第 48 卷 第 3 期（总 331 期）　自然辩证法通讯　Vol.48, No.3 (Serial No.331)
2026 年 3 月　Journal of Dialectics of Nature　March 2026

哥德尔不完全性定理摧毁了希尔伯特纲领吗?

Whether Gödel's Incompleteness Theorems Defeats Hilbert's Program?

陈龙 /CHEN Long

（北京师范大学哲学学院，北京，100875）
(School of Philosophy, Beijing Normal University, Beijing, 100875)

摘　要: 希尔伯特纲领是20世纪20年代希尔伯特为了应对以集合论悖论为代表的数学危机和构造主义的挑战而提出的为古典数学奠基的新方案。它不仅是希尔伯特哲学思想的具体数学体现，也是数学基础中形式主义的重要代表。学界通常认为希尔伯特纲领由于哥德尔不完全性定理的出现而遭受彻底失败，然而也有一些反对意见存在。通过区分对希尔伯特纲领三种不同的可能哲学阐释，可以系统而全面的探讨哥德尔不完全性定理是否对希尔伯特特定认识论意义下的希尔伯特纲领构成致命打击。通过聚焦于“一致性的不可证性”这一核心问题，我们给出肯定的看法，并说明这一论断的哲学含义。

关键词: 希尔伯特纲领　不完全性定理　一致性

Abstract: Hilbert's Program was a foundational proposal advanced by David Hilbert in the 1920s to address the mathematical crisis triggered by set-theoretic paradoxes and the challenges raised by constructivism, aiming to provide a new foundation for classical mathematics. The program not only embodies Hilbert's philosophical ideas in terms of mathematical program, but also stands as a major representative of formalism in the foundation of mathematics. While it is widely accepted in the academia that Hilbert's Program was fundamentally defeated by Gödel's incompleteness theorems, disagreements remain. By distinguishing three possible philosophical interpretations of the program, this paper systematically and comprehensively examines whether Gödel's incompleteness theorems deliver a fatal blow to Hilbert's Program under his specific epistemological interpretations. Moreover, by focusing on the key issue of the unprovability of consistency, we affirm this conclusion and elucidate its philosophical implications.

Key Words: Hilbert's Program; Gödel's Incompleteness theorem; Consistency

中图分类号: N031; O347.4+3　DOI: 10.15994/j.1000–0763.2026.03.007　CSTR: 32281.14.jdn.2026.03.007

引　言

20世纪20年代初，德国数学家希尔伯特（David Hilbert）为应对集合论悖论和构造主义及直觉主义的挑战提出了一个为古典数学奠基的新方案,后被称为“希尔伯特纲领”（Hilbert's Program, HP）。该纲领要求将全部古典数学以公理化方式进行形式化，并证明此数学公理化形式系统的一致性。其核心要点在于：这种一致性证明本身必须仅采用希尔伯特所称的“有穷主义”方法来完成。基于有穷主义推理特有的认识论显明形和直觉可靠性特征，一致性证明将为古典数学提供一劳永逸的正当性依据。HP不仅是希尔伯特自身哲学思想的具体数学体现，也是早期数学基础研究中形式主义最重要

基金项目: 教育部高等人文社会科学重点研究基地重大项目“社会文化视域下的概念与推理研究”（项目编号: 22JJD720021）。
收稿日期: 2025年9月16日
作者简介: 陈　龙（1988–）男，湖南岳阳人，北京师范大学哲学学院讲师，研究方向为数学哲学、逻辑哲学。Email: long.chen@bnu.edu.cn

的代表。然而,正是受HP的启发,哥德尔（Kurt Gödel）发现了算术形式系统的不完全性定理，自此学界普遍认为不完全性定理表明HP无法完全实现。然而，一方面，与普遍认为哥德尔第二不完全性定理摧毁了HP的观点相反，哥德尔在1931年论文中仅谨慎表示："我要明确强调，定理XI［第二不完全性定理］并不否定希尔伯特的形式主义立场。因为该立场仅预设存在仅使用有穷主义证明手段的一致性证明，而可能存在无法在形式系统中表达的有穷主义证明"。（［1］，p.195）至迟在1958年，他明确转变立场："若要证明经典数学甚至经典数论的一致性，必须超越希尔伯特意义下的有穷主义数学框架"。（［2］，p.241）另一方面，由于一致性概念的形式化表述复杂性以及若干一致性断言的可证性，有学者认为第二不完全性定理对HP的打击并非决定性的。[3]在第二节中，我们将首先概述HP,区分其三种不同哲学解释，并在第三节中重点探讨哥德尔定理①是否对希尔伯特特定认识论意义下的HP构成致命打击。通过聚焦于"一致性不可证明性"这一核心问题，我们论证尽管广义下的HP仍是数学哲学领域极具影响力的理论立场，也已成为证明论发展的核心动力，但要符合希尔伯特特定认识论意义下的HP却注定是失败的。后文将详细说明不同解读下失败的具体理由，并说明其哲学含义。

