arXiv:2611.32977v1 [cs.LO] 24 Nov 2026

扩展了素数的 PRESBURGER 算术

JORIS NIEUWVELD

摘要。我们无条件地证明，加入素数谓词后的 Presburger 算术一阶理论是不可判定的。Bateman、Jockusch 和 Woods 于 1993 年在假设 Dickson猜想成立的条件下证明了这一点。我们采用 Karimov、Nieuwveld 和 Ouaknine 的计数器机器框架及其使用“任意排列”的思想；数论方面的依据则来自 Hildebrand 和 Maier 于 1988 年所作的证明： k连续归一化素数间隔的极限点具有正测度。我们使用的是他们在取极限之前得到的离散版本。关键引理是，尽管 Hildebrand 和 Maier 的结果似乎并未给出连续素数间隔的任意排列，但它确实给出了两个有限块连续素数 $p_0 < \cdots < p_k$ 和 $q_0 < \cdots < q_k$ 的差值 $p_i - q_i$ 的任意排列。此外，利用同一个关键引理，我们证明：扩展了将 n 映射到第 n个素数的函数的 Presburger 算术能够定义乘法。

1. 引言

Presburger 算术，即 $\langle \mathbb{N};\ + \rangle$ 的一阶理论，记作 FO($\mathbb{N}$; +)，众所周知是可判定的。自然地，人们开始探讨：可以向 Presburger 算术中加入哪些看似较弱的乘法版本，以获得不可判定性，以及哪些版本能够定义乘法。例如，向 Presburger 算术扩展整除性、互素性或平方数集合的谓词，都可以使乘法变得可定义；参见 Bès 的综述 [3] ，其中介绍了这些及其他例子。一个显而易见的候选问题仍未解决：仅仅能够判断一个数是否为素数，是否足以确立不可判定性？能否定义乘法?

Bateman、Jockusch 和 Woods [2] 利用 Dickson猜想研究了 Presburger 算术的两个扩展，这两个扩展通过素数谓词 P 得到；该猜想有时也被称为 Schinzel的H猜想的线性情形。

猜想 1.1 (Dickson猜想)。设 $a_1, \ldots, a_k \geq 1$ 和 $b_1, \ldots, b_k \in \mathbb{Z}$，使得对于每个素数 p，都存在一个数 x，令 $a_1x + b_1, \ldots, a_kx + b_k$ 中没有任何一个能被 p 整除。那么存在无穷多个 n，使得 $a_1n + b_1, \ldots, a_kn + b_k$ 同时为素数。

这个猜想很强：它蕴含孪生素数猜想。Bateman、Jockusch 和 Woods 证明，假设 Dickson猜想成立，FO($\mathbb{N}$; <,P) 是可判定的（更强的是，这个扩张的单调二阶理论也是可判定的）。若能无条件证明

这一可判定性，就能有效地解决孪生素数猜想。此外，他们证明了 Dickson猜想蕴含在 $\langle \mathbb{N}; +,$ P$\rangle$ 中可以定义乘法，这意味着该理论不可判定。他们证明，对于任意 N，都存在数 a 和 c，使得 $ax + c + 0, ax + c + 2, ax + c + 6, \ldots, ax + c + N^2 + N$ 满足猜想的前提，而其中没有任何数是素数。因此，我们可以得到一个二次多项式的像，从而定义乘法。

假设Dickson猜想成立，Kaplan和Shelah [8] 证明了FO($\mathbb{Z}$; 0、1、 +、P $\cup \{-p : p \in$P$\}$) 是可判定的，且不能定义序。由于在Presburger 算术中，序通过$x \leq y \iff \exists z \in \mathbb{N}$隐含给出： $x + z = y$，因此序似乎是不可判定性所必需的。例如，哥德巴赫

日期：2026年11月。

猜想很容易在FO(ℕ; +、P)中表达；而在Kaplan和Shelah的结构中，自然能够表达的只有一个较弱的版本：每个偶数都是两个素数之和或之差。

关于 FO(ℕ； +，P) 不可判定性的无条件证明仍悬而未决；Bès [3], Cégielski、Richard 和 Vsemirnov [4], Kaplan和Shelah [8], 以及 Karimov、Nieuwveld 和 Ouaknine [10]都讨论过这一问题。我们将在本文中证明其不可判定性。

定理 1.2。 理论 FO(ℕ； +，P) 是不可判定的。

尽管证明了不可判定性，我们并未得到乘法的定义。我们还研究了理论 FO(ℕ； +，pr)，其中 pr(n) 是第 n 个素数；其可判定性曾在 [3, 4] 中受到质疑。因此，pr(1) $= 2$。为使 pr 成为全函数，我们令 pr(0) $= 0$，但这一选择并不重要。

