arXiv:2603.23643v1 [math.FA] 24 Mar 2026

双 Lipschitz 不变量理论中的逼近定理

Jameson Cahill* Joseph W. Iverson† Dustin G. Mixon‡§ Nathan Willey‡

摘要

双利普希茨不变量理论研究将轨道空间低失真嵌入欧几里得空间的问题。迄今为止，仅在少数情况下已知具有最小可能畸变的嵌入，包括：(a) 平面旋转，(b) 实相位恢复，以及 (c) 有限反射群。在本文中，我们证明，对于这三种情况，最小可能畸变均可由“最大滤波器组”与线性变换的复合近似达到。我们的证明分为两个步骤：首先，我们证明只需展示某个利普希茨函数空间的包含关系；其次，我们针对这三种情况采用根本不同的方法证明该包含关系。我们还证明，这些情况与若干相关函数空间的相互作用各不相同，这表明统一处理并非易事。

1 引言

1.1 双利普希茨不变量理论

考虑有限群 $G \leq \mathsf{O}(d)$。其自然作用将欧几里得空间 $\mathbb{R}^d$ 划分为轨道$[x] := G \cdot x$，我们将其视为轨道空间中的点

$$\mathbb{R}^d/G := \{[x] : x \in \mathbb{R}^d\}.$$

赋予 $\mathbb{R}^d/G$ 商度量

$$d_{\mathbb{R}^d/G}([x],[y]) = \min_{g\in G} \|x - gy\|,$$

我们寻求一个 G-i nvariant 映射 $f\colon \mathbb{R}^d \to \mathbb{R}^n$，使诱导映射 $f^{\downarrow}\colon \mathbb{R}^d/G \to \mathbb{R}^n$具有以下性质
对度量的畸变最小 保留该度量。

我们的研究受到群不变信号处理 [6, 5] 和机器学习 [8], 等应用的启发。在这些应用中，映射到欧几里得空间的保度量特征映射使我们能够使用标准的欧几里得工具；例如，请参阅 [10]。由此催生了一套蓬勃发展的双利普希茨不变量 [16, 13, 14, 15], 理论，其中许多研究受到相位恢复领域思想的启发 [17, 1]。

具体而言，我们称 $f^{\downarrow}$ 是**双利普希茨连续的**，如果这些常数

$$\alpha(f^{\downarrow}) := \inf_{\substack{x,y\in\mathbb{R}^d\\ [x]\neq[y]}} \frac{\|f(x)-f(y)\|}{d_{\mathbb{R}^d/G}([x],[y])}, \qquad \beta(f^{\downarrow}) := \sup_{\substack{x,y\in\mathbb{R}^d\\ [x]\neq[y]}} \frac{\|f(x)-f(y)\|}{d_{\mathbb{R}^d/G}([x],[y])}$$

位于 $(0,\ \infty)$ 中。我们希望找到这样的不变量，使其**畸变**

$$\mathrm{dist}(f^{\downarrow}) := \frac{\beta(f^{\downarrow})}{\alpha(f^{\downarrow})}$$

尽可能小，其中我们将 $\mathrm{dist}(f^{\downarrow}) = \infty$解释为：当 $\alpha(f^{\downarrow}) = 0$或 $\beta(f^{\downarrow}) = \infty$时。注意，畸变总是至少为 1，且当且仅当 $f^{\downarrow}$ 是从$\mathbb{R}^d/G$到 $\mathbb{R}^n$乘以一个正标量的等距嵌入时取等号，即 $\alpha(f^{\downarrow}) = \beta(f^{\downarrow})$。如果我们还考虑 G-不变映射

*数学与统计系，北卡罗来纳大学威尔明顿分校，威尔明顿，北卡罗来纳州
†爱荷华州立大学数学系，艾姆斯，爱荷华州
‡俄亥俄州立大学数学系，哥伦布，俄亥俄州§俄亥俄州立大学转化数据分析研究所，哥伦布，俄亥俄州

到无限维希尔伯特空间中，那么最优畸变称为轨道空间 $\mathbb{R}^d/G$的**欧几里得畸变**:

$$c_2(\mathbb{R}^d/G) := \inf_{\substack{G\text{-inv } f\colon \mathbb{R}^d \to H \\ \text{Hilbert space } H}} \operatorname{dist}(f^{\downarrow}).$$

在某些特殊情况下，欧几里得畸变 $c_2(\mathbb{R}^d/G)$ 的精确值已为人所知，甚至也知道如何通过有限维目标空间 $\mathbb{R}^n$ 达到这一畸变。以下列举这一现象的一些重要实例。

例 1。

(a) 平面旋转。 假设 $d = 2$，并令 G 为 SO(2) 中阶为 $r > 1$的有限子群。那么$c_2(\mathbb{R}^2/G) = r \sin \frac{\pi}{2r}$。此外，若定义 $h\colon \mathbb{R}^2 \to \mathbb{R} \times \mathbb{C} \cong \mathbb{R}^3$，使得 $h(0) = 0$ 且

$$h(x) = \|x\| \cdot \left(\cos \tfrac{\pi}{2r},\ (\tfrac{x_1 + \mathrm{i}x_2}{\|x\|})^r \sin \tfrac{\pi}{2r} \right)$$

对于 $x \neq 0$，则 h 是 G-不变量，且 $\operatorname{dist}(h^{\downarrow}) = c_2(\mathbb{R}^2/G)$。（将 $\mathbb{R}^2$ 视为 $\mathbb{C}$后，这一点或许更直观，因为此时第二个坐标对应于将输入的相位提升到第 r次幂。）这一结果最初由 [10], 中的推论 38 和例 18 证明，随后又由 [7]中的定理16进一步推广。

(b) 实相位恢复。 假设 $G = \{\pm I\}$。那么 $c_2(\mathbb{R}^d/G) = \sqrt{2}$。此外，若我们根据算子平方根定义$h\colon \mathbb{R}^d \to \mathbb{R}^{d\times d}_{\mathrm{sym}} \cong \mathbb{R}^{\binom{d+1}{2}}$，则

$$h(x) = \sqrt{xx^{\top}},$$

那么 h 是 G-不变量，且 $\operatorname{dist}(h^{\downarrow}) = c_2(\mathbb{R}^d/G)$。这一结果最早由 [10], 中的推论36和例16确立，随后又由 [7]中的命题17进一步推广。

(c) 反射。 假设 $G \leq \mathrm{O}(d)$是一个有限反射群，也就是说，它是由 $\mathbb{R}^d$中经过原点的超平面上的反射生成的 $\mathrm{O}(d)$ 的有限子群。那么 $c_2(\mathbb{R}^d/G) = 1$。接下来， G中所有反射对应的超平面将 $\mathbb{R}^d$ 分割为称为**魏尔室**的多面锥。若我们固定 G 的一个闭魏尔室 $W \subseteq \mathbb{R}^d$，并定义 $h\colon \mathbb{R}^d \to \mathbb{R}^d$，则

$$h(x) \in [x] \cap W,$$

那么 h 是 G-不变量，且 $\operatorname{dist}(h^{\downarrow}) = c_2(\mathbb{R}^d/G)$。这一结果大概应当被视为众所周知的结论，但例如在 [19]的引理8中就出现过。

在上述三种情形中，我们所描述的最优双 Lipschitz 嵌入 h 具有一种显现出的性质，即**正齐次性**，这意味着

$$h(rx) = rh(x)$$

对于每个 $x \in \mathbb{R}^d$ 和 $r \geq 0$均成立。事实上，迄今为止已知的每一个 $\mathbb{R}^d/G$ 的最优双 Lipschitz 嵌入都表现出这一行为（至少在对目标空间进行平移，使得 $[0] \mapsto 0$之后如此）；关于这些结果的最新综述，参见 [7] 。我们目前还无法解释这一现象。

1.2 最大滤波与主要结果

与上述针对具体情形的研究不同，我们寻求一种更具通用性、能够构造低畸变欧几里得嵌入的方法。一类称为最大滤波的双利普希茨嵌入 [11] 在这一方向上取得了部分进展。具体而言，给定一个**模板** $y \in \mathbb{R}^d$，考虑由**最大滤波器** $\mathbb{R}^d \to \mathbb{R}$定义的

$$x \longmapsto \langle\!\langle [x], [y] \rangle\!\rangle := \max_{g \in G} \langle x, gy \rangle.$$