一、HP及其哲学内涵

1. HP的核心要义

通常认为HP是希尔伯特为了应对有集合论悖论所出现的第三次数学危机所提出的方案，关于HP方案的原始形态及当代发展，有大量文献论述，比如扎克（Richard Zach）、[4]，[5]斯莫林斯基（Craig Smoryński）[6]和叶峰（［7］，Chap.6）等。然而，希尔伯特对数学基础问题的关注早在集合论悖论出现之前就已开始。非欧几何的发现、康托尔集合论的兴起以及"分析的算术化"（arithmetization of analysis）等都是19世纪下半叶基础争论中的焦点问题——这场争论以克罗内克（Leopold Kronecker）为一方，强调数学方法、对象及性质必须具备构造性与可判定性；而以戴德金（Richard Dedekind）和康托尔（Georg Cantor）为代表的另一方认为数学的本质在于不受任何关于方法或对象的先验哲学限制，只要这些方法能导致丰富的数学成果，便不排斥抽象或非构造性方法。希尔伯特无疑属于戴德金–康托尔阵营，他意识到公理化方法对于这场关于数学对象与方法本质争论的至关重要性。在其首部基础著作《几何基础》中，[8]希尔伯特特意选择以欧几里得几何这一最古老且被认为绝对可靠的数学领域作为试验场，展示了形式公理化方法的威力，还详细探讨了后来成为其元数学核心关注的问题——如一致性、完备性和独立性。与公理化方法相伴的是对其一致性的证明需求——即我们必须确保从公理系统出发，通过有限步骤推演绝不会得出矛盾，否则该公理系统既能推导一切，又无法定义任何内容。然而，集合论悖论的出现，似乎预示着那些抽象、非构造性推理潜藏着不确定性与危险。这对康托尔"数学是自由创造"的构想构成了实质威胁。贝奈斯（Paul Bernays）对此情境有过生动描述：在集合论悖论发现的影响下，希尔伯特曾短暂认为克罗内克的观点或许是正确的。但他很快改变了立场。可以说，此时他的目标转变为：通过革新数学观念，运用有穷性这一克罗内克自己的理论武器来反击其限制性主张。（［9］，p.173）

这一革新后的数学观念正是希尔伯特著名的形式数学与非形式内容性元数学之区分。形式数学即公理化数学，它可以通过罗素与怀特海在《数学原理》中的逻辑形式体系这一新工具来严格表述，希尔伯特本人在这一新工具的创立过程中也起到过重要作用。[10]除了形式化

①粗略而言，第一不完全性定理的非形式化表述是：对于任何包含特定算术的形式系统，都存在不可判定命题——即既不能被证明，也不能被否证的陈述。第二不完备性定理表明：对于任何满足特定基本条件的形式系统而言，该系统的一致性无法通过该系统自身可形式化的方法来证明。

的公理化数学，我们还需另一种数学——即旨在为形式化数学提供保障的元数学。正如希尔伯特在首次阐述这一新理念时所言：

> 在这门真正的数学之外，还出现了一种在某种程度上全新的数学——元数学，它通过双重防护来捍卫数学：既使其免受不必要的各种限制，又使其摆脱悖论的困扰。在此元数学中——与纯粹数学中纯粹形式的推理模式相反——我们运用内容性推理（contentual inference）；尤其适用于公理一致性的证明。（［11］，p.1132）

需要特别指出的是，即便希尔伯特强调捍卫或保护真正的数学，但他从未质疑该学科的数学正确性，而仅是从认识论角度追问其严格基础。因此他坚决反对外尔（Hermann Weyl）所谓“数学危机”的隐喻，[12]而只是认可外尔关于构造性的观点：相较于非构造性方法，构造性倾向具有更高程度的明证性与直觉确定性。如果我们轻率地将仅适用于有限情形的直觉方法无限推广至无限整体，便可能导致谬误。这一教训的普遍意义不在于限制数学本身，而在于夯实其基础与证成：“因此我们认识到：若要为数学奠定严格基础，就无法将分析学中常见的推理模式（如排中律）视为逻辑上无问题的。相反，我们的任务正是要查明——为何以及何种程度上——从分析与集合论中这类超穷推理模式的应用总能获得正确结果。对超穷方法的自由运用与完全驾驭必须在有穷主义的疆域内实现！”（［13］，p.1140）