定理1.3。 乘法可以在结构 $\langle \mathbb{N}; +, \mathrm{pr} \rangle$ 中定义。

证明直觉。 我们利用 Karimov、Nieuwveld 和 Ouaknine 最近提出的框架来证明定理 1.2 [10]：只需证明某个结构 模拟计数器机器即可。计数器机器是一种（Minsky）机器，具有有限个状态和计数器；计数器可以执行递增、递减操作并进行零值测试，而其停机问题通常是不可判定的。如果每个计数器机器都能被转换为一个一阶句子，并且该句子成立当且仅当机器停机，则称某个结构模拟计数器机器。停机断言存在一条经过编码的轨迹，即机器的连续构型；其中，初始构型、停机构型以及转移关系都必须能够用一阶语言表达。Karimov 等人非正式地将模拟计数器机器的结构称为丰富 结构，因为它们需要实现某种类型的每一种适当模式。尽管人们预计素数在自然数集合中表现得相当随机，但控制精确的素数构型仍然十分困难。

第一种尝试是考察 Green 和 Tao [5]关于素数中等差数列的结果。Kaplan和Shelah [9, Propostion 3.6] 指出，该结果意味着，对于任意集合$S \subset \{1, \ldots\ldots, k\}$，存在无穷多个数 a, b 使得对于 $0 \leq i \leq k$， $ai + b$ 是素数当且仅当 $i \in S$。也就是说，素数包含每一种有限模式。然而，我们需要乘以参数 a，这在 Presburger 算术中是不可能的。

第二种尝试是考察连续素数间隔，并假定它们具有任意排列。也就是说，对于任意 k 以及 $\{1, \ldots\ldots, k\}$ 的排列 σ，存在一个连续素数的有限序列 $p_0 < \cdots < p_k$，使得 $p_{\sigma(1)} - p_{\sigma(1)-1} < \ldots < p_{\sigma(k)} - p_{\sigma(k)-1}$。在这种情况下，我们只需要不等式和减法，而它们在 $\langle \mathbb{N}; +, \mathrm{P} \rangle$ 中是可定义的，并且 [10] 的论证几乎可以原样适用。

遗憾的是，目前尚不清楚连续素数间隙是否具有任意排列。关于这一问题，有两类结果。首先，单调模式会无穷多次出现 [1],，但据我们所知，尚无将其推广到任何其他固定模式的结果。其次，Hildebrand 和 Maier 得到了一个结果。

定理 1.4 (Hildebrand–Maier [6])。设 $k \geq 1$。那么存在一个常数 $c_k > 0$，使得对于所有充分大的 T，下列集合的极限点的

$$\left\{ \left(\frac{\mathrm{pr}(n+1) - \mathrm{pr}(n)}{\log(n)}, \ldots, \frac{\mathrm{pr}(n+k) - \mathrm{pr}(n+k-1)}{\log(n)} \right) : n \in \mathbb{N}_{\geq 1} \right\} \cap [0, T]^k$$

勒贝格测度至少为 $c_k T^k$。

如果 c_k 大于 $1 - 1/k!$，那么连续素数间隔就会出现任意排列。尽管 c_k 并未明确计算出来，但它似乎小得多，因此我们无法得出存在任意排列的结论。两个极限点之差可以产生任意排列，但这两个极限点可能出现在互不相关的尺度 $\log(n)$上。因此，我们采用离散计数（定理 4.1），这样便可直接比较任意两个序列，即由连续素数组成的有限块。

引理 1.5（双序列引理）。对于每个 $k \geq 1$ 以及 $\{1、\ldots、k\}$ 的每个排列 σ，都存在两个序列

$$p_0 < p_1 < \cdots < p_k \qquad and \qquad q_0 < q_1 < \cdots < q_k$$

具有以下性质。令 $e_i = p_i - q_i$ 和 $\Delta = \max_{1 \leq i \leq k} |e_i - e_0|$。 n

（a）$e_{\sigma(1)} < e_{\sigma(2)} < \ldots < e_{\sigma(k)}$，并且（b）第二个序列中的每个间隔都超过 2Δ：$q_j - q_{j-1} > 2\Delta$ 对于 $1 \leq j \leq k$。

由此引理可知，(a) 项给出了任意排列，而 (b) 项使我们能够确定给定的 p_i 和 q_j 是否具有相同的索引，从而确定它们的差是否用于编码。我们在第 1.2 定理的第 2 节证明中对此作出精确说明。

该结果的函数版本，即定理 1.3，同样可得。考虑两个序列 $(\mathrm{pr}(n+i))_{i=1}^{k}$ 和 $(\mathrm{pr}(r+i))_{i=1}^{k}$，并令 $\psi(i) := \mathrm{pr}(n+1) - \mathrm{pr}(r+1) \geq \mathrm{pr}(n+i) - \mathrm{pr}(r+i)$。根据引理 1.5，存在 r 和 n，使得对于 $1 \leq i \leq k$，当且仅当 i 为平方数时，$\psi(i)$成立。反之，如果 $\psi(4)$ 成立，并且对于所有满足 ψ且为相邻数的 $0 < i_1 < i_2 < i_3$，都有 $i_3 = 2i_2 - i_1 + 2$，则任何满足 ψ 的数都是平方数。因此，乘法是可定义的。