那么，对于 $n = 2d$一般位置的模板 $y_1, \ldots, y_n \in \mathbb{R}^d$，有

$$f\colon x \longmapsto \begin{bmatrix} \langle\!\langle [x], [y_1] \rangle\!\rangle \\ \vdots \\ \langle\!\langle [x], [y_n] \rangle\!\rangle \end{bmatrix}$$

确定一个双利普希茨 **最大滤波器组** $f^{\downarrow}$: $\mathbb{R}^d/G \to \mathbb{R}^n$ [11, 3]。（特别是，对于每个有限子群 $c_2($ $\mathbb{R}^d/G$，欧几里得畸变 $G \leq \mathrm{O}(d)$ 都是有限的。）关于随机模板下此类嵌入的畸变，有明确的一般界 [11, 20], 但人们认为这些界并不紧；在已知 $c_2(\mathbb{R}^d/G$的情形下，通常会存在差距。例如，在例1(b)中，[24] 的主要结果表明，最大滤波器组的下确界失真为

$$\sqrt{\tfrac{\pi}{\pi-2}} \approx 1.659 > 1.414 \approx \sqrt{2} = c_2(\mathbb{R}^d/G).$$

在例 1(c) 中，情况更加极端：其中 $c_2(\mathbb{R}^d/G) = 1$，但最优最大滤波器组的畸变可高达 $\Theta(d)$，具体取决于 G；参见 [19]。

在某些情况下，通过后合成一个线性映射即可立即消除这一差距。下面的例子对此作了明确说明。

例 2. 设 $G \leq \mathrm{O}(3)$ 是与 S_3 同构的反射群，它通过置换坐标作用于 $\mathbb{R}^3$。考虑以下模板

$$y_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad y_2 := \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad y_3 := \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

令 sort : $\mathbb{R}^3 \to \mathbb{R}^3$ 表示将输入坐标按非增序重新排列的函数。那么，使用上述模板进行最大滤波可表示为

$$\begin{bmatrix} \langle\!\langle [x], [y_1] \rangle\!\rangle \\ \langle\!\langle [x], [y_2] \rangle\!\rangle \\ \langle\!\langle [x], [y_3] \rangle\!\rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \mathrm{sort}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right).$$

将上述 3×3 矩阵的逆矩阵作用于等式两边，得到

$$\mathrm{sort}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \langle\!\langle [x], [y_1] \rangle\!\rangle \\ \langle\!\langle [x], [y_2] \rangle\!\rangle \\ \langle\!\langle [x], [y_3] \rangle\!\rangle \end{bmatrix}.$$

值得注意的是，$[x] \mapsto \mathrm{sort}(x)$ 是例 1(c) 中所述最优双 Lipschitz 嵌入的一个实例，其中 W 由 $\mathbb{R}^3$ 中坐标按非增序排列的向量组成。

示例 2 的关键结论是，至少在这个例子中：

> 最优双 *Lipschitz* 嵌入是最大滤波器组的一个线性变换。

或许令人惊讶的是，这一结论似乎在某种意义上可以推广。具体而言，最大滤波器组的线性变换的下确界失真似乎等于欧几里得畸变；也就是说，一般而言，通过后接一个精心选取的线性映射，就可以弥合最大滤波器组带来的差距。我们没有这一结论的一般性证明，但下面给出的是我们的主要结果，它聚焦于示例 1中的三类群。（稍后，我们也会给出其他群的数值结果，这些结果表明该结论或许适用于更一般的情形。）

定理 3 （主要结果）。固定一个群 $G \leq \mathrm{O}(d)$ 以及一个目标维数 n，使得

(a) $d = 2$, G 是 $\mathrm{SO}(2)$ 的非平凡有限子群，且 $n = 3$; *(b)* $G = \{\pm I\}$ 和 $n = \binom{d+1}{2}$，或者

(c) G 是一个有限反射群，且 $n = d$。

对于每个 $\epsilon > 0$，都存在 $m = m(\epsilon)$，以及一个最大滤波器组 Φ: $\mathbb{R}^d/G \to \mathbb{R}^m$ 和一个线性映射L: $\mathbb{R}^m \to \mathbb{R}^n$，使得复合映射 $L \circ \Phi$ 的畸变至多为 $c_2(\mathbb{R}^d/G) + \epsilon$。（事实上，在 *(iii)* 的情况下，可以取 $\epsilon = 0$ 和 $m = d$。）

请注意，在定理 3的每种情形中，n 的取值都与例 h 1中的最优嵌入的目标维数相匹配；事实上，我们的证明依赖于对 h的坐标函数进行逼近。对于非反射群情形，当 m 时，“隐藏层”的维数会任意增大 $\epsilon \to 0$。尤其是，最终线性层实现了显著的降维。我们并非首个通过在不变量后复合一个线性映射来降低目标维数的工作 [3, 4, 2], ，但我们的结果在性质上截然不同。先前的定理证明，尽管进行了降维，单射性仍得以保持，即“没有坏处”。而在这里，我们正是由于降维，才实现了近最优畸变：“大有裨益！”

1.3 Lipschitz 函数类与主要结果的证明

尽管定理 3 本质上探讨的是商空间的欧氏畸变，但我们对这一结果的证明归结为某些 Lipschitz 函数空间之间的包含关系，正如下文定理 5所述。在说明这两个定理之间的联系之前，我们先回顾一些有关 Lipschitz 连续性的基本概念。

设 X 和 Y 为度量空间，其中 X 的基数至少为 2。给定函数 $f\colon X \to Y$，我们定义

$$\alpha(f) := \inf_{\substack{x,y\in X\\ x\neq y}} \frac{d_Y(f(x),f(y))}{d_X(x,y)}, \qquad \beta(f) := \sup_{\substack{x,y\in X\\ x\neq y}} \frac{d_Y(f(x),f(y))}{d_X(x,y)}, \qquad \operatorname{dist}(f) := \frac{\beta(f)}{\alpha(f)},$$

与前文一样，当 $\alpha(f) = 0$ 或 $\beta(f) = \infty$ 时，我们将 $\operatorname{dist}(f) = \infty$ 解释为.....。若 f 满足 $\beta(f) < \infty$，则称其为 Lipschitz 连续函数。对于特定情形 $Y = \mathbb{R}$，Lipschitz 连续函数 $X \to \mathbb{R}$ 构成一个向量空间 $\mathrm{Lip}(X)$。在这个向量空间上，$\beta(\cdot)$ 是一个半范数；若将其视作范数，我们也记作 $\|\cdot\|_{\mathrm{Lip}}$。由中值定理可知，光滑函数 $\mathbb{R}^d \to \mathbb{R}$ 的 Lipschitz 范数等于其梯度的一致范数，这一现象也推广到 Lipschitz 连续函数（见下文命题 10）。注意，当且仅当 f 为常数时，$\|f\|_{\mathrm{Lip}} = 0$，因此 $\|\cdot\|_{\mathrm{Lip}}$ 在 $\mathrm{Lip}(X)$ 对常函数取商所得的商空间上诱导出一个范数。集合 $\mathcal{F} \subseteq \mathrm{Lip}(X)$ 的 Lipschitz 闭包由所有在 $\|\cdot\|_{\mathrm{Lip}}$ 中任意接近 $\mathcal{F}$ 中元素的函数组成。（特别地，对任意 $f \in \mathcal{F}$ 以及任意常函数 c，它都包含 $f + c$。）

根据以下引理（其证明见第 2 节），只要证明最优双 Lipschitz 嵌入的坐标函数属于最大滤波器张成空间的 Lipschitz 闭包，我们便可得出定理 3 。特别是，Lipschitz 范数上的接近意味着畸变也接近。重要的是，一致范数上的接近并不能保证这种性质，而一致范数正是通用逼近定理通常采用的设定。（此外，即使将范围限制为 G-不变的正齐次 Lipschitz 函数，最大滤波器的张成空间通常也不具备通用逼近性质；稍后会进一步讨论。）

引理 4. 设 X 是一个度量空间，G是作用于 X 上的一个群，其作用非传递、保持等距，且具有拓扑闭轨道。令 $\mathcal{F}$ 表示一组 G-不变的 *Lipschitz* 函数 $X \to \mathbb{R}$，并假设 $h\colon X \to \mathbb{R}^n$ 是一个非常值函数，其坐标函数属于 $\mathcal{F}$的 *Lipschitz* 闭包。（特别地，h 是 G-不变的。）那么，对于任意 $\epsilon > 0$，存在 $f\colon X \to \mathbb{R}^n$，其坐标函数属于 $\mathcal{F}$，使得畸变$(f^{\downarrow}) <$ 畸变$(h^{\downarrow}) + \epsilon$。