对希尔伯特而言，唯有基于直观可靠的方法获得超穷形式数学系统的一致性证明，方能实现这种保障。如此一来，数学基础的认识论问题本身便转化为可严格处理的数学问题。希尔伯特视此为自身方法相较其他哲学方案的巨大优越性——关于数学认识论的问题，都可在数学内部解决，而无需引入任何外部要素，诸如克罗内克的“上帝”（赋予其整数概念），庞加莱（Henri Poincaré）获取数学归纳法真理的特殊能力，布劳威尔（Luitzen Egbertus Jan Brouwer）的原始直觉，或罗素与怀特海等逻辑主义者所依赖的“还原公理”之类的内容性假设。

希尔伯特与构造主义者之间的另一关键差异，在于其有穷主义立场这一构成希尔伯特数学哲学贡献核心的理念，既作为构造性方法的替代方案，也可能成为其整个研究纲领的终极归宿：

> 康德早已阐明……数学拥有独立于所有逻辑的可靠内容，因而绝不能仅通过逻辑获得基础。……作为逻辑推理运用与逻辑操作执行的前提，某些超逻辑的具体对象必须预先在我们的表象能力中被给予——这些对象作为先于一切思维的直观直接经验而呈现。若要确保逻辑推理的可靠性，就必须能完全综观这些对象的所有部分，而它们的存在、彼此差异、相继或关联的事实，将与对象本身一同被直观直接给予——既不可被还原为其他事物，也无需任何还原。（［14］，p.376）

2. HP的三种哲学阐释

HP的核心目标——为足够完备的形式数学系统提供有穷主义一致性证明——在数学层面相对明确，但有关其哲学含义却有着多种不同解释路径。我们将依据文本呈现三种不同的重要解读，并在后文中结合哥德尔不完全性定理探讨其理论后果。

（1）一致性作为数学存在与真理的条件。自早期几何学研究开始，一致性问题在希尔伯特的公理化方法构想中始终占据特殊地位。他与戴德金和康托尔共享这样的观点：数学活动应摆脱任何哲学束缚。可以说，一致性本身即是数学存在的理由（raison d’être）。在他于1899年与弗雷格（Gottlob Frege）关于几何学基础的通信中，希尔伯特已明确将一致性视为数学存在与真理的标准。（［15］，p.42）而在其著名的巴黎演讲《数学问题》中讨论算术公理一致性证明的意义时，希尔伯特更加强调：“若能证明：通过有限步骤的逻辑推演，赋予概念的属性绝不会导致矛盾，则我断言该概念的数学存在性即由此得证。就此处讨论的算术实数公理而言，公理一致性的证明同时也就是完备实数系统（或连续统）数学存在性的证明”。（［16］，p.1105）如此，在希尔伯特特有的转

化下，数学对象的存在性这一哲学问题被转变为可精确求解的数学问题——真理问题同样如此。在《数学的新基础》中对HP的首次完整阐述中，他仍坚持一致性对数学真理的根本意义：“由此可见，唯有通过解决分析算术公理一致性问题，才能使这些基础研究获得令人满意的定论。若能完成此项证明，我们便可断言数学命题实乃无可辩驳的终极真理——这种认知（鉴于其普遍哲学意义）对我们具有至高重要性”。（[11]，p.1121）仅关乎符号、公式及机械可推导性的句法的一致性概念如何能导出具有实质内涵的存在与真理概念？我们将在下节结合哥德尔定理，探讨该论题的可信度，并回答在此HP阐释框架下，希尔伯特是否可被认定为形式主义者这一相关问题。

（2）作为保守性纲领的HP。与将一致性证明的意义泛化为数学存在与真理的阐释不同，作为保守性纲领的HP更关注数学中有穷–实在–内容性与超穷–理想–形式性要素的区分，并通过表明理想部分可被消除（或换言之，可还原为实在部分）来回避关于理想领域中真理与存在问题的讨论。若数学的理想部分相对于实在部分具有保守性——即任何借助理想元素证明的实在陈述，均可通过纯粹有穷主义方法得证——则理想部分的使用仅出于实用考量：简化证明过程、统一不同证明等。但若理想部分不具备保守性，则可证明某个有穷主义手段无法证实的实在陈述P。假定所有有穷主义真理皆可证明，那么P必然是一个可被证伪的实在陈述，因而，与可被证明的非P这个否定式一起将导致整个理论不一致。因此，一致性保证了形式化数学系统中理想部分对实在部分的保守性。尽管希尔伯特从未明确将理想对实在陈述（尤其一般有穷陈述）的保守性列为其基础纲领的目标，但其对HP的若干阐述暗示了这种解读可能，例如：

