MAGIC Set theory

lecture 6

David Asperó

University of East Anglia

15 November 2018
Recall:

We defined \( (V_\alpha : \alpha \in \text{Ord}) \) by recursion on \( \text{Ord} \):

- \( V_0 = \emptyset \)
- \( V_{\alpha + 1} = \mathcal{P}(V_\alpha) \)
- \( V_\delta = \bigcup \{ V_\beta : \beta < \delta \} \) if \( \delta \) is a limit ordinal.

\( (V_\alpha : \alpha \in \text{Ord}) \) is called the cumulative hierarchy.
We saw:

**Proposition**

$V_\alpha$ is transitive for every ordinal $\alpha$.

**Proposition**

For all $\alpha < \beta$, $V_\alpha \subseteq V_\beta$. 
Definition
For every $x \in \bigcup_{\alpha \in \text{Ord}} V_\alpha$, let $\text{rank}(x)$ be the first $\alpha$ such that $x \in V_{\alpha+1}$.

Definition
For every set $x$, the transitive closure of $x$, denoted by $TC(x)$, is $\bigcup \{X_n : n < \omega\}$ where

- $X_0 = x$
- $X_{n+1} = \bigcup X_n$

So $TC(x) = x \cup \bigcup x \cup \bigcup \bigcup x \cup \bigcup \bigcup \bigcup \bigcup x \cup \ldots$

[Exercise: $TC(x)$ is the $\subseteq$–least transitive set $y$ such that $x \subseteq y$. In other words, $TC(x) = \bigcap \{y : y$ transitive, $x \subseteq y\}$.]
Definition

\( V \) denotes the class of all sets; that is, \( V = \{ x : x = x \} \).

Definition

\( \text{WF} = \bigcup \{ V_\alpha : \alpha \in \text{Ord} \} \): The class of all \( x \) such that \( x \in V_\alpha \) for some ordinal \( \alpha \).

Note: \( \text{WF} \) is a transitive class: \( y \in x \in V_\alpha \) implies \( y \in V_\alpha \) since \( V_\alpha \) is transitive.

Proposition

If \( x \subseteq \text{WF} \), then there is some \( \alpha \in \text{Ord} \) such that \( x \subseteq V_\alpha \).

Proof.

Define the function \( F(y) = \min \{ \gamma : y \in V_\gamma \} \). By the assumption on \( x \), \( F \) is a well-defined function there. By Replacement \( \{ \gamma : \exists y \in x : \gamma = F(y) \} \) is a set, and by Union it has a supremum \( \alpha \). Therefore \( x \subseteq V_\alpha \). \( \square \)

Corollary

If \( x \subseteq \text{WF} \), then \( x \in \text{WF} \).
Definition
$V$ denotes the class of all sets; that is, $V = \{x : x = x\}$.

Definition
$WF = \bigcup \{V_\alpha : \alpha \in \text{Ord}\}$: The class of all $x$ such that $x \in V_\alpha$ for some ordinal $\alpha$.

Note: $WF$ is a transitive class: $y \in x \in V_\alpha$ implies $y \in V_\alpha$ since $V_\alpha$ is transitive.

Proposition
If $x \subseteq WF$, then there is some $\alpha \in \text{Ord}$ such that $x \subseteq V_\alpha$.

Proof.
Define the function $F(y) = \min \{\gamma | y \in V_\gamma\}$. By the assumption on $x$, $F$ is a well-defined function there. By Replacement $\{\gamma | \exists y \in x : \gamma = F(y)\}$ is a set, and by Union it has a supremum $\alpha$. Therefore $x \subseteq V_\alpha$. 

Corollary
If $x \subseteq WF$, then $x \in WF$. 
Theorem

(ZF) $V = WF$

Proof: Let $x$ be some set, and let $y$ be $TC(x)$. It is enough to show that $y \subseteq WF$, since that implies $x \subseteq WF$ and therefore $x \in WF$. Let $y^* = \{z \in y \mid z \notin WF\}$. If $y^*$ is not empty, by Foundation it has some $\in$-minimal $a$. By transitivity of $y$, $a \subseteq y$ and so if $z \in a$, $z \in WF$. But this means that $a \subseteq WF$ and therefore $a \in WF$.

