arXiv:2602.09011v1 [math.AT] 9 Feb 2026

具有相同贝蒂曲线的条形码计数

Henry Ashley、H˚avard Bakke Bjerkevik、Justin Curry,

Riley Decker、Robert Green*

2026年2月10日

摘要

本文探讨拓扑数据分析（TDA）中的一个重要逆问题：有多少种不同的条形码会产生相同的贝蒂曲线？等价地，给定一个函数

$$\beta\colon [n] = \{1 < \cdots < n\} \to \mathbb{Z}_{\geqslant 0},$$

我们可以用多少种不同的方式，将 β 表示为由区间 $[n]$上的示性函数之和？为回答这个问题，我们将持久同调与 Kostant 分拆函数的研究，以及 Ronald Graham 等人研究的所谓“神奇”杂耍序列的计数组合学联系起来。具体而言，我们证明了我们的逆问题与这另外两个领域中相应命题之间的等价性。从应用和统计的角度来看，我们的工作量化了从持久同调转向持续贝蒂数时，拓扑数据分析流程所造成的信息损失。

1 引言

1.1 背景与先前工作

拓扑数据分析（TDA）为提取和概括数据中的形状提供了一套出色的工具。TDA流程的输入可以是若干数据类型之一：取自 $\mathbb{R}^d$的有限子集、半代数集、流形上的标量场或向量场，等等。TDA流程的输出也可以是若干类概括对象之一：合并树、Reeb 图、持久图/条形码，或分段常数 $\mathbb{Z}$-值函数，例如贝蒂曲线或欧拉曲线。通过固定输入数据的类别和输出对象的类型，人们自然会遇到一个逆问题：对于该类数据集，TDA流程距离成为单射还有多远？由于拓扑必然忽略数据集的某些方面，可以预期该映射并非单射，并且通常具有高维且几何上复杂的纤维[1]。这一点可由拓扑数据分析中一系列扎实的逆问题研究加以证明，其中 [21, 23, 14, 27, 12, 33,18, 16, 30, 29, 22, 31, 38, 7, 6, 15, 37]很可能只是一次不完整的抽样。

*纽约州立大学奥尔巴尼分校数学与统计学系[1]提醒一下，映射 $f\colon Y \to X$ 在点 $x \in X$ 上的纤维是其原像 $f^{-1}(x)$。

本文将既有研究 [14, 27, 22, 31, 15]拓展到TDA对象自身之间的逆问题，这类问题往往更具组合色彩。那些论文主要关心的是理解 π_0，亦即连通分支的集合，是如何细化 H_0，即向量空间，它由连通分支的集合自由生成。在持久性设定下，这引出了如今已被充分理解的从合并树到条形码的映射。本文所考虑的问题在某些方面与之相似，因为我们考察的是用其维数取代向量空间这一过程的信息损失有多大。

1.2 问题陈述

在本文中，我们针对有限索引情形，提出并阐明了一个迄今尚未被研究的两个广受研究的拓扑数据分析对象——条形码和贝蒂曲线——之间的逆问题。具体而言，回顾如下：给定任意函子

$$F\colon [\mathbf{n}] = \{1 < \cdots < n\} \to \mathbf{vect}$$

从含有 n 个元素的全序集到有限维向量空间范畴的函子，都有一个与之对应的维数（或希尔伯特）函数

$$\dim(F)\colon [n] \to \mathbb{Z}_{\geqslant 0},$$

它将每个 $i \in [n]$ 映射为向量空间 $F(i)$ 的维数。我们的逆问题如下：

给定：一个维数函数 $h\colon [n] \to \mathbb{Z}_{\geqslant 0}$。

枚举：所有函子 $F\colon [\mathbf{n}] \to \mathbf{vect}$，按同构分类，满足 维数 $F = h$。

等价地，函子 $F\colon [\mathbf{n}] \to$ **向量空间**是一个类型为 A_n 的箭图表示，其同构类由其分解为区间表示所确定。维数函数以及分解得到的区间多重集，就是我们定义贝蒂曲线和条形码的依据。

定义 1.1。长度为 n 的**贝蒂曲线**是一个函数 $\beta\colon [n] \to \mathbb{Z}_{\geqslant 0}$，或者等价地，是一个有限序列 $(\beta_1, \beta_2, \ldots\ldots, \beta_n)$，其中 $\beta_i = \beta(i)$。我们将 (n) 定义为长度为 n的贝蒂曲线集合。由于 $\beta(i) \in \mathbb{Z}_{\geqslant 0}$ 可以任意取值，**贝蒂**(n) 必然是一个无限集合。

对于 $i \leqslant j$，我们用记号 $[i, j)$ 表示集合 $\{i, i+1, \ldots\ldots, j-1\}$，并将这种形式的非空集合称为**区间**。例如，按照此约定，$[n] = [1, n+1)$。

定义 1.2。**条形码** 是区间多重集。换言之，条形码是一个集合

$$B = \{(I_j, m_j) \mid I_j \text{ 是 } [n] \text{ 中的一个区间，而 } m_j \in \mathbb{Z}_{\geqslant 0} \text{ 表示重数}\}$$

我们将$\mathbf{Barc}(n)$ 定义为区间是 $[n]$的子集的所有条形码组成的集合。

注意，我们可以将贝蒂曲线 β 与条形码 B关联起来：令 $\beta(i)$ 为 B 中包含 i的区间数量。更准确地说，令 $B = \{(I_j, m_j)\}$。我们定义函数 $\varphi\colon \mathbf{Barc}(n) \to \mathbf{Betti}(n)$，按以下方式将条形码关联到贝蒂曲线：

$$\begin{aligned} \varphi(B) &= \beta, \\ \beta(i) &= \sum_{j \mid I_j \cap \{i\} \neq \varnothing} m_j. \end{aligned}$$

注意， φ 是满射：对于 $\beta = (\beta_1, \beta_2, \ldots\ldots, \beta_n)$，我们可以构造 $B \in \varphi^{-1}(\beta)$，令 B 对每个 i都包含 β_i 个 $[i, i+1)$ 的副本。我们定义

$$\mathbf{Barc}(\beta) \coloneqq \varphi^{-1}(\beta).$$

换言之，$\mathbf{Barc}(\beta)$ 是贝蒂曲线为 β 的所有条形码的集合；它刻画了所有 A_n 型箭图表示的维数函数的纤维。因此，我们的逆问题等价于以下问题：

Given: *a Betti curve* $\beta\colon [n] \to \mathbb{Z}_{\geqslant 0}$.