2. 从两个序列到不可判定性

我们在假设引理1.2 的前提下，使用 [10] 的计数器机器框架证明定理1.5。

2.1. 计数器机器模拟。我们回顾 [10, §1.2], 中的概念，并将其专门化为长度为一的索引元组。若结构$\mathcal{M}$ 的定义域为 $\mathbb{N}$，则称其模拟计数器机器，如果存在 $\ell \geq 1$、映射 $\mathrm{Seq}: \mathbb{N}^{\ell} \to \mathbb{N}^{*}$和 $\mathrm{Rep}: \mathbb{N}^{\ell} \to \mathbb{N}^{*}$（分别为编码序列及表示其各项的元素），以及公式 $\mathrm{rep}(\mathbf{x}\ y)$、$\mathrm{cnst}_K(\mathbf{x}\ y)$（可由 K计算得到）、$\mathrm{succ}(\mathbf{x}\ y,\ z)$、$\mathrm{inc}(\mathbf{x}\ y,\ z)$和 $\mathrm{eq}(\mathbf{x}\ y,\ z)$，使得对所有$\mathbf{x} \in \mathbb{N}^{\ell}$:

(1) Seq 是从自然数有限序列到全体有限序列的满射；(2) $|\mathrm{Seq}(\mathbf{x})| = |\mathrm{Rep}(\mathbf{x})|$，且 Rep(x) 的各项互不相同；(3) 表示：当且仅当 y 出现在 Rep(x) 中时，$\mathrm{rep}(\mathbf{x}\ y)$ 成立；(4) 常值 K：当且仅当存在某个 i，使得 $\mathrm{Seq}(\mathbf{x})_i = K$ 且 $y = \mathrm{Rep}(\mathbf{x})_i$ 时，$\mathrm{cnst}_K(\mathbf{x}\ y)$ 成立；(5) 后继：当且仅当存在某个 i，使得 $y = \mathrm{Rep}(\mathbf{x})_i$ 且 $z = \mathrm{Rep}(\mathbf{x})_{i+1}$ 时，$\mathrm{succ}(\mathbf{x}\ y, z)$ 成立；(6) 递增：当且仅当 $y = \mathrm{Rep}(\mathbf{x})_i$、$z = \mathrm{Rep}(\mathbf{x})_j$ 且 $\mathrm{Seq}(\mathbf{x})_j = \mathrm{Seq}(\mathbf{x})_i + 1$ 时，$\mathrm{inc}(\mathbf{x}\ y, z)$ 成立；(7) 相等性：当且仅当 $y = \mathrm{Rep}(\mathbf{x})_i$、$z = \mathrm{Rep}(\mathbf{x})_j$ 且 $\mathrm{Seq}(\mathbf{x})_i = \mathrm{Seq}(\mathbf{x})_j$ 时，$\mathrm{eq}(\mathbf{x}\ y, z)$ 成立。

定理 2.1 ([10, 定理 2])。每个能够模拟计数器机器的结构，其一阶理论都是不可判定的。

该定理的证明将给定的计数器机器转换为一个一阶句子 φ，用以模拟其轨迹。要检验这是否确实是一条有效轨迹，只需比较轨迹中相邻的项。如果计数器机器停机，这条轨迹就是有限的，而某个$\mathbf{x} \in \mathbb{N}^{\ell}$ 可以见证这一点。也就是说，$\exists \mathbf{x} \varphi(\mathbf{x})$成立；对于这个$\mathbf{x}$，Seq(x)恰好是该计数器机器的轨迹，而 Rep(x)则编码了一系列相应的项，我们在一阶公式中对其进行迭代。

在 Presburger 算术中，我们将 $x \leq y$ 写作 $\exists z(x + z = y)$，将 $x < y$ 写作 $x \leq y \wedge x \neq y$，并将满足 $x = z + y$（当 $y \leq x$ 时）的唯一 z 写作 $x - y$。因此，诸如“p 是 $[a, c]$ 中的一个素数”和“p，p' 是连续素数”这样的陈述都是一阶的。

2.2. 将计数器机器应用于素数间隔。我们使用 $\ell = 6$，并记作 $\mathbf{x} = (a, b, c, a', c', \delta)$。对于素数 $p \in [a, c]$ 和 $q \in [a', c']$，若 $|(p - q) - (a - a')| \leq \delta$，则记作 $p \sim_{\mathbf{x}} q$，也就是说，

$$p + a' \leq q + a + \delta \quad \text{and} \quad q + a \leq p + a' + \delta.$$