> 在我的证明论中，我们将超限公理与公式附加于有穷公理之上，正如在复数理论中引入虚数元素，或在几何中引入理想对象。如此操作的动机与成效，在我的证明论中是同样的：超限公理的添加在某种意义上实现了理论的简化与完备化。（[13]，p.1144）

（3）作为可靠性纲领的HP。对HP方案最合理的解读或许是将其视为某种可靠性规划，有穷主义推理相较于超限方法的优势在于：后者丧失直觉明证性，而确定性与可靠性仅存于前者。因此，将有穷领域的确定性与可靠性扩展至超穷领域，即便只是通过数学自身担保的某种方法间接实现，便成为自然选择。这正契合希尔伯特在数学新基础研究中的目标：一劳永逸地消除对数学推理可靠性的普遍质疑。相较于保守性阐释，我们不再先验限制超穷推理可能发挥的作用——即不再要求其结论必须严格吻合有穷主义推理所能证明者，而仅要求通过超穷推理所证命题具备可靠性与内容真理性。这一目标将通过数学严格性实现。

若将有穷主义数学中的全称命题①纳入考量，则可靠性阐释等价于断言：所有可证的Π1句子皆为真，这又显然等同于它们的一致性②。为便于后续讨论，我们可用更形式化的方式表述Π_1-可靠性——即系统T的Π_1-反射原理：对于任意Π1句子φ，T可证明$Prov([\varphi])\rightarrow\varphi$，其中$Prov(x)$为形式可证性谓词，[φ]为φ在合理编码下的哥德尔数。

二、哥德尔定理及其理论关联

本节将依次探讨：在哥德尔不完全性定理的审视下，HP的三种阐释是否仍能成立。

1. 一致性与数学真理及存在性

若将“存在”与“真理”解释为模型中的存在与真值，那么“一致性是否保证存在与真理”的问题便获得更精确的表述：每个（组）一致命题都有模型。这又等价于断言形式系统

---

①这些被逻辑学家称为Π1句子的命题，是以（一个或系列）全称量词开头且主体部分无量词的陈述。某些全称公式必须归属于有穷主义一般命题，因为断言形式系统T一致性的陈述正属此类。

②若某个可证的全称命题为假，则其反例（作为原子命题）必然可被证明。通过全称示例规则，该全称命题亦能推出其反例的否定式，从而导致矛盾与不一致性。

具有完备性——即所有有效式（在所有模型中为真的命题）皆可被证明。一阶逻辑正是这样的系统。然而，哥德尔第一不完全性定理已对此普遍主张提出质疑。以二阶算术系统T为例：假设G是不可判定命题，则G与其否定¬G都与T一致，且各自存在模型——这与“二阶算术具有范畴性”（即所有模型同构）的著名事实直接矛盾。

然而，若不从模型论角度理解存在与真理，我们该如何诠释这种形式主义语境下的存在与真理表述？首先需要明确的是：希尔伯特绝非将数学视为无意义公式游戏的形式主义者，比如，拉姆齐（Frank Ramsey）曾将希尔伯特的形式主义思想描述成将数学视为“一种在纸上按固定规则玩弄无意义符号的游戏，类似于井字棋”。（[17]，p.188）相反，大量证据驳斥这种简单化的形式主义阐释：他热情称数学为“我们最珍贵的宝藏”（[11]，p.1119）将分析学比作“一曲无穷的交响乐”，（[14]，p.373）更著名的是将康托尔集合论誉为不容被驱逐的“天堂”。同时也强调：

> 除了数学价值外，它还具有重大哲学意义。因为这些公式按特定规则展开，体现了我们思维的技艺。这些规则构成一个可被发现并明确陈述的封闭系统。我的证明论的根本理念，无非是描述理解活动的规则，为我们实际思维过程建立记录。（[14]，p.475）

正如物理学家仅要求从自然法则或假设中通过类似公式游戏的纯推理导出某些可观测命题，希尔伯特的证明论同样只需保证从超穷部分推导出的有穷命题能够直接被验证。此类比的关键不在于宣称超穷命题如理论物理术语般毫无意义，而在于强调：正如研究电子或场无需直接观测，超穷命题也无需直接直觉意义。斯坦因（Howard Stein）对此有过精辟总结：

> ［希尔伯特的］核心观点，我认为毋宁是：数学逻各斯（mathematical logos）无需对任何强加的意义标准负责——既非康德或布劳威尔的“直觉”，亦非有穷或可判定性要求，更非任何人的形而上学“本体论”标准；其唯一的“形式”或“法理”责任就是保持一致性（当然，它或许还负有所谓“道德”或“审美”责任：要有用、有趣或优美；但这无法被强制——诗歌不因审查而生。（[18]，p.255）

