4 \# 2rk(\phi) \# 2, m}\right. \sim\phi[x \mapsto 0], \exists x \, \phi \wedge \sim\phi[x \mapsto x+1], \forall x \, \phi.$$

Proof. In order to use I$\forall$, we would like to deduce, for each n,

$$\sim\phi[x \mapsto 0], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \phi[x \mapsto n].$$

We construct this deduction by induction on n. For $n = 0$, we have

$$\dfrac{\begin{array}{c}\vdots\\ \sim\phi[x \mapsto 0], \phi[x \mapsto 0]\end{array}}{\sim\phi[x \mapsto 0], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \phi[x \mapsto 0]}\,\mathsf{W}$$

Suppose we have already deduced

$$\sim\phi[x \mapsto 0], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \phi[x \mapsto n].$$

Then we can construct a deduction

$$\dfrac{\dfrac{\dfrac{\begin{array}{c}\vdots\\ \sim\phi[x \mapsto 0], \phi[x \mapsto n], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])]\end{array}}{\sim\phi[x \mapsto 0], \phi[x \mapsto n], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \phi[x \mapsto n+1]}\,\mathsf{W} \quad \dfrac{\begin{array}{c}\vdots\\ \sim\phi[x \mapsto n+1], \phi[x \mapsto n+1]\end{array}}{\sim\phi[x \mapsto 0], \sim\phi[x \mapsto n+1], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \phi[x \mapsto n+1]}\,\mathsf{W}}{\sim\phi[x \mapsto 0], \phi[x \mapsto n] \wedge (\sim\phi[x \mapsto n+1])], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \phi[x \mapsto n+1]}\,\mathsf{I}\wedge}{\sim\phi[x \mapsto 0], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \phi[x \mapsto n+1]}\,\mathsf{I}\exists}{\sim\phi[x \mapsto 0], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \phi[x \mapsto n+1]}\,\mathsf{C}$$

Since we can do this for each n, we can put them together using the $I\forall$ rule:

$$\dfrac{\cdots \qquad \sim\phi[x \mapsto 0], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \phi[x \mapsto n] \qquad \cdots}{\sim\phi[x \mapsto 0], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \forall n\,\phi}$$

To obtain bounds, we need to count steps. We can inductively count the number of steps in the deductions of

$$\phi[x \mapsto 0], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \phi[x \mapsto n]:$$

it takes $2rk(\phi)\#1$ steps to deduce this when $n = 0$, and when the n-th deduction takes d_n steps, the $n+1$-st takes $\max\{d_n + 4, 2rk(\phi) + 4\} = d_n + 4$ steps. So, in general, the n-th deduction takes at most $4n + 2rk(\phi) + 1$ steps. So we can take

$$Z_\infty \vdash^{\omega\cdot 4\#2rk(\phi)\#2, m} \sim\phi[x \mapsto 0], \exists x\,[\phi \wedge (\sim\phi[x \mapsto x+1])], \forall n\,\phi.$$

□

We would like to refine the previous lemma to actually show that we can deduce the induction axiom of PA^+, in precisely the form we stated it.

Corollary 16.6. *For any ϕ and any $m > \max\{7, 2rk(\phi)\}$,*

$$Z_\infty \vdash^{\omega\cdot 4\#2rk(\phi)\#8, m} \sim\phi[x \mapsto 0] \vee \exists x\,\phi \wedge \sim\phi[x \mapsto x+1] \vee \forall x\,\phi.$$

Proof. We can extend the deduction from the lemma:

$$
\begin{array}{cl}
\vdots & \\
\sim\phi[x \mapsto 0], \exists x\, \phi \wedge \sim\phi[x \mapsto x+1]\, , \forall x\, \phi & \\
\hline
\sim\phi[x \mapsto 0] \vee \exists x\, \phi \wedge \sim\phi[x \mapsto x+1]\; \exists x\, \phi \wedge \sim\phi[x \mapsto x+1]\, , \forall x\, \phi & \mathrm{I}\vee \\
\hline
\sim\phi[x \mapsto 0] \vee \exists x\, \phi \wedge \sim\phi[x \mapsto x+1]\; \sim\phi[x \mapsto 0] \vee \exists x\, \phi \wedge \sim\phi[x \mapsto x+1]\, , \forall x\, \phi & \mathrm{I}\vee \\
\hline
\sim\phi[x \mapsto 0] \vee \exists x\, \phi \wedge \sim\phi[x \mapsto x+1]\, , \forall x\, \phi & \mathrm{C} \\
\hline
\sim\phi[x \mapsto 0] \vee \exists x\, \phi \wedge \sim\phi[x \mapsto x+1] \vee \forall x\, \phi\, , \forall x\, \phi & \mathrm{I}\vee \\
\hline
\sim\phi[x \mapsto 0] \vee \exists x\, \phi \wedge \sim\phi[x \mapsto x+1] \vee \forall x\, \phi\, , \sim\phi[x \mapsto 0] \vee \exists x\, \phi \wedge \sim\phi[x \mapsto x+1] \vee \forall x\, \phi & \mathrm{I}\vee \\
\hline
\sim\phi[x \mapsto 0] \vee \exists x\, \phi \wedge \sim\phi[x \mapsto x+1] \vee \forall x\, \phi & \mathrm{C}
\end{array}
$$

□

This particular pattern is very common in Z_∞: to get from ϕ, ψ to $\phi \vee \psi$ we use I$\vee$ twice, one on each of ϕ and ψ, and then use contraction to combine them into a single copy.

