arXiv:2303.15045v1 [math.LO] 27 Mar 2023

弱集合论中沃彭卡原理的推论

Athanassios Tzouvaras

塞萨洛尼基亚里士多德大学数
学系，希腊塞萨洛尼基 54124 电子
邮箱：tzouvara@math.auth.gr

谨以此文纪念 Petr Vopˇenka

摘要

本文证明，沃彭卡原理（VP）能够在 ZF 的一个弱片段上恢复几乎整个 ZF。具体而言，若 EST 是由外延公理、空集公理、配对公理、并集公理、笛卡尔积公理、 Δ_0-分离公理以及沿 ω 的归纳法构成的理论，那么 EST + VP 可推出无穷公理、替换公理（因而也可推出分离公理）和幂集公理。此结果受到 [17]，中先前结果以及 H. Friedman 的 [8]，的启发；后者区分了 VP 的各种形式。由此可推得，EST+基础公理+VP=ZF + VP，且 EST+基础公理+选择公理 + VP =ZFC + VP。此外，本文还证明，基础公理独立于 ZF–{ 基础公理 }+VP。选择公理是否独立于 ZF +VP 尚未解决。VP 可推出一种非常弱的选择形式，本文还引入了若干类似的其他选择形式。

数学主题分类（2010年）：03E30、03E20、03E65。

关键词：沃彭卡原理（VP）、沃彭卡原理的变体、弱集合论、初等集合论（EST）。

1 引言

沃彭卡原理（下文简称 VP）主要被视为一种（非常）大型基数公理（见 [11]）。早已证明了该原理的其他一些推论，尤其是在范畴论中（见 [1]）。近期，通过对范畴论结果给出新的集合论证明，人们重新对 VP 产生了兴趣（见 Bagaria 和 Brooke-Taylor [2]）。Brooke-Taylor 还在 [5] 中证明了 VP 与几乎所有常见的 ZFC 独立命题的相对一致性，例如 GCH 和菱形原理（见 [5] 的引言）。

上述所有结果所依据的理论都是 ZFC。与之相反，本文旨在揭示 Vopěnka 原理的一种截然不同的能力：

在适当的弱集合论中加入 Vopěnka 原理后，它能够恢复 ZF 最基本的公理，即替换公理（从而也包括分离公理）和幂集公理，以及无穷公理（若缺少后者）。我们通常所说的“弱集合论”是指以下内容。

定义 1.1 弱集合论是不包含幂集公理和替代公理的集合论。

文献中研究过各种弱集合论系统。其中最著名的或许是 Kripke-Platek 理论 KP，它是可容许集理论的基础（见 [6], 第 48 页）。Devlin 基础集合论（BS）是一个比 KP 更弱的系统，见 [6] （参见 [6], 第 36 页）。 [13] 对一系列弱系统进行了广泛而详细的论述，其中包括 Devlin 的 BS。

弱集合论可能包含无穷公理，也可能不包含。此外，它可能是策梅洛-弗兰克尔集合论的一个片段，也可能不是。例如， [16]和 [17] 提出的 LZFC（“局部 ZFC”）理论是弱集合论，能够证明无穷公理，但不是 ZFC 的片段。另一方面，下文第 2 节介绍的系统 EST，作为 VP 的基础理论，是策梅洛-弗兰克尔集合论的一个片段，但不包含无穷公理。然而，EST+VP 可以证明这一公理。值得指出的是，即使加入无穷公理，EST 仍弱于 BS，因为它缺少基础公理。因此，它当然也弱于 KP。

下文将使用 ZFC 中众所周知的公理的常用名称，不再另作解释。这些公理是：外延公理、空集公理、配对公理、并集公理、幂集公理、无穷公理、分离公理、基础公理、替换公理和选择公理。有时使用缩写表示这些公理会更方便，尤其是在定理中。具体来说，我们常用 Ext 表示外延公理， $Pair$ 表示配对公理， Pow 表示幂集公理， Sep 和 Δ_0-Sep 分别表示分离公理与 Δ_0-分离公理， Rep 和 Δ_0-Rep 分别表示替换公理与 Δ_0-替换公理， $Found$ 表示基础公理， Inf 表示无穷公理，AC 表示选择公理。下文还将用到另一个弱公理，即笛卡尔积公理，缩写为 $CartProd$；该公理断言，对于任意集合 x、 y， $x \times y$ 都是一个集合。

VP 恢复替换公理和幂集公理的能力最早是在 [17] 中被注意到的。在那篇论文中，我们证明，若将 VP 加入 LZFC 的一个加强变体中，就能恢复替换公理和幂集公理。后来，我们读到 [8], ，其中给出了 VP 的七种变体，它们看似强度依次递减，但在 ZFC 上仍彼此等价；这时我们意识到， [17] 中用到的并非完整的 VP 原理，而只是一种较弱的形式，在 Friedman 的列表中记作 VP_4 。此外，我们发现，在 [17] 的结果中，若使用 VP 而不是 VP_4, ，该结果适用于理论 LZFC 本身，而非其加强变体。这促使我们将关注点从 LZFC 转向 VP 在 ZF 的弱片段上能够证明什么。

本文内容安排如下。在第 2.1 节中，我们介绍 Vopěnka 原理的各种变体，尤其是一般形式的 Vopěnka 原理和较弱的 Vopěnka 原理$_4$, ，并说明它们的基本区别。此外，我们概述 Vopěnka 原理作为集合存在原理的作用方式。在第 2.2 节中，我们介绍弱理论 EST。在第 2.3 节中，我们证明 Vopěnka 原理可以在 EST 中表达。第 2.4 节给出第 2 节的主要结果，即 EST+Vopěnka 原理证明无穷公理、替换公理和幂集公理。

在第 3 节中，我们证明基础公理独立于策梅洛-弗兰克尔集合论+Vopěnka 原理；也就是说，若策梅洛-弗兰克尔集合论+Vopěnka 原理一致，则策梅洛-弗兰克尔集合论$_0$ + Vopěnka 原理 $+ \neg Found$也一致，其中策梅洛-弗兰克尔集合论$_0$ =ZFC $- \{Found\}$。

第 4 节提出了选择公理是否独立于 ZF+VP 的问题。这个问题仍未解决。具体而言，尚不清楚是否可以

借助对称模型和置换模型解决这个问题。文中还指出，Vopěnka 原理蕴含一种非常弱的选择公理形式。这也为引入其他一些类似的选择公理形式提供了契机；它们相对于 ZF 以及 ZF+VP 的相对强度同样尚未确定。

2 某些弱集合论中的沃彭卡原理

2.1 Vopěnka 原理及其变体

设 $\mathcal{L} = \{\in\}$ 为集合论语言。给定 $\mathcal{L}$中的单自由变元公式 $\phi(x)$，令 X_ϕ 表示 $\phi(x)$ 的外延 $\{x : \phi(x)\}$ 。按照惯例，我们将 X_ϕ 称为“类”。沃彭卡原理是一个对类进行量化的陈述，因此无法在策梅洛-弗兰克尔集合论中表述为单个公理；不过，它可以表述为一个公理模式。显然，对于每个 ϕ，“X_ϕ 是一个真类”都是一阶语句。因此，以下陈述也都是一阶语句：

> (Vopěnka 原理 ϕ) 若 X_ϕ 是由某个一阶语言 L构成的 L-结构的真类，则存在不同的 $\mathcal{M}$、$\mathcal{N} \in X_\phi$，使得 $\mathcal{M} \precsim \mathcal{N}$，

其中 $\mathcal{M} \precsim \mathcal{N}$ 表示存在一个初等嵌入 $f : \mathcal{M} \to \mathcal{N}$。令

$$\mathrm{VP} = \{\mathrm{VP}_\phi : \phi(x) \text{ a formula of } \mathcal{L}\}.$$

[请注意，为避免与集合论的固定语言 $\mathcal{L} = \{\in\}$ 混淆，我们用符号“L”表示 VP$_\phi$ 的陈述中提到的任意一阶语言。当然，$\mathcal{L}$ 是其中一种 L，但它非常特殊。本文始终沿用这一做法。]

我们把上述 VP$_\phi$ 的表述称为“直接”表述。我们还有其逆否表述：“若 X_ϕ 是某个一阶语言 L下的 L结构所构成的类，且不存在不同的 $\mathcal{M}$、$\mathcal{N} \in X_\phi$ 使得$\mathcal{M} \precsim \mathcal{N}$，那么 X_ϕ 不是一个真类。”但“X_ϕ 不是一个真类”恰好意味着 X_ϕ 是一个集合。后一种表述使 VP$_\phi$ 能够充当集合存在原理：它表明，“如果关于ϕ 和 X_ϕ的某些条件成立，那么 X_ϕ 就是一个集合”。

VP 的直接表述适用于 ZF 和 ZFC，在这两种理论中，通常很容易判断其前提是否成立，也就是说，判断一个由结构组成的类 X_ϕ 是否是真类。相比之下，逆否表述使 VP 适用于弱集合论。原因在于，这类理论的论域定义不充分，往往难以确定哪些类 X_ϕ是集合，哪些是真类。因此，在弱集合论中，我们更需要能推出某些类 X_ϕ 是集合的原理。例如，在缺少替换公理和幂集公理的弱集合论中，可以利用 VP 的某些实例证明以下类实际上是集合：$\{x\colon (\exists y \in A)(F_\psi(y) = x)\}$，其中 $F_\psi(y) = x$ 是一个函数关系；或者 $\{x\colon x \subseteq A\}$，对任意集合 A均成立；由此便可证明前述公理。正如上文所说，我们最初是在研究 [17]时注意到这一点的。受该文结果以及 Friedman 的 [8], 启发，我们更系统地探索了上述思路，并证明，事实上，借助 VP 的逆否作用，可以在 ZF 的一个极弱片段中建立 ZF 的三大基石公理：幂集公理、替换公理和无穷公理。

H. Friedman 在 [8] 中给出了 VP 的七种变体，即上述 VP 以及六种较弱的形式 VP$_i$, $1 \le i \le 6$，其强度大致依次递减。每种变体都源于缩小可参与该公理模式的一阶语言的范围，或将“初等嵌入”替换为“嵌入”，或两者兼有。

。具体而言，VP_1 是通过将“ 一阶语言”限制为“有限关系型语言”而从 VP 得到的。VP_2 是通过将“初等嵌入”替换为“嵌入”而从 VP_1 得到的。VP_3 是通过将“有限关系型语言”替换为“具有单个二元运算的语言”而从 VP_2 得到的。VP_4 是通过将“具有单个二元运算的语言”替换为“具有单个二元关系的语言”而从 VP_3 得到的。VP_4 本质上就是 [17], 中使用的变体，因此我们省略 VP 的其他变体。更准确地说，令

（Vopěnka 原理$_{4,\ \phi}$）如果 X_ϕ 是语言 L上带有单个二元关系 R 的 L-结构的真类，那么存在互不相同的$\mathcal{M}$、$\mathcal{N} \in X_\phi$，使得 $\mathcal{M}$ 可嵌入到 $\mathcal{N}$ 中。

另设

$$\mathrm{VP}_4 = \{\mathrm{VP}_{4,\phi} : \phi(x) \text{ is a formula of } \mathcal{L}\}.$$