令 Valid(x) 表示这样一个公式：$a \leq b \leq c$，并且对于每个素数 $p \in [a, c]$，恰好存在一个素数 $q \in [a', c']$ 使得 $p \sim_{\mathbf{x}} q$。对于有效的 $\mathbf{x}$ 和素数 $p \in [a, c]$，我们将此称为 $q$the

匹配 p，并记作 Match(p)。Valid 以及 Match 的图都是在$\langle\mathbb{N};+,\mathsf{P}\rangle$ 中一阶可定义的。令 $v(p)=p-\mathrm{Match}(p)\in\mathbb{Z}$。我们仅在比较中使用 v：$v(p)\le v(p')$ 表示公式 $p+\mathrm{Match}(p')\le p'+\mathrm{Match}(p)$，对于 $<$ 也类似。

命题 2.2。如果引理 *1.5* 成立，则 $\langle\mathbb{N};+,\mathsf{P}\rangle$ 可以模拟计数器机器。

证明。如果 Valid(x) 不成立，则令 Seq(x) 和 Rep(x) 为空序列。否则，将 $(a,\ b]$ 标记 中的素数称为，令 Rep(x) 按递增顺序列出 $(b,\ c]$ 中的素数，并设

$$\mathrm{Seq}(\mathbf{x})_i=\#\{p'\text{ a mark}:v(p')\le v(\mathrm{Rep}(\mathbf{x})_i)\}.$$

对于 Rep(x) 的元素，记作 $p,\ p'$，则公式为：

- rep(x p)：Valid(x)、$b<p\le c$ 和 P(p)；
- cnst$_K$(x p)：rep(x p)，且恰好有 K 个标记 $p\prime$ satisfy $v(p\prime)\le v(p)$；
- succ(x $p,p\prime$)：rep(x p)、rep(x $p\prime$)、$p<p\prime$，且 p 与 $p\prime$之间没有素数；
- inc(x $p,p\prime$)：rep(x p)、rep(x $p\prime$)，且恰有一个标记 $p\prime\prime$ 满足 $v(p)<v(p\prime\prime)\le v(p\prime)$；
- eq(x $p,p\prime$)：rep(x p)、rep(x $p\prime$)，且没有标记 $p\prime\prime$ 满足 $v(p)<v(p\prime\prime)\le v(p\prime)$ 或 $v(p\prime)<v(p\prime\prime)\le v(p)$。

“恰好 K”是一个可由 K计算的一阶陈述。

条目 (2)–(7)。若 Valid(x) 不成立，则每个公式均为假，两个序列也都为空。否则，(2)–(5) 是显然的。对于 (6) 和 (7)，设 $p=\mathrm{Rep}(\mathbf{x})_i$ 和 $p'=\mathrm{Rep}(\mathbf{x})_j$。若 $v(p)\le v(p')$，则 $\mathrm{Seq}(\mathbf{x})_j-\mathrm{Seq}(\mathbf{x})_i$ 是满足 $v(p)<v(p'')\le v(p')$ 的标记 p'' 的数量。反之，若 $v(p)>v(p')$，则 $\mathrm{Seq}(\mathbf{x})_j\le\mathrm{Seq}(\mathbf{x})_i$，且没有标记 p'' 满足 $v(p)<v(p'')\le v(p')$。因此，当且仅当 $\mathrm{Seq}(\mathbf{x})_j=\mathrm{Seq}(\mathbf{x})_i+1$ 时，inc(x p,p') 成立。同理，通过对 p 和 p' 的两种序应用相同论证可得，当且仅当 $\mathrm{Seq}(\mathbf{x})_i=\mathrm{Seq}(\mathbf{x})_j$ 时，eq(x p,p') 成立。

条目 *(1)*。任意满足 Valid(x) 为假的x 都会产生空序列。设 $(t_1,\dots,t_\tau)$ 非空，取 $r\ge\max(1,t_1,\dots,t_\tau)$，并令 $k=r+\tau$。按照键的顺序排列位置 $1,\dots,k$：

$$\kappa(j)=j\ \ (1\le j\le r),\qquad \kappa(r+i)=t_i+\tfrac{i}{\tau+1}\ \ (1\le i\le\tau).$$

这些键互不相同，且恰有 t_i 个前 r 位置的键小于 $\kappa(r+i)$。令 σ按键的递增顺序列出这些位置，并令 $p_0<\cdots<p_k$和 $q_0<\cdots<q_k$为引理1.5 应用于 σ所得的序列。令

$$\mathbf{x}=(p_0,\,p_r,\,p_k,\,q_0,\,q_k,\,\Delta).$$

对于 $0\le i\le k$，有 $p_i\sim_{\mathbf{x}}q_i$，因为 $|(p_i-p_0)-(q_i-q_0)|=|e_i-e_0|\le\Delta$。对于 $j\ne i$，由 (b) 可得 $|q_j-q_i|>2\Delta$，因此 $|(p_i-p_0)-(q_j-q_0)|>2\Delta-\Delta=\Delta$。所以 x 有效且 $\mathrm{Match}(p_i)=q_i$，故 $v(p_i)=e_i$。这些标记为 $p_1,\dots,p_r$，并且 $\mathrm{Rep}(\mathbf{x})=(p_{r+1},\dots,p_k)$。根据 (a)，$e_j\le e_{r+i}$ 成立当且仅当 $\kappa(j)\le\kappa(r+i)$，因此 $\mathrm{Seq}(\mathbf{x})_i=t_i$。□