Finally, we are ready to prove the following.

Theorem 16.7. *If a sentence ϕ can be proven in PA^+ then there is an ordinal $\alpha < \epsilon_0$ and a natural number k so that $Z_\infty \overset{\alpha,k}{\vdash} \phi$.*

Proof Sketch. A complete proof would require filling in the details of what it means to prove something in PA^+—that is, picking some sort of deduction system for first-order logic. Since that would be a digression, we'll focus on the main points, which would be mostly the same no matter what proof system one used.

Let us assume that a proof in PA^+ is a sequence of formulas

$$\phi_0, \phi_1, \dots, \phi_n$$

where each formula is either an axiom or is justified by previous steps and one of the rules of first-order logic. (This is the typical set-up in a "Hilbert-style" deduction system.) What we'd like to do is show, by induction on i, that $Z_\infty \vdash \phi_i$ with appropriate bounds.

These bounds will depend a bit on the specific formulas ϕ_i. We promised that every function symbol we put into the language of arithmetic would be bounded by h_α for some α. Only finitely many symbols appear in the finitely many formulas ϕ_i, so we may choose a single ordinal β_0 larger than any of these α, and with $\tau(\alpha) \leq \tau(\beta_0)$ for each of these ordinals α. We may also take a natural number $r = \max_i\{2rk(\phi_i)\} + 8$. Finally, we can take a value b to be the largest value of any closed term appearing in any ϕ_i

There's an immediate complication: ϕ_i could be an arbitrary formula, while Z_∞ can only deduce sentences. We can address this by replacing each free variable in ϕ_i with a natural number: instead of trying to deduce the formula ϕ_i, we'll try to deduce the sentence

$$\phi_i[x_1 \mapsto m_1] \cdots [x_d \mapsto m_d].$$

In order to get the proof to go through right, we have to be careful about uniformity—the ordinal part of the bound should only depend on i, and

not on the choice of numbers m_j that we substitute in. What we show, by induction on i, is

> Suppose ϕ_i is a formula free variables $x_1, \dots, x_d$. Then for any natural numbers $m_1, \dots, m_d$,
>
> $$Z_\infty \vdash^{\beta_0 \# \omega \cdot 4 \# r \# 7 \# 6i, \max\{b, m_1, \dots, m_d\}} \phi_i[x_1 \mapsto m_1] \cdots [x_d \mapsto m_d].$$

If ϕ_i is an axiom, ϕ_i is either a true universal sentence, so $Z_\infty \vdash^{k,0} \phi_i$ by Lemma 16.1, or ϕ_i is an induction axiom for ψ, so $Z_\infty \vdash^{\omega \cdot 4 \# 2s \# 7, 0} \phi_i$ for some s with $2s + 8 \leq r$ by Corollary 16.6. By Lemma 16.4, we can increase the numeric bounds, so $Z_\infty \vdash^{k, \max\{b, m_1, \dots, m_d\}} \phi_i$ or $Z_\infty \vdash^{\omega \cdot 4 \# 2s \# 7, \max\{b, m_1, \dots, m_d\}} \phi_i$, respectively. Since $k < \beta_0 \# \omega \cdot 4 \# r \# 6i, \max\{b, m_1, \dots, m_d\}$ and $\omega \cdot 3 \# 2s \# 7 < \beta_0 \# \omega \cdot 4 \# r \# 6i, \max\{b, m_1, \dots, m_d\}$, we can apply the weakening rule to conclude $Z_\infty \vdash^{\beta_0 \# \omega \cdot 4 \# r \# 6i, \max\{b, m_1, \dots, m_d\}} \phi_i$.

Here we've used the weakening rule in a trivial way—we have "weakened" the set $\{\phi_i\}$ to the set $\{\phi_i\}$—and have instead found it useful because it lets us weaken the *bound* instead of the conclusion. The weakening rule has been defined in just the right way to make this use possible.

Otherwise, ϕ_i follows from previous steps using some rule of first-order logic, so we need to check that the rules of Z_∞ capture the usual rules of first-order logic. Again, the details depend on the particular choice of rules of first-order logic, but let us consider typical examples of what the quantifier rules might look like, since these are the crucial rules. (In particular, if we made a mistake in the definition of the bounds, it would probably show up here.)

Most systems of first-order logic have some sort of "generalization" rule—something like

> from $\psi_0 \vee \psi_1$ where x does not appear free in ψ_0, conclude $\psi_0 \vee \forall x\, \psi_1$.

(More commonly, a system might have the equivalent rule deducing $(\sim \psi_0) \to \forall x\, \psi_1$ from $(\sim \psi_0) \to \psi_1$.)

