Introduction and Proof of the Goodstein Sequence and Hydra Game

Hermine Tigerschiöld

15 June 2017

This thesis will introduce and prove the theorem of the Hydra Game as well as the theorem surrounding the Goodstein sequence. The two problems which are clearly intertwined will be introduced and solved with the assumption that the reader has no previous knowledge of them and little to no knowledge of ordinal numbers. All results have previously been reached and presented, most notably by Kirby and Paris (Kirby et al, 1982). The aim of the paper is not to break ground but educate. The end of the paper will mention how these theorems show limitations of Peano Arithmetic but will not attempt to prove the statement.

1 Introduction

1.1 The Hydra Game

The Hydra Game tells the slightly modified story of when Hercules fought the horryfing multi headed monster - the Hydra. The Hydra has incredible regenerative powers. Every time Hercules severs one of its heads, several new heads grow from the Hydra's body. To describe the Hydra in a simple graph of segments and nodes we use drawings such as Figures 1 through 5. In Figure 1, the node named A is the root of the Hydra, nodes which are not leaves such as B are part of the Hydra's body, and the leaves including stems such as C are the heads of the Hydra.

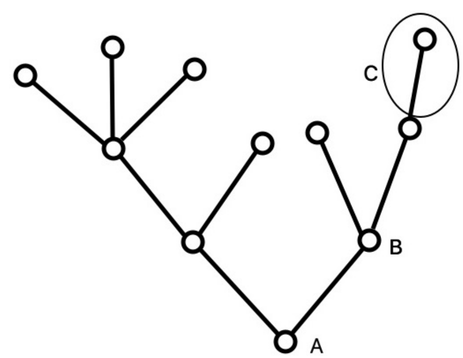

Figure 1: Example of a Hydra.

Now that we have established how to draw the Hydra, it will be easier to describe its regenerating powers:

- Hercules for the n^{th} time severs a head.
- Traverse one node down from where the head was severed.
- The entire part of the Hydra from this node up is multiplied n times, and grows from the relevant node.

The game itself has the following basis: a hydra is decided upon, it can have any number of heads as well as any shape. Then the game itself proceeds, a head is severed followed by regeneration of the Hydra. If in a finite number of blows the Hydra's root is reached the game is won by Hercules, if not the game is won by the Hydra. An example of how the game could proceed follows. The original hydra for the example is portrayed in Figure 2, all heads are numbered. Hercules decides to sever head 1. The regenerative powers will then cause the Hydra to look as in Figure 3. Head 1 is gone and there are three new heads:

7, 8 and 9. Hercules decides that he will next strike head 6. The results can be seen in Figure 4. In two attacks the Hydra's number of heads has increased from 6 to 12.

It is evident that the amount of heads has increased. Does this mean that it is impossible to cut off all heads and eventually win against the Hydra? In this case a strategy consisting of first cutting off the heads furthest away from the body and then working inwards would result in a total sum of 222 nodes and heads at step 15. Another 15 hits and the total sum would have increased to 671. This does not look very promising; as we doubled the hits from 15 to 30 the amount of nodes did not decrease but instead increased three times over. Before discussing this example further we will explore if it is possible to win the game against a smaller hydra.

If Figure 5 we have a complete game against a smaller hydra. The game is won in 5 hits. This suggests that there are at least some games that can be won.

Let us return to the previous example. At step 100 there will be approximately 5000 nodes. Even more terrifying, there will be approximately $2 \cdot 10^6$ nodes at step 170. Still the Hydra shows no sign of decreasing in size. Although numbers in some cases speak for themselves this is not one of those cases. It is noteworthy that despite the Hydra's increase in number of nodes, the maximum distance between a head and the root is the same as it was in the original hydra. This might seem trivial but will show to be of vital importance. It suggests that if all heads positioned 3 nodes away from the root were hit then the Hydra despite becoming very wide in the process would be permanently reduced to being a hydra with heads maximum 2 nodes away from the root. If this method would be applied again the Hydra eventually would be reduced to a hydra of height 1, at which point the Hydra no longer has the ability to regenerate.

It so happens that this course of events is not only possible but will always take place and that all hydras therefore can be beaten, although it might take an extensive amount of time. In the example above, the Hydra can be beat after approximately $10^{2\cdot 10^6}$ steps. This is a very large number. Another very large number is the amount of stars in the sky, and an even larger number is the total amount of stars in the universe. This number, the combination of all stars in all galaxies is approximated to being roughly 10^{22} (ESA, n.d.). This compared to our number $10^{2\cdot 10^6}$, which can also be written $10^{2000000}$. The exponent of the number of hits to win is 9000 times larger than the exponent of the number of stars in the universe. We conclude that although the game in our example can be won it will take an incredibly large number of hits. This suggests that it is possible to win games that seem hopeless and sure enough this is true. With this our first theorem is presented.

Theorem 1.1. *Every strategy against the Hydra is a winning strategy.*

No matter which head Hercules decides to sever and no matter which order he chooses to do it in, he cannot lose. Eventually he will reach the root, at which point the game is won. This is the first of two theorems that will be proven.

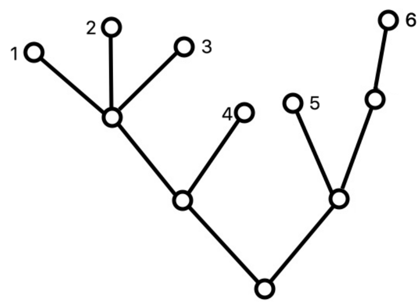

Figure 2: The hydra in the example, $n = 1$

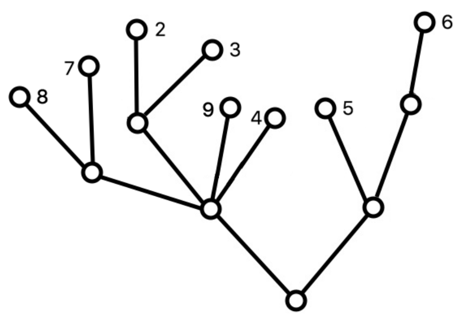

Figure 3: The hydra in the example, $n = 2$

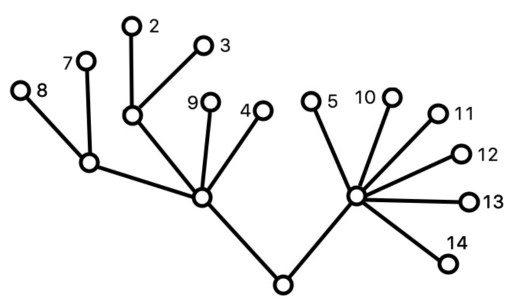

Figure 4: The hydra in the example, $n = 3$

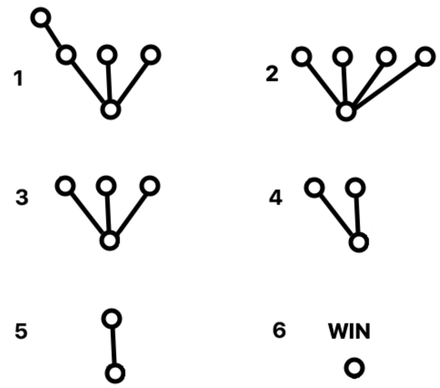

Figure 5: The smaller Hydra, complete game.