正如 Friedman 在 [8], 中所述，当 Vopěnka 原理$_i$以类论的语言表述时，以上七种变体在带选择公理的冯·诺伊曼–哥德尔–伯奈斯类论（NGBC）上等价。这意味着，在 ZF 中以公理模式给出的上述 Vopěnka 原理$_i$ 表述，在 ZFC 上也等价。Vopěnka 原理$_i$ 的等价性本应早已广为人知，至少范畴论学家应当如此。另如 [1], 第 278–279 页《历史评述》所示，Vopěnka 原理的原始表述实际上是 Vopěnka 原理$_4$ ，而不是 Vopěnka 原理。不过，这些变体在较弱理论上未必等价，例如我们下文讨论的理论（关于 Vopěnka 原理与 Vopěnka 原理$_4$ 在 EST 上的区别，见推论 2.16 及其后的评注）。

尽管本文已从较弱的变体 VP_4转向最一般的 VP，但在下文主要结果的证明中，我们仍不需要 VP 的全部一般性。具体而言，下文所用的 VP_4 与 VP 之间的主要区别在于，VP 中涉及的一阶语言可以包含任意集合的常元，而 VP_4, 中涉及的一阶语言则不能。另一方面，就我们的目的而言，VP 中的语言不必包含多于一个二元谓词和一个一元谓词。此外，要推出替换公理和幂集公理，只需 VP 谈论嵌入，而不必谈论初等嵌入。（见下文注 2.17。）

区分了 VP 和 VP_4, 之后，我们现在可以介绍 [17] 的结果；这一结果在很大程度上促成了本文的写作。（尽管在 [17] 中我们提到 VP，但实际用到的是 VP_4。）为表述该结果，我们首先需要 P. Vopˇenka、A. Pultr 和 Z. Hedrl´ın 的一个经典定理 [18], ，我们将其简称为 V-P-H；该定理的证明对 [17]的结果至关重要。

给定一个集合 A 及其二元关系 $R \subseteq A \times A$，我们将有序对 $\langle A, R\rangle$ 称为图。给定一个图 $\langle A, R\rangle$，若映射 $f : A \to A$ 满足：对于所有 $x, y \in A$, $\langle x, y\rangle \in R$， $\langle f(x)\ f(y)\rangle \in R$，则称其为自同态。若图 $\langle A, R\rangle$ 的唯一自同态是恒等关系，则称其为刚性的。

V-P-H 定理如下：

定理 2.1（V-P-H [18]）（ZFC）对于每个无穷集合 A，都存在一个二元关系 $R \subset A \times A$，使得图 $\langle A, R\rangle$ 是刚性的。

显然，前面的定理只涉及具有单个二元关系的结构。此外，刚性指的是一种关于（甚至不存在）自同态，而非嵌入（更不用说初等嵌入）的性质。因此， [17] 中证明的内容本质上是以下定理。

定理 2.2（[17, 定理 6.3]）若 T 是一个满足 LZFC $\subseteq T$ 且 $T \vdash$V-P-H 的理论，则 $T +$VP$_4$ 证明替换公理和幂集公理，即 ZFC $\subseteq T +$VP$_4$。

[此表述与 [17] 的区别仅在于，我们提及的是 VP$_4$，而非 VP。[1]]

在定理 2.2 中，假设理论 T 证明 V-P-H，就保证了每个集合 A 上都存在一个刚性二元关系 R。随后，将 VP$_4$ 应用于适当配备了此类 R 的结构，可以证明 $\mathcal{P}(A)$ 和 F"A 不是真类。变体 VP$_4$ 足以达到这一目的，因为我们只需要使用仅含一个二元关系 R 的语言，并且只需使用自同态，而非嵌入或初等嵌入。

从上文的定理 2.2 推导出本文结果的基本观察是：在缺少 V-P-H 的情况下，可以改用结构语言中的无穷多个常元来保证刚性性质，即为每个元素 $a \in A$ 配备一个常元 c$_a$。VP 允许使用此类语言，但 VP$_4$ 不允许。

2.2 用于表述 Vopěnka 原理的策梅洛-弗兰克尔集合论弱片段

下面我们定义一个称为 EST（初等集合论）的 ZF 弱片段。首先，令

$$\mathrm{EST}_0 = \{\mathit{Ext},\ \mathit{Emptyset},\ \mathit{Pair},\ \mathit{Union},\ \mathit{CartProd},\ \Delta_0\text{-}\mathit{Sep}\}.$$

对于任意集合 x、y，有序对 $\langle x, y\rangle$ 在 EST$_0$ 中按通常方式定义，即定义为集合 $\{\{x\},\{x,y\}\}$。对于任意给定的集合 M、N、$M \times N = \{\langle x, y\rangle :\ x \in M$、$y \in N\}$ 根据 $\mathit{CartProd}$ 是一个集合。M 与 N 之间的二元关系定义为集合 $R \subseteq M \times N$。函数 $f :\ M \to N$ 定义为特殊的二元关系 $f \subseteq M \times N$。全文中，符号 f 表示函数，因此 $(\exists f)(\cdots)$ 是 $(\exists f)(f$ 是一个函数 $\wedge \cdots)$ 的简写。

接下来，我们定义自然数类 ω，它也不必是一个集合。令 $Tr(x)$ 表示谓词"x 是一个传递集"。我们还定义谓词

$$\mathit{Ord}(x) := \mathit{Tr}(x) \wedge (\forall y, z \in x)(y \in z \vee y = z \vee z \in y),$$

$$\mathit{Succ}(x) := (\exists y \in x)(x = y \cup \{y\}),$$

$$\mathit{Nat}(x) := \mathit{Ord}(x) \wedge (\forall y)[(y \in x \vee y = x) \to (y = \emptyset \vee \mathit{Succ}(y))].$$

最后，令

$$\omega = \{x : \mathit{Nat}(x)\}.$$

我们称 ω 为自然数类。在书写$\mathcal{L}$的公式时，通常使用记号 $x \in \omega$较为方便，尽管一般而言 ω 是一个类；它是谓词 $\mathit{Nat}(x)$ 的缩写。特别地，$(\exists x \in \omega)\phi$ 和 $(\forall x \in \omega)\phi$分别表示 $(\exists x)(\mathit{Nat}(x) \wedge \phi)$ 和 $(\forall x)(\mathit{Nat}(x) \to \phi)$。此外，照例，字母 m 和 n 表示 ω的元素。

此外，我们还需要归纳法沿着 ω 成立，也就是说，$X \subseteq \omega$ 的每个非空子类都有一个最小元素。为此，需要增加一条公理。这就是下文给出的归纳模式 $\mathit{Ind}(\omega)$，它与皮亚诺算术的归纳模式几乎完全相同。具体而言，对于每个

[1]此处不需要给出 LZFC 理论的精确定义。只需说明，它是定义 1.1 意义下的一种弱集合论。它不是策梅洛-弗兰克尔集合论的一个片段，但满足无穷公理（意即 ω 是最小的归纳集合）、选择公理、笛卡尔积公理等。为后文之用，我们还指出，LZFC 比下文引入的 EST 理论强得多。

$\mathcal{L}$ 中满足 $X_\phi \subseteq \omega$ 的公式 $\phi(x)$（也就是说，$(\forall x)(\phi(x) \to Nat(x))$），令 $Ind_\phi(\omega)$ 表示公式

$$(Ind_\phi(\omega)) \qquad [\phi(\emptyset) \wedge (\forall x \in \omega)(\phi(x) \to \phi(x \cup \{x\})] \to (\forall x \in \omega)\phi(x).$$

另设

$$Ind(\omega) = \{Ind_\phi(\omega) : X_\phi \subseteq \omega\}.$$

[等价地，$Ind(\omega)$ 表示每个非空的 $X_\phi \subseteq \omega$ 都有一个关于 $\in$的最小元素。] 最后，我们定义

$$\mathrm{EST} = \mathrm{EST}_0 + Ind(\omega).$$

EST 是一种弱理论，以下将用作 Vopěnka 原理的基础理论。

注2.3 上文对 ω 的定义与 [4, 第 468 页及后续内容]中的定义相同（另见 [3]），但有一点不同：在 [4] 中，谓词 $Ord(x)$表示“x 关于 $\in$是良序的”；而在我们对 $Ord(x)$的定义中，“x 关于 $\in$是线性序的”。请注意以下几点：

(i) 这两个性质的区别在于，“线性序”是Δ_0，而“良序”不是（它是 Π_1）。这一事实将在下文定理 2.12 的证明中用到。另一方面，借助 $Ind(\omega)$，可以很容易地证明，对于每个 $x \in \omega$，x 确实关于 $\in$是良序的。

(ii) 如果可以使用基础公理，那么“x 是线性序”和“x 是良序”（相对于 $\in$而言）这两个性质是等价的。

(iii) 然而，即使可以使用基础公理，$\mathrm{EST}_0 + Found$ 也无法证明 Ind（ω）。此外还需要分离公理或替换公理。也就是说，$\mathrm{EST}_0 + Found + Sep \vdash Ind$（$\omega$）。

借助 Ind_ϕ（ω），可以证明关于自然数的所有基本事实。以下是其中一些后文会用到的事实：

事实2.4 (i) 对任意 x、$y \in \omega$，若 $x \in y$，则 $x \subsetneq y$，且 x 是 y的初始段。

(ii) 若 x、$y \in \omega$ 且 $x \subsetneq y,$，则不存在单射 f：$y \to x$。(iii) 对任意 x、$y \in \omega$、$x \in y \vee y \in x \vee x = y$。(iv) 若 $x \in \omega$ 且 $x = y \cup \{y\}$，则 y 是 x 的最大元素。

定义了有序对和自然数之后，就可以像通常那样通过归纳法定义有序 n 元组，其中 $n \in \omega$。具体而言，对于每个 $n > 2$、$\langle x_0, \ldots\ldots, x_{n+1}\rangle = \langle\langle x_0, \ldots\ldots, x_n\rangle, x_{n+1}\rangle$。此外，根据 $CartProd$，对于 $n \in \omega$、$n > 0$ 以及任意 n 元组的集合 M_0、……、M_{n-1}，集合 $M_0 \times \cdots \times M_{n-1} = \{\langle x_0, \ldots, x_{n-1}\rangle : x_i \in M_i\}$ 是一个集合。

下文还需要函数集 $f : n \to M$，记作 ${}^nM$，其中 M 是任意集合，$n \in \omega$ 是任意一个。由于 EST 中没有替换公理，我们将以略有不同的方式定义 ${}^nM$ 的元素。具体而言，虽然一个 $f \in {}^nM$ 通常具有如下形式

$$f = \{\langle 0, x_0\rangle, \ldots, \langle n-1, x_{n-1}\rangle\},$$

显然，我们可以将后一个集合与 n 元组等同起来

$$\langle\langle 0, x_0\rangle, \ldots, \langle n-1, x_{n-1}\rangle\rangle,$$

它是笛卡尔积公理的一个元素

$$(\{0\} \times M) \times \cdots \times (\{n-1\} \times M).$$

正如我们上文所见，后者是一个集合，因此我们可以定义

$$^nM := (\{0\} \times M) \times \cdots \times (\{n-1\} \times M).$$

按照惯例，我们用字母 g、h 表示 ^nM 中的元素，并用 $h(i) = x$ 表示 $\langle i, x\rangle$ 是 n-元组h的第 i 个元素这一事实。