Therefore $y^* = \emptyset$, so $y \subseteq WF$ as wanted. $\square$
The picture of the universe provided by $V = WF$ is a very appealing and very natural one (once one has come across it, at least). This picture of the universe of all sets, and the fact that ZF implies $V = WF$, is the main source of *intrinsic* justifications of the ZF axioms.
Inner models and relativization

Let \((M, \in^M)\) be a submodel, or inner model, defined by a formula \(\Theta(x)\); in other words, \(M = \{a : \Theta(a)\}\) and, for all \(a, b \in M\), \(a \in^M b\) if and only if \(a \in b\) (we usually leave out \(\in^M\) and write \(M\) instead of \((M, \in^M)\)).

(Examples: \(V\), \(WF\), \(L\), \(HOD\), ...).

We define the relativization to \(M\) of a formula \(\varphi(\vec{x})\), to be denoted \(\varphi^M(\vec{x})\), in the following manner.

- \((x \in y)^M\) is \(x \in y\).
- \((x = y)^M\) is \(x = y\).
- \((\varphi_0 \lor \varphi_1)^M\) is \(\varphi_0^M \lor \varphi_1^M\).
- \((\neg \varphi)^M\) is \(\neg(\varphi^M)\).
- \(((\forall x)(\varphi(\vec{x}))^M\) is \(\forall x(\Theta(x) \rightarrow \varphi^M(\vec{x}))\). We may also write \((\forall x \in M)\varphi^M(x)\).
Note: Given a formula $\varphi(x_0, \ldots, x_n)$ and $a_0, \ldots, a_n \in M$, $\varphi^M(a_0, \ldots, a_n)$ holds if and only if $M \models \varphi[a_0, \ldots, a_n]$. [Easy to prove by induction on the complexity of $\varphi$.]

Notation: If $(N, E)$ is a structure in the language of set theory, $M$ is an inner model defined by a formula $\Theta(x)$ possibly with parameters (i.e., $M = (N, E \upharpoonright (M \times M))$), where $M = \{a \in N : (N, E) \models \Theta(a)\}$, and we want / need to emphasise that $M$ is the inner model defined by $\Theta(x)$ as defined within $(N, E)$, then we often write $M^N$ instead of $M$. Example: $WF^M$

Note: For every ordinal $\alpha$, $V^WF_\alpha = V_\alpha$ (here $V_\alpha$ refers to the set, definable from the parameter $\alpha$, with the definition that we have seen). (This note is pertinent under $ZF \setminus \{\text{Foundation}\}$, in which context $V = WF$ might not be true.)
Many facts about the universe $V$ are inherited by reasonable submodels. For example:

**Lemma**

*Suppose $M$ is a transitive set or a transitive proper class. Then $M \models \text{Axiom of Extensionality}$.*

**Proof:** Let $a, b \in M$ and suppose $M \models (\forall x)(x \in a \leftrightarrow x \in b)$

(this of course is shorthand for

$$M \models (\forall x)(x \in y \leftrightarrow x \in z)[\tilde{a}]$$

where $\tilde{a}$ is any assignment sending the variable $y$ to $a$ and the variable $z$ to $b$).
This means that \( a \cap M = b \cap M \).

Since \( M \) is transitive (in \( V \)), every member of \( a \) or of \( b \) is a member of \( M \). It follows that \( a \cap M = a \) and \( b \cap M = b \) and therefore \( a = b \). Hence \( M \models a = b \).

In sum, \( M \) thinks that for all \( y, z \), if \( y \) and \( z \) have the same elements, then they are equal. In other words, \( M \models \text{Axiom of Extensionality} \). \( \Box \)
Lemma
Suppose $M$ is a transitive set or a transitive proper class which is closed under unordered pairs (meaning that for all $a, b \in M$, $\{a, b\} \in M$). Then $M \models$ Axiom of Unordered pairs.

Proof.
Let $c = \{a, b\} \in M$. Check, as in the previous proof, that $M \models (\forall x) x \in c \iff x = a \lor x = b.$
Similarly:

**Lemma**

*Suppose* $M$ *is a transitive set or a transitive proper class. Suppose* $\bigcup a \in M$ *for every* $a \in M$. *Then* $M \models$ *Union set Axiom.*

**Lemma**

*Suppose* $M$ *is a transitive set or a transitive proper class. Suppose for every* $a \in M$ *there is some* $b \in M$ *such that* $b = \mathcal{P}(a) \cap M$. *Then* $M \models$ *Power set axiom.*

[Proofs: Exercises.]