To prove this theorem it is helpful to start by studying the so called Goodstein sequence.

1.2 The Goodstein Sequence

The Goodstein sequence is similar to the Hydra in that it seems to be growing at an alarming rate only to eventually end up at zero. The steps to formulate a Goodstein sequence are described below in the list. An example is explained in parallel.

- An integer m larger than zero is chosen. Example: 111.
- There is also a second integer; n larger than 1. Example: $n = 2$ in the beginning.
- The integer m is described in base n. Example: $111 = 2^6 + 2^5 + 2^3 + 2^2 + 2^1 + 2^0$
- The exponents are then described in base n, and also the exponents of the exponents etc. Example: $111 = 2^{2^2+2} + 2^{2^2+1} + 2^{2+1} + 2^2 + 2 + 1$
- The next number in the sequence, $G_n(m)$, is created by exchanging every number n by $n+1$ and then subtracting 1. Example: $G_2(111) = 3^{3^3+3} + 3^{3^3+1} + 3^{3+1} + 3^3 + 3 + 1 - 1$

The last step of the instruction is then repeated creating a sequence, the Goodstein sequence, $m_0, m_1, m_2, ...$for the number m. A continuation of the example started would result in

$$m = m_0 = 111_0 = 2^{2^2+2} + 2^{2^2+1} + 2^{2+1} + 2^2 + 2 + 1 = 111$$

$$G_2(m_0) = 111_1 = 3^{3^3+3} + 3^{3^3+1} + 3^{3+1} + 3^3 + 3 \sim 2.29 \cdot 10^{14}$$

$$G_3(m_1) = 111_2 = 4^{4^4+4} + 4^{4^4+1} + 4^{4+1} + 4^4 + 3 \sim 3.49 \cdot 10^{158}$$

$$G_4(m_2) = 111_3 = 5^{5^5+5} + 5^{5^5+1} + 5^{5+1} + 5^5 + 2 \sim 5.98 \cdot 10^{2187}$$

It is evident that the sequence grows incredibly quickly, at least for the first few steps. In 4 steps we have arrived at a number that just like the number in the Hydra introduction is many times larger than 10^{22}, the total amount of stars in the universe. We note that 10^{22} is not nearly large enough to describe these numbers, but its difficult to produce a comprehensive larger number. The age of the universe, approximately $14 \cdot 10^9$ years (Redd, 2017), counted in seconds is $4.4 \cdot 10^{17}$ seconds old. This is smaller than 10^{22} but might be easier to grasp. It is the number we would reach if we started counting the integers at Big Bang, one integer every second, until present day. If this was done the size of the number reached would still not be close to the size of $G_4(m_2)$. That means that if the sequence stops growing after the fourth step and every new number in the sequence is simply the previous number subtracted by one and this iteration was performed from the beginning of time until now the number reached in present time by the sequence would barely be smaller than the original number.

But the sequence does not seem to stop growing. Let us assume that the growth pattern continues and the sequence keeps growing, the exponential increasing by roughly a factor of 10 to 15 per step. This increase makes the next numbers of this sequence too large to understand. But before any rash conclusions are drawn from this information it helps to as in the case of the Hydra take a step back and not simply look at the numbers. In the case of the Goodstein sequence it helps to look at the way the sequence is constructed. The first term m is defined using 6 terms. In the next number $G_2(m)$ the term furthest to the left has been exterminated and the result is a total of five terms. The sequence is still growing uncontrollably as every term increases in size but the amount of terms the sequence builds of is still limited to five. It is evident that the term furthest to the right is being reduced and in another two steps at $G_6(m_4)$ it will be eliminated. After that the subtraction will inevitably break down the fourth term. At this point the fourth term will be very large but nonetheless it will happen. This is one way of picturing the eventual decrease of the sequence. The sequence is growing as the terms grow but those terms are not indestructible. They are vulnerable to the slow but steady subtraction of one. This ensures that the following theorem is true.

Theorem 1.2. *For any numbers m and $n > 1$ the Goodstein sequence for m starting at n hits zero.*

The proof of Theorem 1.2 will enable the proof of Theorem 1.1, the two problems are quite similar.

2 Definitions and Background

The proofs in the next section rely on several mathematical concepts and ideas that will now be presented. One of the most basic mathematical tools that will be used is the notion of a set. The word "set" is used frequently in everyday life with a definition similar to the mathematical definition. A set is a collection of elements, but is in itself viewed as a single object. If an element t is a member of the set A then the notation $t \in A$ is used, if t is not an element in A then the notation $t \notin A$ is used (Enderton, 1977, 1). This most basic definition of a set does not have any underlying conditions that the elements are ordered or that they could be compared to one another. An example of a set where some elements can be compared to each other is a *Partially Ordered set.*

Definition 2.1. *Partially Ordered Set* Let L be a set where the elements can be compared. A relation is used to compare two elements, its notation commonly being $\leq$. Let a and b be elements of the set L. Depending on the qualities of a and b the statement $a \leq b$ might be either true or false. A *Partially Ordered Set* is the combination of a set, L, with an partial ordering, $\leq$. A partial ordering is a relation for which the following conditions are fulfilled.

1. *Reflexivity* For all $a \in L$, $a \leq a$.
2. *Anti-symmetry* If $a \leq b$ and $b \leq a$ then $a = b$.
3. *Transistivity* If $a \leq b$ and $b \leq c$ then $a \leq c$.

Where $a, b, c \in L$. The partial ordered set is then written $(L, \leq)$ (Kaplansky, 1977, 9) (Devlin, 1993, 11).

A partially ordered set in layman terms is therefore a set consisting solely of elements that can be compared to one another and where the relation used to do so fulfills the criteria above. It is important to note that for a patrially ordered set there is no criteria guaranteeing that given $a, b \in L$ either $a \leq b$ or $b \leq a$ is true. A basic example of a partially ordered set is $(\mathbb{Z}, \leq)$. The relation "less than or equal to" fulfills all three points above for all integers. An example of the opposite could be $(\mathbb{N}, R)$, where the relation R is defined by

$$\{aRb \mid \exists c \neq 1, c \in \mathbb{Z} \text{ such that } a/c, b/c \in \mathbb{N}, \quad a, b \in \mathbb{N}\}.$$

In this case $(\mathbb{N}, R)$ would not qualify to be a partially ordered set as the ordering does not fulfill the second criteria in the definition and is therefore not a partial ordering. This is evident for the elements 2 and 4, where $2R4$ and $4R2$ are both true despite $2 \neq 4$.