枚举：$\boldsymbol{Barc}(\beta)$ 中的所有条形码 B 。

举一个具体的枚举问题示例，若 $n = 3$ 且已知贝蒂曲线为 $\beta = (2, 3, 2)$，则有 13 种不同的条形码具有这条贝蒂曲线。

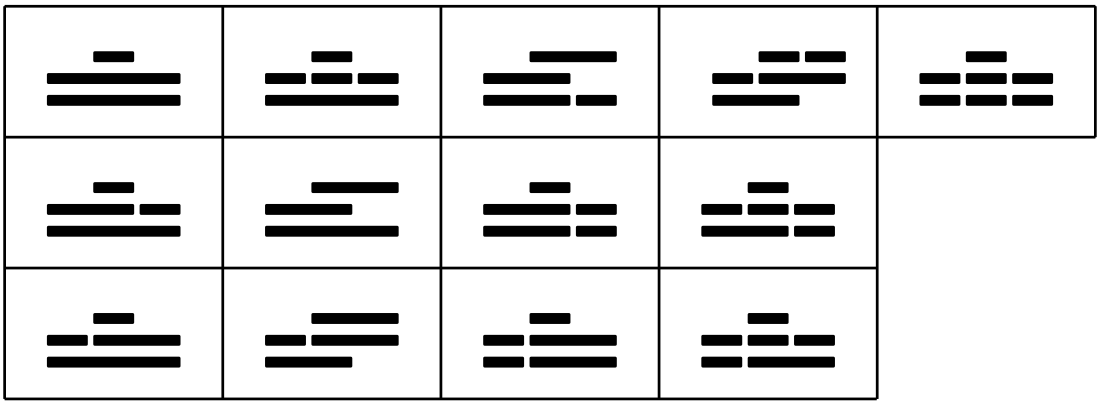

图 1： 贝蒂曲线为 $\beta = (2, 3, 2)$ 的所有条形码表格。此处每条 bar 都是其自身的区间；将高度相同的条形码放在一起仅为便于视觉呈现。

1.3 主要贡献

我们从多个方面为条形码计数问题提供了解法：首先，我们从李代数的视角重新诠释这一问题，并指出在 A_n 型李代数框架下，我们所提问题的答案恰好是 Kostant 分拆函数 K。

n定理 A。 如果 $\mu = \sum_{i=1} \beta_i \alpha_i$ 是 A_n 型李代数的一个权，那么

$$|\mathbf{Barc}(\beta)| = K(\mu).$$

上文中，我们用 α_i 表示类型为 A_n 的根系中的单根 $\mathbf{e}_i - \mathbf{e}_{i+1}$。Kostant 分拆函数是表示论和代数组合学中的一个著名对象；它用于计数李代数的一个权被分拆为正根的非负整数线性组合的方式数。回想 Gabriel 定理时，Kostant 分拆函数的出现就显得十分自然。该定理指出，Dynkin 箭图表示的不可分解表示与相应 Dynkin 图的根系中的正根一一对应。由于域系数的持久模是一个 A_n 箭图

表示，因此条形码与正根组合之间的这种关系并不令人意外。

在第 2.3节中，我们着重讨论这个逆问题的计算方面，并给出一个显式递归公式，用于计算$\mathbf{Barc}(\beta)$的元素个数。Schmidt 和 Bincer [35] 最初是在 Kostant 分拆函数的背景下发现这一公式的；我们为该结果提供了另一种证明，并给出图示解释。

在第3节中，我们建立了条形码与所谓的“魔术多重”杂耍序列 [11] 之间的新型双射。杂耍序列是由杂耍状态组成的序列，而杂耍状态是一个向量 $s = \langle s_1, s_2, \dots, s_h \rangle$，可将其理解为当前空中所有球的高度。如果 $s_3 = 2$，那么空中有两个球，我们将在三个时间步后接住它们，并立即再次将它们抛向空中。杂耍状态向量中允许出现负数——这些负数最终会与被投掷到相应高度的球相抵消。令 $\mathrm{JS}(a, b, n)$ 表示长度为 n、初始状态为 a 且终止状态为 b 的所有魔术杂耍序列构成的集合；令 $\mathrm{JS}(\leqslant n+1, b, n)$ 表示长度为 n、终止状态为 b，且初始状态在前 $n+1$ 个条目之后为零的所有魔术杂耍序列构成的集合。对于贝蒂曲线 β，令 $\delta(\beta)$ 表示满足$\delta(\beta)_i = \beta_i - \beta_{i-1}$的向量，其中我们设定 $\beta_0 = 0$。我们定义一个映射 σ，将条形码映射为杂耍序列，并证明如下结果。

定理 B. 映射 $\sigma\colon \boldsymbol{Barc}(n) \to \boldsymbol{JS}(\leqslant n+1, \langle 0 \rangle, n)$ 是一个双射。此外，对于每条贝蒂曲线 $\beta = (\beta_1, \beta_2, \dots, \beta_n)$，$\sigma$ 限制为一个双射 $\sigma_\beta\colon \boldsymbol{Barc}(\beta) \to \boldsymbol{JS}(\langle \delta(\beta) \rangle, \langle 0 \rangle, n)$。

这便将拓扑数据分析与历史悠久的杂耍数学领域联系起来 [9,10, 13, 19, 8, 4, 36, 34]。我们的双射的基本思想是：条形码中一个在时间 i 出生、在时间 j 消亡的区间，等价于这样一个杂耍动作：在时间 i 将一个球投掷到高度 $j-i$。因此，贝蒂曲线也等价于任意时刻 i空中球的数量。

此外，在 [8], 的基础上，我们将定理A作为定理B的推论重新推导出来。因此，我们工作的最终成果，是构建了拓扑数据分析、杂耍和李代数的组合方面之间的一座现代罗塞塔石碑，如下所示。

TDA对象	杂耍类比	李代数类比
区间模 $\mathbb{k}[i,j)$	在时间 i 投掷至高度 $j-i$	A的正根 $\mathbf{e}_i - \mathbf{e}_j$
条形码	杂耍序列	权的分拆
贝蒂曲线	杂耍序列集合	A_n的权
$\lvert\mathbf{Barc}(\beta)\rvert$	$\lvert\mathbf{JS}(\langle \delta(\beta) \rangle, \langle 0 \rangle, n)\rvert$	$K\left(\sum_{i=1}^{n} \beta_i \alpha_i\right)$

