Nonstandard Analysis and Combinatorial Number Theory

Renling Jin

College of Charleston, SC

Day One: Foundation of Nonstandard Analysis

2023 Fudan Logic Summer School
Shanghai, China, August 7, 2023

OUTLINE:

1. Introduction
2. First-order Logic and Ultrapower of Real Field
3. Ultrapower of Superstructure

OUTLINE:

1. Introduction
2. First-order Logic and Ultrapower of Real Field
3. Ultrapower of Superstructure

OUTLINE:

1. Introduction
2. First-order Logic and Ultrapower of Real Field
3. Ultrapower of Superstructure

OUTLINE:

1. Introduction
2. First-order Logic and Ultrapower of Real Field
3. Ultrapower of Superstructure

Blah Blah Blah ......

Therefore, nonstandard analysis is definitely important and useful.

Blah Blah Blah ......

Therefore, nonstandard analysis is definitely important and useful.

We will limit the breadth of our non-traditional introduction of the first-order logic. We will touch only the part enough for the purpose of this course.

Logical symbols:

- connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$,
- quantifiers: $\forall$, $\exists$,
- equality symbol: $=$,
- variables: x, y, z, $\ldots$.

Non-logical symbols:

- $\mathscr{L} = \{+, \cdot, \leq, 0, 1, P\}_{P \in \mathcal{P}}$ for ordered field or
- $\mathscr{L} = \{\in\}$ for superstructure where $\in$ is a binary relation symbol.

We will limit the breadth of our non-traditional introduction of the first-order logic. We will touch only the part enough for the purpose of this course.

Logical symbols:

- connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$,
- quantifiers: $\forall$, $\exists$,
- equality symbol: $=$,
- variables: x, y, z, $\ldots$.

Non-logical symbols:

- $\mathscr{L} = \{+, \cdot, \leq, 0, 1, P\}_{P \in \mathcal{P}}$ for ordered field or
- $\mathscr{L} = \{\in\}$ for superstructure where $\in$ is a binary relation symbol.

We will limit the breadth of our non-traditional introduction of the first-order logic. We will touch only the part enough for the purpose of this course.

Logical symbols:

- connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$,
- quantifiers: $\forall$, $\exists$,
- equality symbol: $=$,
- variables: x, y, z, ...

Non-logical symbols:

- $\mathscr{L} = \{+, \cdot, \leq, 0, 1, P\}_{P \in \mathcal{P}}$ for ordered field or
- $\mathscr{L} = \{\in\}$ for superstructure where $\in$ is a binary relation symbol.

We will limit the breadth of our non-traditional introduction of the first-order logic. We will touch only the part enough for the purpose of this course.

Logical symbols:

- connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$,
- quantifiers: $\forall$, $\exists$,
- equality symbol: $=$,
- variables: x, y, z, $\ldots$.

Non-logical symbols:

- $\mathscr{L} = \{+, \cdot, \leq, 0, 1, P\}_{P \in \mathcal{P}}$ for ordered field or
- $\mathscr{L} = \{\in\}$ for superstructure where $\in$ is a binary relation symbol.

We will limit the breadth of our non-traditional introduction of the first-order logic. We will touch only the part enough for the purpose of this course.

Logical symbols:

- connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$,
- quantifiers: $\forall$, $\exists$,
- equality symbol: $=$,
- variables: x, y, z, $\ldots$.

Non-logical symbols:

- $\mathscr{L} = \{+, \cdot, \leq, 0, 1, P\}_{P \in \mathcal{P}}$ for ordered field or
- $\mathscr{L} = \{\in\}$ for superstructure where $\in$ is a binary relation symbol.

We will limit the breadth of our non-traditional introduction of the first-order logic. We will touch only the part enough for the purpose of this course.

Logical symbols:

- connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$,
- quantifiers: $\forall$, $\exists$,
- equality symbol: $=$,
- variables: x, y, z, ...

Non-logical symbols:

- $\mathscr{L} = \{+, \cdot, \leq, 0, 1, P\}_{P \in \mathcal{P}}$ for ordered field or
- $\mathscr{L} = \{\in\}$ for superstructure where $\in$ is a binary relation symbol.

We will limit the breadth of our non-traditional introduction of the first-order logic. We will touch only the part enough for the purpose of this course.

Logical symbols:

- connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$,
- quantifiers: $\forall$, $\exists$,
- equality symbol: $=$,
- variables: x, y, z, $\ldots$.

Non-logical symbols:

- $\mathscr{L} = \{+, \cdot, \leq, 0, 1, P\}_{P \in \mathcal{P}}$ for ordered field or
- $\mathscr{L} = \{\in\}$ for superstructure where $\in$ is a binary relation symbol.

We will limit the breadth of our non-traditional introduction of the first-order logic. We will touch only the part enough for the purpose of this course.

Logical symbols:

- connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$,
- quantifiers: $\forall$, $\exists$,
- equality symbol: $=$,
- variables: x, y, z, $\ldots$.

Non-logical symbols:

- $\mathscr{L} = \{+, \cdot, \leq, 0, 1, P\}_{P \in \mathcal{P}}$ for ordered field or
- $\mathscr{L} = \{\in\}$ for superstructure where $\in$ is a binary relation symbol.

We will limit the breadth of our non-traditional introduction of the first-order logic. We will touch only the part enough for the purpose of this course.

Logical symbols:

- connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$,
- quantifiers: $\forall$, $\exists$,
- equality symbol: $=$,
- variables: x, y, z, $\ldots$.

Non-logical symbols:

- $\mathscr{L} = \{+, \cdot, \leq, 0, 1, P\}_{P \in \mathcal{P}}$ for ordered field or
- $\mathscr{L} = \{\in\}$ for superstructure where $\in$ is a binary relation symbol.

In the language of ordered field, the symbols $+$, $\cdot$, $\leq$, $0, 1$, can all be viewed as relation symbols. Hence, Just say $\mathscr{L} := \{P\}_{P \in \mathcal{P}}$ is enough. However, we list $+$, $\cdot$, $\leq$, $0, 1$, separately from $\mathcal{P}$ just for clarity.

For notational simplicity all non-logical symbols considered are relational symbols (note that an n-variable function can be identified with the graph of the function which is an $(n+1)$-dimensional relation and a constant symbol is a 0-dimensional relation symbol).

We do not distinguish each symbol from its intended interpretation. For example, $+$ represents a three dimensional relation symbol in the language of ordered field as well as the actual addition in an ordered field.

In the language of ordered field, the symbols $+$, $\cdot$, $\leq$, $0, 1$, can all be viewed as relation symbols. Hence, Just say $\mathscr{L} := \{P\}_{P \in \mathcal{P}}$ is enough. However, we list $+$, $\cdot$, $\leq$, $0, 1$, separately from $\mathcal{P}$ just for clarity.

For notational simplicity all non-logical symbols considered are relational symbols (note that an n-variable function can be identified with the graph of the function which is an $(n+1)$-dimensional relation and a constant symbol is a 0-dimensional relation symbol).

We do not distinguish each symbol from its intended interpretation. For example, $+$ represents a three dimensional relation symbol in the language of ordered field as well as the actual addition in an ordered field.

In the language of ordered field, the symbols $+$, $\cdot$, $\leq$, $0, 1$, can all be viewed as relation symbols. Hence, Just say $\mathscr{L} := \{P\}_{P\in\mathcal{P}}$ is enough. However, we list $+$, $\cdot$, $\leq$, $0, 1$, separately from $\mathcal{P}$ just for clarity.

For notational simplicity all non-logical symbols considered are relational symbols (note that an n-variable function can be identified with the graph of the function which is an $(n+1)$-dimensional relation and a constant symbol is a 0-dimensional relation symbol).

We do not distinguish each symbol from its intended interpretation. For example, $+$ represents a three dimensional relation symbol in the language of ordered field as well as the actual addition in an ordered field.

Definition (1.1)

An $\mathscr{L}$-formula can be formed inductively by the following two steps ($\mathscr{L}$ will be omitted later on).

1. *Basic step: The atomic formulas are those in the form of $P(\overline{x}, \overline{c})$ where $P \in \mathscr{L} \cup \{=\}$ is a relation symbol with arity m and $\overline{x}$ represents the k-tuple $(x_1, x_2, \ldots, x_k)$ of variables and $\overline{c}$ represents the $m-k$ tuple of constant symbols;*
2. *Inductive step: If φ and ψ are formulas, so are $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \rightarrow \psi$, $\varphi \leftrightarrow \psi$, $\forall x\, \varphi$, and $\exists x\, \varphi$.*

By the complexity of a formula, we mean the number of steps in Definition 1.1 used to form the formula.

Definition (1.1)

An $\mathscr{L}$-formula can be formed inductively by the following two steps ($\mathscr{L}$ will be omitted later on).

1. *Basic step: The atomic formulas are those in the form of $P(\overline{x}, \overline{c})$ where $P \in \mathscr{L} \cup \{=\}$ is a relation symbol with arity m and $\overline{x}$ represents the k-tuple $(x_1, x_2, \ldots, x_k)$ of variables and $\overline{c}$ represents the $m - k$ tuple of constant symbols;*
2. *Inductive step: If φ and ψ are formulas, so are $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \rightarrow \psi$, $\varphi \leftrightarrow \psi$, $\forall x\, \varphi$, and $\exists x\, \varphi$.*

By the complexity of a formula, we mean the number of steps in Definition 1.1 used to form the formula.

Definition (1.1)

An $\mathscr{L}$-formula can be formed inductively by the following two steps ($\mathscr{L}$ will be omitted later on).

1. *Basic step: The atomic formulas are those in the form of $P(\overline{x}, \overline{c})$ where $P \in \mathscr{L} \cup \{=\}$ is a relation symbol with arity m and $\overline{x}$ represents the k-tuple $(x_1, x_2, \ldots, x_k)$ of variables and $\overline{c}$ represents the $m - k$ tuple of constant symbols;*
2. *Inductive step: If φ and ψ are formulas, so are $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \rightarrow \psi$, $\varphi \leftrightarrow \psi$, $\forall x\, \varphi$, and $\exists x\, \varphi$.*

By the complexity of a formula, we mean the number of steps in Definition 1.1 used to form the formula.

Definition (1.1)

An $\mathscr{L}$-formula can be formed inductively by the following two steps ($\mathscr{L}$ will be omitted later on).

1. *Basic step: The atomic formulas are those in the form of $P(\overline{x}, \overline{c})$ where $P \in \mathscr{L} \cup \{=\}$ is a relation symbol with arity m and $\overline{x}$ represents the k-tuple $(x_1, x_2, \ldots, x_k)$ of variables and $\overline{c}$ represents the $m-k$ tuple of constant symbols;*
2. *Inductive step: If φ and ψ are formulas, so are $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \to \psi$, $\varphi \leftrightarrow \psi$, $\forall x\, \varphi$, and $\exists x\, \varphi$.*

By the complexity of a formula, we mean the number of steps in Definition 1.1 used to form the formula.

The sub-formula φ is called the scope of the quantifier $\forall$ or $\exists$ in the formula $\forall x\, \varphi$ or $\exists x\, \varphi$, respectively. The variable x in $\forall x\, \varphi$ or $\exists x\, \varphi$ is called bounded. An occurrence of a variable x is called bounded in an formula φ if it is bounded in a sub-formula $\forall x\, \psi$ or $\exists x\, \psi$ of φ. An occurrence of a variable x is called free in φ if it is not bounded.

We write $\overline{x}$ for a tuple of variables, and write $\varphi(\overline{x})$ to indicate implicitly that all free variables in φ are among the variables in $\overline{x}$.

Definition (1.2)

A model $\mathcal{M} := (M; P^{\mathcal{M}})_{P \in \mathscr{L}}$ contains a nonempty base set M together with the interpretation $P^{\mathcal{M}} \subseteq M^m$ of each relation symbol $P \in \mathscr{L}$ with arity m. We sometimes write $\mathcal{M}$ for a model as well as its base set.

The sub-formula φ is called the scope of the quantifier $\forall$ or $\exists$ in the formula $\forall x\, \varphi$ or $\exists x\, \varphi$, respectively. The variable x in $\forall x\, \varphi$ or $\exists x\, \varphi$ is called bounded. An occurrence of a variable x is called bounded in an formula φ if it is bounded in a sub-formula $\forall x\, \psi$ or $\exists x\, \psi$ of φ. An occurrence of a variable x is called free in φ if it is not bounded.

We write $\overline{x}$ for a tuple of variables, and write $\varphi(\overline{x})$ to indicate implicitly that all free variables in φ are among the variables in $\overline{x}$.

Definition (1.2)

A model $\mathcal{M} := (M; P^{\mathcal{M}})_{P\in\mathscr{L}}$ contains a nonempty base set M together with the interpretation $P^{\mathcal{M}} \subseteq M^m$ of each relation symbol $P \in \mathscr{L}$ with arity m. We sometimes write $\mathcal{M}$ for a model as well as its base set.

The sub-formula φ is called the scope of the quantifier $\forall$ or $\exists$ in the formula $\forall x\, \varphi$ or $\exists x\, \varphi$, respectively. The variable x in $\forall x\, \varphi$ or $\exists x\, \varphi$ is called bounded. An occurrence of a variable x is called bounded in an formula φ if it is bounded in a sub-formula $\forall x\, \psi$ or $\exists x\, \psi$ of φ. An occurrence of a variable x is called free in φ if it is not bounded.

We write $\overline{x}$ for a tuple of variables, and write $\varphi(\overline{x})$ to indicate implicitly that all free variables in φ are among the variables in $\overline{x}$.

Definition (1.2)

A model $\mathcal{M} := (M; P^{\mathcal{M}})_{P \in \mathscr{L}}$ contains a nonempty base set M together with the interpretation $P^{\mathcal{M}} \subseteq M^m$ of each relation symbol $P \in \mathscr{L}$ with arity m. We sometimes write $\mathcal{M}$ for a model as well as its base set.

Definition (1.3)

For each formula $\varphi(\overline{x})$ and a tuple $\overline{a}$ of elements in a model $\mathcal{M}$, define $\mathcal{M} \models \varphi(\overline{a})$, i.e., $\varphi(\overline{a})$ is true in $\mathcal{M}$, inductively on the complexity of the formula:

(1) *φ is an atomic formula $P(\overline{x}, \overline{c})$: $\mathcal{M} \models P(\overline{a}, \overline{c})$ iff $(\overline{a}, \overline{c^{\mathcal{M}}}) \in P^{\mathcal{M}}$;*

(2) *$\mathcal{M} \models \neg\varphi$ iff $\mathcal{M} \not\models \varphi$, i.e., it's not true that $\mathcal{M} \models \varphi$ (so "$\neg$" means "not");*

(3) *$\mathcal{M} \models \varphi \wedge \psi$ iff $\mathcal{M} \models \varphi$ and $\mathcal{M} \models \psi$ (so "$\wedge$" means "and");*

(4) *$\mathcal{M} \models \varphi \vee \psi$ iff $\mathcal{M} \models \varphi$ or $\mathcal{M} \models \psi$ (so "$\vee$" means "or");*

Definition (1.3)

For each formula $\varphi(\overline{x})$ and a tuple $\overline{a}$ of elements in a model $\mathcal{M}$, define $\mathcal{M} \models \varphi(\overline{a})$, i.e., $\varphi(\overline{a})$ is true in $\mathcal{M}$, inductively on the complexity of the formula:

(1) *φ is an atomic formula $P(\overline{x}, \overline{c})$: $\mathcal{M} \models P(\overline{a}, \overline{c})$ iff $(\overline{a}, \overline{c^{\mathcal{M}}}) \in P^{\mathcal{M}}$;*

(2) *$\mathcal{M} \models \neg\varphi$ iff $\mathcal{M} \not\models \varphi$, i.e., it's not true that $\mathcal{M} \models \varphi$ (so "$\neg$" means "not");*

(3) *$\mathcal{M} \models \varphi \wedge \psi$ iff $\mathcal{M} \models \varphi$ and $\mathcal{M} \models \psi$ (so "$\wedge$" means "and");*

(4) *$\mathcal{M} \models \varphi \vee \psi$ iff $\mathcal{M} \models \varphi$ or $\mathcal{M} \models \psi$ (so "$\vee$" means "or");*

Definition (1.3)

For each formula $\varphi(\overline{x})$ and a tuple $\overline{a}$ of elements in a model $\mathcal{M}$, define $\mathcal{M} \models \varphi(\overline{a})$, i.e., $\varphi(\overline{a})$ is true in $\mathcal{M}$, inductively on the complexity of the formula:

(1) *φ is an atomic formula $P(\overline{x}, \overline{c})$: $\mathcal{M} \models P(\overline{a}, \overline{c})$ iff $(\overline{a}, \overline{c^{\mathcal{M}}}) \in P^{\mathcal{M}}$;*

(2) *$\mathcal{M} \models \neg\varphi$ iff $\mathcal{M} \not\models \varphi$, i.e., it's not true that $\mathcal{M} \models \varphi$ (so "$\neg$" means "not");*

(3) *$\mathcal{M} \models \varphi \wedge \psi$ iff $\mathcal{M} \models \varphi$ and $\mathcal{M} \models \psi$ (so "$\wedge$" means "and");*

(4) *$\mathcal{M} \models \varphi \vee \psi$ iff $\mathcal{M} \models \varphi$ or $\mathcal{M} \models \psi$ (so "$\vee$" means "or");*

Definition (1.3)

For each formula $\varphi(\overline{x})$ and a tuple $\overline{a}$ of elements in a model $\mathcal{M}$, define $\mathcal{M} \models \varphi(\overline{a})$, i.e., $\varphi(\overline{a})$ is true in $\mathcal{M}$, inductively on the complexity of the formula:

(1) *φ is an atomic formula $P(\overline{x}, \overline{c})$: $\mathcal{M} \models P(\overline{a}, \overline{c})$ iff $(\overline{a}, \overline{c^{\mathcal{M}}}) \in P^{\mathcal{M}}$;*