The example $(\mathbb{Z}, \leq)$ above is also an example of a *Totally Ordered set.* The totally ordered set is very similar to the partially ordered set but with one more criterion known as the comparability criterion. The criterion must be fulfilled by the partial ordering of the set (Weisstein, n.d.).

- For every $a, b \in L$ either $a \leq b$ or $b \leq a$.

This essentially means that every two elements from the set can be compared to one another. An example of a set that does not fulfill this additional criteria but it a partially ordered set is $(\mathbb{N}, \preccurlyeq)$ where the relation $\preccurlyeq$ is defined by;

$$\{a \preccurlyeq b \mid b/a \in \mathbb{N}, \quad a, b \in \mathbb{N}\}.$$

This relation between a and b does not fulfill the added criteria for a totally ordered set as not all fractions are natural numbers. Let $a = 3$ and $b = 7$ then $a \preccurlyeq b$ is $7/3$ which the relation is not defined for as 7 is not divisible by 3. The other possible comparison is $b \preccurlyeq a$ which is $3/7$, still undefined as 3 is not divisible by 7. From this we can draw the conclusion that neither $a \preccurlyeq b$ nor $b \preccurlyeq a$ can be used to compare a and b which means that the comparability criterion is not fulfilled. Despite the set not being a totally ordered set it fulfills all criteria for a partially ordered set. Using the definition for a partially ordered set it is now possible to define a *Well Ordered set.*

Definition 2.2. *Well Ordered Set* is a partially ordered set where every subset has a least element (Abian, 1965, 142).

The difference between the definition a well ordered set and a partially ordered set might seem negligible but will show to be of vital importance. The least element is defined as the element that if compared, using the partial ordering, to every other element in the set/subset it is smaller than all other elements, with the exception of the element itself. An example of a partially ordered set that is not a well ordered set is $(\mathbb{Z}, \leq)$. The set is defined as

$$\{..., -4, -3, -2, -1, 0, 1, 2, 3, 4, ...\}.$$

This set has no least element, neither has several of its subsets, for example the negative integers form a subset, $\{..., -4, -3, -2, -1\} \in \mathbb{N}$, which lacks a least element. This as opposed to for example the natural numbers, $\mathbb{N} = \{0, 1, 2, 3, 4, ...\}$ which have a clear least element 0, which also is true for all of its subsets making $(\mathbb{N}, \leq)$ a well ordered set.

2.1 Ordinal Numbers

Now that we have defined well ordered sets it is almost time to define ordinal numbers. In order to understand the definition we need to define a notation first.

Definition 2.3. *Initial Segment* Let $(P, \leq)$ be a partially ordered set. For every element a of P, the set of all elements x of P such that $x < a$ is denoted by $I(a)$ and is called the initial segment of P determined by a.

The length of the definition does in this case not correspond to the difficulty of the concept. An initial segment is simply the subset that consists of all elements smaller than the specified element. For example if we have a partially ordered set;

$$(P, \leq) = \{2, 3, 4, 5, 6\}.$$

Then we could define the initial segment $I(4)$ as;

$$I(4) = \{2, 3\}.$$

Now to the definition of ordinal numbers.

Definition 2.4. *Ordinal number* A set is called an Ordinal number if w can be well ordered so that for every element v of w the initial segment $I(v)$ of w is equal to v, ie;

$$I(v) = v \text{ for every } v \in w.$$

(Abian, 1965, 317)

This definition may seem unintuitive, its main benefit is that it communicates the importance of the construction of the ordinal number. This definition will be further explained using an example, but first we will study how natural numbers are most commonly represented as sets. The first number that will be described is the number zero. This quite intuitively is the empty set, ie $0 = \emptyset$. Note that

$$I(0) = \emptyset = 0.$$

The next natural number is the number 1. To represent it we count from 0, which also happens to be the only digit. The number 1 is therefore the set of the elements in zero, ie $1 = \{0\} = \{\emptyset\}$. Also

$$I(1) = \{0\} = \{\emptyset\} = 1.$$

The process is repeated. The number 2 is the set of the elements of 0 and 1.

$$I(2) = \{0, 1\} = \{\emptyset, \{\emptyset\}\} = 2.$$

Next is the number 3, again it is the set of the sets of 0, 1 and 2.

$$I(3) = \{0, 1, 2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} = 3.$$

This process is then repeated to get all natural numbers. This is how the ordinal numbers are built recursively. It is also possible to extend over limits and start counting infinities. The most common example of this will be mentioned further down.

Now we will show how the definition is used in practice. According to the definition it should be possible to choose an element of the set, $5 = \{0, 1, 2, 3, 4\}$, for example 3, to create an initial segment. The set $I(3)$ becomes;

$$I(3) = \{n \in \{0, 1, 2, 3, 4\} \mid n < 3\} = \{0, 1, 2\} = 3.$$

This is clearly the set representation of 3 (Abian, 1965, 317). The ordinal number of a set A, will from now on have the notation $o(A)$.

Alternatively we can define an ordinal number as follows. If we have two well ordered sets they will have the same ordinal number if and only if the sets are order isomorphic. Order isomorphism means that there must exist a function from one set to the other where the function has the properties that it is bijective and preserves order (Kaplansky, 1977, 55). A bijective mapping might be familiar to the reader; it means that there is a function that maps between two sets such that the mapping is both injective and surjective; or in layman terms that all elements are mapped in pairs one-to-one and that all elements are mapped to and from. The definition of an order preserving function is

Definition 2.5. *Order preserving* Let there be two partially ordered sets $(A, \leq)$ and $(B, \leq)$. A function, or mapping, f from A to B is order preserving if for every two elements $a, b \in A$ if;

$$a \leq b \quad \text{implies} \quad f(a) \leq f(b).$$

(Abian, 1965, 284)

This simply states that for two sets to be order isomorphic there must exist a function that is not only bijective but that maps all elements from one set, A, onto the other set, B, in a way that the elements of B are ordered in accordance with its partial order. What is essential to this definition is that there must be a single ordinal that represents each collection of sets that are order isomorphic. This is true and is formulated in Theorem 2.1 further down.

For the purpose of using ordinal numbers it is necessary to define their arithmetic, therefore addition, multiplication as well as exponents of ordinal numbers will now be defined. Due to the well ordering of an ordinal number neither addition nor multiplication of ordinal numbers are commutative, meaning that $a + b \neq b + a$. But before the arithmetics is defined, the smallest *transfinite ordinal*, ω, will be presented. It is the ordinal number of the set of all natural numbers with the usual ordering $\leq$;

$$\omega = \{0, 1, 2, 3, 4, ...\}.$$

Transfinite ordinals are the ordinal numbers which have an infinite number of elements, this as opposed to the finite ordinals, which all have a finite number of elements (Conway, 2000, 271)(Abian, 1965, 317). There are a certain type of transfinite ordinals that have no immediate predecessors; *limit ordinals*. The smallest limit ordinal is ω which has just been defined above, the second is 2ω which will soon be defined (Kaplansky, 1977, 57).