2 背景

2.1 拓扑数据分析的表示论方面

我们在拓扑数据分析中研究的基本对象是**持久模**，即函子 $F\colon [n] \to$ **向量空间**。或者，我们也可以将这些函子视为某种类型的**表示** A_n的实例

Dynkin箭图

$$\bullet \to \bullet \to \cdots \to \bullet,$$

它是通过将每个顶点替换为一个向量空间、将每条箭替换为一个线性变换而得到的。换言之，它具有如下形式

$$V_1 \xrightarrow{\Phi_1} V_2 \xrightarrow{\Phi_2} \dots \xrightarrow{\Phi_{n-1}} V_n.$$

用表示论的语言来说，贝蒂曲线就是一个维数向量，即与某个表示相关联、记录每个指标处向量空间维数的向量。

当这些向量空间是有限维时，得益于以下两个定理，此类表示具有一种特别优良的形式。

定理2.1 (Azumaya-Krull-Remak-Schmidt)。 *[1]* 设 Q 是一个有限箭图，且$\Bbbk$是一个域。每个定义在$\Bbbk$ 上的有限维表示 $\mathbb{V}$ Q 都能唯一地（在同构及不可分解直和项的排列意义下）分解为直和

$$\mathbb{V}^1 \oplus \mathbb{V}^2 \oplus \cdots \oplus \mathbb{V}^m,$$

其中每个 $\mathbb{V}^i$ 都是不可分解的。

定理2.2 (Gabriel 定理，适用于类型 A_n 箭图)。 *[20]* 设 Q 是一个类型为 A_n 的箭图，且$\Bbbk$是一个域。

1. Q 的每一个不可分解的有限维表示都同构于形如[2]的表示之一

$$0 \to \cdots \to 0 \to \Bbbk \xrightarrow{id} \Bbbk \xrightarrow{id} \dots \xrightarrow{id} \Bbbk \to 0 \to \cdots \to 0.$$

2. 此外，这些不可分解表示与 A_n根系的正根之间存在双射，相关构造将在下文回顾。

我们将此类不可分解表示称为**区间表示**，并将$\Bbbk$的第一个指标称为该区间表示的**出生**，将出生之后的第一个零向量空间称为该区间表示的**死亡**。我们用 $\Bbbk[i, j)$ 表示出生为 i 、死亡为 j的区间表示。这些定理给出了我们对条形码的定义。由于每一个 A_n 型箭图表示都可（在同构意义下）分解为区间表示的直和，我们可以用来自这些区间的出生-死亡对多重集完全刻画该表示的同构类型。

拓扑数据分析界常常忽略的是，Gabriel 定理的第二部分给出了与李代数的直接联系，我们现在对此加以说明。

2.2 李代数与Kostant分拆函数

由于我们关注的是 A_n 型李代数，因此我们将定义陈述在这一具体层次上。有关李代数及其表示论的更多背景，请参见 [26]。

[2]本陈述中箭头的方向实际上取决于它们在 Q中的方向。我们选择让 Q 的所有箭头都向前，因为该箭图的表示是由全序集 $\{1 <$索引的持久模。$\dots < n\}$。

定义 2.3。 A_n 型**根系** 是向量的集合

$$\Phi_{A_n} = \{\pm(\mathbf{e}_i - \mathbf{e}_j) : 1 \leqslant i < j \leqslant n+1\},$$

其中，每个$\mathbf{e}_i$ 都是一个标准基向量。 Φ_{A_n} 中的元素称为**根**，我们用$\Phi^+_{A_n} \subset \Phi_{A_n}$ 表示正根集合，也就是形如$\mathbf{e}_i - \mathbf{e}_j$ 且满足$i < j$的根。Φ_{A_n} 的**单根**为

$$\{\alpha_i = \mathbf{e}_i - \mathbf{e}_{i+1} : 1 \leqslant i \leqslant n\}.$$

注意，$\Phi^+_{A_n}$ 中的每个元素都可以唯一地表示为单根的正和。特别地，$\mathbf{e}_i - \mathbf{e}_j = \alpha_i + \cdots + \alpha_{j-1}$。

定义 2.4。若元素 $\mu \in \mathbb{R}^{n+1}$ 位于根格的正锥中，即对于 $z_i \in \mathbb{Z}_{\geqslant 0}$有 $\mu = \sum_{i=1}^{n} z_i \alpha_i$ ，则称其为一个**权**。

虽然权通常有更广义的定义，但这个定义足以满足我们的需求。

定义 2.5。**Kostant 分拆函数** K 是一个函数，用于计数权 μ 可以表示为正根的非负整数线性组合的方式数。

Kostant 分拆函数最初用于建立一个公式，用来计算复半单李代数的不可约表示中某个权的重数 [28],但它也出现在各种组合背景中，例如杂耍序列的研究[8] 以及流多胞体 [32]。

例 2.6. 权 $\mu = \alpha_1 + 2\alpha_2 + \alpha_3$ 可以用 5 种不同的方式表示（每个选取的正根都用括号括起来）：

$$\begin{aligned}
\mu &= (\alpha_1 + \alpha_2 + \alpha_3) + (\alpha_2) \\
&= (\alpha_1 + \alpha_2) + (\alpha_2) + (\alpha_3) \\
&= (\alpha_1) + (\alpha_2) + (\alpha_2 + \alpha_3) \\
&= (\alpha_1 + \alpha_2) + (\alpha_2 + \alpha_3) \\
&= (\alpha_1) + (\alpha_2) + (\alpha_2) + (\alpha_3)
\end{aligned}$$

因此， $K(\mu) = 5$。

现在我们可以证明定理A。

定理 A. 如果 $\mu = \sum_{i=1}^{n} \beta_i \alpha_i$ 是类型为 A_n的李代数的一个权，那么

$$|\mathbf{Barc}(\beta)| = K(\mu).$$

证明。 首先，我们注意到，Gabriel 定理第 2 部分所述的不可分解 A_n 箭图表示与相应根系中的正根之间的双射，由（Hilbert）维数函数 dim 给出。更确切地说，区间表示 $\Bbbk[i, j)$ 被映射为维数向量

$$(0, \ldots, 0, \underbrace{1, 1, \ldots, 1}_{[i,j-1]}, 0, \ldots, 0)$$

在单根基下，即正根 $\alpha_i + \ldots + \alpha_{j-1} = \mathbf{e}_i - \mathbf{e}_j$。由于我们可以将每个正根 $\mathbf{e}_i - \mathbf{e}_j$ 与区间 $[i, j)$ 对应，因此权

$$\mu = \sum_{i=1}^{n} \beta_i \alpha_i$$

可分解为正根之和的方式数，恰好等于$\beta\colon [n] \to \mathbb{Z}_{\geqslant 0}$ 可写成由区间支撑的示性函数之和的方式数。 □