正如引言中已经提到的，EST 即使加入无穷公理，仍弱于 Devlin 的系统 BS。关于下一节将给出的 EST 中的 Sat 谓词的定义，审稿人好意告知我，Devlin 的 BS 虽然强于 EST，但众所周知，它不足以实现其引入目的，包括定义满足谓词。[2]（关于 BS 在这一点上的缺陷及其补救措施，见 [13, §10]。）EST 能够做到而 BS 做不到，主要原因似乎是上文给出的集合 ^nM 的非正统定义。出人意料的是，要使这一切奏效，我们只需对 ^nM 采用这种不同寻常但合理的形式化。

2.3 EST 中 Vopěnka 原理的可表达性

从现在起，我们在 EST 中工作，其语言为 $\mathcal{L} = \{\in\}$。本小节将说明，如何在 EST 中定义 Vopěnka 原理的表述所需的概念。Vopěnka 原理的表述（至少就本文的需要而言）要求如下:

(a) 对每个集合 A，语言 L_A的定义：它至多包含一个一元关系符号和一个二元关系符号，但对于每个 $a \in A$ 都包含一个常元$\mathbf{c}_a$，并且包含无限多个变量。由于缺少替换公理L_A 未必是一个集合。

(b) L_A的项、公式和句子的定义，使这些对象类具有归纳性，从而可以归纳地证明关于这些句法对象的通常事实。同样，项、公式和句子的类未必是集合。

(c) 对每种语言 L_A，L_A-结构的定义。

(d) 满足关系“$\mathcal{M} \models \sigma(x_0, \ldots, x_{n-1})$”的定义，其中 L_A-结构为 $\mathcal{M}$，L_A 中的公式为 $\sigma(v_0, \ldots, v_{n-1})$，其自由变元为$v_0, \ldots, v_{n-1}$，而 n-元组由 M中的元素 $x_0, \ldots, x_{n-1}$ 组成。据此，L_A-结构之间的初等可嵌入关系 $\mathcal{M} \precsim \mathcal{N}$（或简单可嵌入关系）便可立即定义。

下面我们展示如何在 EST 中实现定义 (a)–(d)。由于满足关系 $\mathcal{M} \models \sigma(x_0, \ldots, x_{n-1})$ 的可定义性对于 VP 在 EST 中的可表达性至关重要，因此我们将在下文明确给出所需的定义。

定义 2.5 对每个 $n \in \omega$，令 $v_n := \langle 0, n\rangle$。$v_n$ 是（每种语言中的）第 n 个变量。对每个集合 a，令 $\mathbf{c}_a := \langle 1, a\rangle$。集合 $\mathbf{c}_a$ 称为常元。我们还将（足够多的）逻辑符号 $\equiv$, $\neg$, $\wedge$ 和$\exists$ 分别与 ω 的元素对应起来，如下所示：$\equiv := 2$, $\neg := 3$, $\wedge := 4$, $\exists := 5$。最后，令 $\mathbf{U} := 6$ 且 $\mathbf{R} := 7$。符号 $\equiv$、U 和 R 称为谓词。$\equiv$ 和 R 是二元谓词，而 U 是一元谓词。对每个集合 A，令

$$L_A = \{\equiv, \neg, \wedge, \exists\} \cup \{v_n : n \in \omega\} \cup \{\mathbf{c}_a : a \in A\} \cup \{\mathbf{U}, \mathbf{R}\}.$$

另令

$$V(L_A) = \{v_n : n \in \omega\},$$

[2]关于这一问题的讨论可见于 http://mathoverflow.net/questions/77734/devlins-constructibility-as-a-resource。

$$C(L_A) = \{\mathbf{c}_a : a \in A\},$$
$$Term(L_A) = \{v_n : n \in \omega\} \cup \{\mathbf{c}_a : a \in A\}$$

分别为 L_A 的变量、常元和项类。

定义 2.6 给定任意集合 A, L_A 的原子公式 $AFml(L_A)$ 类定义如下:

$$AFm(L_A) = \{\langle \equiv, t, s\rangle : t, s \in Term(L_A)\} \cup \{\langle \mathbf{U}, t\rangle : t \in Term(L_A)\} \cup$$
$$\{\langle \mathbf{R}, t, s\rangle : t, s \in Term(L_A)\}.$$

[上述编码的含义显而易见。$\langle \equiv$、t、$s\rangle$、$\langle \mathbf{U}\ t\rangle$ 和 $\langle \mathbf{R}\ t$、$s\rangle$ 分别表示公式 $t \equiv s$、$\mathbf{U}(t)$ 和 $\mathbf{R}(t, s)$。]

定义 2.7 L_A 的公式类 $Fml(L_A)$ 定义如下。首先定义谓词 $Fm_{L_A}(x, f, n)$ 如下:

$$Fml_{L_A}(x, f, n) := [n \neq 0 \wedge dom(f) = n \wedge f(n-1) = x \wedge$$
$$(\forall k < n)[(f(k) \in AFml(L_A) \vee (\exists j < k)(f(k) = \langle \neg, f(j)\rangle) \vee$$
$$(\exists j, l < k)(f(k) = \langle \wedge, f(j), f(l)\rangle) \vee$$
$$(\exists j < k)(\exists m \in \omega)(f(k) = \langle \exists, m, f(j)\rangle)]].$$

于是我们令

$$Fml(L_A) = \{x : (\exists f)(\exists n \in \omega) Fml_{L_A}(x, f, n)\}.$$

上述公式定义本质上是第 3] 章 [7, 定义 5.2 中给出的定义。尽管 ω 不一定是集合，这一定义仍然适用。不过，上述 $Fml(L_A)$ 的定义并不完全精确。写作 $Fml_{L_A}(x, f, n)$ 时，f 意指一个定义域为 n 的函数，该函数枚举 x 的子公式，且仅枚举这些子公式。但是，例如，若 $dom(f) = 3$，且 $f(0)$、$f(1)$ 和 $f(2)$ 是原子公式，而 $f(2) = x$，那么根据 2.7，$Fml(x, f, 3)$ 成立，尽管 $f(0)$ 和 $f(1)$ 并非 x 的子公式。因此，必须在 $Fml_{L_A}(x, f, n)$ 的定义中添加一项限制，以防止 f 枚举无关的原子公式。这项限制就是要求 f 的定义域是最小的，具体而言，$dom(f) = |Sub(x)|$（x 的子公式数目）。在上述例子中，枚举某个原子公式子公式的函数，其定义域应为 1，而非 3。通过对 $Fml_{L_A}(x, f, n)$ 的定义作一个简单修改，即可将此要求形式化表达，并由此修改 $Fml(L_A)$ 的定义。[3] 今后我们假定，每当写作 $Fml_{L_A}(x, f, n)$ 时，这项要求都已隐含满足。

2.7 的关键之处在于，它能够支持归纳证明和递归定义。我们用字母 σ、τ 表示 $Fml(L_A)$ 的元素。作为 2.7 的第一个应用，我们为每个 $\sigma \in Fml(L_A)$ 指定一个长度，即某个枚举函数 f 的定义域 n，该枚举函数用于枚举 σ 的子公式；或者，由于这些函数都是最小的，该长度就是子公式 σ 的数量。更一般地，鉴于沿着 ω 进行归纳是有效的，以下结论成立。

[3]修改如下：我们设

$$Fml(L_A) = \{x : (\exists f)(\exists n \in \omega) Fml^*_{L_A}(x, f, n)\},$$

其中

$$Fml^*_{L_A}(x, f, n) := Fml_{L_A}(x, f, n) \wedge (\forall g)(\forall m)[Fml_{L_A}(x, g, m) \to n \leq m].$$

引理 2.8 设 $X \subseteq Fml(L_A)$ 是 $Fml(L_A)$ 的一个子类。若 $AFml(L_A) \subseteq X$ 且 X 关于 $\neg$、 $\wedge$ 和 $\exists$ 封闭，则 $X = Fml(L_A)$。

证明。假设 X 如所述，并设 $Fml(L_A) - X \neq \emptyset$。那么，根据 ω的归纳性质，存在长度最小的 $\sigma \in Fml(L_A) - X$ n。于是根据 $Fml(L_A)$的定义，我们立即得到矛盾。⊣

Given a language L_A as above, L_A-structures are defined as follows.

定义 2.9 对任意集合 A，一个 L_A-结构是一个四元组

$$\mathcal{M} = \langle M, U, R, I \rangle,$$

其中 $U \subseteq M$、 $R \subseteq M \times M$和 I是（集合）函数 I: $A \to M$。我们将M 称为 $\mathcal{M}$ 的定义域，并将 I 称为 L_A的常元赋值。如果 L_A中缺少 U 或 R，则 L_A-结构分别为三元组 $\mathcal{M} = \langle M$、 R、 $I\rangle$ 或$\mathcal{M} = \langle M$、 U、 $I\rangle$。 $\mathcal{M}$ 中 L_A 的非逻辑符号的解释定义如下：$\mathbf{U}^{\mathcal{M}} = U$，$\mathbf{R}^{\mathcal{M}} = R$，对每个$a \in A$，$\mathbf{c}_a^{\mathcal{M}} = I(a)$。另一方面， $\equiv^{\mathcal{M}}$ 是恒等关系。

对于每个 $\sigma \in Fml(L_A)$， σ 的自由变元（有限）集合，记作 $FV(\sigma)$，按照通常的方式通过对 σ 的长度进行归纳（也就是说，沿着 2.7 中的步骤）来定义。此外，对于每个 σ， $|FV(\sigma)| \in \omega$。现在我们来定义满足关系 $Sat(\mathcal{M}, \sigma, e)$，它形式化了关系 $\mathcal{M} \models \sigma(e(0) \ldots, e(n-1))$：对于一个 L_A-结构 $\mathcal{M} = \langle M, U, R, I\rangle$、一个自由变元为 $v_0, \ldots, v_{n-1}$ 的公式 $\sigma(v_0, \ldots, v_{n-1})$，以及一个映射 e: $n \to M$，也就是说， $e \in {}^nM$ （回想一下，根据第 2.2 节末尾的讨论， ${}^nM$ 是一个集合）。下一个定义改编自 [7] 的第 3 章定义 5.4。（回想一下，字母 f, g 总是表示函数。）在下面的定义中，我们假设符号 U 和 R 都出现；如果其中某个符号缺失，则按显而易见的方式修改定义。