Definition 2.6. *Addition* Let the sets A and B have the ordinal numbers $o(A) = \alpha$ and $o(B) = \beta$ respectively. The sum of α and β is written as $\alpha + \beta$ and is defined as the disjoint union $A \sqcup B$. The disjoint set is sorted so that first all elements from A is sorted into the new set, ordered the same way they were originally ordered in A. Followed by all the elements in B, ordered as they were originally ordered in the set B (Hansson et al, 1973, 34).

The disjoint union is very similar to the traditional union but all elements from the two sets A and B are regarded as being of different "types". This means that despite there being a number a in both sets the two a's are seen as different elements and will both be represented in the disjoint union.

Theorem 2.1 is important for the definition of addition.

Theorem 2.1. *For every well ordered set w there exists a unique ordinal number similar to w (Abian, 1965, 324).*

The word similar in this theorem means that there is one ordinal number a well ordered set is order isomorphic to. This means that there is no ambiguity. When a set is created through addition it will be represented through a single ordinal number. Now let A be the set of the natural numbers, ordered using the "less-than-or-equal-to" ordering. The ordinal number $o(A)$ then becomes ω. Let B be the set $\{0\}$. The ordinal number $o(B)$ is then 1. Addition of the form $A + B$ then becomes

$$\omega + 1 = o(A) + o(B) = o(A \sqcup B) = o(\{0, 1, 2, 3, 4, ..., 0\}) = \omega + 1.$$

The addition $o(A) + o(B)$ prior to the operation having been performed can be written as $\omega + 1$, as $o(A) = \omega$ and $o(B) = 1$. Above the addition was then calculated and the result is $\omega + 1$, mainly because of the zero placed last in the disjoint union. The zero communicates that there exists a specific element larger than all other elements in the set; therefore a $+1$ is added onto the ω. This as opposed to addition of the form $B + A$

$$1 + \omega = o(B) + o(A) = o(B \sqcup A) = o(\{0, 0, 1, 2, 3, 4, ...\}) = \omega.$$

In $B \sqcup A$ there is no largest element. This difference between $A \sqcup B$ and $B \sqcup A$ create a large enough distinction to ensure that there could not be a bijective and order preserving mapping between the two sets. As the criterion for the two sets to have the same ordinal number is not fulfilled, the two sets cannot have the same ordinal number. This confirms that the order of the addition is vital - the addition defined, is not commutative. It is worth mentioning that when finite numbers are added the addition becomes commutative, it simply becomes equal to the numbers of elements involved, regardless of order. A short example would be

$$o(\{0, 1, 2, 3\}) + o(\{99, 98, 97\}) = o(\{0, 1, 2, 3, 99, 98, 97\}) = 7$$

$$o(\{99, 98, 97\}) + o(\{0, 1, 2, 3\}) = o(\{99, 98, 97, 0, 1, 2, 3\}) = 7$$

Complications arise when infinities are added.

Multiplication of ordinal numbers is based of the same idea as addition.

Definition 2.7. *Multiplication* Let the sets A and B have the ordinal numbers α and β respectively. The product of α and β is written as $\alpha \cdot \beta$ and is defined as the ordinal number of $A \times B$ with the order $\prec$, that allows $\langle a, b\rangle \prec \langle c, d\rangle$ if and only if either $a \prec c$, or $a = c$ and $b \prec d$ (Hansson et al, 1973, 35).

Similarly to addition, multiplication of ordinal numbers results in a set that is only isomorphic to a single ordinal number, again as a result of Theorem 2.1. If the reader is not familiar with the notation $A \times B$ it is the cartesian product, i.e. it is the set of all combinations of (a, b) where $a \in A$ and $b \in B$ (Kaplansky, 1977, 19). The result of multiplication is therefore described as all combinations of (a, b) ordered with regards first to the first element and then to the second. This ordering is most commonly known as the lexicographical order as it is frequently used to organize words based on letters. For example are the words in dictionaries organized first by the first letter, then second letter, then third etc. Multiplication of ordinal numbers is also not commutative, as seen in the following example. Let $o(A) = \omega$ and $o(B) = o(\{1, 2\}) = 2$ then $o(A) \times o(B)$ becomes

$$o(A) \times o(B) = \{(0,1), (0,2), (1,1), (1,2), (2,1), (2,2), ... \\ .., (n,1), (n,2), (n+1,1), (n+1,2), ...\}.$$

This is isomorphic with the ordinal number ω, which is evident upon further inspection. All elements have defined predecessors and the amount of elements are infinite. The only ordinal number that fulfills these criteria is ω, the first limit ordinal.

Multiplication of the form $o(B) \times o(A)$ results in

$$o(B) \times o(A) = \{(1,0), (1,1), (1,2), (1,3), ...(2,0), (2,1), (2,2), (2,3), ...\}.$$

This is isomorphic with $\omega + \omega$. This becomes evident as we note that all the elements $\{(2,0), (2,1), (2,2), (2,3), \dots\}$ are larger than all the elements $\{(1,0), (1,1), (1,2), (1,3), \dots\}$. This can be thought of as one infinite set being the successor to another, or ω succeeding ω resulting in $\omega + \omega$.

We conclude that $o(A) \times o(B) = \omega \cdot 2 = \omega$ and that $o(B) \times o(A) = 2 \cdot \omega = \omega + \omega$ which confirms that multiplication is not commutative.

The exponent will be defined with the help of multiplication but also using what will be explained later as transfinite induction. This means that the exponent is defined recursively.

Definition 2.8. *Exponent* Let α be an ordinal number. We define the exponent with the aid of transfinite induction. The three criteria are listed below.

- $\alpha^0 = \alpha^\emptyset = 1$,
- If β is not a limit ordinal and therefore can be written as $\beta = \gamma + 1$ then $\alpha^\beta = \alpha \cdot \alpha^\gamma$,
- If β is a limit ordinal then $\alpha^\beta = \bigsqcup_{\gamma<\beta} \alpha^\gamma$.

(Devlin, 1993, 74)

The three criteria can be used recursively to calculate the exponent. Calculating an exponent is very similar to repeatedly multiplying an ordinal number with itself. This means that the cartesian product as opposed to in multiplication is extended, there will be more than two elements in every bracket. An example could be α^3 where $\alpha = \{0, 1, 2\}$.

$$\alpha^3 = \alpha \cdot \alpha \cdot \alpha = \{(0,0), (0,1), ...(2,2)\} \cdot \alpha = \\ \{(0,0,0), (0,0,1), (0,0,2), (0,1,0), ...(2,2,1), (2,2,2)\}$$

This is an example of when the the exponent is not a limit ordinal. If the exponent is a limit ordinal a very similar approach is used but the cartesian product brackets become infinite, there is an infinite set of elements in each.