(2) *$\mathcal{M} \models \neg\varphi$ iff $\mathcal{M} \not\models \varphi$, i.e., it's not true that $\mathcal{M} \models \varphi$ (so "$\neg$" means "not");*

(3) *$\mathcal{M} \models \varphi \wedge \psi$ iff $\mathcal{M} \models \varphi$ and $\mathcal{M} \models \psi$ (so "$\wedge$" means "and");*

(4) *$\mathcal{M} \models \varphi \vee \psi$ iff $\mathcal{M} \models \varphi$ or $\mathcal{M} \models \psi$ (so "$\vee$" means "or");*

Definition (1.3)

For each formula $\varphi(\overline{x})$ *and a tuple* $\overline{a}$ *of elements in a model* $\mathcal{M}$*, define* $\mathcal{M} \models \varphi(\overline{a})$*, i.e.,* $\varphi(\overline{a})$ *is true in* $\mathcal{M}$*, inductively on the complexity of the formula:*

(1) φ *is an atomic formula* $P(\overline{x}, \overline{c})$: $\mathcal{M} \models P(\overline{a}, \overline{c})$ *iff* $(\overline{a}, \overline{c^{\mathcal{M}}}) \in P^{\mathcal{M}}$;

(2) $\mathcal{M} \models \neg\varphi$ *iff* $\mathcal{M} \not\models \varphi$*, i.e., it's not true that* $\mathcal{M} \models \varphi$ *(so "*$\neg$*" means "not");*

(3) $\mathcal{M} \models \varphi \wedge \psi$ *iff* $\mathcal{M} \models \varphi$ *and* $\mathcal{M} \models \psi$ *(so "*$\wedge$*" means "and");*

(4) $\mathcal{M} \models \varphi \vee \psi$ *iff* $\mathcal{M} \models \varphi$ *or* $\mathcal{M} \models \psi$ *(so "*$\vee$*" means "or");*

Definition (1.3)

(5) $\mathcal{M} \models \varphi \to \psi$ *iff* $\mathcal{M} \models \varphi$ *implies* $\mathcal{M} \models \psi$ *(so "$\to$" means "imply");*

(6) $\mathcal{M} \models \varphi \leftrightarrow \psi$ *iff* $\mathcal{M} \models (\varphi \to \psi) \wedge (\psi \to \varphi)$ *(so "$\leftrightarrow$" means "equivalent to");*

(7) $\mathcal{M} \models \forall x\, \varphi(x, \bar{a})$ *iff* $\mathcal{M} \models \varphi(b, \bar{a})$ *for every* $b \in \mathcal{M}$ *(so "$\forall$" means "for every");*

(8) $\mathcal{M} \models \exists x\, \varphi(x, \bar{a})$ *iff* $\mathcal{M} \models \varphi(b, \bar{a})$ *for some* $b \in \mathcal{M}$ *(so "$\exists$" means "for some").*

Note that by (1), (7), and (8) the intended value in $\mathcal{M}$ for a variable x is always an element of $\mathcal{M}$. This is the reason why we call the logic system above the first-order logic.

Definition (1.3)

(5) $\mathcal{M} \models \varphi \rightarrow \psi$ *iff* $\mathcal{M} \models \varphi$ *implies* $\mathcal{M} \models \psi$ *(so "$\rightarrow$" means "imply");*

(6) $\mathcal{M} \models \varphi \leftrightarrow \psi$ *iff* $\mathcal{M} \models (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$ *(so "$\leftrightarrow$" means "equivalent to");*

(7) $\mathcal{M} \models \forall x\, \varphi(x, \bar{a})$ *iff* $\mathcal{M} \models \varphi(b, \bar{a})$ *for every* $b \in \mathcal{M}$ *(so "$\forall$" means "for every");*

(8) $\mathcal{M} \models \exists x\, \varphi(x, \bar{a})$ *iff* $\mathcal{M} \models \varphi(b, \bar{a})$ *for some* $b \in \mathcal{M}$ *(so "$\exists$" means "for some").*

Note that by (1), (7), and (8) the intended value in $\mathcal{M}$ for a variable x is always an element of $\mathcal{M}$. This is the reason why we call the logic system above the first-order logic.

Definition (1.3)

(5) $\mathcal{M} \models \varphi \to \psi$ *iff* $\mathcal{M} \models \varphi$ *implies* $\mathcal{M} \models \psi$ *(so "$\to$" means "imply");*

(6) $\mathcal{M} \models \varphi \leftrightarrow \psi$ *iff* $\mathcal{M} \models (\varphi \to \psi) \wedge (\psi \to \varphi)$ *(so "$\leftrightarrow$" means "equivalent to");*

(7) $\mathcal{M} \models \forall x\, \varphi(x, \bar{a})$ *iff* $\mathcal{M} \models \varphi(b, \bar{a})$ *for every* $b \in \mathcal{M}$ *(so "$\forall$" means "for every");*

(8) $\mathcal{M} \models \exists x\, \varphi(x, \bar{a})$ *iff* $\mathcal{M} \models \varphi(b, \bar{a})$ *for some* $b \in \mathcal{M}$ *(so "$\exists$" means "for some").*

Note that by (1), (7), and (8) the intended value in $\mathcal{M}$ for a variable x is always an element of $\mathcal{M}$. This is the reason why we call the logic system above the first-order logic.

Definition (1.3)

(5) $\mathcal{M} \models \varphi \to \psi$ *iff* $\mathcal{M} \models \varphi$ *implies* $\mathcal{M} \models \psi$ *(so "$\to$" means "imply");*

(6) $\mathcal{M} \models \varphi \leftrightarrow \psi$ *iff* $\mathcal{M} \models (\varphi \to \psi) \wedge (\psi \to \varphi)$ *(so "$\leftrightarrow$" means "equivalent to");*

(7) $\mathcal{M} \models \forall x\, \varphi(x, \bar{a})$ *iff* $\mathcal{M} \models \varphi(b, \bar{a})$ *for every* $b \in \mathcal{M}$ *(so "$\forall$" means "for every");*

(8) $\mathcal{M} \models \exists x\, \varphi(x, \bar{a})$ *iff* $\mathcal{M} \models \varphi(b, \bar{a})$ *for some* $b \in \mathcal{M}$ *(so "$\exists$" means "for some").*

Note that by (1), (7), and (8) the intended value in $\mathcal{M}$ for a variable x is always an element of $\mathcal{M}$. This is the reason why we call the logic system above the first-order logic.

Definition (1.3)

(5) $\mathcal{M} \models \varphi \to \psi$ *iff* $\mathcal{M} \models \varphi$ *implies* $\mathcal{M} \models \psi$ *(so "$\to$" means "imply");*

(6) $\mathcal{M} \models \varphi \leftrightarrow \psi$ *iff* $\mathcal{M} \models (\varphi \to \psi) \wedge (\psi \to \varphi)$ *(so "$\leftrightarrow$" means "equivalent to");*

(7) $\mathcal{M} \models \forall x \, \varphi(x, \bar{a})$ *iff* $\mathcal{M} \models \varphi(b, \bar{a})$ *for every* $b \in \mathcal{M}$ *(so "$\forall$" means "for every");*

(8) $\mathcal{M} \models \exists x \, \varphi(x, \bar{a})$ *iff* $\mathcal{M} \models \varphi(b, \bar{a})$ *for some* $b \in \mathcal{M}$ *(so "$\exists$" means "for some").*

Note that by (1), (7), and (8) the intended value in $\mathcal{M}$ for a variable x is always an element of $\mathcal{M}$. This is the reason why we call the logic system above the first-order logic.

When we say a formula, we mean a first-order formula unless otherwise specified. A formula without free variable is called a sentence. If a model $\mathcal{M}$ is given and every free variable of a formula $\varphi(\overline{x})$ is substituted by an element in $\overline{a}$ in $\mathcal{M}$, we call also $\varphi(\overline{a})$ a sentence or a sentence with parameters $\overline{a}$. So, the truth value of a sentence in a model is always determined.

It is an easy fact that each formula φ is logically equivalent to a formula ψ, i.e., φ and ψ have the same truth value in any model, where ψ does not use any of the symbols $\vee$, $\rightarrow$, $\leftrightarrow$, or $\forall$. Hence, it suffices to consider only the formulas using logic connectives $\neg$, $\wedge$, and quantifier $\exists$ in some of the proofs later on.

When we say a formula, we mean a first-order formula unless otherwise specified. A formula without free variable is called a sentence. If a model $\mathcal{M}$ is given and every free variable of a formula $\varphi(\overline{x})$ is substituted by an element in $\overline{a}$ in $\mathcal{M}$, we call also $\varphi(\overline{a})$ a sentence or a sentence with parameters $\overline{a}$. So, the truth value of a sentence in a model is always determined.

It is an easy fact that each formula φ is logically equivalent to a formula ψ, i.e., φ and ψ have the same truth value in any model, where ψ does not use any of the symbols $\vee$, $\rightarrow$, $\leftrightarrow$, or $\forall$. Hence, it suffices to consider only the formulas using logic connectives $\neg$, $\wedge$, and quantifier $\exists$ in some of the proofs later on.

Example (1.4)

Let $\mathscr{L} = \{+, \cdot, \leq, 0, 1, P\}_{P \in \mathcal{P}}$ *be the language of ordered field and* $\mathcal{R} := (\mathbb{R}; +, \cdot, \leq, 0, 1, P^{\mathcal{R}})_{P \in \mathcal{P}}$ *be the usual real ordered field with some extra relations. Then* $\mathcal{R}$ *is an* $\mathscr{L}$*-model. If* φ *is the sentence*

$$\forall x, y, z\,(x \leq y \to x + z \leq y + z),$$

then $\mathcal{R} \models \varphi$*.*

Note that the sentence above can formally be written as a logic sentence

$$\forall x \forall y \forall z \forall u \forall v\,(\leq (x, y) \wedge +(x, z, u) \wedge +(y, z, v) \to \leq (u, v)).$$

We will use conventional expressions more often than the formal ones. The reader is guaranteed that all conventional expressions can be re-written as formal ones.

Example (1.4)

Let $\mathscr{L} = \{+, \cdot, \leq, 0, 1, P\}_{P \in \mathcal{P}}$ *be the language of ordered field and* $\mathcal{R} := (\mathbb{R}; +, \cdot, \leq, 0, 1, P^{\mathcal{R}})_{P \in \mathcal{P}}$ *be the usual real ordered field with some extra relations. Then* $\mathcal{R}$ *is an* $\mathscr{L}$*-model. If* φ *is the sentence*

$$\forall x, y, z\,(x \leq y \rightarrow x + z \leq y + z),$$

then $\mathcal{R} \models \varphi$.

Note that the sentence above can formally be written as a logic sentence

$$\forall x \forall y \forall z \forall u \forall v\,(\leq (x, y) \wedge +(x, z, u) \wedge +(y, z, v) \rightarrow \leq (u, v)).$$

We will use conventional expressions more often than the formal ones. The reader is guaranteed that all conventional expressions can be re-written as formal ones.

Example (1.4)

Let $\mathscr{L} = \{+, \cdot, \leq, 0, 1, P\}_{P \in \mathcal{P}}$ *be the language of ordered field and* $\mathcal{R} := (\mathbb{R}; +, \cdot, \leq, 0, 1, P^{\mathcal{R}})_{P \in \mathcal{P}}$ *be the usual real ordered field with some extra relations. Then* $\mathcal{R}$ *is an* $\mathscr{L}$*-model. If* φ *is the sentence*

$$\forall x, y, z\, (x \leq y \to x + z \leq y + z),$$

then $\mathcal{R} \models \varphi$*.*

Note that the sentence above can formally be written as a logic sentence

$$\forall x \forall y \forall z \forall u \forall v\, (\leq (x, y) \wedge +(x, z, u) \wedge +(y, z, v) \to \leq (u, v)).$$

We will use conventional expressions more often than the formal ones. The reader is guaranteed that all conventional expressions can be re-written as formal ones.

The following is a familiar sentence in the language of real ordered field which is not first-order because variable X in the sentence takes a set not an element of $\mathbb{R}$ as its value.

Example (1.5)

Let φ be the sentence

$$\forall X \subseteq [0,1]\, \exists \beta\, (\beta \ \textit{ is the least upper bound of } X).$$

then φ is true in $\mathcal{R}$.

We now construct an ultrapower of $\mathcal{R}$. Let $\mathbb{N} := \{0, 1, 2, \ldots\}$ and $\mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}$. If n is a positive integer, let $[n] := \{0, 1, \ldots, n-1\}$.

The following is a familiar sentence in the language of real ordered field which is not first-order because variable X in the sentence takes a set not an element of $\mathbb{R}$ as its value.

Example (1.5)

Let φ be the sentence

$$\forall X \subseteq [0,1]\, \exists \beta\, (\beta \text{ is the least upper bound of } X).$$

then φ is true in $\mathcal{R}$.

We now construct an ultrapower of $\mathcal{R}$. Let $\mathbb{N} := \{0, 1, 2, \ldots\}$ and $\mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}$. If n is a positive integer, let $[n] := \{0, 1, \ldots, n-1\}$.

The following is a familiar sentence in the language of real ordered field which is not first-order because variable X in the sentence takes a set not an element of $\mathbb{R}$ as its value.

Example (1.5)

Let φ be the sentence

$$\forall X \subseteq [0,1]\, \exists \beta\, (\beta \text{ is the least upper bound of } X).$$

then φ is true in $\mathcal{R}$.

We now construct an ultrapower of $\mathcal{R}$. Let $\mathbb{N} := \{0, 1, 2, \ldots\}$ and $\mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}$. If n is a positive integer, let $[n] := \{0, 1, \ldots, n-1\}$.

Definition (1.6)

Let X be an infinite set and $\mathscr{P}$ be the power set operator. A collection $\mathcal{F} \subseteq \mathscr{P}(X)$ is called a non-principal ultrafilter *on X if for any $A, B \subseteq X$*

1. *$\emptyset$ is not in $\mathcal{F}$ and every co-finite subset A of X (i.e., $X \setminus A$ is finite) is in $\mathcal{F}$;*
2. *if A, B are in $\mathcal{F}$, then $A \cap B$ is in $\mathcal{F}$;*
3. *if A is in $\mathcal{F}$ and $A \subseteq B$, then B is in $\mathcal{F}$;*
4. *if A is not in $\mathcal{F}$, then $X \setminus A$ is in $\mathcal{F}$.*

The existence of a non-principal ultrafilter on an infinite set X is guaranteed by the axiom of choice. For simplicity we use only a fixed non-principal ultrafilter $\mathcal{F}$ on $X := \mathbb{N}$. In fact, any non-principal ultrafilter on an infinite set X works as long as it is countably incomplete ($\mathcal{F}$ on a countable set such as $\mathbb{N}$ is trivially countably incomplete.)

Definition (1.6)

Let X be an infinite set and $\mathscr{P}$ be the power set operator. A collection $\mathcal{F} \subseteq \mathscr{P}(X)$ is called a non-principal ultrafilter on X if for any $A, B \subseteq X$

1. *$\emptyset$ is not in $\mathcal{F}$ and every co-finite subset A of X (i.e., $X \setminus A$ is finite) is in $\mathcal{F}$;*
2. *if A, B are in $\mathcal{F}$, then $A \cap B$ is in $\mathcal{F}$;*
3. *if A is in $\mathcal{F}$ and $A \subseteq B$, then B is in $\mathcal{F}$;*
4. *if A is not in $\mathcal{F}$, then $X \setminus A$ is in $\mathcal{F}$.*

The existence of a non-principal ultrafilter on an infinite set X is guaranteed by the axiom of choice. For simplicity we use only a fixed non-principal ultrafilter $\mathcal{F}$ on $X := \mathbb{N}$. In fact, any non-principal ultrafilter on an infinite set X works as long as it is countably incomplete ($\mathcal{F}$ on a countable set such as $\mathbb{N}$ is trivially countably incomplete.)

Definition (1.6)

Let X be an infinite set and $\mathscr{P}$ be the power set operator. A collection $\mathcal{F} \subseteq \mathscr{P}(X)$ is called a non-principal ultrafilter on X if for any $A, B \subseteq X$

1. *$\emptyset$ is not in $\mathcal{F}$ and every co-finite subset A of X (i.e., $X \setminus A$ is finite) is in $\mathcal{F}$;*
2. *if A, B are in $\mathcal{F}$, then $A \cap B$ is in $\mathcal{F}$;*
3. *if A is in $\mathcal{F}$ and $A \subseteq B$, then B is in $\mathcal{F}$;*
4. *if A is not in $\mathcal{F}$, then $X \setminus A$ is in $\mathcal{F}$.*

The existence of a non-principal ultrafilter on an infinite set X is guaranteed by the axiom of choice. For simplicity we use only a fixed non-principal ultrafilter $\mathcal{F}$ on $X := \mathbb{N}$. In fact, any non-principal ultrafilter on an infinite set X works as long as it is countably incomplete ($\mathcal{F}$ on a countable set such as $\mathbb{N}$ is trivially countably incomplete.)

Definition (1.6)

Let X be an infinite set and $\mathscr{P}$ be the power set operator. A collection $\mathcal{F} \subseteq \mathscr{P}(X)$ is called a non-principal ultrafilter on X if for any $A, B \subseteq X$

1. *$\emptyset$ is not in $\mathcal{F}$ and every co-finite subset A of X (i.e., $X \setminus A$ is finite) is in $\mathcal{F}$;*
2. *if A, B are in $\mathcal{F}$, then $A \cap B$ is in $\mathcal{F}$;*
3. *if A is in $\mathcal{F}$ and $A \subseteq B$, then B is in $\mathcal{F}$;*
4. *if A is not in $\mathcal{F}$, then $X \setminus A$ is in $\mathcal{F}$.*

The existence of a non-principal ultrafilter on an infinite set X is guaranteed by the axiom of choice. For simplicity we use only a fixed non-principal ultrafilter $\mathcal{F}$ on $X := \mathbb{N}$. In fact, any non-principal ultrafilter on an infinite set X works as long as it is countably incomplete ($\mathcal{F}$ on a countable set such as $\mathbb{N}$ is trivially countably incomplete.)