令 $\mathrm{Lip_h}(\mathbb{R}^d)$ 表示正齐次 Lipschitz 函数 $\mathbb{R}^d \to \mathbb{R}$ 构成的向量空间，其范数为 $\|\cdot\|_{\mathrm{Lip}}$。（这里，$\|\cdot\|_{\mathrm{Lip}}$ 是范数，而不只是半范数，因为 Lip 中不存在非零常数函数$\mathrm{h}(\mathbb{R}^d$。）这个赋范向量空间是完备的；事实上，它是满足 $f\colon \mathbb{R}^d \to \mathbb{R}$ 且 $f(0) = 0$的 Lipschitz 函数空间的闭子空间；根据 [23]中的命题 2.3(b)，后者是 Banach 空间。我们将看到，主要结果可由引理 4和以下结论推出：

$$h_1, \ldots, h_n \in \overline{\operatorname{span} \mathcal{M}}, \tag{1}$$

其中

- $h_1, \ldots, h_n\colon \mathbb{R}^d \to \mathbb{R}$ 是最优不变量 h的坐标函数，定义见例 1，
- $\mathcal{M}$ 是最大滤波函数集 $\langle\!\langle [\cdot],\ [y] \rangle\!\rangle\colon \mathbb{R}^d \to \mathbb{R}$，其中 $y \in \mathbb{R}^d$，并且
- 对于任意子集 $\mathcal{F} \subseteq \mathrm{Lip_h}(\mathbb{R}^d)$，$\overline{\mathcal{F}}$ 表示其在 $\mathrm{Lip_h}(\mathbb{R}^d)$ 中的拓扑闭包。

我们尚未找到一个统一的论证，来解释例 1中的三种情况为何都满足（1）。有趣的是，这些情况似乎有着本质上的差异，因为其中涉及的对象与其他值得关注的函数空间之间的相互作用取决于具体情况。为明确说明这一点，我们需要引入一些额外的记号：

- $\mathrm{PL_h}$ 是正齐次分段线性函数 $\mathbb{R}^d \to \mathbb{R}$ 的集合，
- $\mathcal{P}$ 是多项式函数 $S^{d-1} \to \mathbb{R}$ 的集合，其中 S^{d-1} 表示 $\mathbb{R}^d$ 中的单位球面，
- 给定一个 G-不变子集 $S \subseteq \mathbb{R}^d$ 和一个函数集 $\mathcal{F}$ $S \to \mathbb{R}$，集合 $\mathcal{F}^G$ 由G-不变的 $\mathcal{F}$中的元素组成，并且

- ext 是将函数 $S^{d-1} \to \mathbb{R}$ 扩展为正齐次函数 $\mathbb{R}^d \to \mathbb{R}$的**齐次延拓**映射，即对于 $r \geq 0$ 和 $u \in S^{d-1}$，$\mathrm{ext}(f)(ru) = r f(u)$。

在这种 记号下，我们将考察 $\mathrm{Lip_h}(\mathbb{R}^d)^G$ 中以下子空间之间的相互关系：

$$\mathrm{span}\{h_i\}_{i=1}^n, \qquad \mathrm{ext}\,\mathcal{P}^G, \qquad \mathrm{span}\,\mathcal{M}, \qquad \mathrm{PL}_{\mathrm{h}}^G .$$

定理 5. 固定输入维度 d 和群 $G \leq \mathrm{O}(d)$。

(a) 若 $d = 2$ 和 G 是 *SO(2)* 的非平凡有限子群，则

$$\mathrm{span}\{h_i\}_{i=1}^n \subsetneq \overline{\mathrm{ext}\,\mathcal{P}^G} \subsetneq \overline{\mathrm{span}\,\mathcal{M}} \qquad \textit{and} \qquad \mathrm{span}\,\mathcal{M} = \mathrm{PL}_{\mathrm{h}}^G .$$

(b) 若 $d > 2$ 和 $G = \{\pm I\}$，则

$$\mathrm{span}\{h_i\}_{i=1}^n \subsetneq \overline{\mathrm{ext}\,\mathcal{P}^G} \subsetneq \overline{\mathrm{span}\,\mathcal{M}} \subsetneq \overline{\mathrm{PL}_{\mathrm{h}}^G}.$$

(c) 若 G是有限反射群，则

$$\mathrm{span}\{h_i\}_{i=1}^n = \mathrm{span}\,\mathcal{M} = \overline{\mathrm{span}\,\mathcal{M}} \subsetneq \overline{\mathrm{PL}_{\mathrm{h}}^G}.$$

此外，若不存在被所有 G 固定的非零向量（即 *G* 是本质的），则有 t

$$\overline{\mathrm{ext}\,\mathcal{P}^G} \cap \overline{\mathrm{span}\,\mathcal{M}} = \{0\}.$$

鉴于 (a)、(b) 和 (c) 彼此差异很大，我们无法给出 (1) 的统一证明。特别是，(b) 和 (c) 表明，“最大滤波器的线性组合可以逼近所有 G-不变的正齐次分段线性函数”并非总是成立，因此我们无法转而研究这个“更简单”的函数类来解释这一现象。此外，(c) 说明，这种包含关系并非总能用“最大滤波器的线性组合可以逼近球面上的所有不变多项式”来解释。事实上，在 (c) 的情况下，最大滤波器张成的函数有限维子空间，因此不存在任何范数能使最大滤波器的张成空间普遍逼近 G-不变的正齐次 Lipschitz 连续函数。

我们将证 明第 5 节中的定理 4。目前，我们先说明它与引理 4 如何结合起来，从而得出定理 3：

定理 *3*的证明。 考虑例 G 中给出的适当的 -不变量正齐次映射 h: $\mathbb{R}^d \to \mathbb{R}^n$。重要的是，$\mathrm{dist}(h^{\downarrow}) = c_2(\mathbb{R}^d/G)$。令 $\{h_i\}_{i=1}^n$ 表示 h 的坐标函数。根据定理 5，每个坐标函数 h_i 都属于 $\mathcal{F}$ 的 Lipschitz 闭包：$=$ 张成空间 $\mathcal{M}$。根据引理 4，存在 f: $\mathbb{R}^d \to \mathbb{R}^n$，其坐标函数属于 $\mathcal{F}$，使得

$$\mathrm{dist}(f^{\downarrow}) < \mathrm{dist}(h^{\downarrow}) + \epsilon = c_2(\mathbb{R}^d/G) + \epsilon.$$

最后，由于 f 的每个坐标函数都是最大滤波器的有限线性组合，因此我们可以将$f^{\downarrow}$ 表示为最大滤波器组 Φ 与线性映射 L的复合。

对于 (c)，我们还可以取 $f = h$，此时 $\epsilon = 0$ 且 $m = d$。事实上，由于定理 5(c) 中的 span $\{h_i\}_{i=1}^n$ =span $\mathcal{M}$ ，可知每个 h_i 都属于 d 维最大滤波器的张成空间$\mathbb{R}^d \to \mathbb{R}$。因此，我们可以将 Φ 的坐标函数取为任意一组最大滤波器的基，然后通过将每个 $h_i^{\downarrow}$ 表示为该基的线性组合来确定 L 。 □

这些结果引出了一些问题，我们留待未来研究。

问题 6。

(a) 定理 3 是否对每个闭子群 $G \leq \mathrm{O}(d)$ 都成立? (b) 每个闭子群 $G \leq \mathrm{O}(d)$ 是否都存在一个有限的目标维数 n，使得存在一个G-不变映射 h: $\mathbb{R}^d \to \mathbb{R}^n$，满足 $\mathrm{dist}(h^{\downarrow}) = c_2(\mathbb{R}^d/G)$? (c) 若 (b) 的答案为“是”，能否令 h 正齐次? (d) 若以上问题的答案均为“是”，是否存在一个统一证明来证明（1）？

问题 7。 对于每个闭子群 $G \leq \mathrm{O}(d)$，刻画属于最大滤波器的张成空间之闭包的函数。

1.4 路线图

在下一节中，我们将证明引理 4。在例 1的情形 (a) 和 (b) 中，事实证明， h 的坐标函数可以表示为最大滤波器的整数线性组合，而这些最大滤波器又属于最大滤波器线性组合的 Lipschitz 连续闭包。我们将在第 3节证明这一点。第 4节专门用于证明定理 5，最后我们将在第 5节给出数值结果。

