MA5219 LOGIC AND FOUNDATION OF MATHEMATICS I

# Lecture 17: Ordinals

Tin Lok Wong

15 October, 2018

In the previous lecture, we saw Presburger's theorem that $\mathrm{Th}(\mathbb{N}, 0, 1, +, <)$ is recursive. This probes into the limits of the First Incompleteness Theorem, in the weak form we stated on page 56: how much arithmetic truth can one algorithmically determine? In these few lectures, we probe into the limits of the Second Incompleteness Theorem: how much do we need to prove the consistency of arithmetic? As an example, we study the consistency of a natural axiomatization of arithmetic originated from Peano, which we now call *Peano arithmetic* and denote by PA.

We will introduce PA in the next lecture, but we can already give a proof of its consistency now. It will be clear that $\mathbb{N} \models \mathrm{PA}$. So $\mathrm{PA} \not\models \bot$ by the definition of semantic entailment. This implies $\mathrm{PA} \nvdash \bot$ by Soundness.

There are at least two defects of this proof. First, one (traditional) reason for obtaining a consistency proof for PA is to guarantee that mathematicians' reasoning about the natural numbers is free from contradictions. The decisive role played by the natural numbers in this proof deems the proof useless for this purpose. Second, this proof tells us very little about the consistency of PA per se. For instance, the same proof also works for $\mathrm{I}\Delta_0(\exp)$, although intuitively $\mathrm{I}\Delta_0(\exp)$ should require 'weaker arguments' because $\mathrm{I}\Delta_0(\exp)$ is much weaker than PA, as we will see.

Nevertheless, one cannot prove the consistency of PA from thin air: we know from the Second Incompleteness Theorem that even PA cannot prove the consistency of PA. So we need in the proof an (evidently true) axiom which transcends PA. For this purpose, we use the axiom of *transfinite induction* on ordinals, following the work of Gentzen. The aim of this lecture is to introduce ordinals and transfinite induction.

We treat the notion of ordinals as *primitive*, i.e., it is not defined in terms of other notions, in the same way we treat the natural numbers. Hence, the notion of ordinals cannot have a rigorous definition. To convey the idea, we can merely describe intuitively what we mean by ordinals, demonstrate with sufficiently many examples, and assert some key properties. Intuitively, an *ordinal* measures how often a discrete process is repeated. (This description originates from Forster.) The following demonstrates how ordinals are generated by repetition:

$$
\begin{aligned}
&0,\ 1,\ 2,\ 3,\ \ldots,\ n,\ \ldots\ \omega,\ \omega+1,\ \omega+2,\ \ldots,\ \omega+n,\ \ldots\ \omega+\omega=\omega 2,\ \omega 2+1,\ \omega 2+2,\ \ldots,\\
&\omega 2+n,\ \ldots\ \omega 2+\omega=\omega 3,\ \ldots\ \omega 4,\ \ldots\ \omega n,\ \ldots\ \omega\omega=\omega^2,\ \omega^2+1,\ \omega^2+2,\ \ldots\ \omega^2+\omega,\ \ldots\\
&\omega^2+\omega 2,\ \ldots\ \omega^2+\omega n,\ \ldots\ \omega^2+\omega^2=\omega^2 2,\ \ldots\ \omega^2 3,\ \ldots\ \omega^2 n,\ \ldots\ \omega^2\omega=\omega^3,\ \ldots\ \omega^n,\ \ldots\ \omega^\omega,\\
&\ldots\ \omega^\omega 2,\ \ldots\ \omega^\omega 3,\ \ldots\ \omega^\omega n,\ \ldots\ \omega^\omega\omega=\omega^{\omega+1},\ \ldots\ \omega^{\omega+2},\ \ldots\ \omega^{\omega 2},\ \ldots\ \omega^{\omega 3},\ \ldots\omega^{\omega n},\ \ldots\\
&\omega^{\omega^2},\ \ldots\ \omega^{\omega^3},\ \ldots\ \omega^{\omega^n},\ \ldots\ \omega^{\omega^\omega},\ \ldots\ \omega^{\omega^{\cdot^{\cdot^{\cdot^\omega}}}}\Big\}\,n\text{-many }\omega\text{'s},\ \ldots\ \omega^{\omega^{\omega^{\cdot^{\cdot^{\cdot}}}}}\Big\}\,\omega\text{-many }\omega\text{'s}=\varepsilon_0,\ \ldots\\
&\omega^{\varepsilon_0},\ \ldots\ \omega^{\omega^{\varepsilon_0}},\ \ldots\ \omega^{\omega^{\cdot^{\cdot^{\cdot^{\varepsilon_0}}}}}\Big\}\,n\text{-many }\omega\text{'s},\ \ldots\ \varepsilon_1,\ \ldots\ \varepsilon_2,\ \ldots\ \varepsilon_n,\ \ldots\ \varepsilon_\omega,\ \ldots\ \varepsilon_{\omega+1},\ \ldots\ \varepsilon_{\omega+2},\\
&\ldots\ \varepsilon_{\omega+3},\ \ldots\ \varepsilon_{\omega 2},\ \ldots\ \varepsilon_{\omega 3},\ \ldots\ \varepsilon_{\omega^2},\ \ldots\ \varepsilon_{\omega^3},\ \ldots,\ \varepsilon_{\omega^\omega},\ \ldots\ \varepsilon_{\omega^{\omega^\omega}},\ \ldots\ \varepsilon_{\omega^{\omega^{\cdot^{\cdot^{\cdot^\omega}}}}}\Big\}\,n\text{-many }\omega\text{'s},\\
&\ldots\ \varepsilon_{\varepsilon_0},\ \ldots\ \varepsilon_{\varepsilon_{\varepsilon_0}},\ \ldots\ \varepsilon_{\varepsilon_{\cdot_{\cdot_{\cdot_{\varepsilon_0}}}}}\Big\}\,n\text{-many }\varepsilon\text{'s},\ \ldots\ \varepsilon_{\varepsilon_{\varepsilon_{\cdot_{\cdot_{\cdot}}}}}\Big\}\,\omega\text{-many }\varepsilon\text{'s},\ \ldots
\end{aligned}
$$

We see from this list that ordinals are linearly ordered. We denote this order by $<$. Since all natural numbers are ordinals, we can view the notion of ordinals as a generalization of the notion of natural numbers. Analogous to the Principle of (Strong) Induction for natural numbers, the ordinals satisfy the Principle of Transfinite Induction.

**Principle of Transfinite Induction (TI).** Let $P(\delta)$ be a property of an ordinal $\delta$. If

$$\forall \text{ ordinals } \alpha{<}\delta\ \ P(\alpha) \quad\Rightarrow\quad P(\delta) \qquad (\mathrm{Prog}_P(\delta))$$

for every ordinal $\delta$, then $P(\delta)$ for every ordinal $\delta$.

One cannot prove the Principle of Transfinite Induction simply because there is nothing to prove it from. However, one can provide an argument explaining why it should be true. The situation is similar to that for the Principle of Induction for the natural numbers.

*Intuition for TI.* We know

$$\begin{array}{lll} & P(0) & \text{by } \mathrm{Prog}_P(0); \\ \therefore & P(1) & \text{by } \mathrm{Prog}_P(1); \\ \therefore & P(2) & \text{by } \mathrm{Prog}_P(2); \\ & \vdots & \\ \therefore & P(n) & \text{by } \mathrm{Prog}_P(n); \\ & \vdots & \\ \therefore & P(\omega) & \text{by } \mathrm{Prog}_P(\omega); \\ \therefore & P(\omega+1) & \text{by } \mathrm{Prog}_P(\omega+1); \\ & \vdots & \end{array}$$

□

The contrapositive of TI is important enough to have a special name.

**Well-Ordering Principle.** Every nonempty collection $C$ of ordinals has a least element.

*Proof.* Suppose $C$ has no least element. If $\delta$ is an ordinal such that

$$\forall \text{ ordinals } \alpha{<}\delta \quad \alpha \notin C,$$

then $\delta \notin C$ because otherwise $\delta = \min C$. So $\delta \notin C$ for all ordinals $\delta$ by TI. □

In this module, we will only deal with ordinals up to $\varepsilon_0 = \sup\{\omega_n : n \in \mathbb{N}\}$, where

$$\omega_n = \omega^{\omega^{\cdot^{\cdot^{\cdot^{\omega}}}}} \Big\} n\text{-many } \omega\text{'s}.$$