我们既可将超穷命题视作与有穷命题同等富有意义，又可将其当作符号游戏中的无意义公式。然而，无论按常规模型论方式，还是如上述这种独立路径来理解意义，其底层假设——一致性——似乎比意义选择本身带来更严峻的难题。

2. 保守性与不可穷尽性

将HP阐释为追求超穷理论T相对于其有穷部分S的保守性之方案，因哥德尔第一不完全性定理也是不可行的。对任何形式化的S，总存在S中不可判定、却能在T辅助下判定的命题，例如T自身的一致性便属此类命题。这种对立情境被哥德尔称为“数学的不可穷尽性”的著名例证。早在1933年，哥德尔便注意到这种奇特状况：我们意图为数学构建一个单一且全面的形式系统，却最终发现存在无限可扩展的系统序列。然而，哥德尔并未将此视为类型论的缺陷或污点，反而认为其“完美契合”（[19]，p.48）第一不完全性定理的预期——即便对于整数理论，也需不断引入新证明方法或公理：

> 这一事实从另一视角看也颇具深意：它表明高阶类型的构建绝非徒劳，即便对于算术命题这类相对简单的结构，其定理证明也必须依赖类型层级的提升。哥德巴赫猜想即是这样的例子。关于形式系统中不可判定命题存在性的一般定理有一个特例：某些算术命题只能通过解析方法予以证明；更有些命题甚至需借助集合论中的无穷大基数等高阶工具。（[19]，p.48）

3. 可证一致性陈述与反射原理

针对哥德尔第二不完全性定理使有穷主义一致性证明成为不可能的观点，存在如下反驳：即使假设形式系统足够完备以包含所有有穷主义证明手段，某些非标准算术系统确实能通过其独特构造方式证明自身的一致性陈述Con，这一现象最早由费弗曼（Solomon Feferman）指出并进行了系统研究。[20]更甚者，对任何

标准算术系统，若其本身一致，我们总能找到一个能证明同样定理且能证明自身一致性的对应系统。然而这些系统因各种缺陷均非真正的形式数学系统，故无法为HP提供规避第二不完全性定理的实质出路。吉亚昆托（Marcus Giaquinto）（［21］，p.188）讨论了若干案例，并详细分析了费弗曼与罗瑟（John Rosser）系统的缺陷——这些系统因证明关系不可判定与切规则（Cut Rule）失效而未能成为真正的形式数学系统。另一方面，类似于这些一致性导向的形式系统，我们也能基于“一致性导向”的证明谓词，在标准算术系统T内构造可证的非标准一致性陈述。设Prf(x, y)为包含算术的某形式系统的标准证明谓词——即这个算术化谓词表示元数学概念“x是以哥德尔数y为末项公式的证明序列的哥德尔数”。经典一致性陈述Con通常定义为∀x(¬Prf(x, ［0=s0］))，其中s表示后继函数,［0=s0］是解释为“0等于1”（即矛盾式）的公式哥德尔数。若T包含某类归纳原理,则可证此标准Con在T内不可证。（［22］，Chap.31）令Prf(x, y)仍表达T的推导关系，莫斯托夫斯基（Andrzej Mostowski）通过此标准谓词巧妙地定义出了一个新的一致性导向证明谓词（［23］，p.24）如下：

$$\mathrm{MPrf}(x, y) \overset{\mathrm{def}}{=} \mathrm{Prf}(x, y) \land \neg \mathrm{Prf}(x, [0=s0])$$

这个新证明谓词MPrf(x, y)的直观含义是：当且仅当x是以y为结论公式的证明序列的哥德尔数且x不是矛盾式0=1的证明时，x才是y的M-证明。假设系统T一致，则任何公式证明序列都不可能以矛盾式结尾——即¬Prf(x, ［0=s0］)对所有x为真，因而可被推导（作为无量词命题）。从外延角度看，MPrf(x, y)与标准证明谓词Prf(x, y)对相同数值有序对的真值判定完全一致。但二者的“内涵”差异在于：MPrf(x, y)的断言远比Prf(x, y)更强。现若基于此一致性导向的新证明关系构造新一致性陈述Con*$\overset{\mathrm{def}}{=}$∀x(¬MPrf(x, ［0=s0］))，则该陈述可轻松从T导出——因其本质上是谓词逻辑的矛盾律定理：

$$\mathrm{Con}^* \leftrightarrow \forall x \neg [(\mathrm{Prf}x, 0=s0) \land \neg \mathrm{Prf}(x, [0=s0])].$$