So suppose ϕ_i is $\psi_0 \vee \forall x\, \psi_1$ and there is some $j < i$ so that ϕ_j is $\psi_0 \vee \psi_1$. Let us also, for notational simplicity, assume that ψ_0 is a sentence and x is the only free variable in ψ_1. So, by the inductive hypothesis, for every n we have $Z_\infty \vdash^{\beta_0 \# \omega \cdot 4 \# r \# 6j, \max\{b, n\}} \psi_0 \vee \psi_1[x \mapsto n]$. Then we have a deduction

$$\dfrac{\cdots \qquad \dfrac{\dfrac{\vdots}{\psi_0 \vee \psi_1[x \mapsto n]}}{\psi_0, \psi_1[x \mapsto n], \psi_0 \vee \psi_1[x \mapsto n]}\,\text{W} \quad \dfrac{\dfrac{\dfrac{\vdots}{\psi_0, \sim \psi_0}}{\psi_0, \psi_1[x \mapsto n], \sim \psi_0}\,\text{W} \quad \dfrac{\dfrac{\vdots}{\psi_1[x \mapsto n], \sim \psi_1[x \mapsto n]}}{\psi_0, \psi_1[x \mapsto n], \sim \psi_1[x \mapsto n]}\,\text{W}}{\psi_0, \psi_1[x \mapsto n], \sim \psi_0 \wedge \sim \psi_1[x \mapsto n]}\,\text{I}\wedge}{\dfrac{\psi_0, \psi_1[x \mapsto n] \;\text{Cut} \qquad \cdots}{\dfrac{\psi_0, \forall x\, \psi_1}{\dfrac{\psi_0 \vee \forall x\, \psi_1, \forall x\, \psi_1}{\dfrac{\psi_0 \vee \forall x\, \psi_1, \psi_0 \vee \forall x\, \psi_1}{\psi_0 \vee \forall x\, \psi_1}\,\text{C}}\,\text{I}}\,\text{I}\vee}\,\text{I}\forall}$$

Counting steps, we get $Z_\infty \vdash^{\beta_0 \# \omega\cdot 4 \# r \# 6i, \max\{b,n\}} \psi_0 \vee \forall x\, \psi_1$, as desired.

Most systems of first-order logic have a rule along the lines of

from $\psi_0 \vee \psi_1[x \mapsto t]$, conclude $\psi_0 \vee \exists x\, \psi_1$.

So suppose ϕ_i is $\psi_0 \vee \exists x\, \psi_1$ and there is some $j < i$ so that ϕ_j is $\psi_0 \vee \psi_1[x \mapsto t]$. We'll miss an essential subtlety if we assume these are sentences, but we can assume there's just one variable besides x, say y. So, by the inductive hypothesis, for each m we have

$$Z_\infty \vdash^{\beta_0 \# \omega\cdot 4 \# r \# 6j, \max\{b,m\}} \psi_0[y \mapsto n] \vee \psi_1[x \mapsto t][y \mapsto m].$$

Notice the order of substitutions—the issue is that t itself might contain the free variable y, and so the value of t might depend on m.

Let us abbreviate $\psi_0[y \mapsto n]$ by ψ_0', $t[y \mapsto m]$ by y', and $\psi_1[y \mapsto m]$ by ψ_1'. Note that $\psi_1[x \mapsto t][y \mapsto m]$ is the same as $\psi_1'[x \mapsto t']$. Then we have a deduction

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\psi_0' \vee \psi_1'[x \mapsto t']}\,\mathrm{W}}{\psi_0', \psi_1'[x \mapsto t'], \psi_0' \vee \psi_1'[x \mapsto t']}\,\mathrm{W} \qquad \dfrac{\dfrac{\vdots}{\psi_0', \sim \psi_0'}\,\mathrm{W}}{\psi_0', \psi_1'[x \mapsto t'], \sim \psi_0'}\,\mathrm{W} \qquad \dfrac{\vdots}{\psi_1'[x \mapsto t'], \sim \psi_1'[x \mapsto t']}\,\mathrm{W}}{\psi_0', \psi_1'[x \mapsto t'], \sim \psi_0' \wedge \sim \psi_1'[x \mapsto t']}\,\mathrm{I}\wedge}{\psi_0', \psi_1'[x \mapsto t']}\,\mathrm{Cut}}{\psi_0', \exists x\, \psi_1'}\,\mathrm{I}\exists}{\psi_0' \vee \exists x\, \psi_1', \exists x\, \psi_1'}\,\mathrm{I}\vee}{\psi_0' \vee \exists x\, \psi_1', \psi_0' \vee \exists x\, \psi_1'}\,\mathrm{I}\vee}{\psi_0' \vee \exists x\, \psi_1'}\,\mathrm{C}$$

Counting steps, we get

$$Z_\infty \vdash^{\beta_0 \# \omega\cdot 4 \# r \# 6i, \max\{b,m\}} \psi_0' \vee \exists x\, \psi_1'.$$

We should take particular care with the I$\exists$ step: we need to know that $t' \leq h_{\beta_0 \# \omega\cdot 4 \# r \# 6i}(m)$ for all m; we know this because we chose β_0 to be large enough to ensure that this was the case, and this was only possible because we promised that the functions which got symbols in the language of arithmetic would be bounded by some h_α. □

Technical Aside

Lemma 16.8. *If $\beta > \alpha$ and $\tau(\alpha) < k$ then $\beta[k] \geq \alpha$.*

Proof. By induction on β, we show that

for all $\alpha < \beta$ with $\tau(\alpha) < k$, $\beta[k] \geq \alpha$.