2.3 递归公式

目前评估 $K(\mu)$ 的大多数方法，都是将问题建模为计算由 μ定义的多面体的整数格点数。最著名的整数格点计数技术是 Barvinok 1994 年算法 [5]。格点计数可以借助**Sage** [24]，来完成，也可以使用**barvinok**[39] 或**LattE** [17]等更专业的软件。由于编码 Kostant 分拆函数的多面体是流多胞体，因此还可以进行更专业的计算 [3, 32]。有关计算 Kostant 分拆函数的方法概述，请参见 [2]。在本节余下部分，我们将介绍并解释一种递归方法来计算 Kostant 分拆函数，该方法适用于某种类型的 A_n 李代数。这种方法的核心技巧之一，是识别“位于”贝蒂曲线“下方”的杨图。

设 $X = (x_1, x_2, \ldots, x_n)$ 与 $Y = (y_1, y_2, \ldots, y_n)$ 为非负整数序列。我们称 $Y \sqsubset X$，如果

1. Y 是非增的 2.
$y_1 = x_1$3. 对所有
$i \neq 1, \quad y_i \leqslant x_i$

设 $X + Y$ 与 $X - Y$ 分别表示有限序列的逐坐标和与逐坐标差。

定理 2.7. *[35, 方程 4.21]* 设 β 为一条贝蒂曲线。下列递推公式成立。

$$|\boldsymbol{Barc}(\beta)| = \sum_{Y \sqsubset \beta} |\boldsymbol{Barc}(\beta - Y)|$$

注意，每个 $\beta - Y$ 都是以零开头的序列，因此可将其视为长度为 $n-1$的贝蒂曲线。我们所使用的记号不同于 Schmidt 与 Bincer 在其文章 [35]中所使用的记号。我们的方程由他们论文中的方程 4.21 所表明，而条件 2 与 3 则分别由其方程 4.20 与 4.22 加以阐明。正如我们将大小为 n 的贝蒂曲线化为对大小为 $n-1$的贝蒂曲线求和，他们将 A_n 型李代数中某一权的分拆函数化为对 A_{n-1} 型李代数中诸权上的分拆函数求和。为方便起见，我们给出该定理的一个简短且自足的证明。

证明。 回忆函数 $\varphi\colon \mathbf{Barc}(n) \to \textbf{贝蒂}(n)$，它将条形码映射为贝蒂曲线，定义见引言。定义 $\sim$ 上的元素之间的等价关系：若 B 与 B' 具有相同的、始于 1 的区间多重集 y，则 $B \sim B'$。我们用 Θ_y 表示相应的等价类。对于 $Y \sqsubset \beta$，存在唯一的区间多重集 Y，使得所有区间都始于 1，且 $\varphi(Y) = Y$。由此可知，$\sim$ 的等价类与满足 $Y \sqsubset \beta$的 Y 集合之间存在双射。因此，为了证明

该引理，只需针对每个 $Y \sqsubset \beta$构造一个双射 $\gamma\colon \Theta_Y \to \varphi^{-1}(\beta - Y)$。令$\gamma(X) = X - Y$。假设两者都有定义，则 γ 的逆映射显然为 $\gamma^{-1}(Z) = Z \sqcup Y$。以下计算表明 γ有定义：

$$\varphi(X - \overline{Y}) = \varphi(X) - \varphi(\overline{Y}) = \beta - Y.$$

对于 Z in $\varphi^{-1}(\beta - Y)$，我们有

$$\varphi(Z \sqcup \overline{Y}) = \varphi(Z) + \varphi(\overline{Y}) = \beta - Y + Y = \beta,$$

so $\gamma^{-1}(Z) \in \varphi^{-1}(\beta)$。由于 Z 没有以 1 开始的区间，因此得到 $\gamma^{-1}(Z) \in \Theta_Y$ $-$. □

直观而言，序列 $Y \sqsubset \beta$ 源自某个条形码中从 1 开始的区间多重集，该条形码的贝蒂曲线为 β。由于恰好有 $\varphi^{-1}(\beta - Y)$ 个条形码满足这一性质，因此该公式对这个集合中的每个元素恰好计数一次。我们可以用图示来理解 $\sqsubset$ 条件：当 Y 是一个杨图，且“包含在” β 之中，并且 Y 和 β 的前几列中的单元格数量相同时， $Y \sqsubset \beta$ 。

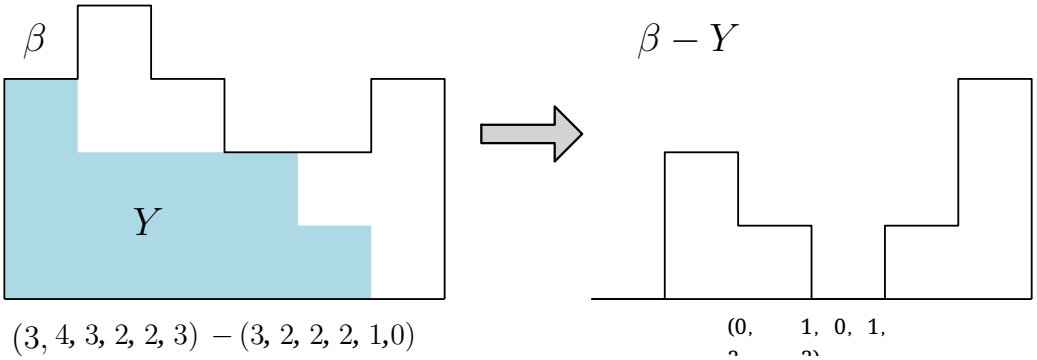

图2： 包含一个序列 $Y \sqsubset \beta$ （左）的贝蒂曲线 β 示例，以及由此得到的贝蒂曲线 $\beta - Y$（右）。

2.4 杂耍序列

对杂耍的数学研究至少可以追溯到20世纪80年代，当时 Claude Shannon 撰写了《杂耍的科学方面》，后来该文作为其文集的一部分出版 [36]。Shannon 的论文涵盖了许多观点，探讨了杂耍的历史、物理、数学和机器人学方面。然而，近期的杂耍数学研究主要源自 [9]，，旨在枚举在各种限制条件下进行理论杂耍的不同方式 [4, 11, 13, 34]。在杂耍研究中，我们考虑杂耍状态（或落地时刻表），即向量，其各项表示若干个球到达杂耍者手中的（离散）时刻。一个杂耍序列是由一系列杂耍状态构成的序列，代表随时间进行的杂耍动作。杂耍序列中的相邻状态之间必须存在所谓的有效转移，其定义因所研究的杂耍变体而异。下文将遵循 [8]给出的魔术杂耍约束。