2 畸变的连续性及引理证明4

在证明引理 4之前，我们先证明几个更简单的引理，它们本身或许也有一定的独立价值；因此，我们以适当的概括程度陈述这些引理。

我们的第一个引理与奇异值的 *Weyl* 不等式有些相似，该不等式指出

$$|\sigma_i(A) - \sigma_i(B)| \le \sigma_1(A - B) \qquad \forall\, i,$$

其中 $\sigma_1(M) \ge \sigma_2(M) \ge \cdots$表示实矩阵 M的奇异值。考虑到高瘦矩阵 M 的最大和最小奇异值等于 $x \mapsto Mx$ 的最优上、下 Lipschitz 界（将其视为欧几里得空间之间的映射时），下面的引理恰当地推广了关于最大和最小奇异值的 Weyl 不等式。

引理 8（Lipschitz 映射的 Weyl 不等式）。给定一个至少含有两个点的度量空间 X 和一个赋范向量空间 V，则对于任意 *Lipschitz* 连续 f、g: $X \to V$，有

$$\max\Big\{\ |\alpha(f) - \alpha(g)|,\ |\beta(f) - \beta(g)|\ \Big\} \le \beta(f - g).$$

证明。令 d 表示 X 上的度量，令 $\|\cdot\|$ 表示 V 上的范数。给定任意 $x \ne y \in X$，记

$$\begin{aligned}\frac{\|f(x) - f(y)\|}{d(x,y)} &= \frac{\|g(x) - g(y) + f(x) - g(x) - f(y) + g(y)\|}{d(x,y)} \\ &= \frac{\|g(x) - g(y) + [f-g](x) - [f-g](y)\|}{d(x,y)}. \end{aligned} \tag{2}$$

由反三角不等式可得（2）的下界:

$$\frac{\|f(x) - f(y)\|}{d(x,y)} \ge \frac{\|g(x) - g(y)\|}{d(x,y)} - \frac{\|[f-g](x) - [f-g](y)\|}{d(x,y)} \ge \alpha(g) - \beta(f - g).$$

由此可得 $\alpha(f) \ge \alpha(g) - \beta(f-g)$。由对称性，还可得 $\alpha(g) \ge \alpha(f) - \beta(f-g)$，将二者结合可得

$$|\alpha(f) - \alpha(g)| \le \beta(f - g).$$

接下来，由三角不等式可得（2）的上界:

$$\frac{\|f(x) - f(y)\|}{d(x,y)} \le \frac{\|g(x) - g(y)\|}{d(x,y)} + \frac{\|[f-g](x) - [f-g](y)\|}{d(x,y)} \le \beta(g) + \beta(f - g).$$

由此可得 $\beta(f) \le \beta(g) + \beta(f-g)$。由对称性，还可得 $\beta(g) \le \beta(f) + \beta(f-g)$，将二者结合可得

$$|\beta(f) - \beta(g)| \le \beta(f - g).$$

□

接下来，我们将 G-不变函数的 Lipschitz 范数与相应函数在商空间上的 Lipschitz 范数联系起来。

引理 9。给定度量空间 X 和 Y，考虑任意一个群 G，它通过等距变换作用于 X，且作用非传递并具有拓扑闭轨道。对于任意 G-不变函数 f: $X \to Y$, $\beta(f^{\downarrow}) = \beta(f)$。

证明。要证明 $\beta(f) \le \beta(f^{\downarrow})$，只需证明

$$\frac{d_Y(f(x), f(y))}{d_X(x,y)} \le \beta(f^{\downarrow}) \qquad \forall\, x, y \in X,\ x \ne y.$$

显然，当 $[x]=[y]$ 时结论成立，因为此时 $f(x)=f(y)$。否则，我们有 $[x]\neq[y]$，此时不等式 $d_{X/G}([x],[y])\leq d_X(x,y)$ 蕴含

$$\frac{d_Y(f(x),f(y))}{d_X(x,y)}\leq\frac{d_Y(f(x),f(y))}{d_{X/G}([x],[y])}\leq\beta(f^{\downarrow}).$$

对于另一边的不等式，任取 $[x]\neq[y]$ in X/G。由于 G in X 的轨道是闭的，存在$g\in G$ 使得 $d_X(gx,y)=d_{X/G}([x],[y])\neq 0$，因此

$$\frac{d_Y(f^{\downarrow}([x]),f^{\downarrow}([y]))}{d_{X/G}([x],[y])}=\frac{d_Y(f(gx),f(y))}{d_X(gx,y)}\leq\beta(f).$$

由此可得 $\beta(f^{\downarrow})\leq\beta(f)$。 □

现在我们可以证明本小节的主要结果了。

引理 4的证明。 任取一个非常值函数 $h\colon X\to\mathbb{R}^n$，其坐标函数属于 $\mathcal{F}$的 Lipschitz 闭包，并固定 $\epsilon>0$。不失一般性，我们可以假设 dist$(h^{\downarrow})$ 有限；特别地，$0<\alpha(h^{\downarrow})\leq\beta(h^{\downarrow})<\infty$。对于每个 $i\in[n]$，令 $h_i\colon X\to\mathbb{R}$ 表示 i的第 h个坐标函数，并选取 $f_i\in\mathcal{F}$ 使得

$$\beta(f_i-h_i)<\frac{1}{\sqrt{n}}\cdot\min\left\{\frac{\epsilon\cdot\alpha(h^{\downarrow})}{1+\operatorname{dist}(h^{\downarrow})+\epsilon},\alpha(h^{\downarrow})\right\}.$$

令 $f\colon X\to\mathbb{R}^n$ 表示这样一个函数：其第 i个坐标函数为 f_i。（值得注意的是，h 和 f 都是G-不变量。）于是，由三角不等式以及引理 8 和 9 可得

$$\beta(f^{\downarrow})\leq\beta(h^{\downarrow})+|\beta(f^{\downarrow})-\beta(h^{\downarrow})|\leq\beta(h^{\downarrow})+\beta(f^{\downarrow}-h^{\downarrow})=\beta(h^{\downarrow})+\beta(f-h),$$

并且，同理，

$$\alpha(f^{\downarrow})\geq\alpha(h^{\downarrow})-|\alpha(f^{\downarrow})-\alpha(h^{\downarrow})|\geq\alpha(h^{\downarrow})-\beta(f^{\downarrow}-h^{\downarrow})=\alpha(h^{\downarrow})-\beta(f-h).$$

为了估计 $\beta(f-h)$，注意对于任意 $x\neq y$ in X，我们有

$$\frac{\|(f-h)(x)-(f-h)(y)\|}{d(x,y)}\leq\frac{\sqrt{n}\cdot\max_i|(f_i-h_i)(x)-(f_i-h_i)(y)|}{d(x,y)}\leq\sqrt{n}\cdot\max_i\beta(f_i-h_i),$$

其中 d 表示 X上的度量。因此，

$$\beta(f-h)\leq\sqrt{n}\cdot\max_i\beta(f_i-h_i)<\min\left\{\frac{\epsilon\cdot\alpha(h^{\downarrow})}{1+\operatorname{dist}(h^{\downarrow})+\epsilon},\alpha(h^{\downarrow})\right\}.$$

特别地，上述最小值中的第二项意味着 $0<\alpha(f^{\downarrow})\leq\beta(f^{\downarrow})<\infty$且$\alpha(h^{\downarrow})-\beta(f-h)>0$，而第一项意味着

$$\operatorname{dist}(f^{\downarrow})=\frac{\beta(f^{\downarrow})}{\alpha(f^{\downarrow})}\leq\frac{\beta(h^{\downarrow})+\beta(f-h)}{\alpha(h^{\downarrow})-\beta(f-h)}<\operatorname{dist}(h^{\downarrow})+\epsilon.$$ □

3 最大滤波器的积分组合

本节中，我们着重讨论例1中的情形 (a) 和 (b)。对于这两种情形，我们证明（1），也就是说，例 h_1, ..., h_n 1 中给出的最优坐标函数属于张成空间 $\mathcal{M}$，即最大滤波器张成空间的 Lipschitz 闭包。在下一小节（具体而言，在定理13中），我们将证明，只需把每个 h_i 表示为最大滤波器的积分组合即可。（这里的积分指的是 $\int$中的含义，而不是 $\mathbb{Z}$。）接着，在第3.2 节和第3.3 节中，我们运用这一思路，分别在例 h_i 1 的情形 (a) 和 (b) 中，将每个（事实上，球面上的每个不变多项式）表示为最大滤波器的积分组合。我们借助调和分析工具来完成这一工作。全文中，ω 表示 S^{d-1}上的 Haar 测度。