定义 2.10 令 $Sat(\mathcal{M}, \sigma, e)$ 表示以下关系：

$$(\exists M, U, R, I)(\exists f, g)(\exists n, m \in \omega)[\mathcal{M} = \langle M, U, R, I\rangle \wedge$$

$$Fml_{L_A}(\sigma, f, n) \wedge |FV(\sigma)| = m \wedge dom(g) = n \wedge e \in g(n-1) \wedge$$

$$(\forall k < n) S(k, f, g, m, \mathcal{M})],$$

其中 $S(k, f, g, m, \mathcal{M})$:=

$$(\exists i, j)[f(k) = \langle \equiv, v_i, v_j\rangle \wedge g(k) = \{h \in {}^mM : h(i) = h(j)\}] \vee$$

$$(\exists i)(\exists a \in A)[f(k) = \langle \equiv, v_i, \mathbf{c}_a\rangle \wedge g(k) = \{h \in {}^mM : h(i) = I(a)\}] \vee$$

$$(\exists j)(\exists a \in A)[f(k) = \langle \equiv, \mathbf{c}_a, v_j\rangle \wedge g(k) = \{h \in {}^mM : h(j) = I(a)\}] \vee$$

$$(\exists a \in A)[f(k) = \langle \equiv, \mathbf{c}_a, \mathbf{c}_a\rangle \wedge g(k) = {}^mM] \vee$$

$$(\exists a \neq b \in A)[f(k) = \langle \equiv, \mathbf{c}_a, \mathbf{c}_b\rangle \wedge g(k) = \emptyset] \vee$$

$$(\exists i)[f(k) = \langle \mathbf{U}, v_i\rangle \wedge g(k) = \{h \in {}^mM : h(i) \in U\}] \vee$$

$$(\exists a \in A)[f(k) = \langle \mathbf{U}, \mathbf{c}_a\rangle \wedge I(a) \in U \wedge g(k) = {}^mM] \vee$$

$$(\exists a \in A)[f(k) = \langle \mathbf{U}, \mathbf{c}_a\rangle \wedge I(a) \notin U \wedge g(k) = \emptyset] \vee$$

$$(\exists i, j)[f(k) = \langle \mathbf{R}, v_i, v_j\rangle \wedge g(k) = \{h \in {}^mM : \langle h(i), h(j)\rangle \in R\}] \vee$$

$$(\exists i)(\exists a \in A)[f(k) = \langle \mathbf{R}, v_i, \mathbf{c}_a\rangle \wedge g(k) = \{h \in {}^mM : \langle h(i), I(a)\rangle \in R\}] \vee$$

$$(\exists j)(\exists a \in A)[f(k) = \langle \mathbf{R}, \mathbf{c}_a, v_j \rangle \wedge g(k) = \{h \in {}^m M : \langle I(a), h(j) \rangle \in R\}] \vee$$
$$(\exists a, b \in A)[f(k) = \langle \mathbf{R}, \mathbf{c}_a, \mathbf{c}_b \rangle \wedge \langle I(a), I(b) \rangle \in R \wedge g(k) = {}^m M] \vee$$
$$(\exists a, b \in A)[f(k) = \langle \mathbf{R}, \mathbf{c}_a, \mathbf{c}_b \rangle \wedge \langle I(a), I(b) \rangle \notin R \wedge g(k) = \emptyset] \vee$$
$$(\exists i)[f(k) = \langle \neg, f(i) \rangle \wedge g(k) = {}^m M - g(i)] \vee$$
$$(\exists i, j)[f(k) = \langle \wedge, f(i), f(j) \rangle \wedge g(k) = g(i) \cap g(j)] \vee$$
$$(\exists i, j)[f(k) = \langle \exists, i, f(j) \rangle \wedge g(k) = \{h \in {}^m M : (\exists x \in M)(h(i/x) \in g(j))\}],$$

（其中在最后一个子句中， $h(i/x)$ 是将 $\langle i, h(i) \rangle$ 替换为 $\langle i, x \rangle$ 后，由 h 得到的 ${}^m M$ 元组。）

在前面的定义中，函数 g 枚举了使 σ 的子公式为真的赋值集合；它在 EST 中可定义，因为一方面 ${}^m M$ 是一个集合，另一方面，在定义的每个子句中， $g(k)$ 是 ${}^m M$ 的一个 Δ_0 子类，因此是一个集合。

在定义了对于一个 L_A 结构 $\mathcal{M}$、一个 L_A 公式 $\sigma(v_0, \ldots, v_{m-1})$（其自由变元均属于 $v_0, \ldots, v_{m-1}$）以及 $x_0, \ldots, x_{m-1} \in M$， $\mathcal{M} \models \sigma(x_0, \ldots, x_{m-1})$ 的含义之后，我们便可照常定义从一个 L_A 结构到另一个结构的初等嵌入。

给定两个 L_A 结构 $\mathcal{M}, \mathcal{N}$，若存在一个一一函数 $f: M \to N$，使得对于 L_A 中任意一个自由变元属于 $v_0, \ldots, v_{m-1}$ 的公式 $\sigma(v_0, \ldots, v_{m-1})$，以及任意的 $x_0, \ldots, x_{m-1} \in M$，我们称 $\mathcal{M}$ 可初等嵌入到 $\mathcal{N}$ 中，记作 $\mathcal{M} \precsim \mathcal{N}$，

$$\mathcal{M} \models \sigma(x_0, \ldots, x_{m-1}) \leftrightarrow \mathcal{N} \models \sigma(f(x_0), \ldots, f(x_{m-1})).$$

令 $f: \mathcal{M} \precsim \mathcal{N}$ 表示这样的事实： f 是从 $\mathcal{M}$到 $\mathcal{N}$的初等嵌入。有时，为了更加精确，我们需要指明所指的语言。这时我们称 $f: M \to N$ 为一个 L_A-初等嵌入，并将其记作

$$f : \mathcal{M} \precsim_{L_A} \mathcal{N}.$$

显然，最后一个关系在 EST 中是可定义的。下面这个简单事实将在后文中反复用到。

事实 2.11 设 $\mathcal{M} = \langle M, U, R, I \rangle, \mathcal{N} = \langle N, Z, S, J \rangle$ 是 L_A-结构，且$f: M \to N$ 是一个 L_A-初等嵌入（或仅仅是一个 L_A-嵌入）。那么 $f \circ I = J$，也就是说，对每个 $a \in A$， $f(I(a)) = J(a)$。等价地，对每个 $a \in A$， $f(\mathbf{c}_a^{\mathcal{M}}) = \mathbf{c}_a^{\mathcal{N}}$。

证明。对每个 L_A-嵌入 $f: \mathcal{M} \to \mathcal{N}$，根据定义， $f(\mathbf{c}_a^{\mathcal{M}}) = \mathbf{c}_a^{\mathcal{N}}$。此外，对每个 $a \in A$， $I(a) = \mathbf{c}_a^{\mathcal{M}}$且 $J(a) = \mathbf{c}_a^{\mathcal{N}}$。因此 $f(I(a)) = J(a)$。⊣

EST 的语言 $\mathcal{L} = \{\in\}$ 只是上述语言 L_A 中的一个特例，即包含 $A = \emptyset$ 和一个二元关系符号 $\in$的语言。因此， $\mathcal{L}$ 的公式类和句子类在 EST 中已经可以定义。

给定 $\mathcal{L} = \{\in\}$中的一个公式 $\phi(x)$，令 X_ϕ 表示 ϕ的外延，尽管在 EST 中， X_ϕ 未必是集合。表达式“X_ϕ 是一个真类”仅表示以下 $\mathcal{L}$-语句：

$$(\forall x)(\exists y)[(\phi(y) \wedge y \notin x) \vee (\neg\phi(y) \wedge y \in x)].$$

[请注意，在策梅洛-弗兰克尔集合论中，由于分离公理，“X_ϕ 是一个真类”只需表述为 $(\forall x)(X_\phi \not\subseteq x)$。但在 EST 中，缺少分离公理，而且 $X_\phi \subseteq x$

并不能推出 X_ϕ 是一个集合，因此“X_ϕ 是一个真类”需要表述为 $(\forall x)(X_\phi \neq x)$。]

给定关于一个自由变量的任意公式 $\phi(x)$ $\mathcal{L}$，Vopěnka 原理的实例 VP$_\phi$ 表述如下:

VP$_\phi$: “对于每个集合 A，如果 X_ϕ 是由 L_A-结构构成的真类，那么存在 $\mathcal{M} \neq \mathcal{N} \in X_\phi$ 和 f，使得 $f:\ \mathcal{M} \precsim_{L_A} \mathcal{N}$”。

显然，VP$_\phi$ 是一个 $\mathcal{L}$-语句，因此该类

$$\mathrm{VP} = \{x : (\exists \phi \in Fml(\mathcal{L}))(x = \mathrm{VP}_\phi)\}$$

在 EST 中是可定义的。这就完成了对 EST 中 VP 模式的描述。

2.4 EST+Vopěnka 原理的推论

我们先证明无穷公理，因为它不依赖于任何形式的替换公理或幂集公理。我们证明，在 EST+Vopěnka 原理中，第 2.2 节定义的类 $\omega = \{x:\ Nat(x)\}$ 是一个集合。

定理 2.12 在 EST $+$ Vopěnka 原理下，$\omega = \{x:\ Nat(x)\}$ 是一个集合。因此 EST $+$ Vopěnka 原理 $\vdash Inf$。

证明。为避免处理 $\emptyset$ 的初等嵌入，令 $\omega^* = \omega - \{0\}$。显然，在 EST 中，$\omega$ 是集合当且仅当 ω^* 是集合。因此，为了导出矛盾，假设 ω^* 是真类。考虑语言 $L_\emptyset = \{\mathrm{R}\}$，其中只有一个二元关系符号 R（也就是说，$L_\emptyset$ 不包含常元 c_a）。对于每个 $x \in \omega^*$，令 $\in_x = \{\langle y, z\rangle \in x \times x:\ y \in z\}$（$\in$ 在 x 上的限制）。由于 $x \times x$ 是一个集合，根据 Δ_0-分离公理，$\in_x$ 也是一个集合。令 $\mathcal{M}_x = \langle x, \in_x\rangle$。每个 $\mathcal{M}_x$ 都是一个 $L_\emptyset$-结构，其中 R 解释为 $\in_x$。再令

$$K = \{\mathcal{M}_x : x \in \omega^*\}.$$

我们声称，当 ω^* 是一个真类时，K 也是一个真类。事实上，假设 K 是一个集合。由于

$$K = \{\langle x, \in_x\rangle : x \in \omega^*\} = \{\{\{x\}, \{x, \in_x\}\} : x \in \omega^*\},$$

显然 $\omega^* \subset \cup^2 K = \cup \cup K$。特别地，$\omega^* = \{x \in \cup^2 K:\ Nat(x)\}$。由于 $\cup^2 K$ 是一个集合，且 $Nat(x)$ 是 Δ_0，根据 Δ_0-分离公理，ω^* 是一个集合，矛盾。

因此，K 是一个 $L_\emptyset$-结构的真类。根据 Vopěnka 原理，存在 $x_0 \neq x_1 \in \omega^*$ 和一个函数 $f:\ x_0 \to x_1$，使得 $f:\ \mathcal{M}_{x_0} \precsim \mathcal{M}_{x_1}$。

但根据事实2.4中的各条款，很容易看出这是错误的。事实上，由于 f 是一一的且 $x_0 \neq x_1$，根据 2.4 (i)、(ii)、(iii)，$x_0 \in x_1$，因此 x_0 是 x_1的真初始段。此外，根据初等性，我们可以对 x_0 作归纳，证明 f 是恒等关系。[$0 = \emptyset$ 是 x_0 和 x_1的第一个元素，因此 $f(0) = 0$。归纳地，若 $f(n) = n$，那么 n的下一个元素应被映射到 $f(n)$的下一个元素，也就是说，$f(n+1) = n+1$。] ω^* 的元素全都是后继序数，因此令 $x_0 = y_0 \cup \{y_0\}$, $x_1 = y_1 \cup \{y_1\}$。由于$x_0 \neq x_1$，我们有 $y_0 \neq y_1$。根据 2.4 (iv)，y_0 是 x_0 的最大元素，且$f(y_0) = y_0$，因为 f是恒等关系；而根据初等性，$f(y_0)$应是 x_1中最大的元素 y_1。但 $y_1 \neq y_0$，矛盾。⊣