定义 2.8. 我们用记号 $\langle s_1, s_2, \ldots, s_h \rangle$ 表示无限长度的零填充向量 $(s_1, s_2, \ldots, s_h, 0, 0, \ldots)$。**杂耍状态**是一个向量$\mathbf{s} = \langle s_1, s_2, \ldots, s_h \rangle$，其中

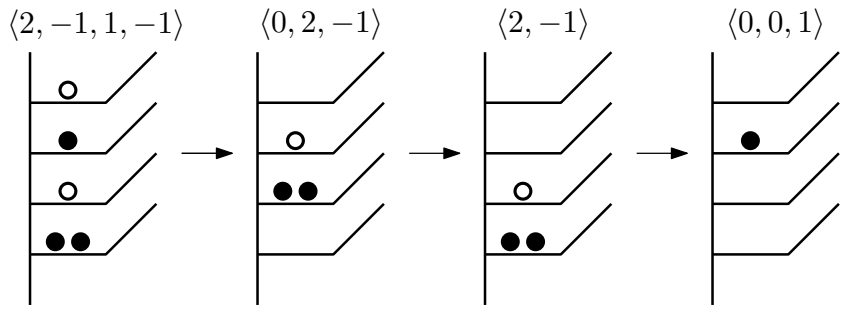

图3： 用“桶”形图表示的魔术杂耍序列。在杂耍序列中，每到下一个状态，所有球都会下降1个桶，除了位于最底层桶中的球，它们会被重新分配。魔术球以白色显示，表示负条目；普通球以黑色显示，表示正条目。当魔术球和普通球位于同一个桶中时，它们会相互抵消。在此例中，时间1时的两个球分别被投掷到高度1和高度2处，前者与一个魔术球抵消。时间2时没有球被投掷。时间3时，最底层的两个球分别被投掷到高度1和高度3处，前者与剩下的那个魔术球抵消。

$h s_i \in \mathbb{Z}$ 且满足 $\sum s_j = m \geqslant 0$。长度为 n 的**魔术杂耍序列**是一个由杂耍状态组成的有序列表 $S = (\mathbf{s}^0, \mathbf{s}^1, \ldots, \mathbf{s}^n)$，其中相邻状态之间存在有效转移。从状态$\mathbf{s}^{i-1} = \langle s_1, s_2, \ldots, s_h\rangle$到状态$\mathbf{s}^i$的**有效转移**发生于$\mathbf{s}^i = \langle s_2 + t_1, s_3 + t_2, \ldots, s_h + t_{h-1}, t_h, \ldots, t_{h'}\rangle$时，其中每个 t_j 都是满足 $\sum_{j=1}^{h'} t_j = s_1$的非负整数。我们称 t_j 为时刻 i 向高度 j投掷的次数。注意， t_j的集合总是相对于从$\mathbf{s}^{i-1}$ 到$\mathbf{s}^i$的转移来定义的，因此时间索引 i 是隐含的。

我们给出以下一个方便的观察，它完整刻画了合法的魔术杂耍序列。

观察 2.9。 令 $S = (\mathbf{s}^0, \mathbf{s}^1, \ldots, \mathbf{s}^n)$ 为一个杂耍状态序列，并记$\mathbf{s}^i = \langle s_1^i, \ldots, s_h^i\rangle$和 $s_k^i = 0$ 分别表示所有 $k > h$和 $k < 1$。注意，通过补零，我们可以为所有状态选择适当的 h 。那么， S是一个合法的魔术杂耍序列，当且仅当

(i) 对所有 $0 \leqslant i, j \leqslant n$，都有 $\sum_{k=1}^{h} s_k^i = \sum_{k=1}^{h} s_k^j$；
(ii) 对所有 $1 \leqslant i \leqslant n$ 和 $1 \leqslant k$，都有 $s_{k+1}^{i-1} \leqslant s_k^i$。

定义2.10。 我们用 $\mathbf{JS}(\mathbf{a}, \mathbf{b}, n)$ 表示长度为 n、初始状态为 $\mathbf{a}$ 且终止状态为 $\mathbf{b}$ 的所有魔术杂耍序列的集合。我们用 $\mathbf{JS}(\leqslant h, \mathbf{b}, n)$ 表示长度为 n 、终止状态为 $\mathbf{b}$ 的所有魔术杂耍序列的集合，但其初始状态可以是形如 $\langle s_1, \ldots, s_h\rangle$ 的任意状态；也就是说，在前 h 项之后，初始状态为零。

的作者 [8] 给出了以下结果，说明魔术杂耍序列与 Kostant 分拆函数之间的关系。

定理 2.11 ([8, 定理 3.8])。 设 μ 是类型为 A_n 的李代数的一个权，且 $(\mu_1, \ldots, \mu_{n+1})$ 是它的标准基向量表示。那么

$$K(\mu) = |\boldsymbol{JS}(\langle \mu_1, \ldots, \mu_n\rangle, \langle \mu_1 + \ldots + \mu_n\rangle, n)|.$$

3 双射：条形码与杂耍序列

定义 3.1. 设 β 是贝蒂曲线。 β的**微分** 记作 $\delta(\beta)\colon [n+1]\to\mathbb{Z}$，是通过以下方式定义的有限序列

$$\delta(\beta)=(\beta_1,\beta_2-\beta_1,\dots,\beta_n-\beta_{n-1},-\beta_n).$$

对于带有关联贝蒂曲线 B 的条形码，我们有时会滥用微分记号，写作 $\delta(B)$ 而不是 $\delta(\varphi(B))$。

定义 3.2。 给定条形码 B和 $i\in\mathbb{N}$，将 B的 **i-截断** 定义为条形码 $\tau_i(B)$：对于 B 中满足 $a\geqslant i+1$的每一条 $[a,b)$，取一条 $[a-i,b-i)$。

简单来说，要构造 $\tau_i(B)$，我们取出 B中的各条，移除左端点不大于 i的条，然后将剩余各条向左移动 i步。

定义 3.3。 对于 $0\leqslant i\leqslant n$，我们记作 $B^i=\tau_i(B)$。我们定义**杂耍映射**

$$\sigma\colon \mathbf{Barc}(n)\to\mathbf{JS}(\leqslant n+1,\langle 0\rangle,n)$$