(Note that in set theory the notation $\omega_n$ usually refers to the $n$th infinite initial ordinal, which is much bigger than our $\omega_n$.) As Cantor showed, every nonzero ordinal less than $\varepsilon_0$ can be expressed uniquely in the normal form

$$\omega^{\alpha_0}k_0 + \omega^{\alpha_1}k_1 + \cdots + \omega^{\alpha_r}k_r,$$

where $\alpha_0, \alpha_1, \ldots, \alpha_r$ is a strictly decreasing sequence of ordinals less than $\varepsilon_0$, and $k_0, k_1, \ldots, k_r \in \mathbb{N} \setminus \{0\}$. As is usually done in proof theory, we identify an ordinal below $\varepsilon_0$ with its Cantor Normal Form representation. This allows us to define these ordinals via their representations. However, one should remember that these representations themselves are *not* ordinals. Analogously, von Neumann provided a representation of ordinals in set theory as transitive sets linearly ordered by $\in$, but one should remember that these representations themselves are not ordinals. Elements of $\varepsilon_0$, as defined below, are finite strings of symbols constructed in a specific way. Nevertheless, the ordinals they represent can have an infinitary side.

**Definition.** Let $\omega_0 = \{0\}$ and $<_0$ be the unique strict linear order on $\omega_0$. Suppose $(\omega_n, <_n)$ is defined. Set

$$\omega_{n+1} = \{0\} \cup \{\omega^{\alpha_0}k_0 + \omega^{\alpha_1}k_1 + \cdots + \omega^{\alpha_r}k_r : \bar{\alpha} \in \omega_n \text{ and } \alpha_0 >_n \alpha_1 >_n \cdots >_n \alpha_r \text{ and } \bar{k} \in \mathbb{N} \setminus \{0\}\}.$$

Let $<_{n+1}$ be the unique strict linear order on $\omega_{n+1}$ satisfying the following properties.

- $0 <_{n+1} \alpha$ for all $\alpha \in \omega_{n+1} \setminus \{0\}$.

- If $\alpha, \beta \in \omega_{n+1}$ and

$$\alpha = \omega^{\alpha_0} k_0 + \omega^{\alpha_1} k_1 + \cdots + \omega^{\alpha_r} k_r \quad \text{and} \quad \beta = \omega^{\beta_0} \ell_0 + \omega^{\beta_1} \ell_1 + \cdots + \omega^{\beta_s} \ell_s,$$

  then $\alpha <_{n+1} \beta$ if and only if one of the two situations below hold:

  - there is $j \leqslant \min\{r, s\}$ such that
    - $\alpha_0 = \beta_0$ and $\alpha_1 = \beta_1$ and ... and $\alpha_{j-1} = \beta_{j-1}$;
    - $k_0 = \ell_0$ and $k_1 = \ell_1$ and ... and $k_{j-1} = \ell_{j-1}$; and
    - $\alpha_j <_n \beta_j$ or ($\alpha_j = \beta_j$ and $k_j < \ell_j$);
  - $r < s$ and $\alpha_0 = \beta_0$ and $\alpha_1 = \beta_1$ and ... and $\alpha_r = \beta_r$ and $k_0 = \ell_0$ and $k_1 = \ell_1$ and ... and $k_r = \ell_r$.

Define $\varepsilon_0 = \bigcup_{n \in \mathbb{N}} \omega_n$ and $<_\omega = < = \bigcup_{n \in \mathbb{N}} <_n$.

The $\omega_n$'s form an increasing chain and the $<_n$'s are compatible with each other. Amongst other consequences, these ensure that $<$ is a linear order on $\varepsilon_0$.

**Lemma 17.1.** $\omega_n \subseteq \omega_{n+1}$ and $<_n = <_{n+1} \restriction \omega_n$ for all $n \in \mathbb{N}$.

*Proof.* Proceed by induction on $n$. Notice $\omega_0 = \{0\} \subseteq \omega_1$ by definition. If $\omega_n \subseteq \omega_{n+1}$, then

$$\begin{aligned} \omega_{n+1} &= \{0\} \cup \{\ldots : \bar{\alpha} \in \omega_n \ldots\} && \text{by the definition of } \omega_{n+1}; \\ &\subseteq \{0\} \cup \{\ldots : \bar{\alpha} \in \omega_{n+1} \ldots\} && \text{as } \omega_n \subseteq \omega_{n+1}; \\ &= \omega_{n+2} && \text{by the definition of } \omega_{n+2}. \end{aligned}$$