Definition (1.6)

Let X be an infinite set and $\mathscr{P}$ be the power set operator. A collection $\mathcal{F} \subseteq \mathscr{P}(X)$ is called a non-principal ultrafilter on X if for any $A, B \subseteq X$

1. *$\emptyset$ is not in $\mathcal{F}$ and every co-finite subset A of X (i.e., $X \setminus A$ is finite) is in $\mathcal{F}$;*
2. *if A, B are in $\mathcal{F}$, then $A \cap B$ is in $\mathcal{F}$;*
3. *if A is in $\mathcal{F}$ and $A \subseteq B$, then B is in $\mathcal{F}$;*
4. *if A is not in $\mathcal{F}$, then $X \setminus A$ is in $\mathcal{F}$.*

The existence of a non-principal ultrafilter on an infinite set X is guaranteed by the axiom of choice. For simplicity we use only a fixed non-principal ultrafilter $\mathcal{F}$ on $X := \mathbb{N}$. In fact, any non-principal ultrafilter on an infinite set X works as long as it is countably incomplete ($\mathcal{F}$ on a countable set such as $\mathbb{N}$ is trivially countably incomplete.)

Definition (1.6)

Let X be an infinite set and $\mathscr{P}$ be the power set operator. A collection $\mathcal{F} \subseteq \mathscr{P}(X)$ is called a non-principal ultrafilter on X if for any $A, B \subseteq X$

1. *$\emptyset$ is not in $\mathcal{F}$ and every co-finite subset A of X (i.e., $X \setminus A$ is finite) is in $\mathcal{F}$;*
2. *if A, B are in $\mathcal{F}$, then $A \cap B$ is in $\mathcal{F}$;*
3. *if A is in $\mathcal{F}$ and $A \subseteq B$, then B is in $\mathcal{F}$;*
4. *if A is not in $\mathcal{F}$, then $X \setminus A$ is in $\mathcal{F}$.*

The existence of a non-principal ultrafilter on an infinite set X is guaranteed by the axiom of choice. For simplicity we use only a fixed non-principal ultrafilter $\mathcal{F}$ on $X := \mathbb{N}$. In fact, any non-principal ultrafilter on an infinite set X works as long as it is countably incomplete ($\mathcal{F}$ on a countable set such as $\mathbb{N}$ is trivially countably incomplete.)

Definition (1.7)

Let $\mathcal{M}$ be a model. Let $\mathcal{M}^{\mathbb{N}}$ be the set of all functions from $\mathbb{N}$ to $\mathcal{M}$. For any $f, g \in \mathcal{M}^{\mathbb{N}}$ define $f \sim g$ iff $\{n \in \mathbb{N} \mid f(n) = g(n)\} \in \mathcal{F}$. The equivalence class of $f \in \mathcal{M}^{\mathbb{N}}$ is the set $[f] := \{g \in \mathcal{M}^{\mathbb{N}} \mid f \sim g\}$. Set $\mathcal{M}^{\mathbb{N}}/\mathcal{F} := \{[f] \mid f \in \mathcal{M}^{\mathbb{N}}\}$. The ultrapower of $\mathcal{M}$ modulo $\mathcal{F}$, denoted by $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$, is a model with the base set $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ and for each relation symbol P, the interpretation of P in $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ is defined by

$$\overline{[f]} \in P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}} \text{ iff } \{n \in \mathbb{N} \mid \overline{f(n)} \in P^{\mathcal{M}}\} \in \mathcal{F}. \quad (1)$$

For each $a \in \mathcal{M}$ let $\phi_a : \mathbb{N} \to \mathcal{M}$ be the constant function with a unique value a. If $\overline{b} = (b_1, b_2, \ldots, b_k)$, we write $\overline{[\phi_b]}$ for $([\phi_{b_1}], [\phi_{b_2}], \ldots, [\phi_{b_k}])$.

Definition (1.7)

Let $\mathcal{M}$ be a model. Let $\mathcal{M}^{\mathbb{N}}$ be the set of all functions from $\mathbb{N}$ to $\mathcal{M}$. For any $f, g \in \mathcal{M}^{\mathbb{N}}$ define $f \sim g$ iff $\{n \in \mathbb{N} \mid f(n) = g(n)\} \in \mathcal{F}$. The equivalence class of $f \in \mathcal{M}^{\mathbb{N}}$ is the set $[f] := \{g \in \mathcal{M}^{\mathbb{N}} \mid f \sim g\}$. Set $\mathcal{M}^{\mathbb{N}}/\mathcal{F} := \{[f] \mid f \in \mathcal{M}^{\mathbb{N}}\}$. The ultrapower of $\mathcal{M}$ modulo $\mathcal{F}$, denoted by $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$, is a model with the base set $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ and for each relation symbol P, the interpretation of P in $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ is defined by

$$\overline{[f]} \in P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}} \text{ iff } \{n \in \mathbb{N} \mid \overline{f(n)} \in P^{\mathcal{M}}\} \in \mathcal{F}. \quad (1)$$

For each $a \in \mathcal{M}$ let $\phi_a : \mathbb{N} \to \mathcal{M}$ be the constant function with a unique value a. If $\overline{b} = (b_1, b_2, \ldots, b_k)$, we write $\overline{[\phi_b]}$ for $([\phi_{b_1}], [\phi_{b_2}], \ldots, [\phi_{b_k}])$.

Definition (1.7)

Let $\mathcal{M}$ be a model. Let $\mathcal{M}^{\mathbb{N}}$ be the set of all functions from $\mathbb{N}$ to $\mathcal{M}$. For any $f, g \in \mathcal{M}^{\mathbb{N}}$ define $f \sim g$ iff $\{n \in \mathbb{N} \mid f(n) = g(n)\} \in \mathcal{F}$. The equivalence class of $f \in \mathcal{M}^{\mathbb{N}}$ is the set $[f] := \{g \in \mathcal{M}^{\mathbb{N}} \mid f \sim g\}$. Set $\mathcal{M}^{\mathbb{N}}/\mathcal{F} := \{[f] \mid f \in \mathcal{M}^{\mathbb{N}}\}$. The ultrapower of $\mathcal{M}$ modulo $\mathcal{F}$, denoted by $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$, is a model with the base set $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ and for each relation symbol P, the interpretation of P in $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ is defined by

$$\overline{[f]} \in P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}} \text{ iff } \{n \in \mathbb{N} \mid \overline{f(n)} \in P^{\mathcal{M}}\} \in \mathcal{F}. \quad (1)$$

For each $a \in \mathcal{M}$ let $\phi_a : \mathbb{N} \to \mathcal{M}$ be the constant function with a unique value a. If $\overline{b} = (b_1, b_2, \ldots, b_k)$, we write $\overline{[\phi_b]}$ for $([\phi_{b_1}], [\phi_{b_2}], \ldots, [\phi_{b_k}])$.

Definition (1.7)

Let $\mathcal{M}$ be a model. Let $\mathcal{M}^{\mathbb{N}}$ be the set of all functions from $\mathbb{N}$ to $\mathcal{M}$. For any $f, g \in \mathcal{M}^{\mathbb{N}}$ define $f \sim g$ iff $\{n \in \mathbb{N} \mid f(n) = g(n)\} \in \mathcal{F}$. The equivalence class of $f \in \mathcal{M}^{\mathbb{N}}$ is the set $[f] := \{g \in \mathcal{M}^{\mathbb{N}} \mid f \sim g\}$. Set $\mathcal{M}^{\mathbb{N}}/\mathcal{F} := \{[f] \mid f \in \mathcal{M}^{\mathbb{N}}\}$. The ultrapower of $\mathcal{M}$ modulo $\mathcal{F}$, denoted by $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$, is a model with the base set $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ and for each relation symbol P, the interpretation of P in $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ is defined by

$$\overline{[f]} \in P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}} \text{ iff } \{n \in \mathbb{N} \mid \overline{f(n)} \in P^{\mathcal{M}}\} \in \mathcal{F}. \quad (1)$$

For each $a \in \mathcal{M}$ let $\phi_a : \mathbb{N} \to \mathcal{M}$ be the constant function with a unique value a. If $\overline{b} = (b_1, b_2, \ldots, b_k)$, we write $\overline{[\phi_b]}$ for $([\phi_{b_1}], [\phi_{b_2}], \ldots, [\phi_{b_k}])$.

Definition (1.7)

Let $\mathcal{M}$ be a model. Let $\mathcal{M}^{\mathbb{N}}$ be the set of all functions from $\mathbb{N}$ to $\mathcal{M}$. For any $f, g \in \mathcal{M}^{\mathbb{N}}$ define $f \sim g$ iff $\{n \in \mathbb{N} \mid f(n) = g(n)\} \in \mathcal{F}$. The equivalence class of $f \in \mathcal{M}^{\mathbb{N}}$ is the set $[f] := \{g \in \mathcal{M}^{\mathbb{N}} \mid f \sim g\}$. Set $\mathcal{M}^{\mathbb{N}}/\mathcal{F} := \{[f] \mid f \in \mathcal{M}^{\mathbb{N}}\}$. The ultrapower of $\mathcal{M}$ modulo $\mathcal{F}$, denoted by $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$, is a model with the base set $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ and for each relation symbol P, the interpretation of P in $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ is defined by

$$\overline{[f]} \in P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}} \text{ iff } \{n \in \mathbb{N} \mid \overline{f(n)} \in P^{\mathcal{M}}\} \in \mathcal{F}. \tag{1}$$

For each $a \in \mathcal{M}$ let $\phi_a : \mathbb{N} \to \mathcal{M}$ be the constant function with a unique value a. If $\overline{b} = (b_1, b_2, \ldots, b_k)$, we write $\overline{[\phi_b]}$ for $([\phi_{b_1}], [\phi_{b_2}], \ldots, [\phi_{b_k}])$.

Definition (1.7)

Let $\mathcal{M}$ be a model. Let $\mathcal{M}^{\mathbb{N}}$ be the set of all functions from $\mathbb{N}$ to $\mathcal{M}$. For any $f, g \in \mathcal{M}^{\mathbb{N}}$ define $f \sim g$ iff $\{n \in \mathbb{N} \mid f(n) = g(n)\} \in \mathcal{F}$. The equivalence class of $f \in \mathcal{M}^{\mathbb{N}}$ is the set $[f] := \{g \in \mathcal{M}^{\mathbb{N}} \mid f \sim g\}$. Set $\mathcal{M}^{\mathbb{N}}/\mathcal{F} := \{[f] \mid f \in \mathcal{M}^{\mathbb{N}}\}$. The ultrapower of $\mathcal{M}$ modulo $\mathcal{F}$, denoted by $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$, is a model with the base set $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ and for each relation symbol P, the interpretation of P in $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ is defined by

$$\overline{[f]} \in P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}} \text{ iff } \{n \in \mathbb{N} \mid \overline{f(n)} \in P^{\mathcal{M}}\} \in \mathcal{F}. \tag{1}$$

For each $a \in \mathcal{M}$ let $\phi_a : \mathbb{N} \to \mathcal{M}$ be the constant function with a unique value a. If $\overline{b} = (b_1, b_2, \ldots, b_k)$, we write $\overline{[\phi_b]}$ for $([\phi_{b_1}], [\phi_{b_2}], \ldots, [\phi_{b_k}])$.

Definition (1.8)

Let $i : \mathcal{M} \to \mathcal{M}^{\mathbb{N}}/\mathcal{F}$ be the function such that $i(a) = [\phi_a]$. The function i is called an elementary embedding associated with the ultrapower construction.

Theorem (1.9, J. Łoś)

Let $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ be the ultrapower of a model $\mathcal{M}$ modulo $\mathcal{F}$. Let $\varphi(\overline{x}, \overline{b})$ be a formula with parameters $\overline{b}$ in $\mathcal{M}$. Then

$$\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \varphi(\overline{[f]}, \overline{[\phi_b]}) \text{ iff } \{n \in \mathbb{N} \mid \mathcal{M} \models \varphi(\overline{f(n)}, \overline{b})\} \in \mathcal{F}.$$

Definition (1.8)

Let $i : \mathcal{M} \to \mathcal{M}^{\mathbb{N}}/\mathcal{F}$ be the function such that $i(a) = [\phi_a]$. The function i is called an elementary embedding associated with the ultrapower construction.

Theorem (1.9, J. Łoś)

Let $\mathcal{M}^{\mathbb{N}}/\mathcal{F}$ be the ultrapower of a model $\mathcal{M}$ modulo $\mathcal{F}$. Let $\varphi(\overline{x}, \overline{b})$ be a formula with parameters $\overline{b}$ in $\mathcal{M}$. Then

$$\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \varphi(\overline{[f]}, \overline{[\phi_b]}) \text{ iff } \{n \in \mathbb{N} \mid \mathcal{M} \models \varphi(\overline{f(n)}, \overline{b})\} \in \mathcal{F}.$$

Proof of Theorem 1.9: If φ is an atomic formula $P(\overline{x}, \overline{c})$, then the theorem follows from the definition of $P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}}$. If φ is $\neg\psi$, then the theorem follows from Part 4 of Definition 1.6 and induction hypothesis for ψ. If φ is $\psi \wedge \chi$, then the theorem follows from Part 2 of Definition 1.6 and induction hypothesis for ψ and χ.

Assume φ is $\exists x\, \psi(x, \overline{y}, \overline{b})$. If

$$A := \{n \in \mathbb{N} \mid \mathcal{M} \models \exists x\, \psi(x, \overline{f(n)}, \overline{b})\} \in \mathcal{F},$$

define a function $g : \mathbb{N} \to \mathcal{M}$ by letting $g(n)$ be any fixed element in $\mathcal{M}$ if $n \notin A$, and $g(n) = a_n$ for some $a_n \in \mathcal{M}$ with $\mathcal{M} \models \psi(a_n, \overline{f(n)}, \overline{b})$ if $n \in A$. Then

$$A \subseteq \{n \in \mathbb{N} \mid \mathcal{M} \models \psi(g(n), \overline{f(n)}, \overline{b})\}.$$

Proof of Theorem 1.9: If φ is an atomic formula $P(\overline{x}, \overline{c})$, then the theorem follows from the definition of $P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}}$. If φ is $\neg\psi$, then the theorem follows from Part 4 of Definition 1.6 and induction hypothesis for ψ. If φ is $\psi \wedge \chi$, then the theorem follows from Part 2 of Definition 1.6 and induction hypothesis for ψ and χ.

Assume φ is $\exists x\, \psi(x, \overline{y}, \overline{b})$. If

$$A := \{n \in \mathbb{N} \mid \mathcal{M} \models \exists x\, \psi(x, \overline{f(n)}, \overline{b})\} \in \mathcal{F},$$

define a function $g : \mathbb{N} \to \mathcal{M}$ by letting $g(n)$ be any fixed element in $\mathcal{M}$ if $n \notin A$, and $g(n) = a_n$ for some $a_n \in \mathcal{M}$ with $\mathcal{M} \models \psi(a_n, \overline{f(n)}, \overline{b})$ if $n \in A$. Then

$$A \subseteq \{n \in \mathbb{N} \mid \mathcal{M} \models \psi(g(n), \overline{f(n)}, \overline{b})\}.$$

Proof of Theorem 1.9: If φ is an atomic formula $P(\overline{x}, \overline{c})$, then the theorem follows from the definition of $P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}}$. If φ is $\neg\psi$, then the theorem follows from Part 4 of Definition 1.6 and induction hypothesis for ψ. If φ is $\psi \wedge \chi$, then the theorem follows from Part 2 of Definition 1.6 and induction hypothesis for ψ and χ.

Assume φ is $\exists x\, \psi(x, \overline{y}, \overline{b})$. If

$$A := \{n \in \mathbb{N} \mid \mathcal{M} \models \exists x\, \psi(x, \overline{f(n)}, \overline{b})\} \in \mathcal{F},$$

define a function $g : \mathbb{N} \to \mathcal{M}$ by letting $g(n)$ be any fixed element in $\mathcal{M}$ if $n \notin A$, and $g(n) = a_n$ for some $a_n \in \mathcal{M}$ with $\mathcal{M} \models \psi(a_n, \overline{f(n)}, \overline{b})$ if $n \in A$. Then

$$A \subseteq \{n \in \mathbb{N} \mid \mathcal{M} \models \psi(g(n), \overline{f(n)}, \overline{b})\}.$$

Proof of Theorem 1.9: If φ is an atomic formula $P(\overline{x}, \overline{c})$, then the theorem follows from the definition of $P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}}$. If φ is $\neg\psi$, then the theorem follows from Part 4 of Definition 1.6 and induction hypothesis for ψ. If φ is $\psi \wedge \chi$, then the theorem follows from Part 2 of Definition 1.6 and induction hypothesis for ψ and χ.

Assume φ is $\exists x\, \psi(x, \overline{y}, \overline{b})$. If

$$A := \{n \in \mathbb{N} \mid \mathcal{M} \models \exists x\, \psi(x, \overline{f(n)}, \overline{b})\} \in \mathcal{F},$$

define a function $g : \mathbb{N} \to \mathcal{M}$ by letting $g(n)$ be any fixed element in $\mathcal{M}$ if $n \notin A$, and $g(n) = a_n$ for some $a_n \in \mathcal{M}$ with $\mathcal{M} \models \psi(a_n, \overline{f(n)}, \overline{b})$ if $n \in A$. Then

$$A \subseteq \{n \in \mathbb{N} \mid \mathcal{M} \models \psi(g(n), \overline{f(n)}, \overline{b})\}.$$

Proof of Theorem 1.9: If φ is an atomic formula $P(\overline{x}, \overline{c})$, then the theorem follows from the definition of $P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}}$. If φ is $\neg\psi$, then the theorem follows from Part 4 of Definition 1.6 and induction hypothesis for ψ. If φ is $\psi \wedge \chi$, then the theorem follows from Part 2 of Definition 1.6 and induction hypothesis for ψ and χ.