它将条形码 B映射到其对应的魔术杂耍序列

$$\sigma(B)\coloneqq(\langle\delta(B^0)\rangle,\dots,\langle\delta(B^n)\rangle).$$

证明其定义良好是定理B的一部分。

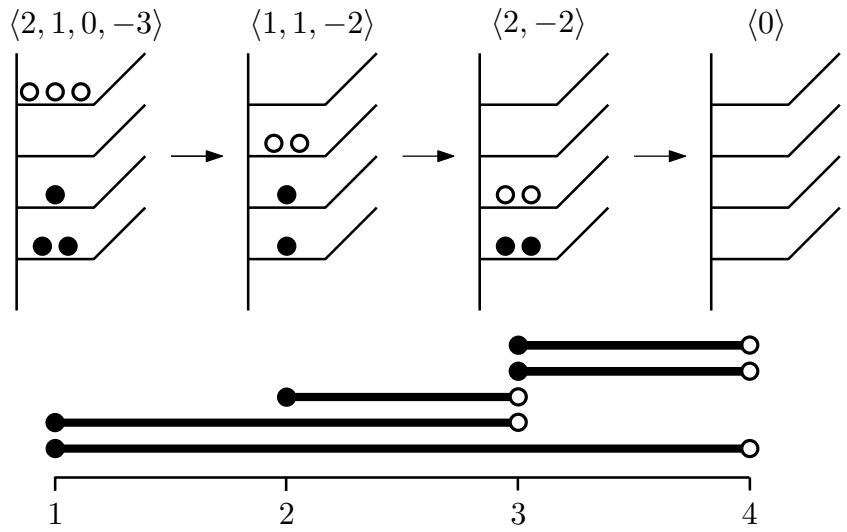

图 4： 在映射 σ 下，一个条形码及其对应的杂耍序列。

引理 3.4。设 B 是一个条形码，其贝蒂曲线为 $\beta=(\beta_1,\beta_2,\dots,\beta_n)$，并设 $b_{i,\,j}$ 表示区间 $[i,j)$ 在 B 中的重数。记 $\delta(B^i)=\langle s_1^i,\ s_2^i,\ \dots,\ s_{n+1}^i\rangle$，并令 $s_k^i=0$ 对于 $k<1$ 和 $k>n+1$。以下三个陈述成立：

(i) $\delta(B^0)=\delta(\beta)$

(ii) B^n 是空条形码，因此 $\delta(B^n)=0$

(iii) $s_j^i-s_{j+1}^{i-1}=b_{i,i+j}$ 对于 $i\in\{1,\dots,n\}$ 以及所有 $j\geqslant 0$。

证明。 观察 (i) 和 (ii) 显然。为说明 (iii)，注意到一个区间在 B 中贡献

- 若从 $i+j$开始，则对 s_j^i 和 s_{j+1}^{i-1} 均作正向贡献，
- 若属于带有 $a \geqslant i$的 $[a, i+j)$ 形式，则对 s_{j+1}^{i-1} 作负向贡献，并且
- 负向地指向 s_j^i，当且仅当它具有 $[a, i+j)$ 的形式，其中 $a \geqslant i+1$。

否则，它对两者都没有贡献。因此，$s_j^i - s_{j+1}^{i-1}$ 恰好是 B 中具有 $[i, i+j)$形式的条数。 □

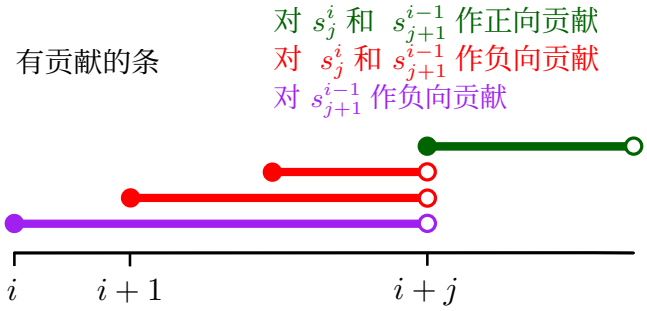

图 5： 展示各种条如何对 s_j^i 和 s_{j+1}^{i-1}作出贡献的示意图。

定理 B。 映射 $\sigma : \boldsymbol{Barc}(n) \to \boldsymbol{JS}(\leqslant n+1, \langle 0\rangle, n)$ 是一个双射。此外，对于每条贝蒂曲线 $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$，$\sigma$ 限制为一个双射 $\sigma_\beta : \boldsymbol{Barc}(\beta) \to \boldsymbol{JS}(\langle \delta(\beta)\rangle, \langle 0\rangle, n)$。

该定理可以用下面的交换图概括。

$$\begin{array}{ccccc} \mathbf{Betti}(n) & \xleftarrow{\varphi} & \mathbf{Barc}(n) & \xrightarrow[\sigma]{\sim} & \mathbf{JS}(\leqslant n+1, \langle 0\rangle, n) \\ \uparrow & & \uparrow & & \uparrow \\ \{\beta\} & \xleftarrow{\varphi} & \mathbf{Barc}(\beta) & \xrightarrow[\sigma_\beta]{\sim} & \mathbf{JS}(\langle \delta(\beta)\rangle, \langle 0\rangle, n) \end{array}$$

证明。 我们首先证明 σ_β 是一个双射。设 B 是一个贝蒂曲线为 β的条形码。首先，我们证明 $\sigma(B)$ 是一个有效的魔术杂耍序列。对于任意 B和任意 i，

$$\sum_j \delta(B^i)_j = 0,$$

因此观察 2.9 (i) 得以满足。此外，观察 2.9 (ii) 正是引理 3.4(iii)，所以 $\sigma(B)$ 是一个有效的魔术杂耍序列。因此，根据引理 3.4 (i) 和 (ii)，$\sigma(B)$ 是 $\mathbf{JS}(\langle \delta(\beta)\rangle, \langle 0\rangle, n)$中的一个定义良好的元素。由于 $\delta(\beta)$ 的长度为 $n+1$，这表明 σ和 σ_β 都有良好的定义。