The orders $<_n$ and $<_{n+1}$ agree on $\omega_n$ because they have the same definition. □

**Convention 17.2.** Let $k \in \mathbb{N} \setminus \{0\}$ and $\alpha \in \varepsilon_0$. Abbreviate

$$\omega^0 k = k \quad \text{and} \quad \omega^{\omega^0} = \omega \quad \text{and} \quad \omega^\alpha 1 = \omega^\alpha.$$

Roughly speaking, the set $\omega_n$ contains those elements of $\varepsilon_0$ in which the tallest tower of exponentials has a stack of strictly less than $n$-many $\omega$'s after abbreviation.

**Example 17.3.** $\omega^{\omega+2} + \omega^2 + \omega + 1 = \omega^{\omega^{\omega^0 1} 1 + \omega^0 2} 1 + \omega^{\omega^0 2} 1 + \omega^{\omega^0 1} 1 + \omega^0 1 \in \omega_3$ because $\omega^0 1, \omega^0 2 \in \omega_1$ and so $\omega^{\omega^0 1} 1 + \omega^0 2, \omega^{\omega^0 2} 1, \omega^{\omega^0 1} 1 \in \omega_2$.

Elements of $\varepsilon_0$ resemble the so-called complete/pure base-$b$ representations of positive integers.

**Definition.** Let $m, b \in \mathbb{N}$ with $m \neq 0$ and $b \geqslant 2$. Then one can uniquely write $m$ as

$$m = b^{m_0} k_0 + b^{m_1} k_1 + \cdots + b^{m_r} k_r,$$

where $m_0, m_1, \ldots, m_r \in \mathbb{N}$ with $m_0 > m_1 > \cdots > m_r$ and $k_0, k_1, \ldots, k_r \in \{1, 2, \ldots, b-1\}$. Every nonzero $m_j$ can also be written uniquely in this form, so that we can repeat this procedure until all the exponents are 0. The result is the *pure base-$b$ representation* of $m$.

**Example 17.4.** The pure base-3 representation of 256 is $3^{3^{3^0 1} 1 + 3^0 2} 1 + 3^{3^0 2} 1 + 3^{3^0 1} 1 + 3^0 1$.

Provided the base $b$ is larger than all the numbers that appear in an element $\alpha \in \varepsilon_0$, one can transform $\alpha$ into a pure base-$b$ representation and back by substitution. This substitution resembles a change of pure bases: if $b, c \in \mathbb{N}$ with $2 \leqslant b \leqslant c$, then replacing each occurrence of $b$ by $c$ in the pure base-$b$ representation of a natural number gives the pure base-$c$ representation of another number.

**Definition.** Let $m, b, c \in \mathbb{N}$ with $2 \leqslant b \leqslant c$. Let $\alpha \in \varepsilon_0$.

- If $m = 0$, then $\mathrm{Sub}^b_c(m) = 0$ and $\mathrm{Sub}^b_\omega(m) = 0$.

- If $m > 0$, then $\mathrm{Sub}_c^b(m)$ is the number obtained from the pure base-$b$ representation of $m$ by replacing each occurrence of $b$ by $c$.
- Similarly, if $m > 0$, then $\mathrm{Sub}_\omega^b(m)$ is the element of $\varepsilon_0$ obtained from the pure base-$b$ representation of $m$ by replacing each occurrence of $b$ by $\omega$.
- If $\alpha = 0$, then $\mathrm{Sub}_b^\omega(0) = 0$ and $\mathrm{C}(\alpha) = 0$.
- If $\alpha \neq 0$, then denote by $\mathrm{C}(\alpha)$ the largest natural number that appears in $\alpha$.
- If $\alpha \neq 0$ and $\mathrm{C}(\alpha) > b$, then $\mathrm{Sub}_b^\omega(\alpha)$ is the number obtained from $\alpha$ by replacing each occurrence of $\omega$ by $b$.