Assume φ is $\exists x\, \psi(x, \overline{y}, \overline{b})$. If

$$A := \{n \in \mathbb{N} \mid \mathcal{M} \models \exists x\, \psi(x, \overline{f(n)}, \overline{b})\} \in \mathcal{F},$$

define a function $g : \mathbb{N} \to \mathcal{M}$ by letting $g(n)$ be any fixed element in $\mathcal{M}$ if $n \notin A$, and $g(n) = a_n$ for some $a_n \in \mathcal{M}$ with $\mathcal{M} \models \psi(a_n, \overline{f(n)}, \overline{b})$ if $n \in A$. Then

$$A \subseteq \{n \in \mathbb{N} \mid \mathcal{M} \models \psi(g(n), \overline{f(n)}, \overline{b})\}.$$

Proof of Theorem 1.9: If φ is an atomic formula $P(\overline{x}, \overline{c})$, then the theorem follows from the definition of $P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}}$. If φ is $\neg\psi$, then the theorem follows from Part 4 of Definition 1.6 and induction hypothesis for ψ. If φ is $\psi \wedge \chi$, then the theorem follows from Part 2 of Definition 1.6 and induction hypothesis for ψ and χ.

Assume φ is $\exists x\, \psi(x, \overline{y}, \overline{b})$. If

$$A := \{n \in \mathbb{N} \mid \mathcal{M} \models \exists x\, \psi(x, \overline{f(n)}, \overline{b})\} \in \mathcal{F},$$

define a function $g : \mathbb{N} \to \mathcal{M}$ by letting $g(n)$ be any fixed element in $\mathcal{M}$ if $n \notin A$, and $g(n) = a_n$ for some $a_n \in \mathcal{M}$ with $\mathcal{M} \models \psi(a_n, \overline{f(n)}, \overline{b})$ if $n \in A$. Then

$$A \subseteq \{n \in \mathbb{N} \mid \mathcal{M} \models \psi(g(n), \overline{f(n)}, \overline{b})\}.$$

Proof of Theorem 1.9: If φ is an atomic formula $P(\overline{x}, \overline{c})$, then the theorem follows from the definition of $P^{\mathcal{M}^{\mathbb{N}}/\mathcal{F}}$. If φ is $\neg\psi$, then the theorem follows from Part 4 of Definition 1.6 and induction hypothesis for ψ. If φ is $\psi \wedge \chi$, then the theorem follows from Part 2 of Definition 1.6 and induction hypothesis for ψ and χ.

Assume φ is $\exists x\, \psi(x, \overline{y}, \overline{b})$. If

$$A := \{n \in \mathbb{N} \mid \mathcal{M} \models \exists x\, \psi(x, \overline{f(n)}, \overline{b})\} \in \mathcal{F},$$

define a function $g : \mathbb{N} \to \mathcal{M}$ by letting $g(n)$ be any fixed element in $\mathcal{M}$ if $n \notin A$, and $g(n) = a_n$ for some $a_n \in \mathcal{M}$ with $\mathcal{M} \models \psi(a_n, \overline{f(n)}, \overline{b})$ if $n \in A$. Then

$$A \subseteq \{n \in \mathbb{N} \mid \mathcal{M} \models \psi(g(n), \overline{f(n)}, \overline{b})\}.$$

By Part 3 of Definition 1.6 and the induction hypothesis on ψ we have $\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \psi([g], \overline{[f]}, \overline{[\phi_b]})$, which implies

$$\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \exists x\, \psi(x, \overline{[f]}, \overline{[\phi_b]}).$$

On the other hand, if $\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \exists x\, \psi(x, \overline{[f]}, \overline{[\phi_b]})$, then there is a $g : \mathbb{N} \to \mathcal{M}$ such that $\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \psi([g], \overline{[f]}, \overline{[\phi_b]})$. By the induction hypothesis for ψ we have

$$B := \{n \in \mathbb{N} \mid \mathcal{M} \models \psi(g(n), \overline{f(n)}, \overline{b})\} \in \mathcal{F}.$$

So, if $n \in B$, we have $\mathcal{M} \models \exists x\, \psi(x, \overline{f(n)}, \overline{b})$. Hence,

$$B \subseteq \{n \in \mathbb{N} \mid \mathcal{M} \models \exists x\, \psi(x, \overline{[f]}, \overline{[\phi_b]})\} \in \mathcal{F}$$

by Part 2 of Definition 1.6. $\square$

By Part 3 of Definition 1.6 and the induction hypothesis on ψ we have $\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \psi([g], \overline{[f]}, \overline{[\phi_b]})$, which implies

$$\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \exists x\, \psi(x, \overline{[f]}, \overline{[\phi_b]}).$$

On the other hand, if $\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \exists x\, \psi(x, \overline{[f]}, \overline{[\phi_b]})$, then there is a $g : \mathbb{N} \to \mathcal{M}$ such that $\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \psi([g], \overline{[f]}, \overline{[\phi_b]})$. By the induction hypothesis for ψ we have

$$B := \{n \in \mathbb{N} \mid \mathcal{M} \models \psi(g(n), \overline{f(n)}, \overline{b})\} \in \mathcal{F}.$$

So, if $n \in B$, we have $\mathcal{M} \models \exists x\, \psi(x, \overline{f(n)}, \overline{b})$. Hence,

$$B \subseteq \{n \in \mathbb{N} \mid \mathcal{M} \models \exists x\, \psi(x, \overline{[f]}, \overline{[\phi_b]})\} \in \mathcal{F}$$

by Part 2 of Definition 1.6. □

By Part 3 of Definition 1.6 and the induction hypothesis on ψ we have $\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \psi([g], \overline{[f]}, \overline{[\phi_b]})$, which implies

$$\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \exists x\, \psi(x, \overline{[f]}, \overline{[\phi_b]}).$$

On the other hand, if $\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \exists x\, \psi(x, \overline{[f]}, \overline{[\phi_b]})$, then there is a $g : \mathbb{N} \to \mathcal{M}$ such that $\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \psi([g], \overline{[f]}, \overline{[\phi_b]})$. By the induction hypothesis for ψ we have

$$B := \{n \in \mathbb{N} \mid \mathcal{M} \models \psi(g(n), \overline{f(n)}, \bar{b})\} \in \mathcal{F}.$$

So, if $n \in B$, we have $\mathcal{M} \models \exists x\, \psi(x, \overline{f(n)}, \bar{b})$. Hence,

$$B \subseteq \{n \in \mathbb{N} \mid \mathcal{M} \models \exists x\, \psi(x, \overline{[f]}, \overline{[\phi_b]})\} \in \mathcal{F}$$

by Part 2 of Definition 1.6. □

By Part 3 of Definition 1.6 and the induction hypothesis on ψ we have $\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \psi([g], \overline{[f]}, \overline{[\phi_b]})$, which implies

$$\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \exists x\, \psi(x, \overline{[f]}, \overline{[\phi_b]}).$$

On the other hand, if $\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \exists x\, \psi(x, \overline{[f]}, \overline{[\phi_b]})$, then there is a $g : \mathbb{N} \to \mathcal{M}$ such that $\mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \psi([g], \overline{[f]}, \overline{[\phi_b]})$. By the induction hypothesis for ψ we have

$$B := \{n \in \mathbb{N} \mid \mathcal{M} \models \psi(g(n), \overline{f(n)}, \overline{b})\} \in \mathcal{F}.$$

So, if $n \in B$, we have $\mathcal{M} \models \exists x\, \psi(x, \overline{f(n)}, \overline{b})$. Hence,

$$B \subseteq \{n \in \mathbb{N} \mid \mathcal{M} \models \exists x\, \psi(x, \overline{[f]}, \overline{[\phi_b]})\} \in \mathcal{F}$$

by Part 2 of Definition 1.6. □

Corollary (1.10)

Let $i : \mathcal{M} \to \mathcal{M}^{\mathbb{N}}/\mathcal{F}$ be the embedding defined in Definition 1.8. For any sentence $\varphi(\overline{b})$ with parameters $\overline{b}$ in $\mathcal{M}$, we have

$$\mathcal{M} \models \varphi(\overline{b}) \text{ iff } \mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \varphi(\overline{i(b)}). \tag{2}$$

Proof: The corollary follows from Theorem 1.9, Part 1 of Definition 1.6, and the fact that the set $\{n \in \mathbb{N} \mid \mathcal{M} \models \varphi(\overline{[\phi_b(n)]})\}$ is either $\mathbb{N}$ or $\emptyset$ depending on whether $\mathcal{M} \models \varphi(\overline{b})$ is true or not. □

The map i satisfying (2) is called an elementary embedding from a model $\mathcal{M}$ to another model $\mathcal{M}' = \mathcal{M}^{\mathbb{N}}/\mathcal{F}$. (2) is also called the transfer principle.

Corollary (1.10)

Let $i : \mathcal{M} \to \mathcal{M}^{\mathbb{N}}/\mathcal{F}$ *be the embedding defined in Definition 1.8. For any sentence* $\varphi(\overline{b})$ *with parameters* $\overline{b}$ *in* $\mathcal{M}$*, we have*

$$\mathcal{M} \models \varphi(\overline{b}) \text{ iff } \mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \varphi(\overline{i(b)}). \tag{2}$$

Proof: The corollary follows from Theorem 1.9, Part 1 of Definition 1.6, and the fact that the set $\{n \in \mathbb{N} \mid \mathcal{M} \models \varphi(\overline{[\phi_b(n)]})\}$ is either $\mathbb{N}$ or $\emptyset$ depending on whether $\mathcal{M} \models \varphi(\overline{b})$ is true or not. □

The map i satisfying (2) is called an elementary embedding from a model $\mathcal{M}$ to another model $\mathcal{M}' = \mathcal{M}^{\mathbb{N}}/\mathcal{F}$. (2) is also called the transfer principle.

Corollary (1.10)

Let $i : \mathcal{M} \to \mathcal{M}^{\mathbb{N}}/\mathcal{F}$ be the embedding defined in Definition 1.8. For any sentence $\varphi(\overline{b})$ with parameters $\overline{b}$ in $\mathcal{M}$, we have

$$\mathcal{M} \models \varphi(\overline{b}) \text{ iff } \mathcal{M}^{\mathbb{N}}/\mathcal{F} \models \varphi(\overline{i(b)}). \quad (2)$$

Proof: The corollary follows from Theorem 1.9, Part 1 of Definition 1.6, and the fact that the set $\{n \in \mathbb{N} \mid \mathcal{M} \models \varphi(\overline{[\phi_b(n)]})\}$ is either $\mathbb{N}$ or $\emptyset$ depending on whether $\mathcal{M} \models \varphi(\overline{b})$ is true or not. □

The map i satisfying (2) is called an elementary embedding from a model $\mathcal{M}$ to another model $\mathcal{M}' = \mathcal{M}^{\mathbb{N}}/\mathcal{F}$. (2) is also called the transfer principle.

In fact, the elementary embedding from $\mathcal{M}$ to $\mathcal{M}'$ can be defined for any two models $\mathcal{M}$ and $\mathcal{M}'$ when (2) is true.

Denote by $\mathcal{M} \preceq \mathcal{M}'$ for the existence of such an elementary embedding from $\mathcal{M}$ to $\mathcal{M}'$. If we want to emphasize that an elementary embedding $i : \mathcal{M} \to \mathcal{M}'$ is not surjective, we can just write $\mathcal{M} \prec \mathcal{M}'$ instead.

The embedding i is often written as $*$ in nonstandard analysis. For example, a nonstandard analyst may write ${}^*A$ more often than $i(A)$.

In fact, the elementary embedding from $\mathcal{M}$ to $\mathcal{M}'$ can be defined for any two models $\mathcal{M}$ and $\mathcal{M}'$ when (2) is true.

Denote by $\mathcal{M} \preceq \mathcal{M}'$ for the existence of such an elementary embedding from $\mathcal{M}$ to $\mathcal{M}'$. If we want to emphasize that an elementary embedding $i : \mathcal{M} \to \mathcal{M}'$ is not surjective, we can just write $\mathcal{M} \prec \mathcal{M}'$ instead.

The embedding i is often written as * in nonstandard analysis. For example, a nonstandard analyst may write $^*A$ more often than $i(A)$.

In fact, the elementary embedding from $\mathcal{M}$ to $\mathcal{M}'$ can be defined for any two models $\mathcal{M}$ and $\mathcal{M}'$ when (2) is true.

Denote by $\mathcal{M} \preceq \mathcal{M}'$ for the existence of such an elementary embedding from $\mathcal{M}$ to $\mathcal{M}'$. If we want to emphasize that an elementary embedding $i : \mathcal{M} \to \mathcal{M}'$ is not surjective, we can just write $\mathcal{M} \prec \mathcal{M}'$ instead.

The embedding i is often written as * in nonstandard analysis. For example, a nonstandard analyst may write $^*A$ more often than $i(A)$.

Example (1.11)

Let $^*\mathcal{R}$ be the ultrapower of the "standard" real ordered field $\mathcal{R}$ modulo $\mathcal{F}$.

(1) *$^*\mathcal{R}$ satisfies the same first-order sentences with parameters from $\mathcal{R}$, in particular, $^*\mathcal{R}$ is an ordered field and contains a copy of $\mathcal{R}$ as its (elementary) sub-model. We call real numbers in $\mathbb{R}$ the standard real numbers.*

(2) *By identifying each $\alpha \in \mathbb{R}$ with $^*\alpha = [\phi_\alpha] \in {}^*\mathbb{R}$, we can assume that $\mathcal{R} \subseteq {}^*\mathcal{R}$.*

(3) *A real $r \in {}^*\mathbb{R}$ is called an infinitesimal, denoted by $r \approx 0$, if $|r| < |\alpha|$ for every non-zero $\alpha \in \mathbb{R}$. Two reals $r_1, r_2 \in {}^*\mathbb{R}$ are said to be infinitesimally close, denoted by $r_1 \approx r_2$, if $r_1 - r_2$ is an infinitesimal.*

Example (1.11)

Let ${}^*\mathcal{R}$ be the ultrapower of the "standard" real ordered field $\mathcal{R}$ modulo $\mathcal{F}$.

(1) *${}^*\mathcal{R}$ satisfies the same first-order sentences with parameters from $\mathcal{R}$, in particular, ${}^*\mathcal{R}$ is an ordered field and contains a copy of $\mathcal{R}$ as its (elementary) sub-model. We call real numbers in $\mathbb{R}$ the standard real numbers.*

(2) *By identifying each $\alpha \in \mathbb{R}$ with ${}^*\alpha = [\phi_\alpha] \in {}^*\mathbb{R}$, we can assume that $\mathcal{R} \subseteq {}^*\mathcal{R}$.*

(3) *A real $r \in {}^*\mathbb{R}$ is called an infinitesimal, denoted by $r \approx 0$, if $|r| < |\alpha|$ for every non-zero $\alpha \in \mathbb{R}$. Two reals $r_1, r_2 \in {}^*\mathbb{R}$ are said to be infinitesimally close, denoted by $r_1 \approx r_2$, if $r_1 - r_2$ is an infinitesimal.*

Example (1.11)

Let $^*\mathcal{R}$ *be the ultrapower of the "standard" real ordered field* $\mathcal{R}$ *modulo* $\mathcal{F}$.

(1) $^*\mathcal{R}$ *satisfies the same first-order sentences with parameters from* $\mathcal{R}$, *in particular,* $^*\mathcal{R}$ *is an ordered field and contains a copy of* $\mathcal{R}$ *as its (elementary) sub-model. We call real numbers in* $\mathbb{R}$ *the standard real numbers.*

(2) *By identifying each* $\alpha \in \mathbb{R}$ *with* $^*\alpha = [\phi_\alpha] \in {}^*\mathbb{R}$, *we can assume that* $\mathcal{R} \subseteq {}^*\mathcal{R}$.

(3) *A real* $r \in {}^*\mathbb{R}$ *is called an infinitesimal, denoted by* $r \approx 0$, *if* $|r| < |\alpha|$ *for every non-zero* $\alpha \in \mathbb{R}$. *Two reals* $r_1, r_2 \in {}^*\mathbb{R}$ *are said to be infinitesimally close, denoted by* $r_1 \approx r_2$, *if* $r_1 - r_2$ *is an infinitesimal.*

Example (1.11)

Let ${}^*\mathcal{R}$ be the ultrapower of the "standard" real ordered field $\mathcal{R}$ modulo $\mathcal{F}$.