Another interesting thing to note is that the ordering used to orgnanize the elements is the lexicographic order mentioned earlier.

Some of the traditional counting rules for exponents are true for ordinal numbers such as;

1. $\alpha^{n+m} = \alpha^n \cdot \alpha^m$
2. $(\alpha^n)^m = \alpha^{n \cdot m}$

But $(\alpha \cdot \beta)^n = \alpha^n \cdot \beta^n$ is most commonly not true.

2.2 Transfinite Induction

The concept of induction must be further developed for use on ordinal numbers. The idea is virtually the same. Let $P(\alpha)$ be a statement about the ordinal α. If it is shown that it is correct for $P(0)$, as well as for all $P(\gamma)$ assuming it is true for $P(\beta)$ where $\gamma > \beta$ we can conclude that the statement $P(\alpha)$ is correct. So far it is identical to traditional induction. But as introduced in the section on ordinal numbers there is a certain type of ordinal called the limit ordinal. Limit ordinals require that the statement is also proved from the transition between λ_1 to λ_2 where λ_2 is a limit ordinal and $\lambda_1 < \lambda_2$.

In order to understand this properly we will study the following example. We shall prove the theorem;

Theorem 2.2. *Every infinite ordinal can be expressed as a sum of a limit ordinal and a finite ordinal. The sum is unique for every ordinal number.*

In order to prove this we shall break the proof down into two parts. First proving that the infinite ordinal can be written as a sum and then the uniqueness of that expression. The first proof is found below written out in stages of induction.

- Base/Limit ordinals: If α is a limit ordinal it can be written as $\alpha + 0$.
- Assumption: The ordinal number β can be written as $\beta = \gamma + n$ where γ is a limit ordinal and n is finite.
- Induction: In order to find the ordinal number $\alpha = \beta + 1$ we combine our expressions for α and β to get $\alpha = \gamma + n + 1$. This expression consists of a limit ordinal γ and a finite ordinal $n + 1$.
- Conclusion: Our assumption and induction-steps confirms the statement to be true, under the circumstances that there is a base step that is correct, which we have also shown. The first part of the theorem is thereby proved.

Now to the second statement; that the expression is unique. This will also be broken down into steps of induction.

- Base/Limit ordinals: $\alpha = \gamma + n$ is only a limit ordinal if $n = 0$. As α can be rewritten as $\alpha = \alpha + 0$ this implies that $\alpha + 0 = \gamma + n$ which in turn implies that $n = 0$ and $\gamma = \alpha$. This confirms that limit ordinals are uniquely expressed.
- Assumption: We assume that the ordinal β is expressed uniquely as $\beta = \gamma + n$.
- Induction: We seek an expression for the ordinal number $\alpha = \beta + 1$. Using our expression for β we can reformulate α into $\alpha = \gamma + n + 1$. We now assume that there is a second way to express α using a limit ordinal and a finite ordinal: $\alpha = \gamma' + k$. We know that α is not a limit ordinal as it is defined as the successor to the ordinal β. This means that $k > 0$ and also that we can express the ordinal β as $\beta = \alpha - 1 = \gamma' + (k - 1)$. As we have an expression for β that we know to be unique of the form $\beta = \gamma + n$ these two forms can be compared. This gives that $\gamma' = \gamma$ and that $n = k - 1$. This confirms that the only way to express α is $\gamma + n + 1$.
- Conclusion: The assumption and induction-steps confirm that these expressions of ordinal sums are unique if there is a base step, which in this case is the step confirming the statement to be true for all limit ordinals.

As we can see from the two proofs the base step doubles as confirmation that the statement is true for all limit ordinals. This as opposed to when working with traditional induction where the base step only needs to confirm that the statement is correct for a single starting point, usually characterized by its location $n = 0$.

2.3 Cantor Normal Form

It is possible to express ordinal α numbers as follows;

$$\alpha = c_1\omega^{\beta_1} + c_2\omega^{\beta_2} + c_3\omega^{\beta_3}...$$

Where c_k are positive integers, β_k are ordinal numbers, ω is the first transfinite ordinal and k is a natural number. This definition of the Cantor Normal form is not complete and as a result does not apply to all ordinals. But as we will not use ordinals of a larger size than this normal form can express the definition will suffice.

3 Proof

The following proof will cover the first as well as the second theorem presented in the introduction. First Theorem 1.2 will be proved. The idea of the proof is to create a similar sequence to the Goodstein sequence. If it can then be shown that that sequence eventually reaches an end it will prove that the Goodstein sequence does as well. For this statement to be true there are several criteria that the second sequence has to fulfill.

Firstly the second sequence has to be bijective to the Goodstein sequence. This is a base criteria that will allow the second sequence to be thought of as similar to the original sequence.

Secondly the second sequence has to have a least element. To ensure this the second sequence will be constructed using ordinal numbers. Ordinal numbers are well ordered sets which means that they by definition have a least elements.

Thirdly the sequence has to be strictly decreasing. As the sequence has a least element due to the second criteria the sequence will with absolute certainty reach a stop if it is also strictly decreasing.

If it is possible to construct such a sequence then the Goodstein sequence is proved to be finite. This is loosely sketched in Figure 6 below. In the figure there are two different sequences presented, $\{a_i\}$ and $\{b_i\}$ which have a bijection between them. The second sequence $\{b_i\}$ terminates at zero which indicates that sequence $\{a_i\}$ also terminates.

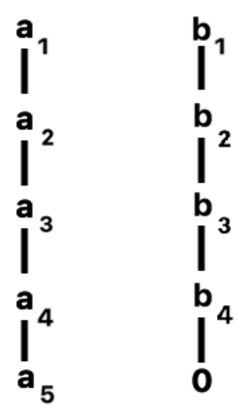

Figure 6: Series $\{a_i\}$ and $\{b_i\}$.

In the case of the Goodstein sequence termination means that it has reached zero as there is no other number that the Goodstein iteration cannot be performed on.

The first step is to more formally define the Goodstein representation of base n. This will then aid us as we attempt to create the sequence of ordinal numbers. The ordinal number will simply be defined by exchanging the n in the base representation of m by ω. We write the base n representation of m to be

$$m = a_k n^k + a_{k-1} n^{k-1} + ... + a_1 n + a_0 \qquad n, m \in \mathbb{N} \text{ and } n > 0$$

Where a_i are the different coefficients of the different n-terms. This is clearly simplified as the exponents are not expressed in base n, this is not done as our general sum representing every term m recursively will describe the exponents.

We will now define the function f, the general sum that is the base representation of m. It can either be used to produce a term m if $x \in \mathbb{N}$ or an ordinal number if $x = \omega$.

$$f^{m,n}(x) = \sum_{i=0}^{k} a_i x^{f^{i,n}(x)}$$

This base representation of m can be validated through induction using the base example $f^{0,n}(x) = 0$. It is possible to use f in order to create terms equivalent of the Goodstein sequence or to produce ordinal numbers, this is demonstrated below.

$$G_n(m) = f^{m,n}(n+1) - 1 \qquad m > 0$$
$$o_n(m) = f^{m,n}(\omega)$$

We will also define $G_n(0) = o_n(0) = 0$. With this we have now defined the second sequence of ordinal numbers discussed in the idea of the proof.