引理3.4 (iii) 保证了 σ的单射性，因为我们可以从 $\sigma(B)$中读出 B 。

为证明 σ_β 的满射性，对于每个 $T = (t^0, t^1, \ldots, t^n) \in \mathrm{JS}(\langle \delta(\beta)\rangle, \langle 0\rangle, n)$，我们找到一个贝蒂曲线为 β 的条形码，使其在 σ 下的像为 T。记 $t^i = \langle t_1^i, t_2^i, \ldots, t_h^i\rangle$，并令 B_T 为区间多重集，其中对于每个 $i \in \{1, \ldots, n\}$ 和每个 $j \geqslant 0$，区间 $[i, i+j)$ 的重数为 $t_j^i - t_{j+1}^{i-1}$（根据观察 2.9 (ii)，该重数非负）。条形码 B_T 的贝蒂曲线为 β，因为 $\langle \delta(B_T)\rangle = \langle \delta(B_T^0)\rangle = \sigma(B_T)^0 = \langle \delta(\beta)\rangle$。记

$\sigma(B_T) = (s^0, s^1, \ldots, s^n)$ 和 $s^i = \langle s^i_1, s^i_2, \ldots, s^i_{n+1}\rangle$。我们有 $\sigma(B_T) = T$：根据引理 3.4 (i)，$s^0 = \delta(\beta) = t^0$,，并且根据引理 3.4 (iii)，对于所有相关的 i 和 j，$s^i_j - s^{i-1}_{j+1} = b_{i,i+j} = t^i_j - t^{i-1}_{j+1}$。由于 s^i 由 s^{i-1} 以及不同 j 对应的值 $s^i_j - s^{i-1}_{j+1}$ 决定，t^i 也同理，因此可得对于所有 i，$s^i = t^i$。因此，$\sigma(B_T) = T$。

我们有 σ 是单射，且对每个 β，σ_β 是双射。因此，只需证明：对于每一个形如 $a = \langle s_1, \ldots, s_{n+1}\rangle$ 且满足 $JS(a, \langle 0\rangle, n) \neq \emptyset$ 的杂耍状态，存在某条贝蒂曲线 $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$，使得我们有 $a = \langle \delta(\beta)\rangle$。我们令 $\beta_i = \sum_{j=1}^{i} s_j$（对于 $1 \leqslant i \leqslant n$），若 $s_{n+1} = -\beta_n$，则得到 $a = \langle \delta(\beta)\rangle$。这等价于 $\sum_{j=1}^{n+1} s_j = 0$。根据 $JS(a, \langle 0\rangle, n) \neq \emptyset$ 的假设，存在一条从 a 开始、在 $\langle 0\rangle$ 结束的有效杂耍序列 S。令 $S^i_{\geqslant j} = \sum_{k=j}^{h} s^i_k$，采用与观察 2.9 中相同的记号。由观察 2.9(i)，对所有 i 有 $S^i_{\geqslant 1} = 0$，因此特别地，$\sum_{j=1}^{n+1} s_j = 0$。

最后，我们需要所有 β_i 都非负，才能构成有效的贝蒂曲线。令 $S \in \mathbf{JS}(a, \langle 0\rangle, n)$，并令 $S^i_{\geqslant j} = \sum_{k=j}^{h} s^i_k$，使用与观察2.9中相同的记号。根据观察2.9(ii)，$S^{i-1}_{\geqslant j+1} \leqslant S^i_{\geqslant j}$ 对于 $1 \leqslant i \leqslant n$ 以及 $j \geqslant 1$。我们有 $S^n_{\geqslant j} = 0$对所有 j都成立，并且我们已经看到 $S^i_{\geqslant 1} = 0$ 对所有 i 都成立。因此，我们得到

$$S^i_{\geqslant j} \leqslant S^{i+1}_{\geqslant j-1} \leqslant \cdots \leqslant 0$$

对所有 i和 j都成立，因此

$$\beta_i = \sum_{j=1}^{i} s_j = S^0_{\geqslant 1} - S^0_{\geqslant i+1} \geqslant 0. \qquad \square$$

利用定理2.11和定理B，即可推论得到 Kostant 分拆函数与条形码计数问题之间的关系。

推论 3.5。若 $\mu = \sum_{i=1}^{n} \beta_i \alpha_i$ 是类型为 A_n的李代数的一个权，则

$$|\mathbf{Barc}(\beta)| = K(\mu).$$

证明。由定理 B，$\mathbf{Barc}(\beta)$ 和 $\mathbf{JS}(\langle \delta(\beta)\rangle, \langle 0\rangle, n)$ 具有相同的基数。$n+1$权 μ也可以写作 $\sum_{i=1} \delta(\beta)_i \mathbf{e}_i$。若令

$$\mathcal{J} := \mathbf{JS}(\langle \delta(\beta)_1, \ldots, \delta(\beta)_n\rangle, \langle \delta(\beta)_1 + \ldots + \delta(\beta)_n\rangle, n),$$

则由定理 2.11，$K(\mu) = |\mathcal{J}|$。注意，由于

$$\delta(\beta)_{n+1} = -\sum_{i=1}^{n} \delta(\beta)_i,$$

如果令 $S \in \mathbf{JS}(\langle \delta(\beta)\rangle, \langle 0\rangle, n)$，那么从 S 的初始状态中删除第 $(n+1)$个条目，会得到 $\mathcal{J}$ 的一个元素。由于该条目不会影响长度为 n的序列中被投掷或接住的球，因此这一删除在 $\mathbf{JS}(\langle \delta(\beta)\rangle, \langle 0\rangle, n)$与 $\mathcal{J}$之间建立了一个双射。 $\square$

4 讨论与未来工作

正如 [25], 所强调的，许多应用往往避免直接使用持久同调，而是使用下游的贝蒂曲线；贝蒂曲线易于定义、稳定，并且适用于平滑或取平均等标准统计技术。因此，我们的主要结果精确量化了使用贝蒂曲线会损失多少信息。正如图 1 所示，简单的贝蒂曲线 $\beta = (2, 3, 2)$ 可以由 13 个不同的条形码得到。再举一例，如果 $\beta = (2, 3, 1, 1, 1)$，计算机计算表明，具有该贝蒂曲线的不同条形码有 32 个。借助本文所建立的拓扑数据分析与代数组合学之间的新联系，许多现有结果很可能都能在这一视角下找到新的应用。最后，由于 Kostant 分拆函数可以针对非常一般的李代数表示进行研究，一个自然的问题是：能否利用这些联系来研究多参数持久性的逆问题。