上述考量使得某些评论者，如德特莱夫森（Michael Detlefsen）[3]对哥德尔第二不完全性定理在HP上的效力产生质疑。然而，更深入的考察将表明：无论某些一致性陈述及其类似陈述的可证性在揭示第一和第二不完全性定理差异方面多么重要，它们都无法完成既定的哲学目标，原因有三：

首先，涉及一致性导向可证性谓词（如上述或其他变体）的可证一致性陈述，其哲学价值非常有限——因为这些陈述是否真正断言一致性仍取决于理论T自身是否一致。若T确实一致，则通过若干技巧构造的证明谓词MPrf(x, y)确实表达了真实的证明关系，并与Prf(x, y)外延重合，此时相关的Con*的确可解读为表达T的一致性。但若T不一致，MPrf便无法表达正统的证明关系，Con*也不能被理解为常规的一致性陈述。以上述情形为例：假设T不一致且m是0=s0证明的哥德尔数，则既有Prf(m, ［0=s0］)又有¬MPrf(m, ［0=s0］)——但在不一致系统中由于爆炸原则①的存在本应一切皆可证。¬MPrf(m, ［0=s0］)仅表明：MPrf(x, y)及其Con*并非我们真正寻求的目标。

其次，与第一点密切相关，这类一致性导向的可证性谓词通常无法满足全部三条希尔伯特–贝奈斯–勒布（HBL）可导性条件——这些条件最初由希尔伯特与伯奈斯提出，[24]他们在该著作中首次完整证明了哥德尔第二不完备性定理；后经勒布（Martin Löb）的改进，[25]条件表述得到了大幅简化。这些条件不仅体现了我们对证明谓词的直觉要求，更是确保与标准证明谓词相关联的系统内一致性陈述不可证的关键。对于任意形式数学系统T，其形式定理谓词“Prov(x)”（定义为Prov(x)$\overset{\mathrm{def}}{=}$∃yPrf(y, x)）需满足的三条条件如下：

1. 若A是T的定理，且［A］表示A的哥德

①爆炸原则（Ex falso quodlibet）说是说从矛盾可推出一切，这是一条经典逻辑中的有效推理规则，虽然有不少非经典逻辑拒斥此原则。

尔数，则Prov([A])也是T的定理。

2. 对T中任意公式A，T可证明“若Prov([A])，则Prov([Prov([A])])”——即条件（1）本身可在T内形式化，表明T“知道”该条件。

3. 对T中任意公式A与B，可证：“若Prov([A])且Prov([A→B])，则Prov([B])”。这实质断言分离规则（modus ponens）规则在T内可形式化。

任何非标准证明谓词都将在某方面无法满足这些条件，从而使得其关联的一致性陈述能在同一系统内可证——这看似违背了哥德尔第二不完全性定理。以著名的罗瑟谓词（Rosser predicate）为例，其定义如下：

$$RPrf(x, y) \stackrel{def}{=} Prfx, y \wedge (\forall z \leq x) \neg Prf(z, Neg(y))$$

其中，Neg(y)表示哥德尔数为y的公式之否定的哥德尔数。RPrf(x, y)直觉含义是：当且仅当x是以y为结论公式的证明序列的哥德尔数且任意小于x的数都不是y的否定的证明，x才是y的R-证明。类似的，当系统T一致时，罗瑟可证性谓词RPrf(x, y)与标准可证性谓词Prf(x, y)具有相同的外延；但当T不一致时，某些公式在罗瑟可证性意义上可能“不可证”，明显和我们关于可证性谓词的直觉相违背。显然，罗瑟可证性也无法满足上述三个条件。对罗瑟证明谓词的主要兴趣在于，它似乎更贴近数学家处理数学证明的实际方式——每当获得某个定理的证明时，他们希望确保在某种预定意义上“之前”从未出现过相反结论的证明。然而，尽管罗瑟方法作为实际数学实践的模型看似合理（即便暂不考虑对证明进行可行排序的实际操作难题），它仍无法为HP提供解决方案，因为它试图挽救的恰恰是HP本身的核心诉求。HP中证明的本质在于形式化的“理想”推理，因此证明必须严格按形式意义理解：即从公理出发，通过逐步应用推理规则最终得到待证定理的公式序列，不容许任何外部元素的介入。就罗瑟证明谓词而言，若T不一致，HP中真实内容性部分中的某些错误定理可能仍具有罗瑟可证性——反例可能出现在证明序列的更晚位置，这将严重违背HP作为可靠性纲领的根本要求。