证明 定理 *1.2*（引理 *1.5*）。将命题2.2 与定理2.1结合 . □

注 2.3。检查在 [10, 定理 2] 的证明中构造的句子可知，它具有 $\exists\mathbf{x}\,[\mathrm{Valid}(\mathbf{x})\wedge\psi(\mathbf{x})]$ 的形式，其中 ψ 是一个 Π_2-公式。因此，FO($\mathbb{N}$; +, P) 的 $\exists^6\forall^*\exists^*$-片段已经是不可判定的，其中 6 来自 $\mathbf{x}\in\mathbb{N}^6$。事实上，可以消去 c'：取$c'=c+a'+\delta-a$，并将 $k+1$ 代入引理 1.5 中的 k，把额外的指标放在 σ的最后，使得 $q_{k+1}-q_k>2\Delta$；则区间 $[a',c']$ 恰好包含证明命题 2.2 时的 q_0、.....、q_k，因为 $c+a'-a=q_k+(e_k-e_0)$，且 q-序列的间隔超过 2Δ。这给出了 $\exists^5\forall^*\exists^*$-片段。我们构造中的内部量词数量没有上界，因为 cnst$_K$ 使用了两种类型的 $K+1$ 量词。

3. 第 n 个素数函数与乘法

接下来我们在假设引理 1.3 的条件下证明定理 1.5。我们首先回顾普特南的一个经典结果 [11]。设 Sq 是表示平方数的谓词。

引理3.1。 结构 $\langle \mathbb{N}; +, \text{平方数} \rangle$ 定义了乘法。

证明。 首先注意到， $x = y^2$ 当且仅当 x 是平方数、 $x + 2y + 1$ 也是平方数，且两者之间不存在平方数。于是，恒等式 $(x+y)^2 = x^2 + y^2 + 2xy$ 表明， $z = xy$ 当且仅当 $(x+y)^2 = x^2 + y^2 + 2z$。□

定理 *1.3*（给定引理 *1.5*）。根据引理 3.1 ，只需定义 平方数。正平方数构成序列 $u_0 = 1$、 $u_1 = 4$、$u_{j+2} = 2u_{j+1} - u_j + 2$，我们将其编码如下。

对于 k， n， $r \in \mathbb{N}$，记 $\mathbf{x} = (k,\ n,\ r)$，并令 $\psi_{\mathbf{x}}(i)$ 表示公式

$$1 \le i \le k \ \wedge\ \mathrm{pr}(n+i) + \mathrm{pr}(r+1) \le \mathrm{pr}(n+1) + \mathrm{pr}(r+i).$$

令 $\varphi(\mathbf{x})$ 为公式

$$\psi_{\mathbf{x}}(4) \wedge \forall i_1, i_2, i_3 : i_1 < i_2 < i_3 \to \Big[\Big(\bigwedge_{j=1}^{3} \psi_{\mathbf{x}}(i_j) \wedge \forall i : i_1 < i < i_3 \to \big(\neg\psi_{\mathbf{x}}(i) \vee i = i_2\big)\Big) \to i_3 + i_1 = 2i_2 + 2\Big].$$

注意，只要满足 $k \ge 1$， $\psi_{\mathbf{x}}(1)$ 就成立；而 $\varphi(\mathbf{x})$ 蕴含 $\psi_{\mathbf{x}}(4)$。若 $i_1 < i_2 < i_3$ 是满足 $\psi_{\mathbf{x}}$ 的数中的连续项，则 $i_3 + i_1 = 2i_2 + 2$。

若 $\varphi(\mathbf{x})$ 成立，则 $k \ge 4$，且每个满足 $\psi_{\mathbf{x}}$ 的数都是平方数。事实上，若 $\psi_{\mathbf{x}}(2)$ 成立，下一个满足 $\psi_{\mathbf{x}}$ 的数将是 5，这与 $\psi_{\mathbf{x}}(4)$ 矛盾。同理， $\psi_{\mathbf{x}}(3)$ 不可能成立。沿着递推关系归纳可知， $\psi_{\mathbf{x}}(i)$ 成立意味着 i 是平方数。