(1) *${}^*\mathcal{R}$ satisfies the same first-order sentences with parameters from $\mathcal{R}$, in particular, ${}^*\mathcal{R}$ is an ordered field and contains a copy of $\mathcal{R}$ as its (elementary) sub-model. We call real numbers in $\mathbb{R}$ the standard real numbers.*

(2) *By identifying each $\alpha \in \mathbb{R}$ with ${}^*\alpha = [\phi_\alpha] \in {}^*\mathbb{R}$, we can assume that $\mathcal{R} \subseteq {}^*\mathcal{R}$.*

(3) *A real $r \in {}^*\mathbb{R}$ is called an infinitesimal, denoted by $r \approx 0$, if $|r| < |\alpha|$ for every non-zero $\alpha \in \mathbb{R}$. Two reals $r_1, r_2 \in {}^*\mathbb{R}$ are said to be infinitesimally close, denoted by $r_1 \approx r_2$, if $r_1 - r_2$ is an infinitesimal.*

Example (1.11)

(4) *If Id* $\in \mathcal{R}^{\mathbb{N}}$ *is the identity function, i.e., Id*$(n) = n$ *for every* $n \in \mathbb{N}$*, then* $[Id] \in {}^*\mathcal{R}$ *and* $[Id] > r$ *for every* $r \in \mathbb{R}$*. So,* ${}^*\mathcal{R}$ *contains numbers larger than every* $r \in \mathbb{R}$*.*

(5) $1/[Id]$ *in* ${}^*\mathcal{R}$ *is a positive infinitesimal;*

(6) *A number* $N \in {}^*\mathbb{N} \setminus \mathbb{N}$ *is called a hyperfinite integer. For example,* $[Id]$ *is a hyperfinite integer. A hyperfinite integer is infinitely large from the standard point of view, but is finite from nonstandard point of view.*

Example (1.11)

(4) *If* $Id \in \mathcal{R}^{\mathbb{N}}$ *is the identity function, i.e.,* $Id(n) = n$ *for every* $n \in \mathbb{N}$*, then* $[Id] \in {}^*\mathcal{R}$ *and* $[Id] > r$ *for every* $r \in \mathbb{R}$*. So,* ${}^*\mathcal{R}$ *contains numbers larger than every* $r \in \mathbb{R}$*.*

(5) $1/[Id]$ *in* ${}^*\mathcal{R}$ *is a positive infinitesimal;*

(6) *A number* $N \in {}^*\mathbb{N} \setminus \mathbb{N}$ *is called a hyperfinite integer. For example,* $[Id]$ *is a hyperfinite integer. A hyperfinite integer is infinitely large from the standard point of view, but is finite from nonstandard point of view.*

Example (1.11)

(4) *If* $Id \in \mathcal{R}^{\mathbb{N}}$ *is the identity function, i.e.,* $Id(n) = n$ *for every* $n \in \mathbb{N}$*, then* $[Id] \in {}^*\mathcal{R}$ *and* $[Id] > r$ *for every* $r \in \mathbb{R}$*. So,* ${}^*\mathcal{R}$ *contains numbers larger than every* $r \in \mathbb{R}$*.*

(5) $1/[Id]$ *in* ${}^*\mathcal{R}$ *is a positive infinitesimal;*

(6) *A number* $N \in {}^*\mathbb{N} \setminus \mathbb{N}$ *is called a hyperfinite integer. For example,* $[Id]$ *is a hyperfinite integer. A hyperfinite integer is infinitely large from the standard point of view, but is finite from nonstandard point of view.*

Proposition (1.12)

A real number $r \in {}^*\mathbb{R}$ is called near standard if $|r| \leq \alpha$ for some $\alpha \in \mathbb{R}$. If r is near standard, then there exists a unique $\beta \in \mathbb{R}$ such that $r \approx \beta$.

Proof. Let $S = \{\gamma \in \mathbb{R} \mid \gamma < r\}$. Then the set $S \subseteq \mathbb{R}$ is bounded above by α. By the completeness property S has a least upper bound β. It is easy to check that $r \approx \beta$. The uniqueness follows from the fact that two distinct standard reals can never be infinitesimally close. □

Proposition (1.12)

A real number $r \in {}^*\mathbb{R}$ is called near standard if $|r| \leq \alpha$ for some $\alpha \in \mathbb{R}$. If r is near standard, then there exists a unique $\beta \in \mathbb{R}$ such that $r \approx \beta$.

Proof: Let $S = \{\gamma \in \mathbb{R} \mid \gamma < r\}$. Then the set $S \subseteq \mathbb{R}$ is bounded above by α. By the completeness property S has a least upper bound β. It is easy to check that $r \approx \beta$. The uniqueness follows from the fact that two distinct standard reals can never be infinitesimally close. □

Example (1.13)

Let S be the set of all infinitesimals in ${}^*\mathcal{R}$. Then S is nonempty and bounded above by 1. Note that S does not have a least upper bound. Indeed, if $\beta > 0$ were the least upper bound of S, then β being infinitesimal would imply 2β being also an infinitesimal which violates β being upper bound of S, and β being non-infinitesimal would imply $\beta/2$ being also a non-infinitesimal which violates β being the least. Either way we have a contradiction.

The example above shows that $\mathcal{R}$ and ${}^*\mathcal{R}$ may not share the same truth beyond the first-order.

Besides the transfer principle, the standard part map is another way to connect ${}^*\mathcal{R}$ to $\mathcal{R}$.

Example (1.13)

Let S be the set of all infinitesimals in $^*\mathcal{R}$. Then S is nonempty and bounded above by 1. Note that S does not have a least upper bound. Indeed, if $\beta > 0$ were the least upper bound of S, then β being infinitesimal would imply 2β being also an infinitesimal which violates β being upper bound of S, and β being non-infinitesimal would imply $\beta/2$ being also a non-infinitesimal which violates β being the least. Either way we have a contradiction.

The example above shows that $\mathcal{R}$ and $^*\mathcal{R}$ may not share the same truth beyond the first-order.

Besides the transfer principle, the standard part map is another way to connect $^*\mathcal{R}$ to $\mathcal{R}$.

Example (1.13)

Let S be the set of all infinitesimals in ${}^*\mathcal{R}$. Then S is nonempty and bounded above by 1. Note that S does not have a least upper bound. Indeed, if $\beta > 0$ were the least upper bound of S, then β being infinitesimal would imply 2β being also an infinitesimal which violates β being upper bound of S, and β being non-infinitesimal would imply $\beta/2$ being also a non-infinitesimal which violates β being the least. Either way we have a contradiction.

The example above shows that $\mathcal{R}$ and ${}^*\mathcal{R}$ may not share the same truth beyond the first-order.

Besides the transfer principle, the standard part map is another way to connect ${}^*\mathcal{R}$ to $\mathcal{R}$.

Definition (1.14)

Let $ns({}^*\mathcal{R})$ be the set of all near standard reals in ${}^*\mathcal{R}$. We define the standard part map *$st : {}^*\mathcal{R} \to \mathbb{R} \cup \{\pm\}$ by letting $st(r) = \alpha$ for every $r \in ns({}^*\mathcal{R})$ where α is the unique number in $\mathbb{R}$ such that $r \approx \alpha$, $st(r) = \infty$ if $r > \alpha$ for every $\alpha \in \mathbb{R}$, and $st(r) = -\infty$ if $r < \alpha$ for every $\alpha \in \mathbb{R}$.*

We would like to present very simple applications of nonstandard analysis to calculus. Note that the arguments in these applications avoid the use of limit process.

Definition (1.15)

Let $s : \mathbb{N} \to \mathbb{R}$ be a standard sequence. The sequence s is convergent *if ${}^*s(N) \approx {}^*s(N')$ for any hyperfinite integers N, N'.*

Definition (1.14)

Let $ns({}^*\mathcal{R})$ be the set of all near standard reals in ${}^*\mathcal{R}$. We define the standard part map $st : {}^*\mathcal{R} \to \mathbb{R} \cup \{\pm\}$ by letting $st(r) = \alpha$ for every $r \in ns({}^*\mathcal{R})$ where α is the unique number in $\mathbb{R}$ such that $r \approx \alpha$, $st(r) = \infty$ if $r > \alpha$ for every $\alpha \in \mathbb{R}$, and $st(r) = -\infty$ if $r < \alpha$ for every $\alpha \in \mathbb{R}$.

We would like to present very simple applications of nonstandard analysis to calculus. Note that the arguments in these applications avoid the use of limit process.

Definition (1.15)

Let $s : \mathbb{N} \to \mathbb{R}$ be a standard sequence. The sequence s is convergent if ${}^*s(N) \approx {}^*s(N')$ for any hyperfinite integers N, N'.

Definition (1.14)

Let $ns({}^*\mathcal{R})$ be the set of all near standard reals in ${}^*\mathcal{R}$. We define the standard part map $st : {}^*\mathcal{R} \to \mathbb{R} \cup \{\pm\}$ by letting $st(r) = \alpha$ for every $r \in ns({}^*\mathcal{R})$ where α is the unique number in $\mathbb{R}$ such that $r \approx \alpha$, $st(r) = \infty$ if $r > \alpha$ for every $\alpha \in \mathbb{R}$, and $st(r) = -\infty$ if $r < \alpha$ for every $\alpha \in \mathbb{R}$.

We would like to present very simple applications of nonstandard analysis to calculus. Note that the arguments in these applications avoid the use of limit process.

Definition (1.15)

Let $s : \mathbb{N} \to \mathbb{R}$ be a standard sequence. The sequence s is convergent if ${}^*s(N) \approx {}^*s(N')$ for any hyperfinite integers N, N'.

Theorem (1.16, Bolzano–Weierstrass)

Every standard bounded sequence contains a convergent subsequence.

Proof: Suppose s is the bounded sequence in $[a, b]$. Let $N \in {}^*\mathbb{N} \setminus \mathbb{N}$. Then ${}^*s(N) \in ns({}^*\mathbb{R})$. Let $L = st({}^*s(N))$. We show that there exists a subsequence s' of s such that s' converges to L.

Theorem (1.16, Bolzano–Weierstrass)

Every standard bounded sequence contains a convergent subsequence.

Proof: Suppose s is the bounded sequence in $[a, b]$. Let $N \in {}^*\mathbb{N} \setminus \mathbb{N}$. Then ${}^*s(N) \in ns({}^*\mathbb{R})$. Let $L = st({}^*s(N))$. We show that there exists a subsequence s' of s such that s' converges to L.

For each $n \in \mathbb{N}$, if there is an $m_n \in \mathbb{N}$ such that $s(m) \notin (L - 1/n, L + 1/n)$ for all $m \geq m_n$ in $\mathbb{N}$, then ${}^*s(m) \notin {}^*(L - 1/n, L + 1/n)$ for any $m \geq m_n$ in ${}^*\mathbb{N}$ by the transfer principle, which contradicts $st({}^*s(N)) = L$ because $N \geq m_n$.

Hence, $(L - 1/n, L + 1/n)$ contains infinitely many terms of s for each $n \in \mathbb{N}$. So, one can choose $m_1 < m_2 < \cdots$ such that $s(m_n) \in (L - 1/n, L + 1/n)$.

Now for any hyperfinite integers $N < N'$ we have ${}^*s(m_N), {}^*s(m_{N'}) \in {}^*(L - 1/N, L + 1/N)$. Hence, ${}^*s(m_N) \approx {}^*s(m_{N'})$.

This shows that the subsequence $s(m_1), s(m_2), \ldots$ is a convergent subsequence of s. □

For each $n \in \mathbb{N}$, if there is an $m_n \in \mathbb{N}$ such that $s(m) \notin (L - 1/n, L + 1/n)$ for all $m \geq m_n$ in $\mathbb{N}$, then ${}^*s(m) \notin {}^*(L - 1/n, L + 1/n)$ for any $m \geq m_n$ in ${}^*\mathbb{N}$ by the transfer principle, which contradicts $st({}^*s(N)) = L$ because $N \geq m_n$.

Hence, $(L - 1/n, L + 1/n)$ contains infinitely many terms of s for each $n \in \mathbb{N}$. So, one can choose $m_1 < m_2 < \cdots$ such that $s(m_n) \in (L - 1/n, L + 1/n)$.

Now for any hyperfinite integers $N < N'$ we have ${}^*s(m_N), {}^*s(m_{N'}) \in {}^*(L - 1/N, L + 1/N)$. Hence, ${}^*s(m_N) \approx {}^*s(m_{N'})$.

This shows that the subsequence $s(m_1), s(m_2), \ldots$ is a convergent subsequence of s. $\square$

For each $n \in \mathbb{N}$, if there is an $m_n \in \mathbb{N}$ such that $s(m) \notin (L - 1/n, L + 1/n)$ for all $m \geq m_n$ in $\mathbb{N}$, then ${}^*s(m) \notin {}^*(L - 1/n, L + 1/n)$ for any $m \geq m_n$ in ${}^*\mathbb{N}$ by the transfer principle, which contradicts $st({}^*s(N)) = L$ because $N \geq m_n$.

Hence, $(L - 1/n, L + 1/n)$ contains infinitely many terms of s for each $n \in \mathbb{N}$. So, one can choose $m_1 < m_2 < \cdots$ such that $s(m_n) \in (L - 1/n, L + 1/n)$.

Now for any hyperfinite integers $N < N'$ we have ${}^*s(m_N), {}^*s(m_{N'}) \in {}^*(L - 1/N, L + 1/N)$. Hence, ${}^*s(m_N) \approx {}^*s(m_{N'})$.

This shows that the subsequence $s(m_1), s(m_2), \ldots$ is a convergent subsequence of s. $\square$

For each $n \in \mathbb{N}$, if there is an $m_n \in \mathbb{N}$ such that $s(m) \notin (L - 1/n, L + 1/n)$ for all $m \geq m_n$ in $\mathbb{N}$, then ${}^*s(m) \notin {}^*(L - 1/n, L + 1/n)$ for any $m \geq m_n$ in ${}^*\mathbb{N}$ by the transfer principle, which contradicts $st({}^*s(N)) = L$ because $N \geq m_n$.

Hence, $(L - 1/n, L + 1/n)$ contains infinitely many terms of s for each $n \in \mathbb{N}$. So, one can choose $m_1 < m_2 < \cdots$ such that $s(m_n) \in (L - 1/n, L + 1/n)$.

Now for any hyperfinite integers $N < N'$ we have ${}^*s(m_N), {}^*s(m_{N'}) \in {}^*(L - 1/N, L + 1/N)$. Hence, ${}^*s(m_N) \approx {}^*s(m_{N'})$.

This shows that the subsequence $s(m_1), s(m_2), \ldots$ is a convergent subsequence of s. $\square$

Definition (1.17)

Let $f : [a, b] \to \mathbb{R}$ be a standard function. Then,

1. *f is continuous at $c \in [a, b]$ if for any $r \in {}^*[a, b]$ we have $r \approx c$ implies ${}^*f(r) \approx f(c)$; f is continuous on $[a, b]$ if f is continuous at every $c \in [a, b]$;*
2. *f is uniformly continuous on $[a, b]$ if $r_1 \approx r_2$ implies ${}^*f(r_1) \approx {}^*f(r_2)$ for any $r_1, r_2 \in {}^*[a, b]$.*

Definition (1.17)

Let $f : [a, b] \to \mathbb{R}$ be a standard function. Then,

1. *f is continuous at $c \in [a, b]$ if for any $r \in {}^*[a, b]$ we have $r \approx c$ implies ${}^*f(r) \approx f(c)$; f is continuous on $[a, b]$ if f is continuous at every $c \in [a, b]$;*
2. *f is uniformly continuous on $[a, b]$ if $r_1 \approx r_2$ implies ${}^*f(r_1) \approx {}^*f(r_2)$ for any $r_1, r_2 \in {}^*[a, b]$.*

Definition (1.17)

Let $f : [a, b] \to \mathbb{R}$ be a standard function. Then,

1. *f is continuous at $c \in [a, b]$ if for any $r \in {}^*[a, b]$ we have $r \approx c$ implies ${}^*f(r) \approx f(c)$; f is continuous on $[a, b]$ if f is continuous at every $c \in [a, b]$;*
2. *f is uniformly continuous on $[a, b]$ if $r_1 \approx r_2$ implies ${}^*f(r_1) \approx {}^*f(r_2)$ for any $r_1, r_2 \in {}^*[a, b]$.*

Theorem (1.18)

If a standard function $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, then f is uniformly continuous on $[a, b]$.

Proof: Suppose that f is continuous on $[a, b]$ but not uniformly continuous on $[a, b]$. Then, there exist $r_1, r_2 \in {}^*[a, b]$ such that $r_1 \approx r_2$ but ${}^*f(r_1) \not\approx {}^*f(r_2)$.

Since $r_1 \approx r_2$ we have $st(r_1) = st(r_2) = c \in [a, b]$. Since $r_1 \approx c \approx r_2$, then ${}^*f(r_1) \approx f(c) \approx {}^*f(r_2)$, which contradicts the assumption that ${}^*f(r_1) \not\approx {}^*f(r_2)$. □

Theorem (1.18)

If a standard function $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, then f is uniformly continuous on $[a, b]$.

Proof: Suppose that f is continuous on $[a, b]$ but not uniformly continuous on $[a, b]$. Then, there exist $r_1, r_2 \in {}^*[a, b]$ such that $r_1 \approx r_2$ but ${}^*f(r_1) \not\approx {}^*f(r_2)$.

Since $r_1 \approx r_2$ we have $st(r_1) = st(r_2) = c \in [a, b]$. Since $r_1 \approx c \approx r_2$, then ${}^*f(r_1) \approx f(c) \approx {}^*f(r_2)$, which contradicts the assumption that ${}^*f(r_1) \not\approx {}^*f(r_2)$. □

Theorem (1.18)

If a standard function $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, then f is uniformly continuous on $[a, b]$.

Proof: Suppose that f is continuous on $[a, b]$ but not uniformly continuous on $[a, b]$. Then, there exist $r_1, r_2 \in {}^*[a, b]$ such that $r_1 \approx r_2$ but ${}^*f(r_1) \not\approx {}^*f(r_2)$.

Since $r_1 \approx r_2$ we have $st(r_1) = st(r_2) = c \in [a, b]$. Since $r_1 \approx c \approx r_2$, then ${}^*f(r_1) \approx f(c) \approx {}^*f(r_2)$, which contradicts the assumption that ${}^*f(r_1) \not\approx {}^*f(r_2)$. □

Definition (1.19)

Let $f : (a, b) \to \mathbb{R}$ be a standard function and $c \in (a, b)$. The function f is differentiable at c if there exists an $\alpha \in \mathbb{R}$ such that

$$f'(c) := st\left(\frac{{}^*f(r) - f(c)}{r - c}\right) = \alpha$$

for any $r \in {}^*\mathbb{R}$ with $r \approx c$ and $r \neq c$.

Given a function $f : X \to Y$, $a \in X$, and $A \subseteq X$, we write $f(a)$ for some element in Y, and write $f[A]$ for the set $\{f(a) \mid a \in A\}$.