接下来我们来证明替换公理。

定理 2.13 (i) EST $+$ VP $\vdash$ Δ_0-替代公理。(ii) EST $+$ VP $+$ Δ_0-$Rep \vdash$ 替代公理。(iii) 因此 EST $+$ VP $\vdash$ 替代公理。

证明。(i) 为证明 Δ_0-Rep，设 $\phi(x,y)$ 是一个 Δ_0 公式，使得 $(\forall x)(\exists! y)\phi(x,y)$。这定义了一个类映射 $F_\phi: V \to V$，满足 $F_\phi(x) = y$ 当且仅当$\phi(x,y)$。固定一个集合 A。只需证明类 $B = F_\phi\text{“}A = \{F_\phi(a) : a \in A\}$ 是一个集合。令 $L_A = \{\mathsf{U}\} \cup \{\mathsf{c}_a : a \in A\}$ 为带有一个一元关系符号 U 和一个常元 c_a（对每个 $a \in A$）的语言。对于每个 $b \in B$，$A \times \{b\}$ 都是一个集合，因此对每个这样的 b，考虑这个 L_A结构

$$\mathcal{M}_b = \langle A \times \{b\}, U_b, I_b \rangle,$$

其中 $U_b \subseteq A \times \{b\}$ 定义如下：对于每个 $a \in A$

$$\langle a, b \rangle \in U_b \iff F_\phi(a) = b.$$

我们有 $U_b = \{\langle a, b\rangle \in A \times \{b\} : \phi(a,b)\}$。根据 $CartProd$，$A \times \{b\}$是一个集合；又因为 ϕ是 Δ_0，所以根据 Δ_0-Sep，U_b是一个解释 U 的集合，即 $\mathsf{U}^{\mathcal{M}_b} = U_b$。常元赋值 $I_b: A \to A \times \{b\}$ 由 $I_b(a) = \langle a, b\rangle$定义，对每个 $a \in A$均如此。根据 Δ_0-Sep，I_b也是一个集合，因为 $I_b = \{\langle x, \langle y, b\rangle\rangle \in A \times (A \times \{b\}) : x = y\}$ 且 $x = y$是 Δ_0。这意味着，对于所有 $a \in A$ 和$b \in B$，$\mathsf{c}_a^{\mathcal{M}_b} = \langle a, b\rangle$。令

$$S = \{\mathcal{M}_b : b \in B\}.$$

只需证明 S 是一个集合。假设确实如此。那么显然存在某个 $n \in \omega$（实际上是 $n = 7$），使得 $B \subset \cup^n S$。此外

$$B = \{y \in \cup^n S : (\exists x \in A)(F_\phi(x) = y)\} = \{y \in \cup^n S : (\exists x \in A)\phi(x, y)\}.$$

由于 $\cup^n S$ 是一个集合，且公式（$\exists x \in A$）ϕ（x, y）是 Δ_0，因此根据Δ_0-分离公理可知，B 是集合。

因此，我们来验证 S 是一个集合。为导出矛盾，假设S 是真类。那么根据 Vopěnka 原理，存在 b、$c \in B$、$b \neq c$以及一个映射f：$A \times \{b\} \to A \times \{b\}$，使得 $f: \mathcal{M}_b \precsim \mathcal{M}_c$。根据初等性，对每个$a \in A$,

$$f(\langle a, b\rangle) = f(\mathbf{c}_a^{\mathcal{M}_b}) = \mathbf{c}_a^{\mathcal{M}_c} = \langle a, c\rangle.$$

另一方面，再次根据初等性，对每个 $a \in A$,

$$F_\phi(a) = b \Leftrightarrow \langle a, b\rangle \in U_b \Leftrightarrow f(\langle a, b\rangle) \in U_c \Leftrightarrow \langle a, c\rangle \in U_c \Leftrightarrow F_\phi(a) = c,$$

这就产生了矛盾，因为 $b \neq c$。

（ii）现在我们在 EST+VP+Δ_0-Rep 中工作，并证明完全替代成立。证明大体上与上面（i）条的证明相似。设$\phi(x, y)$是一个公式，满足（$\forall x$）（$\exists! y$）ϕ(x, y)，并设 $F_\phi(x) = y$当且仅当 $\phi(x, y)$。我们再次固定一个集合 A，并证明若 $B = F_\phi\text{“}A$，则 B 是一个集合。我们像之前一样定义结构 $\mathcal{M}_b$，并令 $S = \{\mathcal{M}_b : b \in B\}$。如（i）中所示，利用 Vopěnka 原理可知 S 不可能是真类。因此 S 是一个集合。与（i）的证明唯一不同之处在于：从 S 是一个集合推断出B 是一个集合。现在可以借助 Δ_0-替代公理推得这一结论：只需注意，映射 $S \ni \mathcal{M}_b \mapsto b \in B$ 显然是Δ_0-可定义且满射的。因此 B 是一个集合。

(iii) 由 (i) 和 (ii) 立即可得。 ⊣

现在我们来证明幂集公理，这一证明基于定理 2.13 的一个子句。

定理 2.14 EST + Vopěnka 原理 + Δ_0-替代公理 $\vdash Pow$。因此，根据定理 2.13 (i)，EST + Vopěnka 原理 $\vdash Pow$。

证明。我们在 EST $+$Vopěnka 原理 $+\Delta_0$-Rep 中进行。固定一个集合 A，并再次令 $L_A =$\{U\}$\cup$\{c$_a$: $a \in A$\}，其中 U 是一个一元关系符号。对于每个 $X \in \mathcal{P}(A)$，考虑这个 L_A-结构

$$\mathcal{M}_X = \langle A, X, id_A \rangle,$$

其中，对每个 $X \subseteq A$，$\mathrm{U}^{\mathcal{M}_X} = X$，且常元赋值是恒等映射 id_A: $A \to A$。注意，根据 Δ_0-Sep，id_A 在 EST 中是一个集合，因为 $id_A = \{\langle x, y\rangle \in A \times A : x = y\}$。因此，对每个 $a \in A$，$\mathrm{c}_a^{\mathcal{M}_X} = a$。为导出矛盾，假设 $\mathcal{P}(A)$ 是一个真类。令

$$K = \{\mathcal{M}_X : X \in \mathcal{P}(A)\}.$$

映射 $K \ni \mathcal{M}_X \mapsto X \in \mathcal{P}(A)$ 显然是 Δ_0，因此根据 Δ_0-替换公理，类 K 也是真类。由 Vopěnka 原理，存在 $X \neq Y \in \mathcal{P}(A)$以及一个初等嵌入 f: $\mathcal{M}_X \to \mathcal{M}_Y$。但对每个$a \in A$，$f(a) = f(\mathrm{c}_a^{\mathcal{M}_X}) = \mathrm{c}_a^{\mathcal{M}_Y} = a$。也就是说，$f = id_A$。另一方面，根据初等性，$f$ 应当将 1-1 X 映到 Y上，因此 $X = Y$，矛盾。$\dashv$

现在，ZF $=$EST$+\{Inf, Pow, Rep, Found\}$。因此，由定理 2.12、2.13 和 2.14，我们立即得到以下结论：

推论 2.15 EST $+ Found +$VP $=$ZF $+$VP，且 EST $+ Found +$AC $+$ VP $=$ ZFC $+$ VP。

如本节开头所述，将 VP$_4$（在定理 2.2 中使用）替换为 VP（在定理 2.12、2.13 和 2.14 中使用），是因为定理 V-P-H 属于定理 2.2 的假设之一，而定理 2.12、2.13 和 2.14 并非如此。考察 [18]，中 V-P-H 的证明可知，它严重依赖选择公理，以及以下两个事实：(a) 对于每个良序集 $\langle x, \leq\rangle$，存在一个（唯一的）序数α，使得 $\langle x, \leq\rangle \cong \langle \alpha, \in\rangle$。(b) 对于每个序数 α，在 α，$\{\beta < \alpha$ 以下存在由可数共尾性的序数组成的集合：cf(β) $= \omega\}$。(a) 需要 Δ_1-替换公理，而 (b) 首先要求 ω 是一个集合，其次要求Σ_1-分离公理。既然具备这些条件，V-P-H 的证明便可顺利进行，因此

$$\mathrm{EST} + \{\mathrm{AC}, \text{“}\omega \text{ is a set”}, \Delta_1\text{-}Rep, \Sigma_1\text{-}Sep\} \vdash \text{V-P-H}.$$

有了 V-P-H，我们就可以完全仿照 2.2 的证明进行推导，只需用 VP$_4$替代 VP（在任意集合 A 上的刚性二元关系 R，即可发挥常元 c$_a$ 在 VP 中的作用）。由此得到以下结果。

推论 2.16 理论

$$\mathrm{EST} + \{\mathrm{AC}, \text{“}\omega \textit{ is a set”}, \Delta_1\text{-}Rep, \Sigma_1\text{-}Sep\} + \mathrm{VP}_4$$

证明了替换公理和幂集公理。

推论 2.16 与上述结果 2.13 和 2.14 形成鲜明对比；这两个结果共同表明，仅凭 EST+VP 就能证明替换公理和幂集公理。这体现了在 EST 上，原理 VP$_4$ 与 VP 表面上强度差异的程度。

注 2.17 还值得一提的是，在定理 2.13 和 2.14 的证明中，我们并未用到 VP 的全部效力。简单检查这些证明即可发现，在这些结果中，我们用到的事实是：对于给定的由 L_A 结构构成的真类 X_ϕ，存在两个不同的结构 $\mathcal{M}$、$\mathcal{N}$，它们属于 X_ϕ

，且存在一个嵌入，但仅有 $f\colon\ \mathcal{M}\to\mathcal{N}$，而非初等嵌入。相比之下，定理 2.12 则要求 f 具备某种初等性。因此，仍可借助以下较弱形式的沃彭卡原理 VP_0，来证明定理 2.13 和 2.14。若将 VP 中的“初等嵌入”替换为“嵌入”（同时语言 L_A仍包含有限多个关系和任意集合的常元），即可得到 VP_0。那么 $\mathrm{EST}+\mathrm{VP}_0$ 可以证明替换公理和幂集公理。

由于 $\mathrm{EST}\ \subset$ LZFC（参见定理 2.2 后的脚注 1），作为定理 2.13 和 2.14 的直接推论，我们得到对 [17] 中定理 6.3 的如下改进：

定理 2.18 LZFC $+$ VP 可证明替换公理和幂集公理，也就是说，ZFC$\subseteq$ LZFC $+$ VP。

显然，这一改进在于：LZFC 不再需要满足 T 才能证明 V-P-H。

另一方面，ZFC+VP 蕴含存在一个可延拓基数构成的真类（参见 [11] 中的引理 20.25 及其证明后紧接着的评注）。由于对每个可延拓基数 κ，$V_\kappa\models$ ZFC，我们立即推得 ZFC $+$ VP $\vdash Loc$(ZFC)，其中 Loc(ZFC) 是 LZFC 的核心公理，意为“每个集合都属于某个 ZFC 的传递模型”。由此可得：