反过来，令 $k \ge 4$，并令 σ 先列出 $\{2, \dots, k\}$ 中的平方数，再列出 1，最后列出 $\{2, \dots, k\}$ 中的非平方数。将引理 1.5 应用于 σ，可得到连续素数序列，因此取适当的 n、 r，使得 $p_i = \mathrm{pr}(n+i)$ 且 $q_i = \mathrm{pr}(r+i)$。根据 (a)， $p_i + q_1 \le p_1 + q_i$ 当且仅当 $i \le k$ 是平方数时成立，因此 $\varphi(k, n, r)$ 成立。

因此，

$$\mathsf{Sq}(i) \iff i = 0 \ \vee\ \exists n, r\ \big(\varphi(4i, n, r) \wedge \psi_{4i,n,r}(i)\big).$$ □

注记3.2. 与注 2.3 中一样，该方法利用希尔伯特第十问题的不可判定性，证明 $\mathrm{FO}(\mathbb{N}; +, \mathrm{pr})$ 的 $\exists^*\forall^*\exists^*$ 片段是不可判定的。由于多项式中的平方数和乘法所需的许多数值都位于外层存在量词中，希尔伯特第十问题并不能直接限定一个量词块中的量词数量。根据 Jones 的结果 [7], ，存在一个显式多项式 $f \in \mathbb{Z}[\mathbf{y}, \mathbf{x}]$ ，使得判定对于给定的 $\mathbf{y}$，是否存在 $\mathbf{x}$ 使 $f(\mathbf{y}, \mathbf{x}) = 0$ 成立，是不可判定的。由于 1 是可用存在量词定义的（$z = 1$ 是唯一满足 $z \ne 2z = \mathrm{pr}(z)$ 的数），我们可以在一个句子中将 $\mathbf{y}$ 的一个分量写成 $1 + 1 + \cdots + 1$。因此，存在 A、 B、 $C \in \mathbb{N}$，使得 $\mathrm{FO}(\mathbb{N};\ +, \mathrm{pr})$ 的 $\exists^A\forall^B\exists^C$ 片段是不可判定的；这里不对其进行计算。

4. 引理的证明1.5

我们先给出 Hildebrand 和 Maier 定理 1.4的离散版本。

对于一个序列 $p_0 < \cdots < p_k$，定义其间隔向量 为 $(p_1 - p_0, \dots, p_k - p_{k-1})$。对于 $x > 1$、 $T > 0$、$N \ge 1$ 和 $\mathbf{n} \in \{1, \dots, N\}^k$，令

$$\mathcal{B}_{\mathbf{n}} = \prod_{j=1}^{k} \left((n_j - 1)\frac{T \log x}{N},\ n_j \frac{T \log x}{N} \right].$$

如果一个盒子包含由 $k+1$ 个连续素数组成的序列的间隔向量，则称该盒子被命中 .

定理 4.1 ([6])。对于 $k \geq 1$，存在 $c \in (0,1]$ 和 $T > 0$，使得对于所有 $N \geq 1$，都有任意大的 x *(取决于 N)*，使得至少 cN^k 个盒子 $\mathcal{B}_{\mathbf{n}}$ 被命中。

这正是 Hildebrand和Maier的论文中的不等式（17）。回顾一下他们的论证：他们构造了一个矩阵，其中包含 $\gg R$ 组由 $k+1$ 个连续素数组成的元组（他们的公式（14）），并证明每个 N^k 个盒子至多包含 $O(R/N^k)$组这样的元组（他们的公式（19））。因此，$\gg N^k$ 个盒子包含这样的元组，从而证明了（17）。这对 x的某个单一取值成立，并且对任意大的 x值也成立。

下一个引理通过鸽巢原理证明：对于给定的向量 $\mathbf{w}$，只要 $\{1,\dots,N\}^k$ 的子集足够大，其中就有两个点之差是 $\mathbf{w}$的一个较小倍数。

引理 4.2. 设 $\mathbf{w} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$，$L \geq 2$，且 $H \subseteq \{1,\dots,N\}^k$满足 $\#H > N^k/L + \|\mathbf{w}\|_1 N^{k-1}$。那么存在 $\mathbf{b} \in H$ 和 $1 \leq t \leq L-1$，使得 $\mathbf{b} + t\mathbf{w} \in H$。

证明。直线 $\{\mathbf{b} + t\mathbf{w} : t \in \mathbb{Z}\}$ 将盒子 $\{1,\dots,N\}^k$划分开来。每条直线 ℓ 与 $\{1,\dots,N\}^k$交于连续的点 $\mathbf{b}_\ell$、$\mathbf{b}_\ell + \mathbf{w}$，…，$\mathbf{b}_\ell + (s_\ell - 1)\mathbf{w}$，其中 $\mathbf{b}_\ell$和 $s_\ell \geq 1$满足相应条件。将每个这样的等差数列切分成包含 L 个连续点的片段，第一个片段可以较短。由于同一片段中的两个点相差 $t\mathbf{w}$ 且满足 $1 \leq t \leq L-1$，因此只需证明片段数少于 $\#\mathcal{H}$。