3.1 主要技术

本小节中，我们介绍本节后续部分将使用的主要技术。首先，我们陈述一个有用的泛化：光滑函数 $f\colon \mathbb{R}^d \to \mathbb{R}$ 的 Lipschitz 范数等于其梯度的 2- 范数的上确界。特别是，这使我们能够运用熟悉的微积分工具来估计 Lipschitz 范数。

命题 10 (强黎曼定理，参见 [23]中的定理 1.41)。 设 $f\colon \mathbb{R}^d \to \mathbb{R}$ 是 *Lipschitz* 函数。那么 f 几乎处处可微，并且

$$\|f\|_{\mathrm{Lip}} = \operatorname*{ess\,sup}_{x\in\mathbb{R}^d} \|\nabla f(x)\|.$$

在 $f\colon \mathbb{R}^d \to \mathbb{R}$ 是最大滤波器这一特殊情形下，命题 10 需要求出 f的梯度，下面的结果给出了这一梯度：

命题 11 (最大滤波器的梯度，参见 [10]中的定理 29)。 给定有限群 $G \leq \mathrm{O}(d)$ 和$y \in \mathbb{R}^d$，最大滤波器 $\langle\!\langle [\cdot], [y] \rangle\!\rangle\colon \mathbb{R}^d \to \mathbb{R}$ 在 $x \in \mathbb{R}^d$ 处可微，当且仅当存在唯一的 $z \in [y]$ 使得 $\langle x, z\rangle = \langle\!\langle [x], [y] \rangle\!\rangle$，此时

$$\nabla \langle\!\langle [\cdot], [y] \rangle\!\rangle(x) = z = \operatorname*{arg\,max}_{u\in[y]} \langle x, u\rangle.$$

接下来，为了确定最大滤波器积分组合的梯度，我们发现可以在积分号下求导，也就是说，所求梯度就是相应的最大滤波器梯度的积分组合：

引理 12. 给定有限群 $G \leq \mathrm{O}(d)$ 和 $q \in L^1(S^{d-1})$，考虑由以下方式定义的 $p\colon \mathbb{R}^d \to \mathbb{R}$

$$p(x) = \int_{S^{d-1}} q(y)\, \langle\!\langle [x], [y] \rangle\!\rangle\, d\omega(y).$$

A对于 G 中每个具有平凡稳定子的 $x \in \mathbb{R}^d$，都有 p 可微，其梯度为 nt

$$\nabla p(x) = \int_{S^{d-1}} q(y) \left(\operatorname*{arg\,max}_{u\in[y]} \langle x, u\rangle \right) d\omega(y).$$

特别地，对于几乎所有 $y \in S^{d-1}$，上述最大值点都是唯一的。

证明。首先，我们证明“特别地”这一部分。考虑由所有使 $\operatorname{arg\,max}_{u\in[y]} \langle x, u\rangle$ 不唯一的 $y \in S^{d-1}$ 构成的集合 B 。对于任意 $y \in B$，存在 $g, h \in G$，满足 $g \neq h$，且$\langle x, gy\rangle = \langle x, hy\rangle$，也就是说，$\langle g^{-1}x, y\rangle = \langle h^{-1}x, y\rangle$。由于 x 在 G 中显然具有平凡稳定子，因此 y 位于 $[x] \subseteq \mathbb{R}^d$ 的某个 Voronoi 单元的边界上。因此， B 包含于 S^{d-1} 与这些边界（有限个）的并集的交集中，而该交集在 S^{d-1} 中是零测集。

接下来，固定一个标准基元素 e_i 和一个非零实数序列 $\{s_n\}_{n=1}^{\infty}$，使得$s_n \to 0$。给定任意 $y \in S^{d-1} \setminus B$，记 $u(y) \in [y]$ 为使$\operatorname{arg\,max}_{u\in[y]} \langle x, u\rangle$达到最大值的最大值点；由于 $y \notin B$，该最大值点唯一。然后，考虑 $\nabla p(x)$的第 i个坐标，只需证明

$$\lim_{n\to\infty} \frac{p(x + s_n e_i) - p(x)}{s_n} = \int_{S^{d-1}} q(y)\, \langle e_i, u(y)\rangle\, d\omega(y),$$

或者等价地，将左侧重新表示后，

$$\lim_{n\to\infty} \int_{S^{d-1}} q(y)\, \frac{\langle\!\langle [x + s_n e_i], [y] \rangle\!\rangle - \langle\!\langle [x], [y] \rangle\!\rangle}{s_n}\, d\omega(y) = \int_{S^{d-1}} q(y)\, \langle e_i, u(y)\rangle\, d\omega(y).$$

我们将通过应用勒贝格控制收敛定理来证明上述恒等式。事实上，只需验证这些函数

$$f_n(y) := q(y)\, \frac{\langle\!\langle [x + s_n e_i], [y] \rangle\!\rangle - \langle\!\langle [x], [y] \rangle\!\rangle}{s_n}$$

具有以下性质：

(i) 对于每个 $y \in S^{d-1}$, $|f_n(y)| \le |q(y)|$，以及 (ii) 对于每个 $y \in S^{d-1} \setminus B$，$\lim_{n\to\infty} f_n(y) = q(y)\,\langle e_i,\ u(y)\rangle$。

对于 (i)，注意到 $x \mapsto \langle\!\langle [x],\ [y] \rangle\!\rangle$ 对任意 $y \in S^{d-1}$ 都是 1-Lipschitz 连续的，因此

$$|f_n(y)| = |q(y)| \cdot \frac{|\langle\!\langle [x + s_n e_i], [y] \rangle\!\rangle - \langle\!\langle [x], [y] \rangle\!\rangle|}{\|(x + s_n e_i) - x\|} \le |q(y)|.$$

对于 (ii)，任取 $y \in S^{d-1} \setminus B$，并回顾一下：$u(y)$ 在轨道 $[y]$上唯一地使线性泛函 $\langle x,\ \cdot\rangle$ 取最大值。在轨道 $[y]$ 对应的 Voronoi 单元中，x 位于包含 $u(y)$ 的开 Voronoi 单元 V内。当 n 足够大时，$x + s_n e_i$ 也位于 V内，此时

$$f_n(y) = q(y)\,\frac{\langle\!\langle [x + s_n e_i], [y] \rangle\!\rangle - \langle\!\langle [x], [y] \rangle\!\rangle}{s_n} = q(y)\,\frac{\langle x + s_n e_i, u(y)\rangle - \langle x, u(y)\rangle}{s_n} = q(y)\,\langle e_i, u(y)\rangle. \qquad \square$$

最后，我们证明本小节的主要结果，即（足够良好的）最大滤波器的整数线性组合属于最大滤波器张成空间的 Lipschitz 闭包：

定理 13。 假设 $G \le \mathrm{O}(d)$ 是有限的，并且给定一个 G-不变的正齐次函数$p\colon\ \mathbb{R}^d \to \mathbb{R}$，假设存在一个 ***Lipschitz*** 函数 $q\colon\ S^{d-1} \to \mathbb{R}$，使得

$$p(x) = \int_{S^{d-1}} q(y)\,\langle\!\langle [x], [y] \rangle\!\rangle\, d\omega(y) \tag{3}$$

对所有 $x \in S^{d-1}$ （因此由正齐次性可知，对所有 $x \in \mathbb{R}^d$）。于是 p 属于最大滤波器 $x \mapsto \langle\!\langle [x], [y] \rangle\!\rangle$ 的张成空间的 ***Lipschitz*** 闭包，其中 $y \in S^{d-1}$，即 $p \in$ 张成空间 $\mathcal{M}$。

证明。 任取 $\epsilon > 0$，并考虑一个划分 $S^{d-1} = \bigsqcup_{k=1}^{n} I_k$，使得每个 I_k 都是可测且连通的，弦距离直径至多为 ϵ。对每个 k，取一个样本点 $y_k \in I_k$。对每个 $x \in \mathbb{R}^d$，我们用黎曼和近似积分组合 $p(x)$：

$$\hat{p}(x) := \sum_{k=1}^{n} q(y_k)\,\langle\!\langle [x], [y_k] \rangle\!\rangle\, |I_k|,$$