引理 2.19 LZFC $\subset$ ZFC $+$ Vopěnka 原理。

由定理 2.18 和引理 2.19 可得：

定理 2.20 LZFC $+$ Vopěnka 原理 $=$ ZFC $+$ Vopěnka 原理。

3 Vopěnka 原理和基础公理

设 $\mathrm{ZF}_0=$ ZF $-\{Found\}$。本小节将证明，在 ZF_0 上，Vopěnka 原理不能推出基础公理。具体而言，以下结论成立：

定理 3.1 如果 ZF $+$ Vopěnka 原理是一致的，那么 ZF_0+ Vopěnka 原理 $+\neg Found$ 也是一致的。将 ZF 换成 ZFC 时，结论同样成立。

证明采用一种众所周知的方法：利用由 V 的可定义置换产生的非标准隶属关系 $\in_\pi$ in V。具体而言，有一个相当广为人知的结果：若 V 是 ZFC 的全集，$\pi\colon V\to V$ 是可定义置换，且 $\in_\pi$ 是由 $x\in_\pi y$iff $x\in\pi(y)$ 定义的二元关系，那么 $V_\pi\models\mathrm{ZFC}_0$（见 [12, 第 IV 章，习题 18]）。为简单起见，以下将 $\langle V,\in_\pi\rangle$ 简记为 V_π，将 $\langle V,\in\rangle$ 简记为 V。要使 V_π 中的基础公理失效，只需取 π，使得对某个 x，有 $x\in\pi(x)$。我们将借此证明下面的定理，并由此推出定理 3.1。

定理 3.2 若 $V\models$ ZF $+$ Vopěnka 原理，则存在一个置换 $\pi\colon V\to V$，使得 $V_\pi\models$ Vopěnka 原理 $+\neg Found$。

我们首先需要一些预备定义和引理。给定一个可定义置换 $\pi\colon V\to V$，对于 $\mathcal{L}=\{\in\}$ 的每个公式ϕ，我们用 ϕ^π 表示如下得到的公式：将 ϕ 中出现的每个原子子公式 $x\in y$ 替换为 $x\in\pi(y)$，所得公式即为 ϕ。通过对 ϕ 的长度作归纳法，很容易验证以下结论。

引理 3.3 (i) 映射 $\phi \mapsto \phi^\pi$ 与逻辑联结词和量词可交换，也就是说，$(\phi \to \psi)^\pi = (\phi^\pi \to \psi^\pi)$，$(\neg\phi)^\pi = \neg\phi^\pi$，以及 $(\forall x\phi)^\pi = (\forall x)\phi^\pi$。(ii) 对每个语句 ϕ，$V_\pi \models \phi$ 当且仅当 $V \models \phi^\pi$。

对于 V 的每个标准概念，例如单元素集、有序对、 n 元组、关系、函数，都有一个相应的、关于 V_π 的 π 概念。例如，π 对是满足 $V_\pi \models$ 的集合 z，其中"z 是一个有序对"。后者成立，当且仅当存在某个 x、y $V_\pi \models z = \langle x、y\rangle$。根据 3.3 (ii)，这等价于 $(z = \langle x, y\rangle)^\pi$ （我们通常只写 ϕ，而不写 $V \models \phi$）。后者也更方便地记作 $z = \langle x, y\rangle^\pi$。类似地， Q 是集合 M 与 N 之间的 π 二元关系这一事实，意味着 $(Q \subseteq M \times N)^\pi$ 为真。由语句 ϕ 表达的 π 概念称为绝对的，如果 $\phi^\pi \leftrightarrow \phi$。在接下来的两个引理中，我们将给出一些关于置换 π 的简单充分条件，以确保某些关键概念是绝对的。

Lemma 3.4 *Suppose that $\pi : V \to V$ fixes all finite sets. Then:*

(i) π-配对是绝对的，也就是说，对所有 x、 y、 $\langle x、 y\rangle^\pi = \langle x、 y\rangle$。(ii) 对所有 M、N、Q，$(Q \subseteq M \times N)^\pi \leftrightarrow \pi(Q) \subseteq \pi(M) \times \pi(N)$。(iii) $[f: M \to N$是一个函数$]^\pi$ 等价于 $\pi(f): \pi(M) \to \pi(N)$ 是一个函数。

证明。（i）假设 π 固定所有有限集。那么 π^{-1} 也固定所有有限集。分析 $(z = \langle x, y\rangle)^\pi := (z = \{\{x\}, \{x, y\}\})^\pi$ 的定义后，很容易看出

$$z = \langle x, y\rangle^\pi \leftrightarrow z = \pi^{-1}(\{\pi^{-1}(\{x\}), \pi^{-1}(\{x, y\})\}). \tag{1}$$

(1) 右侧 π^{-1} 的所有参数都是有限的，因此 π^{-1} 固定它们。所以 (1) 蕴含

$$z = \langle x, y\rangle^\pi \leftrightarrow z = \{\{x\}, \{x, y\}\} = \langle x, y\rangle.$$

(ii) 分析 $(Q \subseteq M \times N)^\pi$ 的定义可知

$$(Q \subseteq M \times N)^\pi \leftrightarrow \pi(Q) \subseteq \{\langle x, y\rangle^\pi : x \in \pi(M), y \in \pi(N)\}. \tag{2}$$

由上面的 (i)，对每一对都有 $\langle x, y\rangle^\pi = \langle x, y\rangle$。因此 (2) 得出

$$(Q \subseteq M \times N)^\pi \leftrightarrow \pi(Q) \subseteq \{\langle x, y\rangle : x \in \pi(M), y \in \pi(N)\} = \pi(M) \times \pi(N).$$

(iii) 容易验证。 ⊣

回想一下，语言 L_A 由符号 R、U 和 c_a 组成，其中 $a \in A$，而一个 L_A-结构是一个四元组 $\mathcal{M} = \langle M, U, R, I\rangle$，其中 $R \subseteq M \times M$,$U \subseteq M$，并且 $I: A \to M$ 是一个映射。

给定一个 L_A 结构 $\mathcal{M} = \langle M, U, R, I\rangle$ 以及一个置换 $\pi: V \to V$，令

$$\mathcal{M}^\pi := \langle \pi(M), \pi(U), \pi(R), \pi(I)\rangle.$$

引理 3.5 假设 $\pi: V \to V$ 固定所有有限集和 ω。

(i) 令 L_A 为 V_π 意义下的一阶语言，其中 $A \in V_\pi$。那么 $\pi(L_A) = L_{\pi(A)}$ 是 V 中的语言。

(ii) 若 σ 是 V_π 意义下的 L_A 公式，则 σ 是 V 中的 $L_{\pi(A)}$ 公式。

(iii) 若 $\mathcal{M}$ 是 V_π 中的 L_A-结构，则 $\mathcal{M}^\pi$ 是 V 中的 $L_{\pi(A)}$-结构。

(iv) 若 $\mathcal{M}$ 是 V_π 中的 L_A-结构，则

$V_\pi \models [x_0, x_1 \in M \wedge (\mathcal{M} \models \mathrm{R}(x_0, x_1))] \leftrightarrow V \models [x_0, x_1 \in \pi(M) \wedge (\mathcal{M}^\pi \models \mathrm{R}(x_0, x_1))]$,对于谓词 U，同理。

证明。（i）回忆一下，根据定义 2.5,

$$V_\pi \models L_A = \{2,3,4,5,6,7\} \cup \{\langle 0,n\rangle : n \in \omega\} \cup \{\langle 1,a\rangle : a \in A\}.$$

鉴于引理 3.3，这等价地写作

$$V \models \pi(L_A) = \{2,3,4,5,6,7\}^\pi \cup \{\langle 0,n\rangle^\pi : n \in \pi(\omega)\} \cup \{\langle 1,a\rangle^\pi : a \in \pi(A)\}.$$

由于 π 固定所有有限集，并且根据引理 3.4， ω、 π-配对是绝对的，$\{2,3,4,5,6,7\}^\pi = \{2,3,4,5,6,7\}$ 且 $\pi(\omega) = \omega$，因此

$$V \models \pi(L_A) = \{2,3,4,5,6,7\} \cup \{\langle 0,n\rangle : n \in \pi(\omega)\} \cup \{\langle 1,a\rangle : a \in \pi(A)\}.$$

但上式右侧显然是语言 $L_{\pi(A)}$，因此 $\pi(L_A) = L_{\pi(A)}$。

（ii）由于对每个 $n \in \omega$， $\pi(\omega) = \omega$ 且 $\pi(n) = n$，所以 ω 在 V_π 中是绝对的。因此，沿着定义 2.7 的步骤对 σ 作简单归纳，即可得出该结论。

（iii） $\mathcal{M} = \langle M、U、R、I\rangle$ 是 V_π 意义下的一个 L_A-结构，这意味着 $(U \subseteq M)^\pi$、$(R \subseteq M \times M)^\pi$，以及 $(I : A \to M$ 是函数$)^\pi$。由于 π 固定所有有限集，根据 3.4，这些事实分别转化为 V，即 $\pi(U) \subseteq \pi(M)$、$\pi(R) \subseteq \pi(M) \times \pi(M)$，以及 $\pi(I) : \pi(A) \to \pi(M)$ 是函数。但这意味着 $\langle \pi(M)\ \pi(U)\ \pi(R)\ \pi(I)\rangle$，也就是说， $\mathcal{M}^\pi$ 是一个 $L_{\pi(A)}$-结构。

(iv) 设 $\mathcal{M} = \langle M, U, R, I\rangle$ 是 V_π 中的一个 L_A-结构，并设

$$V_\pi \models [x_0, x_1 \in M \wedge (\mathcal{M} \models \mathbf{R}(x_0, x_1))].$$

显然，这等价地可写为

$$V_\pi \models [x_0, x_1 \in M \wedge \langle x_0, x_1\rangle \in R)].$$

其在 V 中的翻译为

$$V \models [x_0, x_1 \in \pi(M) \wedge \langle x_0, x_1\rangle^\pi \in \pi(R))].$$

由于根据关于 π 的条件， $\langle x_0, x_1\rangle^\pi = \langle x_0, x_1\rangle$，后者也等价于

$$V \models [x_0, x_1 \in \pi(M) \wedge (\mathcal{M}^\pi \models \mathbf{R}(x_0, x_1))].$$