片段数至多为 N^k/L（至多是那些不包含 $\mathbf{b}_\ell$的片段数）加上与 $\{1,\dots,N\}^k$ 相交的直线数（即包含某个 $\mathbf{b}_\ell$的片段数）。点 $\mathbf{b}_\ell$满足 $\mathbf{b}_\ell - \mathbf{w} \notin \{1,\dots,N\}^k$，因此 $b_{\ell,j} \leq w_j$对某个满足 $w_j > 0$的 j 成立，或者 $b_{\ell,j} > N + w_j$ 对某个满足 $w_j < 0$的 j成立。这种情况至多发生 $\sum_{j=1}^k |w_j| N^{k-1} = \|\mathbf{w}\|_1 N^{k-1}$次，因此片段数至多为 $N^k/L + \|\mathbf{w}\|_1 N^{k-1} < \#\mathcal{H}$。□

我们现在可以证明引理 1.5。

引理 1.5 的证明。固定 $k \geq 1$ 和 σ 的一个排列 $\{1,\dots,k\}$，令 $s(i) = \sigma^{-1}(i)$ 表示 $1 \leq i \leq k$ 的所需秩，并设 $s(0) = 0$。其思路是找到两个序列 $p_0 < \cdots < p_k$ 和 $q_0 < \cdots < q_k$，其中对于 $0 \leq i \leq k$，我们令 $e_i = p_i - q_i$。然后我们确保每个 $e_i - e_0$ 都在一个公共缩放因子意义下近似于 $4ks(i)$，且在相同尺度下误差小于 k。也就是说，间隔向量 $(p_i - p_{i-1})_{i=1}^k$ 和 $(q_i - q_{i-1})_{i=1}^k$ 大致相差 $\mathbf{w}$ 的某个倍数，其中 $w_j = 4k(s(j) - s(j-1))$。为找到这样的序列，我们对定理 4.1 中命中的盒子应用引理 4.2。令

$$c \in (0,1] \text{ 如定理 4.1, } L \text{ 一个整数，满足 } L > \tfrac{4}{c}\text{，并且 } N \text{ 一个整数，满足 } \frac{c}{4k} N \geq \max(8k^2 L, \|\mathbf{w}\|_1).$$

取定理 4.1中的 T 和 x，令 $\mathcal{H} \subset \{1,\dots,N\}^k$ 为命中盒 $\mathcal{B}_{\mathbf{n}}$ 的指标 $\mathbf{n}$集合，令$\widetilde{\mathcal{H}}$ 表示满足 $n_j \geq \frac{c}{4k}N + 1$（对所有 j）的那些指标，并将 $u = T^{\frac{x}{\log N}}$ 记为这些盒子的边长。于是，$\#\mathcal{H} \geq cN^k$。由于 $1 < \frac{c}{4k}N$，至多有 $k \cdot (\frac{c}{4k}N + 1) N^{k-1} < \frac{c}{2}N^k$ 个盒子具有某个 $n_j < \frac{c}{4k}N + 1$，因此

$$\#\widetilde{\mathcal{H}} > \#\mathcal{H} - \tfrac{c}{2}N^k \geq \tfrac{c}{4}N^k + \tfrac{c}{4}N^k > N^k/L + \|\mathbf{w}\|_1 N^{k-1}.$$

引理 4.2 给出 $\mathbf{b}$、$\mathbf{b} + t\mathbf{w} \in H$，满足 $1 \leq t \leq L-1$。取一个间隔向量为 $(q_j - q_{j-1})_{j=1}^k$、属于 $B_{\mathbf{b}}$ 的序列 $q_0 < \cdots < q_k$，以及一个间隔向量为 $(p_j - p_{j-1})_{j=1}^k$、属于 $B_{\mathbf{b}+t\mathbf{w}}$ 的序列 $p_0 < \cdots < p_k$。由于边长为 u 的盒子在坐标 j 上彼此相隔 $tw_j = t \cdot 4k(s(j) - s(j-1))$ 步，因此我们有

$$\left| e_j - e_{j-1} - 4tuk(s(j) - s(j-1)) \right| < u.$$

对 $j \leq i$ 求和得到两个望远镜求和式，从而可得

$$\left| (e_i - e_0) - 4kt\, s(i)\, u \right| < iu \leq ku.$$

如果 $s(i') = s(i) + 1$，那么 $(e_{i'} - e_0) - (e_i - e_0) > 4ktu - 2ku \geq 2ku > 0$，这就证明了 (a)。对于 (b)，$\Delta = \max_i |e_i - e_0| < 4k^2 tu + ku < 4k^2 Lu \leq \frac{c}{8k} Nu$，而 q-序列的每个间隔都至少为 $\frac{c}{4k}Nu > 2\Delta$。□