Theorem (1.20, Chain Rule)

If the standard function $f : (a, b) \to \mathbb{R}$ is differentiable at $c \in (a, b)$, $f[(a, b)] \subseteq (\alpha, \beta)$, and a standard function $g : (\alpha, \beta) \to \mathbb{R}$ is differentiable at $f(c)$, then $g(f(x)) : (a, b) \to \mathbb{R}$ is differentiable at c and $(g(f(x)))'_c = g'(f(c))f'(c)$.

Definition (1.19)

Let $f : (a,b) \to \mathbb{R}$ *be a standard function and* $c \in (a,b)$. *The function* f *is differentiable at* c *if there exists an* $\alpha \in \mathbb{R}$ *such that*

$$f'(c) := st\left(\frac{{}^*f(r) - f(c)}{r - c}\right) = \alpha$$

for any $r \in {}^*\mathbb{R}$ *with* $r \approx c$ *and* $r \neq c$.

Given a function $f : X \to Y$, $a \in X$, and $A \subseteq X$, we write $f(a)$ for some element in Y, and write $f[A]$ for the set $\{f(a) \mid a \in A\}$.

Theorem (1.20, Chain Rule)

If the standard function $f : (a,b) \to \mathbb{R}$ *is differentiable at* $c \in (a,b)$, $f[(a,b)] \subseteq (\alpha,\beta)$, *and a standard function* $g : (\alpha,\beta) \to \mathbb{R}$ *is differentiable at* $f(c)$, *then* $g(f(x)) : (a,b) \to \mathbb{R}$ *is differentiable at* c *and* $(g(f(x)))'_c = g'(f(c))f'(c)$.

Definition (1.19)

Let $f : (a, b) \to \mathbb{R}$ be a standard function and $c \in (a, b)$. The function f is differentiable at c if there exists an $\alpha \in \mathbb{R}$ such that

$$f'(c) := st\left(\frac{{}^*f(r) - f(c)}{r - c}\right) = \alpha$$

for any $r \in {}^*\mathbb{R}$ with $r \approx c$ and $r \neq c$.

Given a function $f : X \to Y$, $a \in X$, and $A \subseteq X$, we write $f(a)$ for some element in Y, and write $f[A]$ for the set $\{f(a) \mid a \in A\}$.

Theorem (1.20, Chain Rule)

If the standard function $f : (a, b) \to \mathbb{R}$ is differentiable at $c \in (a, b)$, $f[(a, b)] \subseteq (\alpha, \beta)$, and a standard function $g : (\alpha, \beta) \to \mathbb{R}$ is differentiable at $f(c)$, then $g(f(x)) : (a, b) \to \mathbb{R}$ is differentiable at c and $(g(f(x)))'_c = g'(f(c))f'(c)$.

Proof of Chain Rule: Given any $r \approx c$ and $r \neq c$. Since f is differentiable at c, there is an infinitesimal t_1 such that that ${}^*f(r) - f(c) = (f'(c) + t_1)(r - c)$. Since g is differentiable at $f(c)$ and ${}^*f(r) \approx f(c)$ there is another infinitesimal t_2 such that ${}^*(g(f(r)) - g(f(c)) = (g'(f(c)) + t_2)({}^*f(r) - f(c))$. Hence, we have

$$\begin{aligned} {}^*(g(f(r)) - g(f(c))) &= (g'(f(c)) + t_2)({}^*f(r) - f(c)) \\ &= (g'(f(c)) + t_2)(f'(c) + t_1)(r - c) \\ &= (g'(f(c))f'(c) + t_2 f'(c) + t_1 g'(f(c)) + t_2 t_1)(r - c), \end{aligned}$$

which implies

$$\begin{aligned} st\left(\frac{g(f(r)) - g(f(c))}{r - c}\right) \\ &= st\left(g'(f(c))f'(c) + t_2 f'(c) + t_1 g'(f(c)) + t_2 t_1\right) \\ &= g'(f(c))f'(c). \end{aligned}$$

Proof of Chain Rule: Given any $r \approx c$ and $r \neq c$. Since f is differentiable at c, there is an infinitesimal t_1 such that that ${}^*f(r) - f(c) = (f'(c) + t_1)(r - c)$. Since g is differentiable at $f(c)$ and ${}^*f(r) \approx f(c)$ there is another infinitesimal t_2 such that ${}^*(g(f(r)) - g(f(c)) = (g'(f(c)) + t_2)({}^*f(r) - f(c))$. Hence, we have

$$\begin{aligned} {}^*(g(f(r)) - g(f(c))) &= (g'(f(c)) + t_2)({}^*f(r) - f(c)) \\ &= (g'(f(c)) + t_2)(f'(c) + t_1)(r - c) \\ &= (g'(f(c))f'(c) + t_2 f'(c) + t_1 g'(f(c)) + t_2 t_1)(r - c), \end{aligned}$$

which implies

$$\begin{aligned} & st\left(\frac{g(f(r)) - g(f(c))}{r - c}\right) \\ &= st\left(g'(f(c))f'(c) + t_2 f'(c) + t_1 g'(f(c)) + t_2 t_1\right) \\ &= g'(f(c))f'(c). \end{aligned}$$

Proof of Chain Rule: Given any $r \approx c$ and $r \neq c$. Since f is differentiable at c, there is an infinitesimal t_1 such that that ${}^*f(r) - f(c) = (f'(c) + t_1)(r - c)$. Since g is differentiable at $f(c)$ and ${}^*f(r) \approx f(c)$ there is another infinitesimal t_2 such that ${}^*(g(f(r)) - g(f(c)) = (g'(f(c)) + t_2)({}^*f(r) - f(c))$. Hence, we have

$$\begin{aligned} {}^*(g(f(r)) - g(f(c))) &= (g'(f(c)) + t_2)({}^*f(r) - f(c)) \\ &= (g'(f(c)) + t_2)(f'(c) + t_1)(r - c) \\ &= (g'(f(c))f'(c) + t_2 f'(c) + t_1 g'(f(c)) + t_2 t_1)(r - c), \end{aligned}$$

which implies

$$\begin{aligned} & st\left(\frac{g(f(r)) - g(f(c))}{r - c}\right) \\ &= st\left(g'(f(c))f'(c) + t_2 f'(c) + t_1 g'(f(c)) + t_2 t_1\right) \\ &= g'(f(c))f'(c). \end{aligned}$$

Proof of Chain Rule: Given any $r \approx c$ and $r \neq c$. Since f is differentiable at c, there is an infinitesimal t_1 such that that ${}^*f(r) - f(c) = (f'(c) + t_1)(r - c)$. Since g is differentiable at $f(c)$ and ${}^*f(r) \approx f(c)$ there is another infinitesimal t_2 such that ${}^*(g(f(r)) - g(f(c)) = (g'(f(c)) + t_2)({}^*f(r) - f(c))$. Hence, we have

$$\begin{aligned} {}^*(g(f(r)) - g(f(c))) &= (g'(f(c)) + t_2)({}^*f(r) - f(c)) \\ &= (g'(f(c)) + t_2)(f'(c) + t_1)(r - c) \\ &= (g'(f(c))f'(c) + t_2 f'(c) + t_1 g'(f(c)) + t_2 t_1)(r - c), \end{aligned}$$

which implies

$$\begin{aligned} & st\left(\frac{g(f(r)) - g(f(c))}{r - c}\right) \\ &= st\left(g'(f(c))f'(c) + t_2 f'(c) + t_1 g'(f(c)) + t_2 t_1\right) \\ &= g'(f(c))f'(c). \end{aligned}$$

To deal with integration we encounter the integral operator which is a linear functional. Therefore, it cannot be handled in ${}^*\mathcal{R}$. We need a structure not only containing functions from $\mathbb{R}^n$ to $\mathbb{R}$ but also containing the functions of the functions, the functions of the functions of the functions, etc. This is why we introduce another model called superstructure to deal with this and many other needs in the next subsection.

The use of superstructure and its elementary extension as the model of nonstandard analysis started by Robinson and Zakon.

Fix a sufficiently large positive integer $\mathfrak{n}$, say $\mathfrak{n} = 100$. Let $\mathscr{L} := \{\in\}$ contain only one binary relation symbol. Given an infinite set X of urelements, i.e., elements without members, the *superstructure* on X, denoted by $\mathcal{V}(X)$, is an $\mathscr{L}$-model $(V(X); \in)$ where $V(X)$ is defined inductively by letting

$$V(X, 0) := X, \ \ V(X, n+1) := V(X, n) \cup \mathscr{P}(V(X, n))$$

for every $n < 2\mathfrak{n}$, $V(X) = V(X; 2\mathfrak{n})$, and letting $\in$ be the true set theoretic membership relation on $V(X)$ as the interpretation of the symbol $\in$ in $\mathscr{L}$. For notational convenience, we don't distinguish $\mathcal{V}(X)$ for the model from the base set of the model. We write also $\mathcal{V}(X, n)$ for both $(V(X, n); \in)$ and $V(X, n)$.

The use of superstructure and its elementary extension as the model of nonstandard analysis started by Robinson and Zakon.

Fix a sufficiently large positive integer $\mathfrak{n}$, say $\mathfrak{n} = 100$. Let $\mathscr{L} := \{\in\}$ contain only one binary relation symbol. Given an infinite set X of urelements, i.e., elements without members, the *superstructure* on X, denoted by $\mathcal{V}(X)$, is an $\mathscr{L}$-model $(V(X); \in)$ where $V(X)$ is defined inductively by letting

$$V(X,0) := X,\ \ V(X,n+1) := V(X,n) \cup \mathscr{P}(V(X,n))$$

for every $n < 2\mathfrak{n}$, $V(X) = V(X; 2\mathfrak{n})$, and letting $\in$ be the true set theoretic membership relation on $V(X)$ as the interpretation of the symbol $\in$ in $\mathscr{L}$. For notational convenience, we don't distinguish $\mathcal{V}(X)$ for the model from the base set of the model. We write also $\mathcal{V}(X,n)$ for both $(V(X,n); \in)$ and $V(X,n)$.

The use of superstructure and its elementary extension as the model of nonstandard analysis started by Robinson and Zakon.

Fix a sufficiently large positive integer $\mathfrak{n}$, say $\mathfrak{n} = 100$. Let $\mathscr{L} := \{\in\}$ contain only one binary relation symbol. Given an infinite set X of urelements, i.e., elements without members, the *superstructure* on X, denoted by $\mathcal{V}(X)$, is an $\mathscr{L}$-model $(V(X); \in)$ where $V(X)$ is defined inductively by letting

$$V(X, 0) := X, \;\; V(X, n+1) := V(X, n) \cup \mathscr{P}(V(X, n))$$

for every $n < 2\mathfrak{n}$, $V(X) = V(X; 2\mathfrak{n})$, and letting $\in$ be the true set theoretic membership relation on $V(X)$ as the interpretation of the symbol $\in$ in $\mathscr{L}$. For notational convenience, we don't distinguish $\mathcal{V}(X)$ for the model from the base set of the model. We write also $\mathcal{V}(X, n)$ for both $(V(X, n); \in)$ and $V(X, n)$.

The use of superstructure and its elementary extension as the model of nonstandard analysis started by Robinson and Zakon.

Fix a sufficiently large positive integer $\mathfrak{n}$, say $\mathfrak{n} = 100$. Let $\mathscr{L} := \{\in\}$ contain only one binary relation symbol. Given an infinite set X of urelements, i.e., elements without members, the *superstructure* on X, denoted by $\mathcal{V}(X)$, is an $\mathscr{L}$-model $(V(X); \in)$ where $V(X)$ is defined inductively by letting

$$V(X, 0) := X, \;\; V(X, n+1) := V(X, n) \cup \mathscr{P}(V(X, n))$$

for every $n < 2\mathfrak{n}$, $V(X) = V(X; 2\mathfrak{n})$, and letting $\in$ be the true set theoretic membership relation on $V(X)$ as the interpretation of the symbol $\in$ in $\mathscr{L}$. For notational convenience, we don't distinguish $\mathcal{V}(X)$ for the model from the base set of the model. We write also $\mathcal{V}(X, n)$ for both $(V(X, n); \in)$ and $V(X, n)$.

The use of superstructure and its elementary extension as the model of nonstandard analysis started by Robinson and Zakon.

Fix a sufficiently large positive integer $\mathfrak{n}$, say $\mathfrak{n} = 100$. Let $\mathscr{L} := \{\in\}$ contain only one binary relation symbol. Given an infinite set X of urelements, i.e., elements without members, the *superstructure* on X, denoted by $\mathcal{V}(X)$, is an $\mathscr{L}$-model $(V(X); \in)$ where $V(X)$ is defined inductively by letting

$$V(X,0) := X, \;\; V(X,n+1) := V(X,n) \cup \mathscr{P}(V(X,n))$$

for every $n < 2\mathfrak{n}$, $V(X) = V(X; 2\mathfrak{n})$, and letting $\in$ be the true set theoretic membership relation on $V(X)$ as the interpretation of the symbol $\in$ in $\mathscr{L}$. For notational convenience, we don't distinguish $\mathcal{V}(X)$ for the model from the base set of the model. We write also $\mathcal{V}(X,n)$ for both $(V(X,n); \in)$ and $V(X,n)$.

The use of superstructure and its elementary extension as the model of nonstandard analysis started by Robinson and Zakon.

Fix a sufficiently large positive integer $\mathfrak{n}$, say $\mathfrak{n} = 100$. Let $\mathscr{L} := \{\in\}$ contain only one binary relation symbol. Given an infinite set X of urelements, i.e., elements without members, the *superstructure* on X, denoted by $\mathcal{V}(X)$, is an $\mathscr{L}$-model $(V(X); \in)$ where $V(X)$ is defined inductively by letting

$$V(X,0) := X, \ \ V(X,n+1) := V(X,n) \cup \mathscr{P}(V(X,n))$$

for every $n < 2\mathfrak{n}$, $V(X) = V(X; 2\mathfrak{n})$, and letting $\in$ be the true set theoretic membership relation on $V(X)$ as the interpretation of the symbol $\in$ in $\mathscr{L}$. For notational convenience, we don't distinguish $\mathcal{V}(X)$ for the model from the base set of the model. We write also $\mathcal{V}(X,n)$ for both $(V(X,n); \in)$ and $V(X,n)$.

The use of superstructure and its elementary extension as the model of nonstandard analysis started by Robinson and Zakon.

Fix a sufficiently large positive integer $\mathfrak{n}$, say $\mathfrak{n} = 100$. Let $\mathscr{L} := \{\in\}$ contain only one binary relation symbol. Given an infinite set X of urelements, i.e., elements without members, the *superstructure* on X, denoted by $\mathcal{V}(X)$, is an $\mathscr{L}$-model $(V(X); \in)$ where $V(X)$ is defined inductively by letting

$$V(X,0) := X,\ \ V(X,n+1) := V(X,n) \cup \mathscr{P}(V(X,n))$$

for every $n < 2\mathfrak{n}$, $V(X) = V(X; 2\mathfrak{n})$, and letting $\in$ be the true set theoretic membership relation on $V(X)$ as the interpretation of the symbol $\in$ in $\mathscr{L}$. For notational convenience, we don't distinguish $\mathcal{V}(X)$ for the model from the base set of the model. We write also $\mathcal{V}(X,n)$ for both $(V(X,n); \in)$ and $V(X,n)$.

For each element $a \in \mathcal{V}(X)$ we define the rank of a, denoted by $\mathrm{rank}(a)$, by that if $a \in X$, then $\mathrm{rank}(a) = 0$ and if $a \in \mathcal{V}(X, n+1) \setminus \mathcal{V}(X, n)$ for some number $n < 2\mathfrak{n}$, then $\mathrm{rank}(a) = n + 1$. The rank function on $\mathcal{V}(X)$ is bounded by $2\mathfrak{n}$ and is definable in $\mathcal{V}(X)$ by a first-order formula.

We assume always that $\mathbb{N} \subseteq \mathbb{R} \subseteq X$. For simplicity, set $X = \mathbb{R}$. Note that all standard mathematical objects mentioned in the lecture notes have ranks below $\mathfrak{n} = 100$. We set the highest rank to be $2\mathfrak{n}$ instead of $\mathfrak{n}$ for convenience.

Note that an ordered pair (a, b) of real numbers $a, b \in \mathbb{R}$ can be viewed as the set $\{\{a\}, \{a, b\}\} \in \mathcal{V}(\mathbb{R}, 2)$ and a function $f : D \subseteq \mathbb{R} \to \mathbb{R}$ can be viewed as a set of ordered pairs in $\mathcal{V}(\mathbb{R}, 2)$. Hence, $f \in \mathcal{V}(\mathbb{R}, 3)$. A linear functional L on functions from $\mathbb{R}$ to $\mathbb{R}$ is a set of pairs $(f, r) = \{\{f\}, \{f, r\}\} \in \mathcal{V}(\mathbb{R}, 5)$. Hence, $L \in \mathcal{V}(\mathbb{R}, 6)$. Note also that the ultrafilter $\mathcal{F}$ on $\mathbb{N}$ is in $\mathcal{V}(\mathbb{R}, 3)$.

For each element $a \in \mathcal{V}(X)$ we define the rank of a, denoted by $\mathrm{rank}(a)$, by that if $a \in X$, then $\mathrm{rank}(a) = 0$ and if $a \in \mathcal{V}(X, n+1) \setminus \mathcal{V}(X, n)$ for some number $n < 2\mathfrak{n}$, then $\mathrm{rank}(a) = n + 1$. The rank function on $\mathcal{V}(X)$ is bounded by $2\mathfrak{n}$ and is definable in $\mathcal{V}(X)$ by a first-order formula.

We assume always that $\mathbb{N} \subseteq \mathbb{R} \subseteq X$. For simplicity, set $X = \mathbb{R}$. Note that all standard mathematical objects mentioned in the lecture notes have ranks below $\mathfrak{n} = 100$. We set the highest rank to be $2\mathfrak{n}$ instead of $\mathfrak{n}$ for convenience.