其中我们将 $|\cdot| := \omega(\cdot)$ 记作简写。特别地，$\hat{}p\colon\ \mathbb{R}^d \to \mathbb{R}$ 属于张成空间 $\mathcal{M}$。只需证明$\|\hat{}p - p\|_{\mathrm{Lip}} = O(\epsilon)$。

注意，由于 q 连续（因此绝对可积），且对于每个单位向量 y，$x \mapsto \langle\!\langle [x],\ [y] \rangle\!\rangle$都是 1-Lipschitz 连续的，所以 p 是 Lipschitz 连续的：

$$|p(x) - p(x')| \le \int_{S^{d-1}} |q(y)|\,|\langle\!\langle [x], [y] \rangle\!\rangle - \langle\!\langle [x'], [y] \rangle\!\rangle|\, d\omega(y) \le \|q\|_{L^1(S^{d-1})} \cdot \|x - x'\|.$$

与此同时，$\hat{}p$是 Lipschitz 连续的，因为它是有限个（Lipschitz 连续的）最大滤波器的线性组合 $x \mapsto \langle\!\langle [x],\ [y_k] \rangle\!\rangle$。接着，命题 10 说明 p 和 $\hat{}p$ 几乎处处可微，而且，

$$\|\hat{p} - p\|_{\mathrm{Lip}} = \operatorname*{ess\,sup}_{x \in \mathbb{R}^d} \|\nabla \hat{p}(x) - \nabla p(x)\|. \tag{4}$$

为了找出一个集合，以便在其上估计上述本质上确界，考虑子集 $S \subseteq \mathbb{R}^d$，它由所有满足以下条件的点组成：对每个 $k \in [n]$，$\langle\!\langle [\cdot], [y_k] \rangle\!\rangle$ 都可微；并且这些点在 G中显然是平凡稳定的。我们断言，$\mathbb{R}^d$中几乎每个点都属于 S。事实上，只有当存在 $u \ne u'$ 属于 $[y_k]$，使得 $\langle z, u\rangle = \langle z, u'\rangle$（见命题 11）时，$\langle\!\langle [\cdot], [y_k] \rangle\!\rangle$ 才会在 $z \in \mathbb{R}^d$ 处不可微；此时，z 位于 $u - u'$的正交补中。因此，$\langle\!\langle [\cdot], [y_k] \rangle\!\rangle$ 的不可微点都落在有限个超平面的并集中，该并集是零测集。此外，被非恒等元素 $g \in G$ 稳定的点位于真子空间 $\ker(g - I)$中，该子空间是零测集。由于根据假设 G是有限的，因此 $\mathbb{R}^d$ 中几乎每个点在 G中都是平凡稳定的。断言得证。

现在任取 $x \in S$。记

$$v_x(y) := q(y)\,\operatorname*{arg\,max}_{u \in [y]} \langle x, u\rangle,$$

引理 12 和命题 11分别表明 p 和 $\hat{}p$ 在 x处均可微，其中

$$\nabla p(x) = \int_{S^{d-1}} v_x(y)\, d\omega(y), \qquad \nabla \hat{p}(x) := \sum_{k=1}^{n} v_x(y_k)\, |I_k|.$$

特别是，

$$\|\nabla \hat{p}(x) - \nabla p(x)\| = \left\| \sum_{k=1}^{n} \int_{I_k} \big(v_x(y) - v_x(y_k)\big)\, d\omega(y) \right\| \le \sum_{k=1}^{n} \int_{I_k} \big\|v_x(y) - v_x(y_k)\big\|\, d\omega(y). \tag{5}$$

我们将把（5）中的各项分为两类，并采用不同的策略进行估计。具体而言，令 $A \subseteq [n]$ 表示满足以下条件的指标集合： $k \in [n]$ 使得 I_k 位于 $\mathbb{R}^d$ 中相对于轨道 $[x]$的某个开 Voronoi 单元内；令 $B := [n] \setminus A$ 表示其余指标。我们将看到，（ $\sum_{k\in A}$ 5）中的部分较小，因为 q 是 Lipschitz 连续的，且每个 I_k 的直径都很小；而 $\sum_{k\in B}$ 部分较小，因为 q 有界，且 $\bigcup_{k\in B} I_k$ 的测度很小。

设 $k \in A$，并令 V表示 $\mathbb{R}^d$ 中相对于轨道 $[x]$ 、且包含I_k的开 Voronoi 单元。存在群元素 $g \in G$ ，使得对于每个 $y \in I_k \subseteq V$，都有 gy 位于包含 x的 Voronoi 单元内。因此， $v_x(y) = q(y)gy$，于是

$$\begin{aligned} \|v_x(y) - v_x(y_k)\| &= \|q(y)gy - q(y_k)gy_k\| \\ &\le \|q(y)gy - q(y_k)gy\| + \|q(y_k)gy - q(y_k)gy_k\| \\ &\le |q(y) - q(y_k)| + |q(y_k)|\|y - y_k\| \\ &\le \|q\|_{\mathrm{Lip}} \cdot \epsilon + \|q\|_\infty \cdot \epsilon. \end{aligned}$$

由此可得

$$\sum_{k\in A} \int_{I_k} \big\|v_x(y) - v_x(y_k)\big\|\, d\omega(y) \le \sum_{k\in A} |I_k| \cdot \big(\|q\|_{\mathrm{Lip}} + \|q\|_\infty\big) \cdot \epsilon \le \omega_{d-1} \cdot \big(\|q\|_{\mathrm{Lip}} + \|q\|_\infty\big) \cdot \epsilon,$$

其中 ω_{d-1} 表示 S^{d-1}的表面积。

对于 $k \in B$ 情形，注意到

$$\sum_{k\in B} \int_{I_k} \big\|v_x(y) - v_x(y_k)\big\|\, d\omega(y) \le \left|\bigcup_{k\in B} I_k\right| \cdot \sup_{\substack{k\in[n] \\ y\in S^{d-1}}} \|v_x(y) - v_x(y_k)\| \le \left|\bigcup_{k\in B} I_k\right| \cdot 2\|q\|_\infty,$$

其 余下的一步由三角不等式得出：对于 $k \in [n]$ 和 $y \in S^{d-1}$，有

$$\|v_x(y) - v_x(y_k)\| \le \|v_x(y)\| + \|v_x(y_k)\| \le 2\|q\|_\infty.$$

接下来只需估计 $\bigcup_{k\in B} I_k$的测度；为此，我们通过估计一个合适的超集的测度来完成。

为此，考虑任意 $k \in B$。根据假设， I_k 是连通的，因此它与构成相对于 $[x]$的 Voronoi 单元各个面的线性超平面 $\{H_j\}_{j=1}^m$ 中的某一个有非平凡交集。（值得注意的是，无论 x如何， $m \le \binom{|G|}{2}$都成立。）特别地，根据假设，由于 I_k 的直径至多为 ϵ ，它位于某个 H_j的 ϵ加厚版本中。同时，这个加厚版本 H_j 的球面测度可通过余面积公式，先对平行于 H_j 的“纬线”积分得到：

$$\big|(H_j + B_0(\epsilon)) \cap S^{d-1}\big| = \omega_{d-2} \int_{-\epsilon}^{\epsilon} (1 - t^2)^{\frac{d-3}{2}}\, dt \le 2\epsilon\omega_{d-2},$$

其中， $\omega_{d-2} > 0$ 量化了赤道 $S^{d-2} \cong H_j \cap S^{d-1} \subseteq S^{d-1}$ 的贡献；特别是，它不依赖于 j。因此

$$\left|\bigcup_{k\in B} I_k\right| \le \sum_{j=1}^{m} \big|(H_j + B_0(\epsilon)) \cap S^{d-1}\big| \le \binom{|G|}{2} \cdot 2\epsilon\omega_{d-2}.$$

综上所述，我们得到

$$\begin{aligned}\|\nabla\hat{p}(x)-\nabla p(x)\| &\le \sum_{k\in A}\int_{I_k}\big\|v_x(y)-v_x(y_k)\big\|\,d\omega(y)+\sum_{k\in B}\int_{I_k}\big\|v_x(y)-v_x(y_k)\big\|\,d\omega(y)\\ &\le \Big(|S^{d-1}|\cdot\big(\|q\|_{\mathrm{Lip}}+\|q\|_\infty\big)+\tbinom{|G|}{2}\cdot 2\omega_{d-2}\cdot 2\|q\|_\infty\Big)\cdot\epsilon.\end{aligned}$$