For the statement to be non-trivial, $\beta > 0$, so

$$\beta = \omega^{\beta_1} \# \cdots \# \omega^{\beta_s}.$$

If $\alpha = 0$, the statement is immediate, so assume $\alpha > 0$ and therefore

$$\alpha = \omega^{\alpha_1} \# \cdots \# \omega^{\alpha_r}.$$

If $\beta = \alpha\#\beta'$—that is, if $s > r$ and $\beta_i = \alpha_i$ for all $i \leq r$—then $\beta[k] = \alpha\#\beta'[k] \geq \alpha$ and we are done. So consider the case where there is some $i < r$ with $\alpha_i < \beta_i$. If the least such i is $< s$ then

$$\beta[k] = \omega^{\beta_1}\# \cdots \omega^{\beta_{s-1}}\#\omega^{\beta_s[k]}\# \cdots \omega^{\beta_s[k]}$$

which is still $> \alpha$, so again we are done.

So the remaining case is when

$$\beta = \omega^{\alpha_1}\# \cdots \omega^{\alpha_{s-1}}\#\omega^{\beta_s}$$

where $\beta_s > \alpha_s$, and we must show that $\omega^{\beta_s}[k] \geq \omega^{\alpha_s}\# \cdots \#\omega^{\alpha_r}$. (Indeed, we will show a strict inequality.)

Since $\beta_s > \alpha_s$, also $\beta_s > 0$, so we have

$$\omega^{\beta_s}[k] = \underbrace{\omega^{\beta_s[k]}\# \cdots \omega^{\beta_s[k]}}_{k \text{ times}}.$$

By the inductive hypothesis, $\beta_s[k] \geq \alpha_s$. If $\beta_s[k] > \alpha_s$, we are done. So suppose $\beta_s[k] = \alpha_s$. Because $\tau(\alpha) < k$, there must be some $j < k$ so that either $\alpha_{s+j} < \alpha_s$ or $s + j = r + 1$ (and therefore α_{s+j} is undefined). Either way,

$$\omega^{\beta_s}[k] > \omega^{\alpha_s}\# \cdots \#\omega^{\alpha_r}.$$

□

Lemma 16.9. *If $\alpha < \beta$ and $\tau(\alpha) < k$ then there is some m so that $\beta[k][k+1]\cdots[k+m] = \alpha$.*

Proof. By induction on β. Let $\alpha < \beta$ be given with $\tau(\alpha) < k$. By the previous lemma, $\beta[k] \geq \alpha$. If $\beta[k] = \alpha$, we are done. Otherwise $\alpha < \beta[k] < \beta$, so by the inductive hypothesis applied to $\alpha < \beta[k]$ and $k + 1 > \tau(\alpha)$, there is some m so that $\beta[k][k+1]\cdots[k+m] = \alpha$. □

Lemma 16.10. *If $\alpha < \beta$ and $\tau(\alpha) < k$ then $h_\alpha(k) < h_\beta(k)$.*

Proof. Recall that, by Lemma 6.4, $h_\alpha(k)$ is the value $k + m$ so that $\alpha[k][k+1]\cdots[k+m] = 0$ and $h_\beta(k)$ is the value $k+m'$ so that $\beta[k][k+1]\cdots[k+m'] = 0$. By the previous lemma, there is some $m_0 \leq m'$ so that $\alpha = \beta[k][k+1]\cdots[k+m_0]$; starting at this value $k + m_0$, we can compare the sequences, observing that $\alpha[k]\cdots[k+j] \leq \alpha[k+m_0]\cdots[k+m_0+j]$ for all j. In particular, this means $h_\alpha(k) < h_\alpha(k + m_0) = h_\beta(k)$. □

17. Restricting Cuts

Deductions which do not use the cut rule are called *cut-free*, and have a special role. Suppose that we have a deduction without cuts, and consider what this deduction looks like working "up" from the root Γ towards the leaves. Every rule besides the Cut rule has the property that every sentence appearing in the premises of the rule is a sub-sentence (possibly after replacing variables with numbers) of a sentence in the conclusion: when we read up from the leaves, no genuinely new sentences appear. So if we know there is a

cut-free deduction of Γ, we can actually guess what it is: essentially the only rules which could be used are the various I rules, which can only be used to break down sentences in Γ to smaller ones.

(The Cut rule totally changes this picture: we could derive Γ from Γ, ϕ and $\Gamma, \sim\phi$, where ϕ is some totally new, unexpected sentence which has no obvious relation to the sentences in Γ.)

Theorem 17.1. *It is not the case that* $Z_\infty \left|\frac{\alpha,k}{}\right. \forall x \exists y\, g_y(x) = 0$ *with a cut-free deduction for any* $\alpha < \epsilon_0$.