Note that an ordered pair (a, b) of real numbers $a, b \in \mathbb{R}$ can be viewed as the set $\{\{a\}, \{a, b\}\} \in \mathcal{V}(\mathbb{R}, 2)$ and a function $f : D \subseteq \mathbb{R} \to \mathbb{R}$ can be viewed as a set of ordered pairs in $\mathcal{V}(\mathbb{R}, 2)$. Hence, $f \in \mathcal{V}(\mathbb{R}, 3)$. A linear functional L on functions from $\mathbb{R}$ to $\mathbb{R}$ is a set of pairs $(f, r) = \{\{f\}, \{f, r\}\} \in \mathcal{V}(\mathbb{R}, 5)$. Hence, $L \in \mathcal{V}(\mathbb{R}, 6)$. Note also that the ultrafilter $\mathcal{F}$ on $\mathbb{N}$ is in $\mathcal{V}(\mathbb{R}, 3)$.

For each element $a \in \mathcal{V}(X)$ we define the rank of a, denoted by rank(a), by that if $a \in X$, then rank$(a) = 0$ and if $a \in \mathcal{V}(X, n+1) \setminus \mathcal{V}(X, n)$ for some number $n < 2\mathfrak{n}$, then rank$(a) = n + 1$. The rank function on $\mathcal{V}(X)$ is bounded by $2\mathfrak{n}$ and is definable in $\mathcal{V}(X)$ by a first-order formula.

We assume always that $\mathbb{N} \subseteq \mathbb{R} \subseteq X$. For simplicity, set $X = \mathbb{R}$. Note that all standard mathematical objects mentioned in the lecture notes have ranks below $\mathfrak{n} = 100$. We set the highest rank to be $2\mathfrak{n}$ instead of $\mathfrak{n}$ for convenience.

Note that an ordered pair (a, b) of real numbers $a, b \in \mathbb{R}$ can be viewed as the set $\{\{a\}, \{a, b\}\} \in \mathcal{V}(\mathbb{R}, 2)$ and a function $f : D \subseteq \mathbb{R} \to \mathbb{R}$ can be viewed as a set of ordered pairs in $\mathcal{V}(\mathbb{R}, 2)$. Hence, $f \in \mathcal{V}(\mathbb{R}, 3)$. A linear functional L on functions from $\mathbb{R}$ to $\mathbb{R}$ is a set of pairs $(f, r) = \{\{f\}, \{f, r\}\} \in \mathcal{V}(\mathbb{R}, 5)$. Hence, $L \in \mathcal{V}(\mathbb{R}, 6)$. Note also that the ultrafilter $\mathcal{F}$ on $\mathbb{N}$ is in $\mathcal{V}(\mathbb{R}, 3)$.

For each element $a \in \mathcal{V}(X)$ we define the rank of a, denoted by $\text{rank}(a)$, by that if $a \in X$, then $\text{rank}(a) = 0$ and if $a \in \mathcal{V}(X, n+1) \setminus \mathcal{V}(X, n)$ for some number $n < 2\mathfrak{n}$, then $\text{rank}(a) = n+1$. The rank function on $\mathcal{V}(X)$ is bounded by $2\mathfrak{n}$ and is definable in $\mathcal{V}(X)$ by a first-order formula.

We assume always that $\mathbb{N} \subseteq \mathbb{R} \subseteq X$. For simplicity, set $X = \mathbb{R}$. Note that all standard mathematical objects mentioned in the lecture notes have ranks below $\mathfrak{n} = 100$. We set the highest rank to be $2\mathfrak{n}$ instead of $\mathfrak{n}$ for convenience.

Note that an ordered pair (a, b) of real numbers $a, b \in \mathbb{R}$ can be viewed as the set $\{\{a\}, \{a, b\}\} \in \mathcal{V}(\mathbb{R}, 2)$ and a function $f : D \subseteq \mathbb{R} \to \mathbb{R}$ can be viewed as a set of ordered pairs in $\mathcal{V}(\mathbb{R}, 2)$. Hence, $f \in \mathcal{V}(\mathbb{R}, 3)$. A linear functional L on functions from $\mathbb{R}$ to $\mathbb{R}$ is a set of pairs $(f, r) = \{\{f\}, \{f, r\}\} \in \mathcal{V}(\mathbb{R}, 5)$. Hence, $L \in \mathcal{V}(\mathbb{R}, 6)$. Note also that the ultrafilter $\mathcal{F}$ on $\mathbb{N}$ is in $\mathcal{V}(\mathbb{R}, 3)$.

The superstructure $\mathcal{V}(\mathbb{R})$ is often called the standard universe, which means that all discussion of ordinary mathematical problems at the moment can be conducted in $\mathcal{V}(\mathbb{R})$. Since X is always $\mathbb{R}$ in the notes we omit $\mathbb{R}$ and write $\mathcal{V}$ for $\mathcal{V}(\mathbb{R})$. Recall that $\mathcal{F}$ is a non-principal ultrafilter on $\mathbb{N}$.

Definition (1.21)

The ultrapower of $\mathcal{V}$ modulo $\mathcal{F}$, denoted by ${}^*\mathcal{V}$, is the model

$$({}^*V; {}^*\in),$$

where the base set is ${}^*V = V(X)^{\mathbb{N}}/\mathcal{F}$ and the interpretation ${}^*\in$ of the binary relation symbol $\in$ is defined by letting $[f] \,{}^*\!\in [g]$ iff $\{n \in \mathbb{N} \mid f(n) \in g(n)\} \in \mathcal{F}$ for any $[f], [g] \in V(X)^{\mathbb{N}}/\mathcal{F}$. Let ${}^* : \mathcal{V} \to {}^*\mathcal{V}$ be the elementary embedding, i.e., ${}^*a := [\phi_a]$ for every $a \in \mathcal{V}$.

The superstructure $\mathcal{V}(\mathbb{R})$ is often called the standard universe, which means that all discussion of ordinary mathematical problems at the moment can be conducted in $\mathcal{V}(\mathbb{R})$. Since X is always $\mathbb{R}$ in the notes we omit $\mathbb{R}$ and write $\mathcal{V}$ for $\mathcal{V}(\mathbb{R})$. Recall that $\mathcal{F}$ is a non-principal ultrafilter on $\mathbb{N}$.

Definition (1.21)

The ultrapower of $\mathcal{V}$ modulo $\mathcal{F}$, denoted by ${}^*\mathcal{V}$, is the model

$$({}^*V; {}^*\!\in),$$

where the base set is ${}^*V = V(X)^{\mathbb{N}}/\mathcal{F}$ and the interpretation ${}^*\!\in$ of the binary relation symbol $\in$ is defined by letting $[f] \,{}^*\!\in [g]$ iff $\{n \in \mathbb{N} \mid f(n) \in g(n)\} \in \mathcal{F}$ for any $[f], [g] \in V(X)^{\mathbb{N}}/\mathcal{F}$. Let ${}^* : \mathcal{V} \to {}^*\mathcal{V}$ be the elementary embedding, i.e., ${}^*a := [\phi_a]$ for every $a \in \mathcal{V}$.

The superstructure $\mathcal{V}(\mathbb{R})$ is often called the standard universe, which means that all discussion of ordinary mathematical problems at the moment can be conducted in $\mathcal{V}(\mathbb{R})$. Since X is always $\mathbb{R}$ in the notes we omit $\mathbb{R}$ and write $\mathcal{V}$ for $\mathcal{V}(\mathbb{R})$. Recall that $\mathcal{F}$ is a non-principal ultrafilter on $\mathbb{N}$.

Definition (1.21)

The ultrapower of $\mathcal{V}$ modulo $\mathcal{F}$, denoted by ${}^*\mathcal{V}$, is the model

$$({}^*V; {}^*\in),$$

where the base set is ${}^*V = V(X)^{\mathbb{N}}/\mathcal{F}$ and the interpretation ${}^*\in$ of the binary relation symbol $\in$ is defined by letting $[f] \,{}^*\!\in [g]$ iff $\{n \in \mathbb{N} \mid f(n) \in g(n)\} \in \mathcal{F}$ for any $[f], [g] \in V(X)^{\mathbb{N}}/\mathcal{F}$. Let ${}^* : \mathcal{V} \to {}^*\mathcal{V}$ be the elementary embedding, i.e., ${}^*a := [\phi_a]$ for every $a \in \mathcal{V}$.

The superstructure $\mathcal{V}(\mathbb{R})$ is often called the standard universe, which means that all discussion of ordinary mathematical problems at the moment can be conducted in $\mathcal{V}(\mathbb{R})$. Since X is always $\mathbb{R}$ in the notes we omit $\mathbb{R}$ and write $\mathcal{V}$ for $\mathcal{V}(\mathbb{R})$. Recall that $\mathcal{F}$ is a non-principal ultrafilter on $\mathbb{N}$.

Definition (1.21)

The ultrapower of $\mathcal{V}$ modulo $\mathcal{F}$, denoted by ${}^*\mathcal{V}$, is the model

$$({}^*V; {}^*\!\in),$$

where the base set is ${}^*V = V(X)^{\mathbb{N}}/\mathcal{F}$ and the interpretation ${}^*\!\in$ of the binary relation symbol $\in$ is defined by letting $[f] \,{}^*\!\!\in [g]$ iff $\{n \in \mathbb{N} \mid f(n) \in g(n)\} \in \mathcal{F}$ for any $[f], [g] \in V(X)^{\mathbb{N}}/\mathcal{F}$. Let ${}^* : \mathcal{V} \to {}^*\mathcal{V}$ be the elementary embedding, i.e., ${}^*a := [\phi_a]$ for every $a \in \mathcal{V}$.

Note that the real ordered field $\mathcal{R}$ is in $\mathcal{V}$. Hence, ${}^*\mathcal{R}$ is in ${}^*\mathcal{V}$.

The model ${}^*\mathcal{V}$ is called a nonstandard universe, or a nonstandard elementary extension of the standard universe $\mathcal{V}$.

One of the advantages of using nonstandard methods is to replace a limit argument, which has a higher set theoretic complexity in the standard model, by an infinitesimal argument, which has a lower set theoretic complexity in a nonstandard model. The reader is encouraged to treat all $r \in {}^*\mathbb{R}$ as urelements instead of equivalence classes of functions from $\mathbb{N}$ to $\mathbb{R}$ to take this advantage and treat ${}^*\in$ as a real membership relation.

Note that the real ordered field $\mathcal{R}$ is in $\mathcal{V}$. Hence, ${}^*\mathcal{R}$ is in ${}^*\mathcal{V}$.

The model ${}^*\mathcal{V}$ is called a nonstandard universe, or a nonstandard elementary extension of the standard universe $\mathcal{V}$.

One of the advantages of using nonstandard methods is to replace a limit argument, which has a higher set theoretic complexity in the standard model, by an infinitesimal argument, which has a lower set theoretic complexity in a nonstandard model. The reader is encouraged to treat all $r \in {}^*\mathbb{R}$ as urelements instead of equivalence classes of functions from $\mathbb{N}$ to $\mathbb{R}$ to take this advantage and treat ${}^*\!\in$ as a real membership relation.

Note that the real ordered field $\mathcal{R}$ is in $\mathcal{V}$. Hence, ${}^*\mathcal{R}$ is in ${}^*\mathcal{V}$.

The model ${}^*\mathcal{V}$ is called a nonstandard universe, or a nonstandard elementary extension of the standard universe $\mathcal{V}$.

One of the advantages of using nonstandard methods is to replace a limit argument, which has a higher set theoretic complexity in the standard model, by an infinitesimal argument, which has a lower set theoretic complexity in a nonstandard model. The reader is encouraged to treat all $r \in {}^*\mathbb{R}$ as urelements instead of equivalence classes of functions from $\mathbb{N}$ to $\mathbb{R}$ to take this advantage and treat ${}^*\in$ as a real membership relation.

Formally, assume that elements in ${}^*\mathbb{R}$ are urelements with rank 0. Let $\mathscr{M}$ be the Mostowski collapsing map on ${}^*\mathcal{V}$, i.e., $\mathscr{M}(a) = a$ for every $a \in {}^*\mathbb{R}$ and

$$\mathscr{M}(b) := \{\mathscr{M}(a) \mid a \,{}^*\!\!\in b\}$$

for every $b \in {}^*\mathcal{V} \setminus {}^*\mathbb{R}$. Then $\mathscr{M}$ is an injection and $a \,{}^*\!\!\in b$ iff $\mathscr{M}(a) \in \mathscr{M}(b)$. If one identifies ${}^*\mathcal{V}$ with the image of ${}^*\mathcal{V}$ under $\mathscr{M}$, one can pretend that ${}^*\!\!\in$ is the true membership relation and consider ${}^*\mathcal{V}$ as a subset of the superstructure $\mathcal{V}({}^*\mathbb{R})$. Hence, we can drop the upper-left superscript $*$ from ${}^*\!\!\in$ for notational convenience.

Similar to the elements in $\mathcal{V}$, the rank function can also be defined for elements in ${}^*\mathcal{V}$. Every element in ${}^*\mathbb{R}$ has rank 0. It is easy to check that every element in ${}^*\mathcal{V}(\mathbb{R}, n+1) \setminus {}^*\mathcal{V}(\mathbb{R}, n)$ has the rank $n+1$ for $n \in [2\mathfrak{n}]$.

Formally, assume that elements in ${}^*\mathbb{R}$ are urelements with rank 0. Let $\mathscr{M}$ be the Mostowski collapsing map on ${}^*\mathcal{V}$, i.e., $\mathscr{M}(a) = a$ for every $a \in {}^*\mathbb{R}$ and

$$\mathscr{M}(b) := \{\mathscr{M}(a) \mid a \,{}^*\!\!\in b\}$$

for every $b \in {}^*\mathcal{V} \setminus {}^*\mathbb{R}$. Then $\mathscr{M}$ is an injection and $a \,{}^*\!\!\in b$ iff $\mathscr{M}(a) \in \mathscr{M}(b)$. If one identifies ${}^*\mathcal{V}$ with the image of ${}^*\mathcal{V}$ under $\mathscr{M}$, one can pretend that ${}^*\!\!\in$ is the true membership relation and consider ${}^*\mathcal{V}$ as a subset of the superstructure $\mathcal{V}({}^*\mathbb{R})$. Hence, we can drop the upper-left superscript $*$ from ${}^*\!\!\in$ for notational convenience.

Similar to the elements in $\mathcal{V}$, the rank function can also be defined for elements in ${}^*\mathcal{V}$. Every element in ${}^*\mathbb{R}$ has rank 0. It is easy to check that every element in ${}^*\mathcal{V}(\mathbb{R}, n+1) \setminus {}^*\mathcal{V}(\mathbb{R}, n)$ has the rank $n+1$ for $n \in [2n]$.

Formally, assume that elements in ${}^*\mathbb{R}$ are urelements with rank 0. Let $\mathscr{M}$ be the Mostowski collapsing map on ${}^*\mathcal{V}$, i.e., $\mathscr{M}(a) = a$ for every $a \in {}^*\mathbb{R}$ and

$$\mathscr{M}(b) := \{\mathscr{M}(a) \mid a \,{}^*\!\!\in b\}$$

for every $b \in {}^*\mathcal{V} \setminus {}^*\mathbb{R}$. Then $\mathscr{M}$ is an injection and $a \,{}^*\!\!\in b$ iff $\mathscr{M}(a) \in \mathscr{M}(b)$. If one identifies ${}^*\mathcal{V}$ with the image of ${}^*\mathcal{V}$ under $\mathscr{M}$, one can pretend that ${}^*\!\!\in$ is the true membership relation and consider ${}^*\mathcal{V}$ as a subset of the superstructure $\mathcal{V}({}^*\mathbb{R})$. Hence, we can drop the upper-left superscript $*$ from ${}^*\!\!\in$ for notational convenience.

Similar to the elements in $\mathcal{V}$, the rank function can also be defined for elements in ${}^*\mathcal{V}$. Every element in ${}^*\mathbb{R}$ has rank 0. It is easy to check that every element in ${}^*\mathcal{V}(\mathbb{R}, n+1) \setminus {}^*\mathcal{V}(\mathbb{R}, n)$ has the rank $n+1$ for $n \in [2\mathfrak{n}]$.

Formally, assume that elements in ${}^*\mathbb{R}$ are urelements with rank 0. Let $\mathscr{M}$ be the Mostowski collapsing map on ${}^*\mathcal{V}$, i.e., $\mathscr{M}(a) = a$ for every $a \in {}^*\mathbb{R}$ and

$$\mathscr{M}(b) := \{\mathscr{M}(a) \mid a \,{}^*\!\!\in b\}$$

for every $b \in {}^*\mathcal{V} \setminus {}^*\mathbb{R}$. Then $\mathscr{M}$ is an injection and $a \,{}^*\!\!\in b$ iff $\mathscr{M}(a) \in \mathscr{M}(b)$. If one identifies ${}^*\mathcal{V}$ with the image of ${}^*\mathcal{V}$ under $\mathscr{M}$, one can pretend that ${}^*\!\!\in$ is the true membership relation and consider ${}^*\mathcal{V}$ as a subset of the superstructure $\mathcal{V}({}^*\mathbb{R})$. Hence, we can drop the upper-left superscript $*$ from ${}^*\!\!\in$ for notational convenience.

Similar to the elements in $\mathcal{V}$, the rank function can also be defined for elements in ${}^*\mathcal{V}$. Every element in ${}^*\mathbb{R}$ has rank 0. It is easy to check that every element in ${}^*\mathcal{V}(\mathbb{R}, n+1) \setminus {}^*\mathcal{V}(\mathbb{R}, n)$ has the rank $n+1$ for $n \in [2\mathfrak{n}]$.