We will now define an operation $\langle\alpha\rangle(n)$ on ordinal numbers. It is this operation that will allow us to show that the sequence of ordinal numbers is bijective to the Goodstein sequence and that it is strictly decreasing. The operation is defined for $n \in \mathbb{N}$ and $\alpha < \epsilon_0$ and is defined by transfinite induction on α;

$$\langle 0\rangle(n) = 0, \quad \langle\beta + 1\rangle(n) = \beta$$

and also for $\delta > 0$,

$$\langle(\beta+1)\omega^{\delta}\rangle(n) = \beta\omega^{\delta} + n\omega^{\langle\delta\rangle(n)} + \langle\omega^{\langle\delta\rangle(n)}\rangle(n)$$

This operation corresponds to a reduction in complexity, which is synonymous with the operation decreasing the ordinal. The best way of understanding the operation is likely through a thorough example that will be presented after the smaller examples below. This first operation will demonstrate how an ordinal is reduced to a natural number.

$$\langle\omega\rangle(n) = 0 \cdot \omega + n\omega^0 + \langle\omega^0\rangle(n) = 0 + n + \langle 1\rangle(n) = n + 0 = n$$

Above we demonstrated that a lone ω will be reduced into a finite number n. This finite number depends on how many times the operation has been performed. In the case above it has been performed n times. This leads us to our next example which is very similar to the first.

$$\langle 2\omega\rangle(n) = \langle(1+1)\omega\rangle(n) = 1 \cdot \omega + n\omega^0 + \langle\omega^0\rangle(n) =$$
$$= \omega + n + \langle 1\rangle(n) = \omega + n + 0 = \omega + n$$

This example strengthens our previous assumption that the operation has the ability to reduce infinite ordinal numbers to finite numbers. In this case first one of the ordinal numbers is reduced to a finite number while the other is left unaffected. The last example will be derived from the second basic description of the operation.

$$\langle\omega + k\rangle(n) = \omega + k - 1$$

The reduction above is a consequence of the second basic defining characteristic of our operation, and also similar to a characteristic of the Goodstein sequence. We know that if there is a finite number present, the operation's function is that it reduces that finite number by one.

To summarize it is clear that the operation is strictly reducing the input. Now that certain examples have been presented along with the general idea we will now attempt to reduce the ordinal $\omega^\omega + 3\omega^2 + \omega$, starting at $n = 1$.

$$\begin{aligned}
\langle \omega^\omega + 3\omega^2 + \omega \rangle(1) &= \omega^\omega + 3\omega^2 + \langle \omega \rangle(1) = \omega^\omega + 3\omega^2 + 1 \\
\langle \omega^\omega + 3\omega^2 + 1 \rangle(2) &= \omega^\omega + 3\omega^2 + 1 - 1 = \omega^\omega + 3\omega^2 \\
\langle \omega^\omega + 3\omega^2 \rangle(3) &= \omega^\omega + \langle (2+1)\omega^2 \rangle(3) = \\
&= \omega^\omega + 2\omega^2 + 3\omega^{\langle 2 \rangle(3)} + \langle \omega^{\langle 2 \rangle(3)} \rangle(3) = \omega^\omega + 2\omega^2 + 3\omega + 3
\end{aligned}$$

Another tree steps will reduce the sequence to $2\omega^\omega + \omega^2 + 3\omega$.

$$\langle \omega^\omega + 2\omega^2 + 3\omega \rangle(6) = \omega^\omega + 2\omega^2 + \langle (2+1)\omega \rangle(6) = \omega^\omega + 2\omega^2 + 2\omega + 6$$

Another six steps will reduce the sequence to $\omega^\omega + 2\omega^2 + 2\omega$. The last term 2ω will be reduced in the same manner showed in the last step above. When this is done what remains of the sequence is $\omega^\omega + 2\omega^2$. This can be compared to when $n = 2$ and the sequence was $\omega^\omega + \omega^2 3$. It is apparent that one of the ω^2 terms has been reduced. As the operation keeps being performed the sequence will become ω^ω. We will assume that this happens when $n = k$. This following reduction is quite extensive

$$\begin{aligned}
&\langle \omega^\omega \rangle(k) = \\
&\quad \langle (0+1)\omega^\omega \rangle(k) = \\
&\quad 0 \cdot \omega^\omega + k\omega^{\langle \omega \rangle(k)} + \langle \omega^{\langle \omega \rangle(k)} \rangle(k) = \\
&\quad k\omega^k + \langle (0+1)\omega^k \rangle(k) = \\
&\quad k\omega^k + 0 \cdot \omega^k + k\omega^{\langle k \rangle(k)} + \langle \omega^{\langle k \rangle(k)} \rangle(k) = \\
&\quad k\omega^k + k\omega^{k-1} + \langle (0+1)\omega^{k-1} \rangle(k) = \\
&\quad k\omega^k + k\omega^{k-1} + 0 \cdot \omega^{k-1} + k\omega^{\langle k-1 \rangle(k)} + \langle \omega^{\langle k-1 \rangle(k)} \rangle(k) = \\
&\quad k\omega^k + k\omega^{k-1} + k\omega^{k-2} + \langle (0+1)\omega^{k-2} \rangle(k) = \\
&\quad k\omega^k + k\omega^{k-1} + k\omega^{k-2} + 0 \cdot \omega^{k-2} + k\omega^{\langle k-2 \rangle(k)} + \langle \omega^{\langle k-2 \rangle(k)} \rangle(k) = \\
&\quad k\omega^k + k\omega^{k-1} + k\omega^{k-2} + k\omega^{k-3} + \langle (0+1)\omega^{k-3} \rangle(k) = \\
&\qquad \qquad \ldots\ldots\ldots \\
&\qquad \qquad = k\omega^k + k\omega^{k-1} + k\omega^{k-2} + k\omega^{k-3} + \ldots + k\omega^2 + k\omega + k
\end{aligned}$$

The result should be inspected, the most important observation is that the highest complexity term $\omega^k k$ is of lower complexity than ω^ω, the original term which the operation was performed on. This is the function of the operation, it lowers the complexity of an ordinal number and performed enough times will reduce it to zero.

The change of base that is custom when working with the traditional Goodstein sequence is also represented in the operation. This can be shown by comparing the result above with rewriting a Goodstein sequence term into another base. Let the our term be $7^7 - 1$, in ordinal numbers this would correspond to $\langle \omega^\omega \rangle(6)$. The operation would reduce the ordinal, using the result above, to;

$$6\omega^6 + 6\omega^5 + 6\omega^4 + 6\omega^3 + 6\omega^2 + 6\omega + 6.$$

Rewriting $7^7 - 1$ into base 7 becomes;

$$6 \cdot 7^6 + 6 \cdot 7^5 + 6 \cdot 7^4 + 6 \cdot 7^3 + 6 \cdot 7^2 + 6 \cdot 7 + 6.$$

The two expressions show clear similarities, the exponents in both cases are reduced compared to the original numbers.