Proof. We have to prove a slightly stronger statement by induction on α: if every $\phi \in \Gamma$ is either:

- $\forall x \exists y\, g_y(x) = 0$,
- $\exists y\, g_y(n) = 0$ for some $n \leq k$ with $\mathcal{G}(n) > h_\alpha(k)$, or
- $g_t(n) = 0$ for some closed term t with value $< \mathcal{G}(n)$,

then it is not the case that $Z_\infty \left|\frac{\alpha,k}{}\right. \Gamma$ with a cut-free deduction.

Suppose the claim holds for $\beta < \alpha$ but that $Z_\infty \left|\frac{\alpha,k}{}\right. \Gamma$ with a cut-free deduction. Consider the rule justifying this deduction; the justification cannot be the rule True because the only atomic formulas in Γ have the form $g_m(n) = 0$ for $m < \mathcal{G}(n)$, and are therefore false. If the justification is W then we would have to have $Z_\infty \left|\frac{\beta,k}{}\right. \Delta$ for some $\Delta \subseteq \Gamma$ and $\beta < \alpha$ with $\tau(\beta) < k$. Since $h_\beta(k) < h_\alpha(k)$, this violates the inductive hypothesis. If the justification is C, we would have to have $Z_\infty \left|\frac{\beta,k}{}\right. \Gamma, \phi$ for some $\phi \in \Gamma$, and this violates the inductive hypothesis for the same reason.

The justification cannot be $\mathsf{I}\bigwedge$ or $\mathsf{I}\bigvee$, because there is no conjunction or disjunction in Γ for those rules to introduce.

If the justification were $\mathsf{I}\exists$ then we would have $\Gamma = \Gamma' \cup \{\exists y\, g_y(n) = 0\}$ for some n with $\mathcal{G}(n) > h_\alpha(k)$, and $Z_\infty \left|\frac{\beta,k}{}\right. \Gamma' \cup \{g_t(n) = 0\}$ where the value of t is $\leq h_\alpha(k) < \mathcal{G}(n)$. Therefore the inductive hypothesis applies to $\Gamma' \cup \{g_t(n) = 0\}$, so we cannot have $Z_\infty \left|\frac{\beta,k}{}\right. \Gamma' \cup \{g_t(n) = 0\}$.

If the justification were $\mathsf{I}\forall$ then we would have $\Gamma = \Gamma' \cup \{\forall x \exists y\, g_y(x) = 0\}$, and for every n, we must have $Z_\infty \left|\frac{\beta_n,\max\{k,n\}}{}\right. \Gamma' \cup \{\exists y\, g_y(n) = 0\}$. Choose $n \geq k$ large enough that $\mathcal{G}(n) > h_\alpha(n) > h_{\beta_n}(n)$. Then we must have $Z_\infty \left|\frac{\beta_n,n}{}\right. \Gamma' \cup \{\exists y\, g_y(n) = 0\}$, again contradicting the inductive hypothesis. □

18. Rank Bounds on Deductions

More generally, we want to keep track of the complexity of the uses of the cut rule that appear in our deductions. We will write $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma$ when $Z_\infty \left|\frac{\alpha,k}{}\right. \Gamma$ and there is a deduction demonstrating this in which the cut rule is only used on formulas of rank $< c$.

Note that if $Z_\infty \vdash_0^{\alpha,k} \Gamma$ then the Cut rule cannot appear at all in the deduction: these are exactly the cut-free deductions, because any Cut rule would involve a formula with rank < 0, which does not exist.

A deduction does not have to have a finite bound on its cuts: consider a $\forall I$ rule where the n-th premise has bound n. Nonetheless, we have:

Theorem 18.1. *If a sentence ϕ can be proven in PA^+ then there is an ordinal $\alpha < \epsilon_0$ and natural numbers k and c so that $Z_\infty \vdash_c^{\alpha,k} \phi$.*

Proof Sketch. As in the proof of Theorem **??**, checking that the only cuts come from the induction axioms we used in PA^+. Since we only use finitely many induction axioms, the cuts have a finite rank. □

19. Cut Elimination

First, we need a lemma observing that we can weaken a conclusion *without* changing the ordinal bounds.

Lemma 19.1. *If $Z_\infty \vdash_c^{\alpha,k} \Gamma$ and $\Gamma \subseteq \Delta$ then $Z_\infty \vdash_c^{\alpha,k} \Delta$.*

Proof. By induction on α, considering cases based on the justification.

If the justification is True then there is a formula $\phi \in \Gamma$ which is a true atomic or negated atomic formula. Since $\phi \in \Delta$, also $Z_\infty \vdash_c^{\alpha,k} \Delta$.

In all other cases, the result follows by using the same justification and the inductive hypothesis. For instance, if the justification is $\mathsf{I}\vee$, so $Z_\infty \vdash_c^{\alpha,k} \Gamma, \phi \vee \psi$ is derived from $Z_\infty \vdash_c^{\beta,k} \Gamma, \phi$, the inductive hypothesis tells us that $Z_\infty \vdash_c^{\beta,\kappa} \Delta, \phi$, and the same $\mathsf{I}\vee$ rule implies $Z_\infty \vdash_c^{\alpha,k} \Delta, \phi \vee \psi$.