⊣

引理 3.6 设 π 是一个固定所有有限集和 ω 的置换，并设 $\mathcal{M} = \langle M, U, R, I\rangle$, $\mathcal{N} = \langle N, Z, S, J\rangle$ 是 V_π 中的 L_A-结构。那么对于任意 f,

$$V_\pi \models [f : \mathcal{M} \precsim_{L_A} \mathcal{N}]$$

iff

$$V \models [\pi(f) : \mathcal{M}^\pi \precsim_{L_{\pi(A)}} \mathcal{N}^\pi].$$

证明。我们简要证明“→”这一方向。另一方向的证明类似。假设 π 如所述，$\mathcal{M}, \mathcal{N}$ 是 V_π 和$V_\pi \models [f$ 中的 L_A-结构： $\mathcal{M} \precsim_{L_A} \mathcal{N}]$，也就是说， $f : M \to N$ 是一个 L_A-初等嵌入。根据引理 3.5， $\mathcal{M}^\pi, \mathcal{N}^\pi$ 是 V中的 $L_{\pi(A)}$-结构。我们必须证明$\pi(f) : \pi(M) \to \pi(N)$是一个 $L_{\pi(A)}$-初等嵌入。我们必须证明，对于每个 $L_{\pi(A)}$ 公式 $\sigma(v_0, \ldots, v_{n-1})$以及任意的 $x_0, \ldots, x_{n-1} \in \pi(M)$,

$$\mathcal{M}^\pi \models \sigma(x_0, \ldots, x_{n-1}) \leftrightarrow \mathcal{N}^\pi \models \sigma(\pi(f)(x_0), \ldots, \pi(f)(x_{n-1})).$$

这可通过对 σ的长度作常规归纳来证明。这里只证明上述结论对 $L_{\pi(A)}$的原子句子 $\mathbf{R}(x_0, x_1)$成立的情形。这相当于证明：若 $x_0, x_1 \in \pi(M)$且 $y_0, y_1 \in \pi(N)$，则（在 V 中）：

$$\langle x_0, y_0\rangle \in \pi(f) \wedge \langle x_1, y_1\rangle \in \pi(f) \rightarrow$$
$$(\mathcal{M}^\pi \models \mathbf{R}(x_0, x_1) \leftrightarrow \mathcal{N}^\pi \models \mathbf{R}(y_0, y_1)). \tag{3}$$

但根据我们的假设 $V_\pi \models [f\colon \mathcal{M} \precsim_{L_A} \mathcal{N}]$，对于所有 $x_0, x_1 \in M$和 $y_0, y_1 \in N$，我们有：

$$V_\pi \models [\langle x_0, y_0\rangle \in f \wedge \langle x_1, y_1\rangle \in f \rightarrow$$
$$(\mathcal{M} \models \mathbf{R}(x_0, x_1) \leftrightarrow \mathcal{N} \models \mathbf{R}(y_0, y_1)). \tag{4}$$

由引理 3.3 和 3.5 (iv)（这由我们关于 π 的条件保证成立），(3) 只是将 (4) 翻译为 V。归纳法的其他步骤都是常规的。至此，引理的证明完成。⊣

现在我们可以证明 3.2 了。

定理 3.2 的证明。设 $V \models$ ZF +Vopěnka 原理。取 V 的一个置换 π，它固定所有有限集和 ω。那么上文的引理 3.4、3.5 和 3.6 均成立。再假设存在某些集合 X, Y，使得 $X \in Y$, π 交换 X 和 Y，即 $\pi(X) = Y$, $\pi(Y) = X$，因此基础公理在 V_π 中不成立。剩下只需证明 $V_\pi \models$Vopěnka 原理。

令 L_A 为 V_π意义下的一种语言。集合 A本质上包含该语言的常元，可以是任意的，因此我们特别可以有 $A = X$ 或 $A = Y$。因此一般而言，$\pi(A) \neq A$。令 $\phi(x)$为 $\mathcal{L} = \{\in\}$ 的一个公式，使得

$$V_\pi \models \text{ “}X_\phi \text{ is a proper class of } L_A\text{-structures”}. \tag{5}$$

我们必须证明

$$V_\pi \models (\exists \mathcal{M} \neq \mathcal{N} \in X_\phi)(\mathcal{M} \precsim_{L_A} \mathcal{N}). \tag{6}$$

现在，由 (5) 可知，X_ϕ 是 V_π中的一个真类，即

$$V_\pi \models (\forall x)(\exists y)(\phi(y) \wedge y \notin x), \tag{7}$$

并且

$$V_\pi \models (\forall x)[\phi(x) \rightarrow x \text{ is an } L_A\text{-structure}). \tag{8}$$

由 (7) 可知

$$V \models (\forall x)(\exists y)(\phi^\pi(y) \wedge y \notin \pi(x)), \tag{9}$$

即，X_{ϕ^π} 是 V中的一个真类。此外，如果 $\mathcal{M} \in X_{\phi^\pi}$，那么 $V_\pi \models \phi(\mathcal{M})$，因此由 (8)，$\mathcal{M}$ 是 V_π中的一个 L_A-结构。根据引理 3.5 (iii)，$\mathcal{M}^\pi$ 是 V中的一个 $L_{\pi(A)}$-结构。因此

$$\mathcal{M} \in X_{\phi^\pi} \rightarrow \mathcal{M}^\pi \text{ is an } L_{\pi(A)}\text{-structure}. \tag{10}$$

考虑由 $\mathcal{L}$ 定义的公式 $\psi(x)$

$$\psi(x) := (\exists \mathcal{M})(x = \mathcal{M}^\pi \wedge \phi^\pi(\mathcal{M})).$$

那么显然，对于每个 $\mathcal{M}$

$$V \models \phi^\pi(\mathcal{M}) \iff V_\pi \models \phi(\mathcal{M}) \iff V \models \psi(\mathcal{M}^\pi). \tag{11}$$

由 (10) 和 (11) 可知， X_ψ 的元素是 $L_{\pi(A)}$-结构。此外，函数对应 $X_{\phi^\pi} \ni \mathcal{M} \mapsto \mathcal{M}^\pi \in X_\psi$ 是一一对应的，因此 X_ψ 是一个真类，因为 X_{ϕ^π} 也是如此。由于 Vopěnka 原理在 V 中成立，我们有

$$V \models (\exists x, y)(\psi(x) \wedge \psi(y) \wedge x \neq y \wedge x \precsim_{L_{\pi(A)}} y). \tag{12}$$

选取两个不同的 X_ψ 结构 $\mathcal{M}^\pi$和 $\mathcal{N}^\pi$ ，使得

$$\mathcal{M}^\pi \precsim_{L_{\pi(A)}} \mathcal{N}^\pi.$$

那么，根据引理 3.6，可知 $V_\pi \models \mathcal{M} \precsim_{L_A} \mathcal{N}$ ，并且 $\mathcal{M} \neq \mathcal{N} \in X_{\phi^\pi}$。因此

$$V_\pi \models \mathcal{M} \neq \mathcal{N} \in X_\phi \wedge \mathcal{M} \precsim_{L_A} \mathcal{N}.$$

但这正是所需的结论（6）。证明完毕。 ⊣

4 Vopěnka 原理和选择公理

关于 Vopěnka 原理与 ZFC 公理，尚未解决的问题是 Vopěnka 原理与选择公理之间的关系，具体如下:

问题 4.1 假设 ZFC + Vopěnka 原理是一致的，那么选择公理是否独立于策梅洛-弗兰克尔集合论 +Vopěnka 原理?

下面我们作两点评注：一点关于选择公理的独立性，另一点关于相反方向。

4.1 Vopěnka 原理与对称模型

我们猜想，选择公理独立于 ZF+Vopěnka 原理。然而，要证明这一点，最自然的方法似乎是运用 ZFA 的置换模型技术，或 ZF 的对称泛型模型技术；这些都是否定选择公理的标准工具。关于该方法的技术细节，可参见[11]。此外，若要全面了解各种对称模型及其应用，可以参阅 [9, 第三部分]。

这些方法必然会引出以下步骤。我们从一个满足 V ZFA+选择公理+ Vopěnka 原理（或 ZFC+Vopěnka 原理）的模型开始，并选取一个对称模型 $HS \subset V$，以证明 $HS \models$ Vopěnka 原理 + ¬选择公理。假设已经有 $HS \models \neg$ 选择公理，剩下的就是证明 $HS \models$ Vopěnka 原理。设 X_ϕ 是由 $A \in HS$个 L_A-结构组成的真类，其含义见 HS。我们必须证明

$$HS \models (\exists x, y)(\phi(x) \wedge \phi(y) \wedge x \neq y \wedge x \precsim y). \tag{13}$$

由于 HS 是 V 的内子模型，按照 V 的含义， $X_{\phi^{HS}}$ 是由 L_A 结构组成的真类，其中 ϕ^{HS} 是将 ϕ 相对化到 HS 的通常方式。由于 VP 在 V 中成立，我们有

$$V \models (\exists x, y)(\phi^{HS}(x) \wedge \phi^{HS}(y) \wedge x \neq y \wedge x \precsim y). \tag{14}$$

因此，证明选择公理独立性的关键在于证明 (13) 可以由 (14) 推导出来。

事实上，要证明推导式 (14)→(13) 是一个颇具挑战的问题，无法通过“简单的方法”解决。所谓简单的方法，就是通过以下形式的蕴含来推导 (14)→(13)：对于任意语言L_A 以及任意 L_A-结构 $\mathcal{M}, \mathcal{N}$，在 HS 中，

$$V \models \mathcal{M} \precsim \mathcal{N} \ \rightarrow \ HS \models \mathcal{M} \precsim \mathcal{N}. \tag{15}$$

显然，如果 (15) 对每个 L_A都成立，那么蕴含式 (14)→ (13) 也成立。但一般而言，(15) 并不成立。因为如果取 $L_A = \emptyset$，那么 L_A-结构就是集合，而初等嵌入就是简单的单射。因此，(15) 特别会推出：对于所有 $x, y \in HS$

$$V \models |x| \leq |y| \ \rightarrow \ HS \models |x| \leq |y|. \tag{16}$$

但由于 $V \models$选择公理，

$$V \models (\forall x, y)(|x| \leq |y| \vee |y| \leq |x|).$$

因此由 (16) 可得，

$$HS \models (\forall x, y)(|x| \leq |y| \vee |y| \leq |x|).$$

最后一句话表明，在 HS 中，所有集合的基数都是可比较的；众所周知，这等价于选择公理（例如参见 [14, 定理 3.1]）。因此 $HS \models$选择公理，而这是错误的!