由于 $x\in S$ 是任意的，由（4）可得 $\|\hat{}p-p\|_{\mathrm{Lip}}=O(\epsilon)$，证毕。 □

3.2 应用 1：平面旋转

本小节中，我们在1）的特殊情形下证明（ h_1，……， h_n 由例1(a) 给出。根据定理13，只需将每个 h_i 表示为最大滤波器的适当良好的积分组合。考虑到一些微妙之处（我们稍后会处理），这可由下述结论推出，因为每个 h_i 在G上限制时都是一个不变多项式： S^1

定理 14。 取一个非平凡有限子群 $G\le\mathrm{SO}(2)$。对于每个 G-不变多项式 p，存在一个 G-不变多项式 q，使得

$$p(x)=\int_{S^1}q(y)\,\langle\!\langle [x],[y]\rangle\!\rangle\,d\omega(y)$$

对所有 $x\in S^1$均成立。

证明。 不失一般性，取 ω 为 S^1上的均匀概率测度。记$r:=|G|>1$，并令 $x=(\cos\theta,\ \sin\theta)$，则 $\langle\!\langle [x],[e_1]\rangle\!\rangle$ 由函数 $f\colon\mathbb{R}\to\mathbb{R}$ 表示，该函数定义为 $f(\theta)=\cos\theta$，其中 $\theta\in[-\frac{\pi}{r},\frac{\pi}{r}]$，并以 $\frac{2\pi}{r}$为周期延拓。值得注意的是，若令 $y=(\cos\varphi,\ \sin\varphi)$，则 $\langle\!\langle [x],\ [y]\rangle\!\rangle=f(\theta-\varphi)$。接下来， $g(\theta)=p(\cos\theta,\ \sin\theta)$定义了一个$\frac{2\pi}{r}$-周期三角多项式$g\colon\mathbb{R}\to\mathbb{R}$，我们要寻找一个 $\frac{2\pi}{r}$-周期三角多项式 $c\colon\mathbb{R}\to\mathbb{R}$，使得

$$g(\theta)=\frac{1}{2\pi}\int_0^{2\pi}c(\varphi)\,f(\theta-\varphi)\,d\varphi\qquad\forall\,\theta\in\mathbb{R}.$$

我们的任务在傅里叶域中会简化：

$$\hat{g}(k)=\hat{c}(k)\cdot\hat{f}(k)\qquad\forall\,k\in\mathbb{Z}.$$

这提示我们可以如下选择 c:

$$\hat{c}(k)=\begin{cases}\dfrac{\hat{g}(k)}{\hat{f}(k)} & \text{if } \hat{f}(k)\neq 0\\ 0 & \text{else.}\end{cases}\tag{6}$$

ˆ 这一选择成立，因为 $f(k)\neq 0$ 当且仅当 r 整除 k。事实上，“仅当”方向可由 f 是 $\frac{2\pi}{r}$-周期函数这一事实推出；而“当”方向则可通过计算傅里叶级数（需要进行一些繁琐的积分，并借助积化和差公式）得到：

$$\hat{f}(rm)=(-1)^{m+1}\cdot\frac{\frac{r}{\pi}\sin(\frac{\pi}{r})}{(rm)^2-1}\qquad\forall\,m\in\mathbb{Z}.$$ □

为了说明定理 14，我们将其应用于最简单的情形，以表示最优坐标函数 h_1， h_2和 h_3，其中 G 的阶为 2。（从历史上看，这个例子标志着本文的开端；借助计算机实验，我们成功地将最优坐标函数近似表示为最大滤波器的线性组合，而由此出现的神秘积分恒等式促使我们进一步探究。）

例 15. 在 $|G|=2$ 的特殊情形下，即 $G=\{\pm I\}$， $\mathbb{R}^2/G$ 的最小畸变欧几里得嵌入由下式的齐次延拓给出

$$x\mapsto xx^\top=\begin{bmatrix}x_1^2 & x_1x_2\\ x_1x_2 & x_2^2\end{bmatrix}=\begin{bmatrix}\cos^2\theta & \cos\theta\sin\theta\\ \cos\theta\sin\theta & \sin^2\theta\end{bmatrix}$$

当 $x = \left[\overset{\theta}{\text{余弦}\theta}\right]_{\text{正弦}}$ 满足 $\theta \in [0, 2\pi)$。(参见例 1(b)，并注意，在这种 $d = 2$ 情形下，该嵌入可视为与例 1(a) 中的嵌入等距等价。) 我们将把这些坐标函数表示为最大滤波器的整数线性组合。在这种情况下，$\left[\overset{\theta}{\text{余弦}\theta}\right]_{\text{正弦}}$ 与模板 $\left[\overset{\varphi}{\text{余弦}\varphi}\right]_{\text{正弦}}$ 对应的最大滤波器为

$$f(\theta - \varphi) := |\cos(\theta - \varphi)|.$$

这与上文定理的证明中的记号一致 14，因为 $f\colon \mathbb{R} \to \mathbb{R}$ 是 π-周期的，且 $f(\theta) = \cos\theta$ 适用于 $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$。我们按照方法（6）推导出每个坐标函数（证明中的 c）对应的系数函数。首先，我们利用标准三角恒等式得到傅里叶级数

$$\begin{aligned}
\cos^2\theta &= \tfrac{1}{2} + \tfrac{1}{4}e^{i2\theta} + \tfrac{1}{4}e^{-i2\theta},\\
\sin^2\theta &= \tfrac{1}{2} - \tfrac{1}{4}e^{i2\theta} - \tfrac{1}{4}e^{-i2\theta},\\
\cos\theta\sin\theta &= \quad - \tfrac{i}{4}e^{i2\theta} + \tfrac{i}{4}e^{-i2\theta}.
\end{aligned}$$

与此同时，定理的证明中最后一个展示式 14 给出了 f 的傅里叶级数：

$$f(\theta) = \frac{2}{\pi}\sum_{m=-\infty}^{\infty}\frac{(-1)^m}{1-4m^2}\cdot e^{i2m\theta}.$$

令 c 表示 $\cos^2\theta$ 的系数函数，我们应用公式（6）得到 c 的傅里叶级数：

$$c(\theta) = \sum_{k=-\infty}^{\infty}\hat{c}(k)\cdot e^{ik\theta} = \tfrac{\pi}{4} + \tfrac{3\pi}{8}e^{i2\theta} + \tfrac{3\pi}{8}e^{-i2\theta} = \tfrac{\pi}{4} + \tfrac{3\pi}{4}\cos 2\theta,$$

由此得到这个（非平凡的！）恒等式

$$\cos^2\theta = \frac{1}{2\pi}\int_0^{2\pi} c(\varphi)f(\theta-\varphi)d\varphi = \int_0^{\pi}\left(\tfrac{1}{4} + \tfrac{3}{4}\cos 2\varphi\right)|\cos(\theta-\varphi)|\,d\varphi.$$

类似的计算可得其他坐标函数的表达式：

$$\begin{aligned}
\sin^2\theta &= \int_0^{\pi}\left(\tfrac{1}{4} - \tfrac{3}{4}\cos 2\varphi\right)|\cos(\theta-\varphi)|\,d\varphi,\\
\cos\theta\sin\theta &= \int_0^{\pi}\tfrac{3}{4}\sin 2\varphi\,|\cos(\theta-\varphi)|\,d\varphi.
\end{aligned}$$

3.3 应用 2：实相位恢复

接下来，我们在特殊情形下证明（1），其中 $h_1, \dots, h_n$ 由例 1(b) 给出。同样，根据定理 13，只需将每个 h_i 表示为最大滤波器的适当良好的积分组合。由于 $G = \{\pm I\}$，最大滤波器 $\langle\!\langle [x], [y] \rangle\!\rangle$ 就是 $|\langle x, y\rangle|$。与之前一样，将每个 h_i 限制在球面上时，它都是一个 G-不变（即偶）多项式。下面给出本小节的主要结果：