**Note:** There are situations in which there are transitive models $M$ of fragments of ZFC, or even of all of ZFC, and some $a \in M$ such that $\mathcal{P}(a)^M$ is strictly included in $\mathcal{P}(a)$ (i.e., there are subsets $b$ of $a$ such that $b \notin M$).
Lemma
Suppose $M$ is a transitive set or a transitive proper class. If $\omega \in M$, then $M \models \text{Infinity}$.

Proof idea: As in the previous proofs. The point is that $M$ recognises $\emptyset$ correctly, recognises correctly that something is an ordinal, and recognises correctly that something is the successor of an ordinal.

We say that the notion of ordinal is absolute with respect to transitive models. It is possible to identify large families of properties that are absolute with respect to transitive models by virtue of their being definable by syntactically ‘simple’ formulas (from the point of view of their quantifiers):
A formula \( \varphi \) is *restricted* if all its quantifiers are restricted. In other words, they appear in subformulas of the form 
\[
(\forall x)(x \in y \implies \psi) \quad \text{and} \quad (\exists x)(x \in y \land \psi).
\]

We generally abbreviate the above by 
\[
(\forall x \in y)\psi \quad \text{and} \quad (\exists x \in y)\psi,
\]
respectively.

The following general fact is not difficult.

**Fact**

*Suppose* \( \varphi(\vec{x}) \) *is a restricted formula and* \( M \) *is transitive. The following are equivalent for every* \( \vec{a} \in M^{<\omega} \).

1. \( \varphi(\vec{a}) \)
2. \( M \models \varphi(\vec{a}) \)
**Note**: The notion of finiteness is also absolute with respect to transitive models but, on the other hand, the notion of countability is highly *non–absolute* with respect to transitive models:

There are transitive models $M$ and $a \in M$ such that

$$M \models a \text{ is uncountable}$$

but there is a bijection $f : \omega \to a$, so $a$ is countable in $V$. The problem of course is that $f$ is not in $M$.

We will soon see that there are transitive models of (fragments of) ZFC such that all their sets are countable in $V$. And even the whole model can be countable in $V$. 
**Note:** The notion of finiteness is also absolute with respect to transitive models but, on the other hand, the notion of countability is highly *non–absolute* with respect to transitive models:

There are transitive models $M$ and $a \in M$ such that

$$M \models a \text{ is uncountable}$$

but there is a bijection $f : \omega \rightarrow a$, so $a$ is countable in $V$. The problem of course is that $f$ is not in $M$.

We will soon see that there are transitive models of (fragments of) ZFC such that all their sets are countable in $V$. And even the whole model can be countable in $V$. 
The notion of choice function is also absolute with respect to transitive models: If $M$ is transitive, $a \in M$ consists of nonempty sets and $f \in M$, then $f$ is a choice function for $a$ if and only if $M \models "f \text{ is a choice function for } a"$. Hence:

**Lemma**

Let $M$ be a transitive set or a transitive proper class. Suppose for every $a \in M$ consisting of nonempty sets there is a choice function $f$ for $a$, $f \in M$. Then $M \models \text{AC}$. 
Lemma
Let $M$ be a transitive set or a transitive proper class. Suppose $b \in M$ whenever $a \in M$ and $b \subseteq a$ is definable over $M$, possibly from parameters (in other words, $b = \{ c : c \in a, M \models \varphi(c, \bar{p}) \}$ for some parameters $\bar{p} \in M$). Then $M \models \text{Separation}$.

Lemma
Let $M$ be a transitive set or a transitive proper class. Suppose $F[a] \in M$ whenever $a \in M$, and $F$ is a class–function over $M$ (in other words, if $F$ is definable by a formula $\varphi(x, y, \bar{z})$ which, over $M$ is functional, $\bar{p} \in M$, and $a \in M$, then $\{ c : (\exists b \in a)M \models \varphi(b, c, \bar{p}) \} \in M$). Then $M \models \text{Replacement}$.

These last three lemmas can be seen to follow from the general Fact on absoluteness of facts expressible with a restricted formula.
Our first relative consistency proof:
Con(ZF \{\text{Foundation}\}) implies Con(ZF)

Theorem
Let $M \models ZF \setminus \{\text{Foundation}\}$. Then $M \models \sigma^{WF^M}$ for every $\sigma \in ZF$. Hence, $WF^M \models ZF$.

Proof: By the previous lemmas and the construction of $(V_\alpha : \alpha \in \text{Ord})$, $WF^M \models \sigma$ for every axiom $\sigma$ of $ZF \setminus \{\text{Foundation}\}$.