参考文献

[1] Gorˆo Azumaya. 关于我那篇涉及 Krull-Remak-Schmidt 定理的论文的勘误与补充。名古屋数学杂志，1:117–124，1950。[2] M Welleda Baldoni、Matthias Beck、Charles Cochet 和 Michele Vergne。经典根系的多面体体积计算与分拆函数。离散与计算几何，35(4):551–595，2006。[3] Welleda Baldoni 和 Mich`ele Vergne。Kostant 分拆函数与流多胞体。变换群，13:447–469，2008。[4] Esther Banaian、Steve Butler、Christopher Cox、Jeffrey Davis、Jacob Landgraf 和 Scarlitte Ponce。素数杂耍模式计数。图与组合，32(5):1675–1688，2016年5月。[5] Alexander I. Barvinok。当维数固定时，计算多面体中整数点数量的多项式时间算法。运筹学数学，19(4):769–779，1994。[6] David Beers。树与曲线的拓扑学：理论与应用。博士论文，牛津大学，2023。[7] David Beers 和 Jacob Leygonie。树的持久同调纤维。*arXiv* 预印本 *arXiv:2303.16176*，2023。[8] Carolina Benedetti、Christopher R. H. Hanusa、Pamela E. Harris、Alejandro H. Morales 和 Anthony Simpson。Kostant 分拆函数与魔术多重杂耍序列。组合学年刊，24(3):439–473，2020年6月。[9] Joe Buhler、David Eisenbud、Ron Graham 和 Colin Wright。杂耍落球与下降。《美国数学月刊》，101(6):507，1994年6月。[10] Steve Butler、Fan Chung、Jay Cummings 和 Ron Graham。杂耍牌序列。组合学杂志，8(3):507–539，2017。[11] Steve Butler 和 Ronald L. Graham。多路杂耍序列的枚举。组合学年刊，13:413–424，2008。[12] Michael J Catanzaro、Justin M Curry、Brittany Terese Fasy、J¯anis Lazovskis、Greg Malen、Hans Riess、Bei Wang 和 Matthew Zabka。持久性 Morse 函数的模空间。《应用与计算拓扑学杂志》，4(3):353–385，2020。

[13] Fan Chung 和 Ron Graham。原始杂耍序列。《美国数学月刊》，115(3):185–194，2008。[14] Justin Curry。区间上函数的持久性映射的纤维。《应用与计算拓扑学杂志》，2:301–321，2018。[15] Justin Curry、Jordan DeSha、Ad´elie Garin、Kathryn Hess、Lida Kanari 和 Brendan Mallery。从树到条形码再回到树 II：拓扑逆问题的组合与概率方面。《计算几何》，116:102031，2024。[16] Justin Curry、Sayan Mukherjee 和 Katharine Turner。多少个方向可以确定一个形状，以及关于两个拓扑变换的其他充分性结果。《美国数学会会刊》B辑，9(32):1006–1043，2022。[17] Jes´us A. De Loera、Raymond Hemmecke、Jeremiah Tauzer 和 Ruriko Yoshida。有理凸多胞体中的有效格点计数。《符号计算杂志》，38:1273–1302，2004。[18] Jordan DeSha。拓扑数据分析中拓扑摘要的逆问题。博士论文，2021。[19] Richard Ehrenborg 和 Margaret Readdy。杂耍及其在 q-类比中的应用。《离散数学》，157(1–3):107–125，1996年10月。[20] Peter Gabriel。不可分解表示 I。《数学手稿》，6(1):71–103，1972。[21] Marcio Gameiro、Yasuaki Hiraoka 和 Ippei Obayashi。通过持久图延拓点云。《物理学D：非线性现象》，334:118–132，2016。[22] Ad´elie Eliane Garin。从树到条形码再回到树：拓扑逆问题的组合、概率与几何研究。博士论文，2022。[23] Ellen Gasparovic、Maria Gommel、Emilie Purvine、Radmila Sazdanovic、Bei Wang、Yusu Wang 和 Lori Ziegelmeier。度量图的一维内在 Čech 持久图的完整刻画。《计算拓扑研究》，第33–56页，2018。[24]Pamela E. Harris、Erik Insko 和 Anthony Simpson。计算单李代数表示的权重 q-重数。《工程、通信与计算中的应用代数》，29:351–362，2017。[25] Felix Hensel、Michael Moor 和 Bastian Rieck。拓扑机器学习方法综述。《人工智能前沿》，4:681108，2021。[26] James E Humphreys。《李代数与表示论导论》，第9卷。Springer Science & Business Media，2012。[27] Lida Kanari、Ad´elie Garin 和 Kathryn Hess。从树到条形码再回到树：理论与统计视角。《算法》，13(12):335，2020。[28] Bertram Kostant。权重重数的一个公式。《美国国家科学院院刊》，44(6):588–589，1958。[29] Jacob Leygonie。持久同调的微分与纤维。博士论文，牛津大学，2022。[30] Jacob Leygonie 和 Ulrike Tillmann。单纯复形的持久同调纤维。《纯粹与应用代数杂志》，226(12):107099，2022。

[31] Brendan Mallery、Ad´elie Garin 和 Justin Curry。持久性映射的格论视角。*arXiv* 预印本 *arXiv:2203.00643*，2022 年。[32] Karola M´esz´aros 和 Alejandro H Morales。带符号图的流多胞体与 Kostant 分拆函数。《国际数学研究通讯》，2015(3):830–871，2015 年。[33] Steve Oudot 和 Elchanan Solomon。拓扑持久性的逆问题。载于《拓扑数据分析：*Abel* 研讨会 *2018*》，第 405–433 页。Springer，2020 年。[34] Burkard Polster。杂耍数学。Springer，2003 年。[35] Jeffrey R Schmidt 和 Adam M Bincer。单李代数的 Kostant 分拆函数。《数学物理学杂志》，25(8):2367–2373，1984 年。[36] Claude E Shannon。杂耍的科学方面，1993 年。[37] Philip Smith 和 Vitaliy Kurlin。具有相同或平凡一维持久性的有限度量空间的泛族。《应用与计算拓扑学杂志》，8(3):839–855，2024 年。[38] Elchanan Solomon、Alexander Wagner 和 Paul Bendich。从几何到拓扑：分布式持久性的逆定理。《计算几何杂志》，14(2):172–196，2023 年。[39] Sven Verdoolaege、Rachid Seghir、Kristof Beyls、Vincent Loechner 和 Maurice Bruynooghe。Ehrhart 多项式的解析计算：实现更多编译器分析与优化。载于《*2004* 年嵌入式系统编译器、体系结构与综合国际会议论文集》，第 248–258 页，2004 年。