Having defined the operation we will now move on to a lemma. The point of this lemma is to confirm that there is a bijection between the ordinal sequence created by the operation and the Goodstein Sequence.

Lemma 3.1. *(i) For $m \geqslant 0$, $n > 1$, if $o_{n+1}(m_k) = \alpha$ then $o_{n+1}(m_{k+1}) = \langle \alpha \rangle(n)$.*

(ii) For $n > 1$, $\langle o_n(m) \rangle(n) = o_{n+1}(G_n(m))$.

Proof of Lemma 3.1

(i) First we express m in the base $n+1$ representation.

$$m = a_p(n+1)^{f^{p,n+1}(n+1)} + a_{p-1}(n+1)^{f^{p-1,n+1}(n+1)} + ... + a_0(n+1)^{f^{0,n+1}(n+1)}$$

where $0 \leq a_i \leq n$. Now we will let j be as small as can be without $a_j \neq 0$. This represents the termination of terms trough subtraction. If $j = 0$ then m likely is the first element of the Goodstein sequence. As $j = 0$ is a very simple case we will further assume that $j > 0$. Our last assumption is that the result holds for all m' where $0 < m' < m$. We will now observe two representations of what will show to be the same ordinal number. First the ordinal representation of m_{k+1}.

$$o_{n+1}(m_{k+1}) = \Big(\sum_{i=j+1}^{p} a_i \omega^{f^{i,n+1}(\omega)} \Big) + (a_j - 1)\omega^{f^{j,n+1}(\omega)} + \\ o_{n+1}(n \cdot (n+1)^{f^{j,n+1}(n+1)-1}) + o_{n+1}\big((n+1)^{f^{j,n+1}(n+1)-1} - 1\big)$$

The sum above is the sum of the yet untouched terms making up the ordinal number. The remainder of the terms is the result of a subtraction of one. Now the previously defined operation on the ordinal number $\alpha = o_{n+1}(m)$.

$$\langle\alpha\rangle(n) = \Big(\sum_{i=j+1}^{p} a_i\omega^{f^{i,n+1}(\omega)}\Big) + (a_j - 1)\omega^{f^{j,n+1}(\omega)} + n\omega^{\langle f^{i,n+1}(\omega)(n)} + \langle\omega^{\langle f^{i,n+1}(\omega)\rangle(n)}\rangle(n)$$

Using induction it becomes simple to prove that these are equal, showing that the operation has the same consequence as a step of the Goodstein sequence.

(ii) We now let $m = \sum_{i=j}^{p} b_i n^{f^{i,n}(n)}$ where $0 \leq b_i < n$ and $b_j \neq 0$. In the case where $j = 0$ we understand see that $\langle o_n(m)\rangle(n) = o_{n+1}(G_n(m))$ so we will therefore assume that $j > 0$. This means that

$$\langle o_n(m)\rangle(n) = \Big(\sum_{i=j+1}^{p} b_i\omega^{f^{i,n+1}(\omega)}\Big) + (b_j - 1)\omega^{f^{j,n+1}(\omega)} + n\omega^{\langle f^{i,n+1}(\omega)(n)} + \langle\omega^{\langle f^{i,n+1}(\omega)\rangle(n)}\rangle(n)$$

and

$$o_{n+1}(G_n(m)) = \Big(\sum_{i=j+1}^{p} b_i\omega^{f^{i,n+1}(\omega)}\Big) + o_{n+1}\big((n+1)^{f^{j,n+1}(n+1)}b_j - 1\big) =$$

$$\Big(\sum_{i=j+1}^{p} b_i\omega^{f^{i,n+1}(\omega)}\Big) + (b_j - 1)\omega^{f^{j,n+1}(\omega)} + o_{n+1}\big((n+1)^{f^{j,n+1}(n+1)-1}n\big) + o_{n+1}\big((n+1)^{f^{j,n+1}(n+1)-1} - 1\big)$$

From (i) we know that

$$o_{n+1}\big((n+1)^{f^{j,n+1}(n+1)-1}n\big) = n\omega^{\langle f^{i,n+1}(\omega)(n)}$$

and also that

$$o_{n+1}\big((n+1)^{f^{j,n+1}(n+1)-1} - 1\big) = \langle\omega^{\langle f^{i,n+1}(\omega)\rangle(n)}\rangle(n),$$

which finally proves the lemma.

The first part of the lemma simply restates what is described in the text above; the operation reduces the Goodstein Sequence ordinal's complexity, e.g. through subtraction. The second part of the lemma builds of the first part to develop the idea that the Goodstein sequence is similar to the ordinal sequence. More specifically it confirms that the following two numbers are equal;

- The resulting number created by the operation performed on $o_n(m)$.

- The ordinal number of the Goodstein sequence number $G_n(m)$.

The lemma has thereby proved bijectivity between the two sequences. Thereby the lemma confirms that the operation has the properties that we need;

- It decreases the ordinals complexity.
- There is bijectivity between the two sequences.

This indicates that for each Goodstein Sequence $b_0, b_1, b_2, ...$ there is a corresponding sequence of ordinals $o_n(b_0), o_{n+1}(b_1), o_{n+2}(b_2), ...$ for which the operation can be used.

To simplify text we will use the notation $\langle\alpha\rangle(n_1, n_2, ..., n_k)$ for $\langle...\langle\langle\alpha\rangle(n_1)\rangle(n_2)...\rangle(n_k)$. With this notation a sequence of ordinal numbers in a Goodstein sequence can be written as below

$$o_n(b_0) = \alpha, \quad \langle\alpha\rangle(n), \quad \langle\alpha\rangle(n, n+1), \quad \langle\alpha\rangle(n, n+1, n+2), \quad ...$$

It is known from Lemma 3.1 that this must be a decreasing sequence due to the characteristics of the operation. Therefore for any ordinal α and $n \in \mathbb{N}$ the following is true;

$$\langle\alpha\rangle(n) < \alpha$$

As ordinal numbers are well ordered sets a strictly decreasing ordinal number sequence will eventually be reduced to the least element; zero. It has now been shown that a sequence with a bijection to the Goodstein sequence reduces to zero, which therefore ensures that the Goodstein sequence does as well. Theorem 2.2 is thereby proved.

Now that the base is set it is time to move on to Theorem 2.1. The two problems are very similar, the idea for the first proof will be reused for the second proof. To be able to reuse the idea there are two things that must be done;

- The hydra needs to be reformulated into an ordinal number that symbolises its complexity.
- An operation simulating the cutting of a head and the resulting regeneration needs to be defined.