[Go through them one by one and check that $WF$ is closed under the relevant operation, then apply the relevant lemma.]
To see that $M \models \text{Foundation}^{WF}$ holds, let us work in $M$: Let $X \in WF$, and let $b \in X$ such that $\text{rank}(b) = \min\{\text{rank}(x) : x \in X\}$.

By definition of rank and definition of $WF = \bigcup_{\alpha \in \text{Ord}} V_\alpha$ there is then no $y \in X$ such that $y \in b$. Since $WF$ is transitive, this means that there is no $y \in WF \cap X$ such that $WF \models y \in b$. Hence, $WF$ thinks that the restriction of $\in$ to $X$ is well–founded. Since this is true for all $X \in WF$,

$$WF \models \text{Foundation}$$
Corollary
If $ZF \setminus \{\text{Foundation}\}$ is consistent, then $ZF$ is consistent.

Proof.
Suppose $ZF \setminus \{\text{Foundation}\}$. By the completeness theorem we may find a model $M \models ZF \setminus \{\text{Foundation}\}$. Let $M' = WF^M$. By the theorem $M' \models ZF$. Hence, $ZF$ has a model and therefore it is consistent. □

Remark
By exactly the same argument, if $M \models ZFC \setminus \{\text{Foundation}\}$, then $M \models \sigma^{WF^M}$ for every $\sigma \in ZFC$. Hence, $\text{Con}(ZFC \setminus \{\text{Foundation}\})$ implies $\text{Con}(ZFC)$. 
Similar relative consistency results:

One can define “the constructible universe” \( L \):

- \( L_0 = \emptyset \)
- \( L_{\alpha+1} = \text{Def}(L_\alpha) \), where \( \text{Def}(L_\alpha) \) is the set of all subsets of \( L_\alpha \) definable over \( L_\alpha \) possibly with parameters, i.e., the collection of all sets of the form
  \[
  \{ b \in L_\alpha : L_\alpha \models \varphi(b, a_0, \ldots, a_{n-1}) \}
  \]
  for some formula \( \varphi(x, \bar{x}) \) and \( a_0, \ldots, a_{n-1} \in L_\alpha \).
- \( L_\delta = \bigcup_{\alpha < \delta} L_\alpha \) if \( \delta > 0 \) is a limit ordinal.

\( L \) is then \( \bigcup_{\alpha \in \text{Ord}} L_\alpha \).

This construction is due to Gödel. He proved that if we do this construction in \( \text{ZF} \), then \( L \models \text{ZF} \) but also \( L \models \text{AC} \) and \( L \models \text{CH} \).

We will take a closer look at these results soon.
Similar relative consistency results:

One can define “the constructible universe” $L$:

- $L_0 = \emptyset$
- $L_{\alpha+1} = \text{Def}(L_\alpha)$, where $\text{Def}(L_\alpha)$ is the set of all subsets of $L_\alpha$ definable over $L_\alpha$ possibly with parameters, i.e., the collection of all sets of the form
  \[ \{ b \in L_\alpha : L_\alpha \models \varphi(b, a_0, \ldots, a_{n-1}) \} \]
  for some formula $\varphi(x, \bar{x})$ and $a_0, \ldots, a_{n-1} \in L_\alpha$.
- $L_\delta = \bigcup_{\alpha < \delta} L_\alpha$ if $\delta > 0$ is a limit ordinal.

$L$ is then $\bigcup_{\alpha \in \text{Ord}} L_\alpha$.

This construction is due to Gödel. He proved that if we do this construction in ZF, then $L \models ZF$ but also $L \models AC$ and $L \models CH$.

We will take a closer look at these results soon.
The above results imply that if $\text{ZF}$ is consistent, then $\text{ZFC}$ is also consistent, and in fact also $\text{ZFC}+\text{CH}$. Linking this to the implication we have seen we thus have that if $\text{ZF} \setminus \{\text{Foundation}\}$ is consistent, then so is $\text{ZFC}+\text{CH}$.

These relative consistency proofs proceed by building suitable inner models.\(^1\)

Most relative consistency proofs proceed, on the other hand, by building suitable outer models of some given ground model. The construction of these outer models is done with the forcing method. This is an extremely powerful method in set theory. Unfortunately, we will not be able to say much about it here.

\(^1\)What set–theorists understand by “inner models” are usually much more complicated than $\text{WF}$ or $\text{L}$. However, the construction of $\text{L}$ is in fact the paradigm for most of these more complicated constructions.
The above results imply that if $ZF$ is consistent, then $ZFC$ is also consistent, and in fact also $ZFC+CH$. Linking this to the implication we have seen we thus have that if $ZF \setminus \{ \text{Foundation} \}$ is consistent, then so is $ZFC+CH$.