Formally, assume that elements in ${}^*\mathbb{R}$ are urelements with rank 0. Let $\mathscr{M}$ be the Mostowski collapsing map on ${}^*\mathcal{V}$, i.e., $\mathscr{M}(a) = a$ for every $a \in {}^*\mathbb{R}$ and

$$\mathscr{M}(b) := \{\mathscr{M}(a) \mid a \,{}^*\!\!\in b\}$$

for every $b \in {}^*\mathcal{V} \setminus {}^*\mathbb{R}$. Then $\mathscr{M}$ is an injection and $a \,{}^*\!\!\in b$ iff $\mathscr{M}(a) \in \mathscr{M}(b)$. If one identifies ${}^*\mathcal{V}$ with the image of ${}^*\mathcal{V}$ under $\mathscr{M}$, one can pretend that ${}^*\!\!\in$ is the true membership relation and consider ${}^*\mathcal{V}$ as a subset of the superstructure $\mathcal{V}({}^*\mathbb{R})$. Hence, we can drop the upper-left superscript $*$ from ${}^*\!\!\in$ for notational convenience.

Similar to the elements in $\mathcal{V}$, the rank function can also be defined for elements in ${}^*\mathcal{V}$. Every element in ${}^*\mathbb{R}$ has rank 0. It is easy to check that every element in ${}^*\mathcal{V}(\mathbb{R}, n+1) \setminus {}^*\mathcal{V}(\mathbb{R}, n)$ has the rank $n+1$ for $n \in [2\mathfrak{n}]$.

Formally, assume that elements in ${}^*\mathbb{R}$ are urelements with rank 0. Let $\mathscr{M}$ be the Mostowski collapsing map on ${}^*\mathcal{V}$, i.e., $\mathscr{M}(a) = a$ for every $a \in {}^*\mathbb{R}$ and

$$\mathscr{M}(b) := \{\mathscr{M}(a) \mid a \,{}^*\!\!\in b\}$$

for every $b \in {}^*\mathcal{V} \setminus {}^*\mathbb{R}$. Then $\mathscr{M}$ is an injection and $a \,{}^*\!\!\in b$ iff $\mathscr{M}(a) \in \mathscr{M}(b)$. If one identifies ${}^*\mathcal{V}$ with the image of ${}^*\mathcal{V}$ under $\mathscr{M}$, one can pretend that ${}^*\!\!\in$ is the true membership relation and consider ${}^*\mathcal{V}$ as a subset of the superstructure $\mathcal{V}({}^*\mathbb{R})$. Hence, we can drop the upper-left superscript $*$ from ${}^*\!\!\in$ for notational convenience.

Similar to the elements in $\mathcal{V}$, the rank function can also be defined for elements in ${}^*\mathcal{V}$. Every element in ${}^*\mathbb{R}$ has rank 0. It is easy to check that every element in ${}^*\mathcal{V}(\mathbb{R}, n+1) \setminus {}^*\mathcal{V}(\mathbb{R}, n)$ has the rank $n+1$ for $n \in [2\mathfrak{n}]$.

Fix a hyperfinite integer K. Let

$$\Gamma := \left\{ \frac{z}{K} \mid z \in {}^*\mathbb{Z} \right\}$$

and $\Delta t = 1/K$. An element $A \in {}^*\mathcal{V}$ which happens to be a set, function, relation, etc. is called an internal set, function, relation, etc., respectively.

Given an interval $[a, b]$ in $\mathcal{V}$, an internal set $T \subseteq [a, b]$ is called a set of tag points (with respect to Γ) if T contains exactly one element in each subinterval $[c, d]$ where $c, d \in \Gamma \cap [a, b]$ with $d - c = \Delta t$.

Fix a hyperfinite integer K. Let

$$\Gamma := \left\{ \frac{z}{K} \mid z \in {}^*\mathbb{Z} \right\}$$

and $\Delta t = 1/K$. An element $A \in {}^*\mathcal{V}$ which happens to be a set, function, relation, etc. is called an internal set, function, relation, etc., respectively.

Given an interval $[a, b]$ in $\mathcal{V}$, an internal set $T \subseteq [a, b]$ is called a set of tag points (with respect to Γ) if T contains exactly one element in each subinterval $[c, d]$ where $c, d \in \Gamma \cap [a, b]$ with $d - c = \Delta t$.

Fix a hyperfinite integer K. Let

$$\Gamma := \left\{ \frac{z}{K} \mid z \in {}^*\mathbb{Z} \right\}$$

and $\Delta t = 1/K$. An element $A \in {}^*\mathcal{V}$ which happens to be a set, function, relation, etc. is called an internal set, function, relation, etc., respectively.

Given an interval $[a, b]$ in $\mathcal{V}$, an internal set $T \subseteq [a, b]$ is called a set of tag points (with respect to Γ) if T contains exactly one element in each subinterval $[c, d]$ where $c, d \in \Gamma \cap [a, b]$ with $d - c = \Delta t$.

Definition (1.22)

A standard bounded function $f : [a, b] \to \mathbb{R}$ is said to be Riemann integrable on $[a, b]$ if for any two internal sets $T, T' \subseteq {}^*[a, b]$ of tag points, we have

$$\sum_{t \in T} {}^*f(t)\Delta t \approx \sum_{t \in T'} {}^*f(t)\Delta t.$$

If f is Riemann integrable on $[a, b]$, define the integration of f on $[a, b]$ by

$$\int_a^b f(x)dx := st\left(\sum_{t \in T} {}^*f(t)\Delta t\right)$$

for some internal set $T \subseteq {}^*[a, b]$ of tag points.

Definition (1.22)

A standard bounded function $f : [a, b] \to \mathbb{R}$ is said to be Riemann integrable on $[a, b]$ if for any two internal sets $T, T' \subseteq {}^*[a, b]$ of tag points, we have

$$\sum_{t \in T} {}^*f(t)\Delta t \approx \sum_{t \in T'} {}^*f(t)\Delta t.$$

If f is Riemann integrable on $[a, b]$, define the integration of f on $[a, b]$ by

$$\int_a^b f(x)dx := st\left(\sum_{t \in T} {}^*f(t)\Delta t\right)$$

for some internal set $T \subseteq {}^*[a, b]$ of tag points.

Definition (1.22)

A standard bounded function $f : [a, b] \to \mathbb{R}$ is said to be Riemann integrable on $[a, b]$ if for any two internal sets $T, T' \subseteq {}^*[a, b]$ of tag points, we have

$$\sum_{t \in T} {}^*f(t)\Delta t \approx \sum_{t \in T'} {}^*f(t)\Delta t.$$

If f is Riemann integrable on $[a, b]$, define the integration of f on $[a, b]$ by

$$\int_a^b f(x)dx := st\left(\sum_{t \in T} {}^*f(t)\Delta t\right)$$

for some internal set $T \subseteq {}^*[a, b]$ of tag points.

Theorem (1.23)

Given a standard bounded continuous function $g : [0, 1] \times \mathbb{R} \to \mathbb{R}$ and a number $\alpha \in \mathbb{R}$, there exists a standard function $y : [0, 1] \to \mathbb{R}$ satisfying Lipschitz condition such that

$$y(x) = \alpha + \int_0^x g(s, y(s))ds \tag{3}$$

for every $x \in [0, 1]$.

Proof: Let $B \in \mathbb{R}$ be a bound of g and $\Gamma \cap [0, 1] = \{t_0 < t_1 < \cdots < t_N\}$. Define inductively on $n \leq N$ such that $Y(t_0) = \alpha$ and

$$Y(t_{n+1}) = \alpha + \sum_{i=0}^{n} {}^*g(t_i, Y(t_i))\Delta t \tag{4}$$

for $n \in [N]$.

Theorem (1.23)

Given a standard bounded continuous function $g : [0,1] \times \mathbb{R} \to \mathbb{R}$ and a number $\alpha \in \mathbb{R}$, there exists a standard function $y : [0,1] \to \mathbb{R}$ satisfying Lipschitz condition such that

$$y(x) = \alpha + \int_0^x g(s, y(s))ds \tag{3}$$

for every $x \in [0,1]$.

Proof: Let $B \in \mathbb{R}$ be a bound of g and $\Gamma \cap [0,1] = \{t_0 < t_1 < \cdots < t_N\}$. Define inductively on $n \leq N$ such that $Y(t_0) = \alpha$ and

$$Y(t_{n+1}) = \alpha + \sum_{i=0}^{n} {}^*g(t_i, Y(t_i))\Delta t \tag{4}$$

for $n \in [N]$.

Theorem (1.23)

Given a standard bounded continuous function $g : [0,1] \times \mathbb{R} \to \mathbb{R}$ *and a number* $\alpha \in \mathbb{R}$, *there exists a standard function* $y : [0,1] \to \mathbb{R}$ *satisfying Lipschitz condition such that*

$$y(x) = \alpha + \int_0^x g(s, y(s))ds \tag{3}$$

for every $x \in [0,1]$.

Proof: Let $B \in \mathbb{R}$ be a bound of g and $\Gamma \cap [0,1] = \{t_0 < t_1 < \cdots < t_N\}$. Define inductively on $n \leq N$ such that $Y(t_0) = \alpha$ and

$$Y(t_{n+1}) = \alpha + \sum_{i=0}^{n} {}^*g(t_i, Y(t_i))\Delta t \tag{4}$$

for $n \in [N]$.

Note that $Y(t_n) \in ns({}^*\mathbb{R})$ because g being bounded by $B \in \mathbb{R}$ implies ${}^*g$ being bounded by B by the transfer principle. Let $y(0) = st(Y(t_0))$. For each $x \in (0,1]$ let $y(x) = st(Y(x^-))$ where x^- is the largest $t_n \leq x$ in $\Gamma \cap [0,1]$. It is easy to see that y satisfies Lipschitz condition on $[0,1]$.

Indeed, if $0 \leq z_1 \leq z_2 \leq 1$ are standard, then

$$\sum_{z_1^- < t_i \leq z_2^-} \Delta t \approx z_2 - z_1 \text{ and}$$

$$|y(z_2) - y(z_1)| = \left| st\left(\sum_{z_1^- < t_i \leq z_2^-} {}^*g(t_i, Y(t_i))\Delta t \right) \right| \leq B(z_2 - z_1).$$

Note that $Y(t_n) \in ns({}^*\mathbb{R})$ because g being bounded by $B \in \mathbb{R}$ implies ${}^*g$ being bounded by B by the transfer principle. Let $y(0) = st(Y(t_0))$. For each $x \in (0, 1]$ let $y(x) = st(Y(x^-))$ where x^- is the largest $t_n \leq x$ in $\Gamma \cap [0, 1]$. It is easy to see that y satisfies Lipschitz condition on $[0, 1]$.

Indeed, if $0 \leq z_1 \leq z_2 \leq 1$ are standard, then

$$\sum_{z_1^- < t_i \leq z_2^-} \Delta t \approx z_2 - z_1 \text{ and}$$

$$|y(z_2) - y(z_1)| = \left| st\left(\sum_{z_1^- < t_i \leq z_2^-} {}^*g(t_i, Y(t_i))\Delta t \right) \right| \leq B(z_2 - z_1).$$

Note that $Y(t_n) \in ns({}^*\mathbb{R})$ because g being bounded by $B \in \mathbb{R}$ implies ${}^*g$ being bounded by B by the transfer principle. Let $y(0) = st(Y(t_0))$. For each $x \in (0,1]$ let $y(x) = st(Y(x^-))$ where x^- is the largest $t_n \leq x$ in $\Gamma \cap [0,1]$. It is easy to see that y satisfies Lipschitz condition on $[0,1]$.

Indeed, if $0 \leq z_1 \leq z_2 \leq 1$ are standard, then

$$\sum_{z_1^- < t_i \leq z_2^-} \Delta t \approx z_2 - z_1 \text{ and}$$

$$|y(z_2) - y(z_1)| = \left| st\left(\sum_{z_1^- < t_i \leq z_2^-} {}^*g(t_i, Y(t_i))\Delta t \right) \right| \leq B(z_2 - z_1).$$

Note that $Y(t_n) \in ns({}^*\mathbb{R})$ because g being bounded by $B \in \mathbb{R}$ implies ${}^*g$ being bounded by B by the transfer principle. Let $y(0) = st(Y(t_0))$. For each $x \in (0, 1]$ let $y(x) = st(Y(x^-))$ where x^- is the largest $t_n \leq x$ in $\Gamma \cap [0, 1]$. It is easy to see that y satisfies Lipschitz condition on $[0, 1]$.

Indeed, if $0 \leq z_1 \leq z_2 \leq 1$ are standard, then

$$\sum_{z_1^- < t_i \leq z_2^-} \Delta t \approx z_2 - z_1 \text{ and}$$

$$|y(z_2) - y(z_1)| = \left| st\left(\sum_{z_1^- < t_i \leq z_2^-} {}^*g(t_i, Y(t_i))\Delta t \right) \right| \leq B(z_2 - z_1).$$

Similarly, $|Y(t_i) - Y(t_j)| \leq B|t_j - t_i|$ for $t_i \leq t_j$ in $\Gamma \cap [0, 1]$.

We show that y satisfies (3).

By Definition 1.22 the integral at the right side of (3) is infinitesimally close to $\alpha + \sum_{i=0}^{n} {}^*g(t_i, {}^*y(t_i))\Delta t$ and the left side of (3) is infinitesimally close to $\alpha + \sum_{i=0}^{n} {}^*g(t_i, Y(t_i))\Delta t$. Hence, it suffices to show that

$$\sum_{i=0}^{n} \left({}^*g(t_i, {}^*y(t_i)) - {}^*g(t_i, Y(t_i))\right) \Delta t \approx 0.$$

If $st(t_i) = \beta$, then $Y(t_i) \approx Y(\beta^-) \approx y(\beta) \approx {}^*y(t_i)$. Hence, $\eta(t_i) := {}^*g(t_i, {}^*y(t_i)) - {}^*g(t_i, Y(t_i)) \approx 0$ by the continuity of g. This verifies (3). □

Similarly, $|Y(t_i) - Y(t_j)| \leq B|t_j - t_i|$ for $t_i \leq t_j$ in $\Gamma \cap [0,1]$.

We show that y satisfies (3).

By Definition 1.22 the integral at the right side of (3) is infinitesimally close to $\alpha + \sum_{i=0}^{n} {}^*g(t_i, {}^*y(t_i))\Delta t$ and the left side of (3) is infinitesimally close to $\alpha + \sum_{i=0}^{n} {}^*g(t_i, Y(t_i))\Delta t$. Hence, it suffices to show that

$$\sum_{i=0}^{n} \left({}^*g(t_i, {}^*y(t_i)) - {}^*g(t_i, Y(t_i))\right) \Delta t \approx 0.$$

If $st(t_i) = \beta$, then $Y(t_i) \approx Y(\beta^-) \approx y(\beta) \approx {}^*y(t_i)$. Hence, $\eta(t_i) := {}^*g(t_i, {}^*y(t_i)) - {}^*g(t_i, Y(t_i)) \approx 0$ by the continuity of g. This verifies (3). □

Similarly, $|Y(t_i) - Y(t_j)| \leq B|t_j - t_i|$ for $t_i \leq t_j$ in $\Gamma \cap [0,1]$.

We show that y satisfies (3).

By Definition 1.22 the integral at the right side of (3) is infinitesimally close to $\alpha + \sum_{i=0}^{n} {}^*g(t_i, {}^*y(t_i))\Delta t$ and the left side of (3) is infinitesimally close to $\alpha + \sum_{i=0}^{n} {}^*g(t_i, Y(t_i))\Delta t$. Hence, it suffices to show that

$$\sum_{i=0}^{n} \left({}^*g(t_i, {}^*y(t_i)) - {}^*g(t_i, Y(t_i))\right) \Delta t \approx 0.$$

If $st(t_i) = \beta$, then $Y(t_i) \approx Y(\beta^-) \approx y(\beta) \approx {}^*y(t_i)$. Hence, $\eta(t_i) := {}^*g(t_i, {}^*y(t_i)) - {}^*g(t_i, Y(t_i)) \approx 0$ by the continuity of g. This verifies (3). □

Similarly, $|Y(t_i) - Y(t_j)| \leq B|t_j - t_i|$ for $t_i \leq t_j$ in $\Gamma \cap [0,1]$.

We show that y satisfies (3).

By Definition 1.22 the integral at the right side of (3) is infinitesimally close to $\alpha + \sum_{i=0}^{n} {}^*g(t_i, {}^*y(t_i))\Delta t$ and the left side of (3) is infinitesimally close to $\alpha + \sum_{i=0}^{n} {}^*g(t_i, Y(t_i))\Delta t$. Hence, it suffices to show that

$$\sum_{i=0}^{n} \left({}^*g(t_i, {}^*y(t_i)) - {}^*g(t_i, Y(t_i))\right) \Delta t \approx 0.$$

If $st(t_i) = \beta$, then $Y(t_i) \approx Y(\beta^-) \approx y(\beta) \approx {}^*y(t_i)$. Hence, $\eta(t_i) := {}^*g(t_i, {}^*y(t_i)) - {}^*g(t_i, Y(t_i)) \approx 0$ by the continuity of g. This verifies (3). □

Similarly, $|Y(t_i) - Y(t_j)| \leq B|t_j - t_i|$ for $t_i \leq t_j$ in $\Gamma \cap [0,1]$.

We show that y satisfies (3).

By Definition 1.22 the integral at the right side of (3) is infinitesimally close to $\alpha + \sum_{i=0}^{n} {}^*g(t_i, {}^*y(t_i))\Delta t$ and the left side of (3) is infinitesimally close to $\alpha + \sum_{i=0}^{n} {}^*g(t_i, Y(t_i))\Delta t$. Hence, it suffices to show that

$$\sum_{i=0}^{n} \left({}^*g(t_i, {}^*y(t_i)) - {}^*g(t_i, Y(t_i))\right) \Delta t \approx 0.$$

If $st(t_i) = \beta$, then $Y(t_i) \approx Y(\beta^-) \approx y(\beta) \approx {}^*y(t_i)$. Hence, $\eta(t_i) := {}^*g(t_i, {}^*y(t_i)) - {}^*g(t_i, Y(t_i)) \approx 0$ by the continuity of g. This verifies (3). □

The End of Day One

Thank you for your attention.