最后且可能最重要的是，我们可以绕过寻找“正确”的可证性谓词（进而“正确”的一致性陈述）的问题，直接论证HP的不完全性。哥德尔在1972年笔记《关于同一系统中一致性不可证性的最优最普遍版本》（[26]，p.305）中区分了两种一致性概念：内一致性：命题与其否定命题均不可证（即标准的一致性陈述Con）；外一致性：“可证方程在原始递归项上的等式演算规则仅产生正确的数值方程”（对应Π1反射原理）。对于常规系统（其一致性陈述源自满足HBL条件的证明谓词），这两种一致性是等价的。（[22]，Chap.36）但即便证明谓词不满足HBL条件，只要系统T实际一致，Π1反射原理仍无法在T内导出——这与那些非标准的、关注一致性的陈述形成鲜明对比。其本质原因可视为哥德尔第一不完全性定理的直接推论：取标准不可判定句G，由于G满足G↔¬Prov([G])，根据一致性条件，G在T中不可证。由于G本身是Π1语句，若Π1反射原理在T中成立，则Prov([G])→G也应在T中可证。这与G↔¬Prov([G]))结合将导致G在T中可证，产生矛盾。因此，仅凭一致性加上若干宽松条件（如对角化引理的可证性①），就足以使Π1反射原理成为任何系统T都无法实现的目标——而这恰恰是HP中为系统S的超限公理提供正当性证明所需的核心要素。

结　论

简而言之，无论我们对希尔伯特原始纲领作出何种合理的哲学阐释，哥德尔不完全性定理都注定其无法实现。然而，这并不是否定该纲领为数学基础研究带来的哲学价值，也不妨

①这正是对角化引理的核心内容：对于语言中可表达的任意开放公式Φ(x，总存在一个语句P，使得P↔Φ([P])可在形式系统中得证。

碍后人对其可能进行的扩展或修正。更全面而综合的评价可见于哥德尔自身的论述：

> 已经证明的仅仅是：希尔伯特所设想的特定认识论目标无法实现。该目标旨在基于如同初等算术般具体且直接明证的依据，来证明经典数学公理的一致性。然而，若从纯数学视角审视这一局面，基于适当选择的更强元数学前提的一致性证明同样具有重要价值，这些证明能引向对数学证明理论结构的深刻洞见。（[9]，p.277）

事实上，广义化的HP在后续证明论发展中确实催生了重要的数学成果，并极大地深化了我们对数学证明结构的理解。尽管这些描述性成果的哲学意义或许较为有限，且未能如希尔伯特预期的那般直接明晰，但它们仍可能为未来的哲学思考提供素材。正如费弗曼恰当地指出：

> 总体而言，本研究所呈现的这类成果［还原性证明论成果］，其价值在于能更精准地评判各种数学哲学立场——如有穷主义、可定义主义、构造主义及集合论实在论——的支持或反对依据。无论人们是否出于本体论和/或认识论原因认真对待其中某种哲学立场，重要的是要明确：数学的哪些部分最终能在相应哲学基础上获得辩护，而哪些部分不能。（［27］，p.207）

换言之，一方面，人们无需成为彻底的柏拉图主义者亦可接受经典数学的主体内容——只要这些内容能被归约为更具构造性的部分；另一方面，非柏拉图主义者也能更清晰地认识到其原则所无法获取的内容，并准备好作出相应的理论牺牲。只要认识论立场能够被精确表述，便可通过数学严谨性的论证予以否定性反驳或肯定性辩护，这一理念不仅与哥德尔对哲学立场精确性和严密性的追求完全吻合，也是数学哲学的特定视角能带给一般哲学研究的重要启示。

［参 考 文 献］

[1] Gödel, K. 'On Formally Undecidable Propositions of Principia Mathematica and Related Systems I', 1931'[A], Feferman, S., Dawson, J., Moore, G. et al. (Eds.) *Kurt Gödel, Collected Works I.* [C], Oxford: Oxford University Press, 1986, 144–195.

[2] Gödel, K. 'On a Hitherto Unutilized Extension of the Finitary Standpoint'[A], Feferman, S., Dawson, J., Goldfarb, W., et al. (Eds.) *Kurt Gödel, Collected Works, II.* [C], Oxford: Oxford University Press, 1958, 241–253.

[3] Detlefsen, M. *Hilbert's Program: An Essay on Mathematical Instrumentalism* [M]. Berlin: Reidel, 1986.

[4] Zach, R. 'Hilbert's Program'[EB/OL]. The Stanford Encyclopedia of Philosophy, 2003. https://plato.stanford.edu/entries/hilbert-program/. 2025–08–30.