These relative consistency proofs proceed by building suitable inner models.$^1$

Most relative consistency proofs proceed, on the other hand, by building suitable outer models of some given ground model. The construction of these outer models is done with the forcing method. This is an extremely powerful method in set theory. Unfortunately, we will not be able to say much about it here.

$^1$What set–theorists understand by “inner models” are usually much more complicated than $WF$ or $L$. However, the construction of $L$ is in fact the paradigm for most of these more complicated constructions.
Given formulas $\psi_0, \psi_1$, we say that $\psi_0$ is a subformula of $\psi_1$ if and only if

- $\psi_0 = \psi_1$, or
- there is a subformula $\psi$ of $\psi_1$ such that $\psi_0$ is a subformula of $\psi$, or
- either $\psi_1 = \neg \psi_0$ or $\psi_1 = \psi_0 \lor \psi$ for some $\psi$ or $\psi_1 = \psi \land \psi_0$ for some $\psi$ or $\psi_1 = (\exists x)(\varphi_0)$ for some variable $x$. 
Theorem

(ZF) (Lévy–Montague Reflection) Let $\varphi(x_0, \ldots, x_{n-1})$ be a formula in the language of set theory. There is a proper class $C_\varphi$ of ordinals such that:

1. for all $\alpha \in C_\varphi$ and all $a_0, \ldots, a_{n-1} \in V_\alpha$, 
   
   $$V_\alpha \models \varphi(a_0, \ldots, a_{n-1})$$

   if and only if $\varphi(a_0, \ldots, a_{n-1})$ is true (we may also write $V \models \varphi(a_0, \ldots, a_{n-1})$). We write $V_\alpha \prec_\varphi V$.

2. For every $\beta \in \text{Ord}$ there is some $\alpha \in C_\varphi$ such that $\beta < \alpha$ ($C_\varphi$ is unbounded).

3. For every limit ordinal $\beta$, if $\sup(C_\varphi \cap \beta) = \beta$, then $\beta \in C_\varphi$ ($C_\varphi$ is closed).
Proof: Let $(\varphi_i : i < n)$ list all subformulas of $\varphi$ in such a way that if $\varphi_i$ is a subformula of $\varphi_{i'}$, then $i < i'$.

We build a $\subseteq$–decreasing sequence $C_{\varphi_i}$ of closed and unbounded proper classes of ordinals (for $i < n$) as follows.

Let $i < n$ and suppose $C_{\varphi_{i-1}}$ defined if $i > 0$. If $i = 0$ (so $\varphi_i$ is an atomic formula) then $C_{\varphi_i} = \text{Ord}$.

If $i > 0$ but $\varphi_i \neq (\exists x)(\varphi_{i'})$ for any $i' < i$ and any variable $x$, then $C_{\varphi_i} = C_{\varphi_{i-1}}$. 
Finally, suppose $i > 0$ and $\varphi_i = (\exists x)\varphi_{i'}(x, x_0, \ldots, x_{m-1})$ for some $i' < i$ and some variable $x$. We define $C_{\varphi_i} = \{ \alpha_\xi : \xi \in \text{Ord} \}$, where $(\alpha_\xi)_{\xi \in \text{Ord}}$ is a strictly increasing and continuous – i.e., at limit stages $\xi$ we let $\alpha_\xi = \sup_{\xi' < \xi} \alpha_{\xi'}$ – sequence of ordinals defined as follows:

Given an ordinal $\xi$, if $\alpha_\xi$ has been defined let $\beta_0^\xi = \alpha_\xi$ and let $\beta_1^\xi$ be the least $\beta > \alpha_\xi$, $\beta \in C_{\varphi_{i-1}}$, such that for all $a_0, \ldots, a_{m-1} \in V_{\alpha_\xi}$, if there is some $b$ such that $\varphi_{i'}(b, a_0, \ldots, a_{m-1})$, then there is some $b \in V_\beta$ such that $\varphi_{i'}(b, a_0, \ldots, a_{m-1})$.