**Example 17.5.** Let $\alpha$ be the element of $\varepsilon_0$ considered in Example 17.3. Then by Example 17.4,

$$\mathrm{Sub}_4^3(256) = 4^{4^{4^0 1}1+4^0 2}1 + 4^{4^0 2}1 + 4^{4^0 1}1 + 4^0 1 = 4117;$$

$$\mathrm{Sub}_\omega^3(256) = \omega^{\omega^{\omega^0 1}1+\omega^0 2}1 + \omega^{\omega^0 2}1 + \omega^{\omega^0 1}1 + \omega^0 1 = \alpha; \text{ and}$$

$$\mathrm{Sub}_3^\omega(\alpha) = 3^{3^{3^0 1}1+3^0 2}1 + 3^{3^0 2}1 + 3^{3^0 1}1 + 3^0 1 = 256.$$

**Lemma 17.6.** Let $m, n, b, c \in \mathbb{N}$ and $\alpha, \beta \in \varepsilon_0$ such that $2 \leqslant b \leqslant c$ and $\mathrm{C}(\alpha), \mathrm{C}(\beta) < b$. Then

(1) $\mathrm{Sub}_\omega^c(\mathrm{Sub}_c^b(m)) = \mathrm{Sub}_\omega^b(m)$;

(2) $\mathrm{Sub}_b^\omega(\mathrm{Sub}_\omega^b(m)) = m$;

(3) $\mathrm{Sub}_\omega^b(\mathrm{Sub}_b^\omega(\alpha)) = \alpha$;

(4) $\alpha < \beta$ if and only if $\mathrm{Sub}_b^\omega(\alpha) < \mathrm{Sub}_b^\omega(\beta)$; and

(5) $m < n$ if and only if $\mathrm{Sub}_\omega^b(m) < \mathrm{Sub}_\omega^b(n)$.

*Proof sketch.* Part (2) is true because $\mathrm{C}(\mathrm{Sub}_\omega^b(m)) < b$ by the definition of pure base-$b$ representations. Part (3) is true because, as $\mathrm{C}(\alpha) > b$, replacing each occurrence of $\omega$ by $b$ in $\alpha$ as in the definition of $\mathrm{Sub}_b^\omega(\alpha)$ results in the pure base-$b$ representation of a number. Part (1) is true for a similar reason. Parts (4) and (5) are true because one can define the order on natural numbers in terms of pure base-$b$ representations in the same way as we define the order on $\varepsilon_0$. □

Exploiting this connection between ordinals less than $\varepsilon_0$ and pure base-$b$ representations of natural numbers for sufficiently large $b \in \mathbb{N}$, Goodstein proved the following theorem. This theorem may be surprisingly at first sight because it says that the growth induced by a successive increase in pure bases can always be annihilated by successively subtracting 1 in the long run.

**Theorem 17.7** (Goodstein)**.** The following are equivalent.

(i) Every nonempty subset of $\varepsilon_0$ has a least element.

(ii) For every $m \in \mathbb{N}$ and every nondecreasing function $F\colon \mathbb{N} \to \mathbb{N}$ with $F(0) \geqslant 2$, if we define $a_0 = m$ and, for every $j \in \mathbb{N}$,

$$a_{j+1} = \begin{cases} \mathrm{Sub}_{F(j+1)}^{F(j)}(a_j) - 1, & \text{if } a_j \neq 0; \\ 0, & \text{otherwise,} \end{cases}$$

then $\lim_{j\to\infty} a_j = 0$.

**Assignment 17.8.** Prove (i) ⇒ (ii) in Theorem 17.7. (Hint: consider $\{\mathrm{Sub}_\omega^{F(j)}(a_j) : j \in \mathbb{N}\}$.) [5 points]

By the Well-Ordering Principle, we know that condition (i) in Theorem 17.7 is true. Therefore, asserting the equivalence of (i) and (ii) is simply a roundabout way of asserting the truth of (ii). We formulate Theorem 17.7 in this roundabout way because, as the reader can check, the proof of this equivalence uses only very restricted tools, and such tools are not strong enough to establish the Well-Ordering Principle for ordinals less than $\varepsilon_0$. This observation gives extra information: in the coming lectures, we will use (an arithmetization of) the Well-Ordering Principle for ordinals less than $\varepsilon_0$ to prove the consistency of Peano Arithmetic. By the Second Incompleteness Theorem, it will then follow that neither (i) nor (ii) is provable in Peano Arithmetic when suitably arithmetized.