The other cases are similar. □

Next, we need a series of three results called the *inversion theorems* which show that we can simplify certain conclusions without changing the ordinal bounds. There are three versions—one for false atomic statements, one for conjunctions, and one for $\forall$ statements.

Theorem 19.2. *Suppose $Z_\infty \vdash_c^{\alpha,k} \Gamma, \phi$ where ϕ is a false atomic or negated atomic formula. Then $Z_\infty \vdash_c^{\alpha,k} \Gamma$.*

Proof. By induction on α.

Consider a deduction of Γ. We split into cases depending on the last rule of this deduction.

First, consider the cases where the last rule might have been responsible for introducing the formula ϕ into the sequent.

If the justification is True then there's a true atomic or negated atomic formula in Γ, so we also have $Z_\infty \vdash_c^{\alpha,k} \Gamma$.

If the justification is weakening, we have either $Z_\infty \vdash_c^{\beta,k} \Delta$ or $Z_\infty \vdash_c^{\beta,k} \Delta, \phi$ where $\Delta \subseteq \Gamma$. In the second case, we can apply the inductive hypothesis to get $Z_\infty \vdash_c^{\beta,k} \Delta$, and then the same weakening rule justifies $Z_\infty \vdash_c^{\alpha,k} \Gamma$.

If the justification is contraction where $Z_\infty \left|\frac{\beta,k}{c}\right. \Gamma, \phi, \phi$ then we can apply the inductive hypothesis twice, once to get $Z_\infty \left|\frac{\beta,k}{c}\right. \Gamma, \phi$ and again to get $Z_\infty \left|\frac{\beta,k}{c}\right. \Gamma$, and then use a trivial weakening rule to get $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma$.

In any other case, ϕ is incidental the content of the step, and the result will follow directly from the inductive hypothesis. For a representative example, suppose that the justification is $\mathsf{I}\wedge$, so there is a some formula $\psi_0 \wedge \psi_1 \in \Gamma$ so that $Z_\infty \left|\frac{\beta,k}{c}\right. \Gamma', \psi_0, \phi$ and $Z_\infty \left|\frac{\gamma,k}{c}\right. \Gamma'', \psi_1, \phi$. By the inductive hypothesis applied to each of these pieces, $Z_\infty \left|\frac{\beta,k}{c}\right. \Gamma', \psi_0$ and $Z_\infty \left|\frac{\gamma,k}{c}\right. \Gamma', \psi_1$, so $\mathsf{I}\wedge$ justifies $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma$.

The cases where the justification is $\mathsf{I}\bigvee$, $\mathsf{I}\forall$, $\mathsf{I}\exists$, contraction on a formula other than ϕ, or Cut are similar. □

Theorem 19.3 ($\wedge$ Inversion)**.** *Suppose $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma, \phi \wedge \psi$. Then $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma, \phi$ and $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma, \psi$.*

Proof. By induction on α, considering cases for the justification. We only show that $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma, \phi$, since showing $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma, \psi$ is entirely symmetric.

Again, the main cases is where $\phi \wedge \psi$ is the main formula. Suppose the justification is $\mathsf{I}\bigwedge$ where $\phi\wedge\psi$ was the main formula. Then there is a $k' \leq k$ so that either $Z_\infty \left|\frac{\beta,k}{c}\right. \Gamma, \phi$, and then we can use weakening to get $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma, \phi$.

If the last step is a weakening rule, we have either $Z_\infty \left|\frac{\beta,k}{c}\right. \Delta$ or $Z_\infty \left|\frac{\beta,k}{c}\right. \Delta, \phi \wedge \psi$ where $\Delta \subseteq \Gamma$. In the first case, we can just use weakening to get $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma, \phi$. In the second case, we use the inductive hypothesis to get $Z_\infty \left|\frac{\beta,k}{c}\right. \Delta, \phi$, and then weakening to get $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma, \phi$.

If the last step is contraction on $\phi \wedge \psi$ we have $Z_\infty \left|\frac{\beta,k}{c}\right. \Gamma, \phi \wedge \psi, \phi \wedge \psi$. Using the inductive hypothesis twice, we get $Z_\infty \left|\frac{\beta,k}{c}\right. \Gamma, \phi, \phi$, and then we can use contraction to get $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma, \phi$.

In all the other cases, the result follows by using the same justification and the inductive hypothesis. For instance, if the last step is $I\vee$ inferring $Z_\infty \left|\frac{\alpha,k}{c}\right. \Delta, \gamma_0 \vee \gamma_1, \phi\wedge\psi$ from $Z_\infty \left|\frac{\beta,k}{c}\right. \Delta, \gamma_i, \phi\wedge\psi$, we may apply the inductive hypothesis to get $Z_\infty \left|\frac{\beta,k}{c}\right. \Delta, \gamma_i, \phi$ and then use $I\vee$ to get $Z_\infty \left|\frac{\alpha,k}{c}\right. \Delta, \gamma_0 \vee \gamma_1, \phi$.