5. 讨论

定理 1.2 的证明并未表明乘法可在 $\langle \mathbb{N}; +, \mathsf{P}\rangle$ 中定义，而定理 1.3 表明它可在 $\langle \mathbb{N}; +, \mathrm{pr}\rangle$ 中定义。关键区别在于，函数版本允许对素数索引进行算术运算。因此，我们预计本文中的技术无法确定乘法在 $\langle \mathbb{N}; +, \mathsf{P}\rangle$ 中的可定义性，还需要有关素数之间线性关系的更精确信息，例如Dickson猜想所涉及的那些关系。

根据注记2.3和3.2，FO($\mathbb{N}$; +, P) 和 FO($\mathbb{N}$; +, pr) 的 $\exists^*\forall^*\exists^*$片段都是不可判定的。我们预计这些结果并非最优，并且这两个理论的 $\exists^*\forall^*$片段也都是不可判定的。例如，假设 Dickson猜想，对于每个自然数的有限序列 $(t_i)_{i=1}^{\tau}$，都存在一个数 n，使得 $\mathrm{pr}(n+2i+1)-\mathrm{pr}(n+2i) = 2t_i+2$ 对于 $1 \le i \le \tau$。利用第2.1节中的技术，我们可以在 FO($\mathbb{N}$; +, pr) 的 $\exists^*\forall^*$片段中编码给定的计数器机器，从而得到不可判定性。同时，我们预计，在假设 Dickson猜想的情况下，FO($\mathbb{N}$; +, P) 的存在片段又是可判定的。

虽然序在 Presburger 算术中是隐含的，但它也可以在结构 $\langle \mathbb{Z};\ +, \mathsf{P}\rangle$ 中定义：根据 Tao 的一项结果 [12], ，每个不小于 3 的奇数都是至多 5 个素数之和，因此每个非负偶数都是至多 6 个素数之和。因此，$x \ge 0$ 当且仅当 $2x$ 是至多 6 个素数之和。由此，我们可以在结构 $\langle \mathbb{Z}; +, \mathsf{P}\rangle$ 中还原自然数，使其一阶理论不可判定。序的可定义性似乎是必要的，因为 Kaplan 和 Shelah [8] 证明，在假设 Dickson猜想成立的情况下，FO($\mathbb{Z}; 0, 1, +, \mathsf{P} \cup \{-p\colon\ p \in \mathsf{P}\}$) 是可判定的。

致谢

作者获得了牛津大学 Glasstone慈善基金的支持，[Violette 和 Samuel Glasstone 科学研究奖学金 2025]。

AI 声明。本注使用 Claude Opus 5.5 撰写。具体而言，在提示其参考引理 1.5和Hildebrand和Maier的论文 [6], 后，它给出了第 4节所呈现证明的初稿。证明的其余部分由我们自行完成，不过 AI 帮助我们构思并润色了论证。此外，我们还使用 AI 校对本注并提出修改建议。作者对结果的正确性承担全部责任。

参考文献

1. W. D. Banks、T. Freiberg 和 C. L. Turnage-Butterbaugh，《元组中的连续素数》，Acta Arith. 167 (2015), 261–266。2. P. T. Bateman、C. G. Jockusch 和 A. R. Woods，《带有素数谓词的理论的可判定性与不可判定性》，J. Symbolic Logic 58 (1993), 672–687。3. A. Bès，《算术可定义性综述》，Bull. Belg. Math. Soc. Simon Stevin (2001), no. suppl., 1–54, Suppl.（向 Maurice Boffa 致敬）。4. P. Cégielski、D. Richard 和 M. Vsemirnov，《素数加法理论》，Fund. Inform. 81 (2007), 83–96。5. B. Green 和 T. Tao，《素数中的线性方程》，Ann. of Math. (2) 171 (2010), 1753–1850。6. A. Hildebrand 和 H. Maier，《素数间隔》，Proc. Amer. Math. Soc. 104 (1988), 1–9。7. J. P. Jones，《通用丢番图方程》，J. Symbolic Logic 47 (1982), 549–571。8. I. Kaplan 和 S. Shelah，《带有素数的整数理论的可判定性与分类》，J. Symbolic Logic 82 (2017), 1041–1050。9. Itay Kaplan 和 Saharon Shelah，《带有素数的整数理论的可判定性与分类》，Journal of Symbolic Logic 82 (2017), no. 3, 1041–1050。10. T. Karimov、J. Nieuwveld 和 J. Ouaknine，《丰富序列与算术理论的可判定性》，2026年，预印本，arXiv:2609.20415。11. H. Putnam，《可判定性与本质不可判定性》，J. Symbolic Logic 22 (1957), 39–54。12. T. Tao，《每个大于 1 的奇数都能表示为至多五个素数之和》，Math. Comp. 83 (2014), 997–1038。

牛津大学计算机科学系