The first step is quite simple. All nodes that correspond to a head is given the number zero. Traverse from the head one step closer to the root; these nodes become the ordinal ω raised to the power of the heads above it. This step is then repeated; traversing closer to the root and having the new ω node be raised to the power of the ordinal above. The final ordinal number of the hydra is then the number at the root. An example of a hydra and its ordinal is shown in Figure 5.

The next step is to formulate an operation that represents the act of severing a head and the following regeneration. We will denote this operation $[\alpha](n)$ but we will abstain from giving a strict definition. The operation has the property

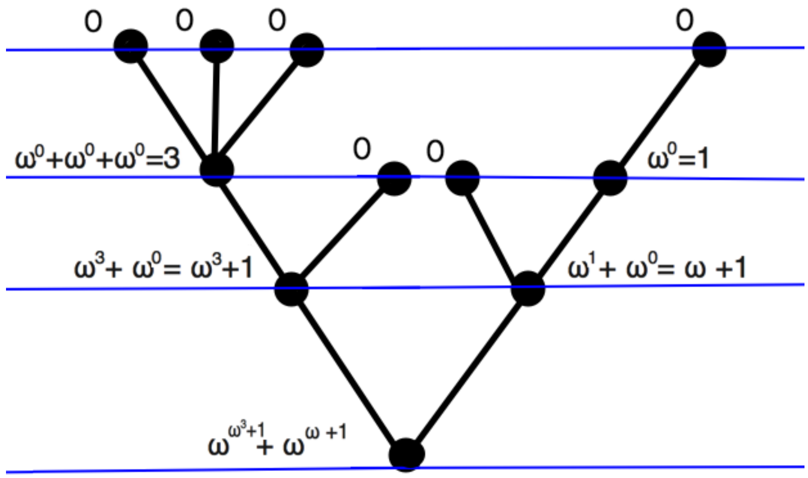

Figure 7: A hydra with its ordinal number presented at its root; $\omega^{\omega^3+1} + \omega^{\omega+1}$

that it decreases the complexity of the ordinal similarly to the operation defined earlier for the proof of Theorem 2.2. This implies that despite the increase of the number of heads the ordinal will become "simpler"; of a lower level. An example of a reduction of the hydra in Figure 7 is shown in Figure 8.

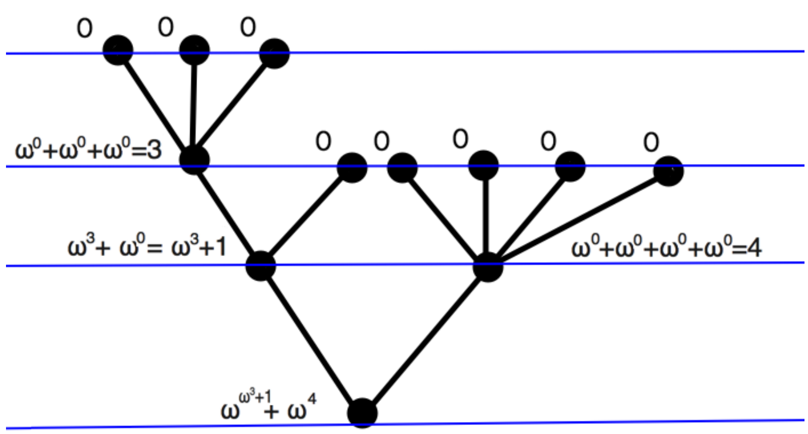

Figure 8: The hydra in Figure 7 after a blow. The ordinal is now $\omega^{\omega^3+1} + \omega^4$

The ordinal changed from $\omega^{\omega^3+1}+\omega^{\omega+1}$ to $\omega^{\omega^3+1}+\omega^4$. More specifically the term $\omega^{\omega+1}$ changed into ω^4. This is a clear example of change in complexity. The ordinal ω^ω is of higher complexity than ω^n. This can also be expressed as

$$[\alpha](n) < \alpha.$$

With the aid of the proof of Theorem 1.2 it is now apparent that the Hydra will eventually be reduced to zero. Hercules will always win the battle, and thereby Theorem 1.1 is proved.

4 Peano Arithmetic

4.1 Peano Arithemtic

There are a group of important properties on the natrual numbers that together are called the Peano Axioms. These are (Devlin, 1993, 108);

(i) 0 is part of the natural numbers.

(ii) For every natrual number the sucessor number S(n) is also a natrual number. (Can also be expressed as n^+).

(iii) The sucessor number to any natural number is not zero.

(iv) If the succesor numbers of n and m are equal, then m and n are equal.

(v) If all natural numbers are elements in a set, zero is part in that set, the succesor number of n is always present in the set if n is; then the set can be said to be the natural numbers.

These five axioms can define the natural numbers and based on this all theory of rational, real and complex numbers can be created. It should be noted that Set Theory has all the tools to define the axioms above and is therefore one of the most basic mathematical disciplines.

The proofs in the section above are more important than for the obvious reason that they prove the presented theorems. The model of Peano Arithmetic is an insufficient basis for these proofs, in fact neither of Theorem 1.1 nor Theorem 1.2 can be proved using the principles of Peano Arithmetic alone. It is required to use a stronger system, such as second order arithmetic. One of the characteristics of second order arithmetic is that it allows quantification. This is the idea that allows successor numbers and as a result enables limit ordinals and their successors. Note that in the Peano Arithmetic system the theorems would not be disproved but there would be insufficient basis to produce a formal proof.

References

[1] Kirby, L., Paris, J. (1982). *Accessible Independence Results for Peano Arithmetic* London Mathematical Society, vol. 14, no. 4, July 1982, pp. 285–293.

[2] European Space Agency (n.d). *How many stars are there in the Universe?* `http://www.esa.int/Our_Activities/Space_Science/Herschel/How_many_stars_are_there_in_the_Universe`

[3] Redd, N. T. (2017, June 7). *How Old is the Universe* `https://www.space.com/24054-how-old-is-the-universe.html`

[4] Enderton, H. B. (1977). *Elements of set theory* Boston: Academic Press.

[5] Kaplansky, I. (1977). *Set Theory and Metric Spaces* (2nd ed.) New York: AMS Chelsea Publishing.

[6] Devlin, K. J. (1993). *The joy of sets : fundamentals of contemporary set theory* (2nd ed.) New York: Springer-Vlg.

[7] Weisstein, E. W. (n.d). *Totally Ordered Set* `http://mathworld.wolfram.com/TotallyOrderedSet.html`

[8] Abian, A. (1965). *The theory of sets and transfinite arithmetic* Philadelphia: Saunder.

[9] Conway, J. H. (2000). *Boken om tal* Lund: Studentlitteratur.

[10] Hansson, B., Gärdenfors, P. (1973). *Mängd och tal : en introduktion till teorin för kardinal- och ordinaltal* Lund: Gleerup.