The cases were the justification is $\mathsf{I}\wedge$ introducing something other than $\phi \wedge \psi$, True, $\mathsf{I}\forall$, $\mathsf{I}\exists$, contraction on a different formula, or Cut are similar. □

Theorem 19.4 ($\forall$ Inversion)**.** *Suppose $Z_\infty \left|\frac{\alpha,k}{c}\right. \Gamma, \forall x\, \phi$. Then for any n, $Z_\infty \left|\frac{\alpha,\max\{k,n\}}{c}\right. \Gamma, \phi[x \mapsto n]$.*

Proof. By induction on α, considering the cases for the justification.

Once again, the main case is that the last step introduce $\forall x\, \phi$. If the last step is $\mathsf{I}\forall$ with main formula $\forall x\, \phi$ then we have $Z_\infty \left|\frac{\beta_n,\max\{k,n\}}{c}\right. \Gamma, \phi[x \mapsto n]$. Therefore by using weakening, we get $Z_\infty \left|\frac{\alpha,\max\{k,n\}}{c}\right. \Gamma, \phi[x \mapsto n]$.

If the the justification is weakening or contraction, we use the same argument as for $\wedge$ inversion. In all other cases, the result follows by using the same justification and the inductive hypothesis.

□

Theorem 19.5. *Suppose* $rk(\phi) < c$, $rk(\psi) < c$, $Z_\infty \vdash_c^{\alpha,k} \Gamma, \phi$, $Z_\infty \vdash_c^{\alpha,k} \Gamma, \psi$, *and* $Z_\infty \vdash_c^{\beta,k} \Delta, \sim\phi \vee \sim\psi, \sim\phi \vee \sim\psi, \ldots, \sim\phi \vee \sim\psi$.

Then $Z_\infty \vdash_c^{\alpha\#\beta,2k} \Gamma, \Delta$.

Proof. By induction on β, considering the justification for $Z_\infty \vdash_n^{\beta,k} \Delta, \sim\phi \vee \sim\psi, \sim\phi \vee \sim\psi, \ldots, \sim\phi \vee \sim\psi$.

The main case is $\mathsf{I}\vee$ introducing $\sim\phi\vee\sim\psi$: in this case there is a $\beta' < \beta$ with $\tau(\beta') < k$ so that $Z_\infty \vdash_c^{\beta',k} \Delta, \sim\phi \vee \sim\psi, \sim\phi \vee \sim\psi, \ldots, \sim\phi \vee \sim\psi, \sim\phi$ (the case where $Z_\infty \vdash_c^{\beta',k} \Delta, \ldots, \sim\psi$ is symmetric). By the inductive hypothesis, we have $Z_\infty \vdash_c^{\alpha\#\beta',2k} \Gamma, \Delta, \sim\phi$. Since $\tau(\alpha) < k$ and $\tau(\beta') < k$, also $\tau(\alpha\#\beta') < 2k$. Since, in addition, $\alpha\#\beta' < \alpha\#\beta$ and $rk(\phi) < c$, we may apply the cut rule with ϕ to obtain $Z_\infty \vdash_c^{\alpha\#\beta,2k} \Gamma, \Delta$.

If the rule is contraction on $\sim\phi \vee \sim\psi$, the claim follows immediately by the inductive hypothesis.

The other cases follow immediately from the inductive hypothesis as in previous arguments. □

Theorem 19.6. *Suppose* $rk(\phi) < c$, *for every* n, $Z_\infty \vdash_c^{\alpha,\max\{k,n\}} \Gamma, \phi[x \mapsto n]$, *and* $Z_\infty \vdash_c^{\beta,k} \Delta, \exists x \sim\phi, \exists x \sim\phi, \ldots, \exists x \sim\phi$. *Then* $Z_\infty \vdash_c^{\alpha\#\beta,h_{\beta\#\omega}(k)} \Gamma, \Delta$.

Proof. By induction on β, considering the justification for $Z_\infty \vdash_n^{\beta,k} \Delta, \exists x \sim\phi, \exists x \sim\phi, \ldots, \exists x \sim\phi$.

The main case is $\mathsf{I}\exists$ introducing $\exists x \sim\phi$: in this case there is a $\beta' < \beta$ with $\tau(\beta') < k$ and an $n \leq h_\beta(k)$ with $Z_\infty \vdash_c^{\beta',k} \Delta, \exists x \sim\phi, \exists x \sim\phi, \ldots, \sim\phi[x \mapsto n]$. By the inductive hypothesis and Lemma 16.4, we have $Z_\infty \vdash_c^{\alpha\#\beta',h_{\beta\#\omega}(k)} \Gamma, \Delta, \sim\phi[x \mapsto n]$. We may apply a cut with $\phi[x \mapsto n]$ to get $Z_\infty \vdash_c^{\alpha\#\beta,h_{\beta\#\omega}(k)} \Gamma, \Delta$. (Adding ω ensures that $h_{\beta\#\omega}(k) > 2k > \tau(\alpha\#\beta')$.)

If the rule is contraction on $\exists x \sim\phi$, the claim follows immediately by the inductive hypothesis.