In general, if $\beta_n^\xi$ has been defined, let $\beta_{n+1}^\xi$ be the least $\beta > \beta_n^\xi$, $\beta \in C_{\varphi_{i-1}}$, such that for all $a_0, \ldots, a_{m-1} \in V_{\beta_n^\xi}$, if there is some $b$ such that $\varphi_{i'}(b, a_0, \ldots, a_{m-1})$, then there is some $b \in V_\beta$ such that $\varphi_{i'}(b, a_0, \ldots, a_{m-1})$. 
Let $\beta = \sup_{n<\omega} \beta_n^\xi$, and note that $\beta \in C_{\varphi_{i-1}}$ since $C_{\varphi_{i-1}}$ is closed. Let $\alpha_{\xi+1} = \beta$.

The construction of $\alpha_0$ of course is as above, starting from $0$ instead of $\alpha_\xi$.

Since $V_\alpha \preceq_{\varphi_i} V$ for all $\alpha \in C_{\varphi_{i-1}}$, it follows that $V_\alpha \preceq_{\varphi_i} V$ for all $\alpha \in C_{\varphi_i}$:

Suppose $\alpha \in C_{\varphi_i}$ and $a_0, \ldots, a_{m-1} \in V_\alpha$. If there is some $b \in V_\alpha$ such that $V_\alpha \models \varphi_{i'}(b, a_0, \ldots, a_{m-1})$, then of course $V \models \varphi_{i'}(b, a_0, \ldots, a_{m-1})$ since $V_\alpha \preceq_{\varphi_i} V$. And if $V \models (\exists x)\varphi_{i'}(x, a_0, \ldots, a_{m-1})$, then by construction of $C_{\varphi_i}$ there is some $b \in V_\alpha$ such that $V \models \varphi_{i'}(b, a_0, \ldots, a_{m-1})$. But then $V_\alpha \models \varphi_{i'}(b, a_0, \ldots, a_{m-1})$ since $V_\alpha \preceq_{\varphi_i} V$. □
Corollary

Suppose \( \varphi_0, \ldots, \varphi_n \) are finitely many formulas in the language of set theory. Then there is a proper class \( C \) of ordinals such that:

1. For every \( \alpha \in C \) and every \( j \leq n \), \( V_\alpha \preceq \varphi_j V \).
2. For every \( \beta \in \text{Ord} \) there is some \( \alpha \in C \) such that \( \beta < \alpha \).
3. For every limit ordinal \( \beta \), if \( \sup(C \cap \beta) = \beta \), then \( \beta \in C \).

Proof.

Simply note that \( C_{\varphi_0} \cap \ldots \cap C_{\varphi_n} \) is a closed and unbounded class of ordinals. It is clearly closed. To see that it is unbounded, given \( \beta \in \text{Ord} \) let \( (\alpha_i)_{i<\omega} \) be such that \( \alpha_0 > \beta \) and such that for all \( i < \omega \) and all \( j \leq n \) there is some \( \gamma \in C_{\varphi_j} \), \( \alpha_i < \gamma < \alpha_{i+1} \). Then, for every \( j \leq n \), \( \sup_{i<\omega} \alpha_i \) is a limit of ordinals in \( C_{\varphi_j} \) and therefore in \( C_{\varphi_j} \). \(\square\)
**Moral**: The universe cannot be characterised in a first order way:
Every first order fact about it is true about some initial segment of it.

Does reflection hold for infinite sets of sentences in the language of set theory? For all of them?

**Warning**: ‘Reflection for all sentences’ cannot be expressed in the first order language of set theory.

(Tarski’s undefinability of truth: There is no formula $\varphi(x)$ such that for every sentence $\sigma$, $(\mathbb{N}, +, \cdot) \models \sigma$ if and only if $(\mathbb{N}, +, \cdot) \models \varphi(\#\sigma)$, where $\#\sigma$ is a code for $\sigma$.)
Moral: The universe cannot be characterised in a first order way:
Every first order fact about it is true about some initial segment of it.

Does reflection hold for infinite sets of sentences in the language of set theory? For all of them?

Warning: ‘Reflection for all sentences’ cannot be expressed in the first order language of set theory.

(Tarski’s undefinability of truth: There is no formula $\varphi(x)$ such that for every sentence $\sigma$, $(\mathbb{N}, +, \cdot) \models \sigma$ if and only if $(\mathbb{N}, +, \cdot) \models \varphi(\#\sigma)$, where $\#\sigma$ is a code for $\sigma$.)
Remark
Most large cardinal notions in the lower part of the large cardinal hierarchy can be motivated via considerations of generalisations / extensions of reflection phenomena of the same flavour as the Lévy–Montague reflection scheme.²

²By, for example, considering ‘reflection of higher order statements’, etc.