[5] Zach, R. 'Hilbert's Program Then and Now'[A], Gabbay, D., Woods, J. (Eds.) *Philosophy of Logic, 5.* [C], Amsterdam: Elsevier, 2006, 411–447.

[6] Smoryński, C. 'Hilbert's Programme'[J]. *CWI Quarterly*, 1988, 1(4): 3–59.

[7] 叶峰. 二十世纪数学哲学：一个自然主义者的新描述［M］. 北京：北京大学出版社，2010.

[8] Hilbert, D. 'Grundlagen der Geometrie'[EB/OL]. Leipzig: B. G. Teubner, 1899.

[9] Reid, C. *Hilbert* [M]. Berlin: Springer-Verlag, 1970.

[10] Sieg, W. 'Hilbert's Programs: 1917–1922'[J]. *The Bulletin of Symbolic Logic* 1999, 5(1): 1–44.

[11] Hilbert, D. 'The New Grounding of Mathematics'[A], Ewald, W. (Ed.) *From Kant to Hilbert: A Source Book in the Foundations 2.* [C], Oxford: Oxford University Press, 1922, 1115–1134.

[12] Weyl, H. 'On the New Foundatioal Crisis of Mathematics'[A], Mancosu, P. (Ed.) *From Brouwer To Hilbert: The Debate on the Foundations of Mathematics in the 1920s* [C], Oxford: Oxford University Press, 1921, 86–118.

[13] Hilbert, D. 'The Logical Foundations of Mathematics'[A], Ewald, W. (Ed.) *From Kant to Hilbert: A Source Book in the Foundations 2.* [C], Oxford: Oxford University Press. 1923, 1134–1148.

[14] Hilbert, D. 'On the Infinite'[A], van Heijenoort, J. (Ed.) *From Frege to Gödel: A Source Book in Mathematical Logic* [C], Cambridge: Harvard University Press, 1925, 367–392.

[15] Hilbert, D. 'Letter to Frege, 29.12.1899'[A], Gabriel, G., McGuinness, B., Kaal, H., et al. (Eds.) *Philosophical and Mathematical Correspondence* [C], Oxford: Basil Blackwell, 1980, 38–43.

[16] Hilbert, D. 'Mathematical Problems'[A], Ewald, W. (Ed.) *From Kant to Hilbert: A Source Book in the Foundations 2.*[C], Oxford: Oxford University Press, 1900, 1096–1104.

[17] Ramsey, F. P. 'Mathematical Logic'[J]. *The Mathematical Gazette*, 1926, 13(184): 185–194.

[18] Stein, H. 'Logos, Logic, and Logistiké: Some Philosophical Remarks on Nineteenth-Century Transformation of Mathematics'[A], Aspray, W., Kitcher, P. (Eds.) *History and Philosophy of Modern Mathematics*[C], Minneapolis: University of Minnesota Press, 1988, 238–259.

[19] Gödel, K. 'The Present Situation in the Foundations of Mathematics'[A], Feferman S., Dawson, J., Goldfarb, W., et al. (Eds.) *Kurt Gödel, Collected Works, III.*[C], Oxford: Oxford University Press, 1933, 45–53.

[20] Feferman, S. 'Arithmetization of Metamathematics in a General Setting'[J]. *Fundamenta Mathematicae*, 1960, 49(1): 35–92.

[21] Giaquinto, M. *The Search for Certainty: A Philosophical Account of Foundations of Mathematics*[M]. Oxford: Clarendon Press, 2002.

[22] Smith, P. *An Introduction to Gödel's Theorems*[M]. 2nd ed. Cambridge: Cambridge University Press, 2013.

[23] Mostowski, A. *Thirty Years of Foundational Studies*[M]. Oxford: Basil Blackwell, 1966.

[24] Hilbert, D., Bernays, P. *Grundlagen der Mathematik Vol. 2*[M]. Berlin: Springer-Verlag, 1939.

[25] Löb, M. H. 'Solution of a Problem of Leon Henkin'[J]. *The Journal of Symbolic Logic*, 1955, 20(2): 115–118.

[26] Gödel, K. 'Some Remarks on the Undecidability Results'[A], Feferman S., Dawson, J., Goldfarb, W., et al. (Eds.) *Kurt Gödel, Collected Works, II.*[C], Oxford: Oxford University Press, 1972, 305–307.

[27] Feferman, S. 'What Rests on What? The Proof-Theoretic Analysis of Mathematics'[A], Feferman, S. (Ed.) *In the Light of Logic*[C], Oxford: Oxford University Press, 1993, 187–208.

［责任编辑 王巍 谭笑］