The other cases follow immediately from the inductive hypothesis as in previous arguments. □

Theorem 19.7. *Suppose that* $Z_\infty \vdash_{c+1}^{\alpha,k} \Gamma$ *then* $Z_\infty \vdash_c^{\omega^\alpha,h_{\omega^\alpha}(k)} \Gamma$.

Proof. By induction on α. The main case is when the justification is a Cut on a formula of rank exactly c. In this case we have $Z_\infty \vdash_{c+1}^{\beta,k} \Gamma, \phi$ and $Z_\infty \vdash_{c+1}^{\gamma,k} \Gamma, \sim\phi$. By the inductive hypothesis, $Z_\infty \vdash_c^{\omega^\beta,h_{\omega^\beta}(k)} \Gamma, \phi$ and $Z_\infty \vdash_{c+1}^{\omega^\gamma,h_{\omega^\gamma}(k)} \Gamma, \sim\phi$.

If ϕ is atomic then one of ϕ or $\sim\phi$ is false. For instance, if ϕ is false then by Theorem 19.2, we have $Z_\infty \left|\frac{\omega^\beta, h_{\omega^\beta}(k)}{c+1}\right. \Gamma$, and so by weakening we have $Z_\infty \left|\frac{\omega^\alpha, h_{\omega^\alpha}(k)}{c}\right. \Gamma$.

If ϕ is $\psi_0 \wedge \psi_1$ then by Theorem 19.3 we have $Z_\infty \left|\frac{\omega^\beta, h_{\omega^\beta}(k)}{c}\right. \Gamma, \psi_0$, $Z_\infty \left|\frac{\omega^\beta, h_{\omega^\beta}(k)}{c}\right. \Gamma, \psi_1$. We also have $Z_\infty \left|\frac{\omega^\gamma, h_{\omega^\gamma}(k)}{c}\right. \Gamma, \sim\psi_0 \vee \sim\psi_1$. By Theorem 19.5, we have $Z_\infty \left|\frac{\omega^\beta \# \omega^\gamma, \max\{h_{\omega^\beta}(k), h_{\omega^\gamma}(k)\}}{c}\right. \Gamma$, and then by weakening we get $Z_\infty \left|\frac{\omega^\alpha, h_{\omega^\alpha}(k)}{c}\right. \Gamma$.

The case where ϕ is $\psi_0 \vee \psi_1$ is symmetric, since $\sim\phi$ is $\sim\psi_0 \wedge \sim\psi_1$.

The case where ϕ is $\forall x\, \psi$ is similar. We have $Z_\infty \left|\frac{\omega^\beta, h_{\omega^\beta}(k)}{c}\right. \Gamma, \forall x\, \psi$ and $Z_\infty \left|\frac{\omega^\gamma, h_{\omega^\gamma}(k)}{c}\right. \Gamma, \exists x \sim\psi$. By Theorem 19.4, we have $Z_\infty \left|\frac{\omega^\beta, \max\{h_{\omega^\beta}(k), n\}}{c}\right.$ $\Gamma, \psi[x \mapsto n]$ for all n. Then by Theorem 19.6, we have $Z_\infty \left|\frac{\omega^\beta \# \omega^\gamma, h_{\omega^\beta \# \omega}(h_{\omega^\gamma}(k))}{c}\right.$ Γ. Finally, by weakening we have $Z_\infty \left|\frac{\omega^\alpha, h_{\omega^\alpha}(k)}{c}\right. \Gamma$.

When ϕ is $\exists x\, \phi$ then we symmetrically have $\sim\phi$ is $\forall x \sim\psi$.

In all other cases, the result follows immediately from IH. □

Definition 19.8. For any c, we define ω_c^α inductively by:

- $\omega_0^\alpha = \alpha$,
- $\omega_{c+1}^\alpha = \omega^{\omega_c^\alpha}$.

Theorem 19.9. *If $Z_\infty \left|\frac{\alpha, k}{c}\right. \Gamma$ then $Z_\infty \left|\frac{\omega_c^\alpha, m}{0}\right. \Gamma$ for some m.*

Proof. Apply the previous lemma c times. □

20. Conclusion

Putting this together:

Theorem 20.1. *Goodstein's Theorem cannot be proven from the axioms of PA^+.*

Proof. If Goodstein's Theorem could be proven from the axioms of PA^+, by Theorem 16.7, we would have $Z_\infty \left|\frac{\alpha, k}{}\right. \forall x \exists y\, g_y(x) = 0$. Inspection of the proof of Theorem 16.7 shows that the only cuts in this deduction are over formulas appearing in the original (finite) proof from PA^+, so in particular, there is a bound c so that $Z_\infty \left|\frac{\alpha, k}{c}\right. \forall x \exists y\, g_y(x) = 0$.

By Theorem 19.9, it follows that $Z_\infty \left|\frac{\omega_c^\alpha, m}{0}\right. \forall x \exists y\, g_y(x) = 0$. Together with Theorem 17.1, this gives the desired contradiction, so Goodstein's Theorem cannot be proven from the axioms of PA^+. □