综上所述：若要对问题 4.1 作肯定回答，就必须找到一个对称模型 HS，并为蕴含关系 (14)→(13) 给出一个非平凡的证明，其中 $L_A \in HS$ 可以是任意对象，L_A 结构的真类 X_ϕ 也可以是任意的。

4.2 与 Vopěnka 原理相关的选择公理弱形式

现在我们来看看相反的方向。尽管 ZF+Vopěnka 原理不太可能证明选择公理，但或许可以推出选择公理的较弱形式。

在前面的评注中，我们提到了选择公理与这样一个事实之间众所周知的等价关系：任意两个集合的基数都可比较。选择公理的这一表述可以自然地弱化，而其中一种极端的弱化形式是 Vopěnka 原理的推论。

为便于讨论，我们约定，在策梅洛-弗兰克尔集合论中，如果两个集合 x、y 的基数可比较，也就是说，$|x| \leq |y|$ 或 $|y| \leq |x|$，那么这两个集合就是可比较的。否则，就称它们不可比较。对于每个只有一个自由变量的公式 $\phi(x)$，考虑以下可比较性公理：

$(Comp_\phi)$ *If X_ϕ is a proper class of sets, then it contains at least two comparable elements.*

下面的事实表明，Vopěnka 原理在一定程度上依赖于选择公理。

事实 4.2 对于每个性质 $\phi(x)$，策梅洛-弗兰克尔集合论 +Vopěnka 原理 $\vdash Comp_\phi$。

证明。设 X_ϕ 是一个由集合构成的真类。正如上文评注 1 中的论证所述，X_ϕ 可以看作是一个由 L-结构构成的真类，其对象为 $L = \emptyset$，其中一一映射就是初等嵌入。因此，根据 Vopěnka 原理，存在 $x \neq y \in X_\phi$ 使得 $|x| \leq |y|$。⊣

一旦定义了 $Comp_\phi$，其他类似的公理便自然出现。对于每个标准基数[4] $\kappa \geq 2$，考虑以下可比较性公理：

$$(Comp_\kappa) \qquad (\forall x)[|x| = \kappa \rightarrow (\exists y \neq z \in x)(y, z \text{ are comparable})].$$

[$Comp_\kappa$ 断言，每个基数为 κ 的集合都包含不同且可比较的元素。] $Comp_\kappa$ 随着 κ 增大而逐渐变弱。此外，$Comp_\phi$看起来像是 $Comp_\kappa$的“最弱极限”，不过，除蕴含关系外

[4]我们所说的标准基数是指初始序数。

$Comp_n \to Comp_\phi$、 $n \in \omega$显然在策梅洛-弗兰克尔集合论中成立，但当时尚不清楚哪些 $\kappa \geq \omega$ （如果存在的话） $Comp_\kappa \to Comp_\phi$ 也成立。

然而，就在不久前，Lefteris Tachtsis [15] 证明了 $Comp_\omega$ 蕴含（在策梅洛-弗兰克尔集合论之上）有限集与戴德金有限集相同（其中，若 X 满足 $\aleph_0 \not\leq |X|$，或者等价地，不存在单射 f： $X \to X$ 使得 f“$X \subsetneq X$，则称其为戴德金有限）。令 F 和 DF 分别表示有限集类与戴德金有限集类。众所周知，在策梅洛-弗兰克尔集合论之上，F $\subseteq$ DF，而选择公理蕴含 F =DF。

定理 4.3（Tachtsis [15]）在策梅洛-弗兰克尔集合论之上， $Comp_\omega$ 蕴含 F=DF。

由 4.3 可得以下结论。

命题 4.4 在策梅洛-弗兰克尔集合论中，以下结论成立：对于每个 ϕ，若 X_ϕ 是真类，则 $Comp_\omega \to Comp_\phi$。

证明。假设 $Comp_\omega$ 成立，且 X_ϕ 是真类。记 X 为 X_ϕ。鉴于 $Comp_\omega$，要证明 $Comp_\phi$ 成立，只需证明存在集合 $x \subset X$ 使得 $|x| = \aleph_0$。按归纳方式定义函数 f： $\omega \to On$ 如下： $f(0) = 0$， $f(n+1) = \min\{\beta > f(n) : X \cap V_{f(n)} \subsetneq X \cap V_\beta\}$。由于 X 是真类，对于每个 $n \in \omega$， f 都有定义。（否则，存在 $k \in \omega$，使得对每个 $\beta > f(k)$，$X \cap V_{f(k)} = X \cap V_\beta$。但这样一来 $X = X \cap V = X \cap V_{f(k)}$，所以 X 是一个集合。）令 $y_n = X \cap V_{f(n)}$。每个 y_n 都是集合，且 $y_n \subsetneq y_{n+1}$。因此显然 $\bigcup_n y_n$ 是一个无限集。由定理 4.3，$\bigcup_n y_n$ 是戴德金无限的，所以存在$x \subseteq \bigcup_n y_n \subset X$ 使得 $|x| = \aleph_0$。 $\dashv$

事实上，公理 $Comp_\kappa$ 并非全新。 $Comp_2$ 是前文提到的选择公理的等价表述，即任意两个集合都可比较。此外， $Comp_n$这些公理（对于 $2 \leq n < \omega$）也见于 [14, 第 22], 页，称为 T3(n)，并归于 A. Tarski（ 1964）。此外， [14, 定理 3.4]还给出了一个重要结果：对于每个 $2 \leq n < \omega$, $Comp_n$ 都等价于选择公理。[5]

对于每个 $\kappa \geq 2$， $\neg Comp_\kappa$的相容性，以及对于某个 ϕ， $\neg Comp_\phi$的相容性，都可由 [10]的以下结果证明：

定理 4.5（[10, 定理 11.1]）设 V是一个 ZFA + 选择公理的模型，含有一个由原子组成的集合 A ；又设 $\langle I, \preceq \rangle$ 是 V 中的一个偏序集，满足 $|A| = |I| \cdot \aleph_0$。那么存在一个满足以下条件的置换模型 $HS \subset V$：

$$\textit{There exists a family of sets } \{S_i : i \in I\} \textit{ such that for all } i, j \in I,$$

$$i \preceq j \leftrightarrow |S_i| \leq |S_j|. \tag{17}$$

置换模型的方法同样适用于原子构成真类而非集合的情况（见第 [10, 139] 页）。以这种模型为起点，不难将 4.5 加强如下：

定理 4.6 设 V 是 ZFA +选择公理 的一个模型，其中现在 A 是原子的真类。再设 $\langle I, \preceq \rangle$ 是一个偏序真类，使得对每个 $i \in I$, $\{j : j \preceq i\}$ 是一个集合，且 $A = \{a_{in} : i \in I, n \in \omega\}$。那么存在一个置换模型 $HS \subset V$ 和一个真类 $\{S_i : i \in I\} \subset HS$，使得 (17) 成立。

[5]感谢 Lefteris Tachtsis 提醒我注意到这篇参考文献。

证明。（概要）根据假设，对每个 $i \in I$，$\hat{}i = \{j \in I : \ j \preceq i\}$ 是一个集合，且 $i \preceq j \leftrightarrow \hat{}i \subseteq \hat{}j$。因此，$\langle I, \preceq \rangle$ 嵌入于 $\langle \mathcal{P}(I) \subseteq \rangle$ 中，其中 $\mathcal{P}(I)$ 是 I 的子集类，因此只需证明存在一个类$\{S_x : \ x \subset I\}$，使得对所有 $x, y \subset I$，$x \subseteq y \leftrightarrow |S_x| \leq |S_y|$。根据假设$A = \{a_{in} : \ i \in I, n \in \omega\}$。对每个 $x \in \mathcal{P}(I)$，令 $S_x = \{a_{in} : \ i \in x, n \in \omega\}$。考虑由满足 $\pi(S_{\{i\}}) = S_{\{i\}}$ 的 π 对 A 的置换组成的类群 $\mathcal{G}$；也就是说，对每个 $i \in I$ 和每个 $n \in \omega$，存在 $m \in \omega$，使得 $\pi(a_{in}) = a_{im}$($\mathcal{G}$ 是编码类的类)。再令 $\mathcal{F}$ 为由 A 的有限子集所成理想生成的 $\mathcal{G}$ 的（适当编码的）子群滤子。然后证明过程与上文定理 4.5 相同。⊣

现在令 $\langle I、 \preceq \rangle$ 为反链，即 $i \preceq i \leftrightarrow i = j$，由此可将前述定理的推论得出：对于特定的 ϕ，$\neg Comp_\kappa$ 和 $\neg Comp_\phi$ 的相对一致性。

推论 4.7 (i) 设 $V \models$ ZFA $+$选择公理，且有一个包含 A 个原子的集合，并设 κ 为满足 $|A| = \kappa \cdot \aleph_0$ 的基数。则存在一个置换模型 $HS \subset V$，使得 $HS \models \neg Comp_\kappa$。

(ii) 设 $V \models$ ZFA $+$选择公理，且有一个真类的原子 $A = \{a_{in} : \ i \in I, n \in \omega\}$。则存在一个置换模型 $HS \subset V$ 和一个类 $X_\phi = \{S_i : \ i \in I\}$ 属于 HS，使得 $HS \models \neg Comp_\phi$。

证明。(i) 令 I 为满足 $|I| = \kappa$ 的反链，该断言便由定理 4.5 立即得出。

（ii）类似地，对于一个反链 I，由定理 4.6 可得；该反链是一个真类，且 $X_\phi = \{S_i : \ i \in I\}$。⊣

最后，我们提出两个问题，探讨可比较性公理的相对强度及其与 Vopěnka 原理的关系。

问题 4.8 （1）对于 $\omega \leq \kappa < \lambda$，$Comp_\kappa \rightarrow Comp_\lambda$ 中的任一蕴含是否可逆？无论是在策梅洛-弗兰克尔集合论中，还是在策梅洛-弗兰克尔集合论 + Vopěnka 原理中？

（2）策梅洛-弗兰克尔集合论 +Vopěnka 原理是否能证明 $Comp_\kappa$，其中 $\kappa \geq 2$ 是否成立？

致谢。衷心感谢匿名审稿人对稿件的仔细审阅，以及提出的诸多更正和宝贵建议。

参考文献

[1] J. Ad´amek 和 J. Rosick´y，《局部可表示范畴与可接近范畴》，伦敦数学会讲义丛书，第189卷，剑桥大学出版社，1994年。[2] J. Bagaria 和 A. Brooke-Taylor，《关于余极限与初等嵌入》，J. Symb. Logic 78 (2012)，第2期，562–578。[3] S. Baratella 和 R. Ferro，《含有选择公理否定的集合理论》，Math. Logic. Quart. 39 (1993)，338–352。[4] J.L Bell 和 M. Machover，《数理逻辑教程》，North Holland P.C.，1977年。[5]A.D. Brooke-Taylor，《沃彭卡原理的不可破坏性》，Arch. Math. Logic 50 (2011)，515–529。

[6]K.J. Devlin，《可构造性》，《数学逻辑展望》，Spinger-Verlag，1984年。[7] F.D.Drake，《集合论：大基数导论》，Noth Holland P.C.，1974年。[8]H. Friedman（2005年），FOM档案251：嵌入公理（http://www.cs.nyu.edu/piper mail/fom/2005-August/009023.html).

[9] P. Howard 和 J.E. Rubin，《选择公理的推论》，《数学调查与专著》第59卷，美国数学学会，1998年。[10] T. Jech，《选择公理》，North-Holland P.C.，1973年。[11] T. Jech，《集合论：第三千年版》，施普林格，2003年。[12] K. Kunen，《集合论：独立性证明导论》，North Holland，阿姆斯特丹，1983年。[13] A.R.D. Mathias，《Gandy、Jensen和Devlin的弱系统》，载于：《集合论》，巴塞罗那数学研究中心，2003—2004年，Joan Bagaria 和 Stevo Todorcevic（编），《数学新趋势》，Birkhäuser，2006年，第149–224页。[14] H. Rubin 和 J.E. Rubin，《选择公理的等价命题》第二卷，North Holland P.C.，1985年。[15] E. Tachtsis，$Comp_\omega$ 蕴含 F =DF，未发表。[16]A. Tzouvaras，《公理的局部化》，《数学逻辑档案》49（2010年），第5期，571–601页；以及《〈公理的局部化〉勘误》，《数学逻辑档案》50（2011年），第3期，513页。[17] A. Tzouvaras，《局部ZFC中的大型传递模型》，《数学逻辑档案》53（2014年），第3–4期，233–260页。[18] P. Vopˇenka、A. Pultr 和 Z. Hedrl´ín，《任意集合上都存在刚性关系》，《卡罗来纳大学数学评论》6（1965年），149–155页。