定理16. 取 $d > 2$ 和 $G = \{\pm I\} \le \mathrm{O}(d)$。对于任意 G-不变（即偶）多项式 p，存在一个 G-不变多项式 q，使得

$$p(x) = \int_{S^{d-1}} q(y)\,|\langle x, y\rangle|\,d\omega(y)$$

对所有 $x \in S^{d-1}$ 均成立。

我们对定理16 的证明沿用了我们证明定理 14时的相同思路，但在这种情况下，正确的傅里叶级数概念由球谐函数理论给出；我们推荐参阅 Stein 和 Weiss 的第四章第二节 [22]。我们始终假设 $d > 2$，以避免处理性质截然不同的球面 S^1；后者则由定理14涵盖。回顾分解

$$L^2(S^{d-1}) = \bigoplus_{k=0}^{\infty}\mathcal{H}_k,$$

其中， $\mathcal{H}_k$ 表示 S^{d-1}上次数为 k 的齐次调和多项式函数子空间。由于 $\mathcal{H}_k$ 是有限维的，对于每个 $x \in S^{d-1}$，由 $p \mapsto p(x)$定义的求值映射 $\mathcal{H}_k \to \mathbb{C}$ 都可以表示为与某个 $Z_x^{(k)} \in \mathcal{H}_k$的内积：

$$p(x) = \langle p, Z_x^{(k)} \rangle_{L^2(S^{d-1})} \tag{7}$$

这个表示函数 $Z_x^{(k)}$称为次数为 k 、极点为 x的带极球谐函数。事实证明， $Z_x^{(k)}$是实值的，而且对于每个 $g \in \mathrm{SO}(d)$，都有 $Z_{gx}^{(k)}(gy) = Z_x^{(k)}(y)$。在相差一个比例因子的意义下， $x \mapsto Z_x^{(k)}$是唯一满足这种 $\mathrm{SO}(d)$不变性的函数 $S^{d-1} \to \mathcal{H}_k$：

命题17 (第四章 [22, 中的推论2.13])。 固定 $d > 2$ 和 $k \geq 0$，并考虑任意一个F： $S^{d-1} \to \mathcal{H}_k$ 满足

$$F(gx)(gy) = F(x)(y) \qquad \forall\, x, y \in S^{d-1},\ g \in \mathrm{SO}(d).$$

那么存在 $c \in \mathbb{C}$，使得对于每个 $x \in S^{d-1}$，都有 $F(x) = cZ_x^{(k)}$。

定理 16的证明。 首先注意，对于每个非负整数 k， $\mathcal{H}_k$ 是 d 的 SO 不变子空间，而 $L^2(S^{d-1})$ 的正交投影映射 π_k： $L^2(S^{d-1}) \to \mathcal{H}_k$ 是 $\mathrm{SO}(d)$-等变的。因此，对于每个 $x \in S^{d-1}$，函数 $F_k(x)$: $= \pi_k\big(|\langle x, \cdot \rangle|\big) \in \mathcal{H}_k$ 满足

$$F_k(gx)(gy) = \pi_k\big(|\langle gx, \cdot \rangle|\big)(gy) = \big(g^{-1} \cdot \pi_k\big(|\langle gx, \cdot \rangle|\big)\big)(y) = \pi_k\big(|\langle gx, g\,\cdot \rangle|\big)(y) = \pi_k\big(|\langle x, \cdot \rangle|\big)(y) = F_k(x)(y)$$

对每个 $y \in S^{d-1}$ 和 $g \in \mathrm{SO}(d)$。根据命题 17，存在一个序列 $\{c_k\}_{k=0}^{\infty}$，使得对每个 $x \in S^{d-1}$，

$$|\langle x, \cdot \rangle| = \sum_{k=0}^{\infty} \pi_k\big(|\langle x, \cdot \rangle|\big) = \sum_{k=0}^{\infty} F_k(x) = \sum_{k=0}^{\infty} c_k Z_x^{(k)},$$

在 $L^2(S^{d-1})$ 中收敛。此外，每个 c_k 都是实数，因为函数 $|\langle x, \cdot \rangle|$ 以及两两正交的函数 $Z_x^{(k)}$都取实值。

接下来，给定一个偶多项式函数 p，我们可以将 $p = \sum_k p_k$分解，其中每个 p_k 都属于 $\mathcal{H}_k$。由于这是多项式函数之间的等式，因此该和是有限的。此外，由于 p 是偶函数，因此当 k 为奇数时，p_k 为零。因此，不失一般性，我们可以假设 $p \in \mathcal{H}_k$ 且 $k = 2m$ 。于是，对每个 $x \in S^{d-1}$，都有

$$\int_{S^{d-1}} p(y)\, |\langle x, y \rangle|\, d\omega(y) = \big\langle p, |\langle x, \cdot \rangle| \big\rangle_{L^2(S^{d-1})} = \langle p, c_k Z_x^{(k)} \rangle_{L^2(S^{d-1})} = c_k \langle p, Z_x^{(k)} \rangle_{L^2(S^{d-1})} = c_k p(x),$$

其中最后一步用到了再生性质（7）。因此，我们可以取 q： $= p/c_k$，前提是证明 $c_k \neq 0$。这是证明余下部分的重点。

固定某个 $x \in S^{d-1}$。由于 $c_k Z_x^{(k)}$是 $|\langle x, \cdot \rangle|$ 在 $\mathcal{H}_k$上的投影，因此它也是$|\langle x, \cdot \rangle|$ 在 $Z_x^{(k)}$的张成空间上的投影，于是

$$c_k = \frac{\big\langle |\langle x, \cdot \rangle|, Z_x^{(k)} \big\rangle_{L^2(S^{d-1})}}{\|Z_x^{(k)}\|^2_{L^2(S^{d-1})}} = \Big\langle |\langle x, \cdot \rangle|, \tfrac{1}{Z_x^{(k)}(x)} Z_x^{(k)} \Big\rangle_{L^2(S^{d-1})},$$

其中最后一步由再生性质（7）得出，即 $\|Z_x^{(k)}\|^2_{L^2(S^{d-1})} = Z_x^{(k)}(x)$。由于 $|\langle x, \cdot \rangle|$ 和 $\frac{1}{Z_x^{(k)}(x)} Z_x^{(k)}$在与极点 x正交的“纬线”上均为常数，应用光滑余面积公式 后，所求内积可化为对 $y \in S^{d-1}$的 x分量 $t \in [-1, 1]$ 的积分：

$$c_k = \omega_{d-2} \int_{-1}^{1} |t|\, G_k(t)\, (1 - t^2)^{\frac{d-3}{2}}\, dt,$$

其中， $\omega_{d-2} > 0$ 表示赤道 S^{d-2}的适当贡献测度， G_k 表示次数为 k的 *Gegenbauer* 多项式； 参见 [22, 第四章] 中的定理 2.14。序列 $\{G_j\}_{j=0}^{\infty}$由以下方式唯一确定：迭代地令 G_j 为任意一个次数为 j 的实值多项式，且

它关于 $[-1,1]$上的测度 $(1-t^2)^{\frac{d-3}{2}}\,dt$ 与 $\{G_i\}_{i<j}$ 正交，然后归一化，使得$G_j(1)=1$。

最后，我们利用 $k=2m$ 是偶函数这一事实得出结论。由于 G_{2m} 是偶函数，我们可以写成

$$c_{2m}=\omega_{d-2}\int_0^1 G_{2m}(t)\,(1-t^2)^{\frac{d-3}{2}}\,2t\,dt.$$

$m=0$ 的情形很简单：$G_0(t)=1$，因此 $c_0>0$。现在假设 $m\geq 1$。令 $\alpha:=\frac{d-2}{2}$，应用罗德里格公式：

$$G_{2m}(t)\,(1-t^2)^{\alpha-\frac{1}{2}}=\frac{\Gamma(\alpha+\frac{1}{2})}{4^m\Gamma(\alpha+2m+\frac{1}{2})}\cdot F^{(2m)}(t),\qquad F(t):=(1-t^2)^{2m+\alpha-\frac{1}{2}}.$$

那么 c_{2m} 是下列积分的正倍数，而该积分可通过分部积分求解：

$$\int_0^1 t\,F^{(2m)}(t)\,dt=F^{(2m-1)}(1)-F^{(2m-2)}(1)+F^{(2m-2)}(0).$$

由于 $d\geq 3$，对于每个 $k\leq 2m-1$，$F(t)$ 的 k 阶导数都含有 $1-t^2$ 的正次幂。特别地，$F^{(2m-1)}(1)=F^{(2m-2)}(1)=0$。接下来，定义 $f(u):=(1-u)^{2m+\alpha-\frac{1}{2}}$，则

$$F'(t)=2t\,f'(t^2),\qquad F''(t)=(2t)^2\,f''(t^2)+2f'(t^2).$$

继续归纳，我们便得到

$$F^{(2k)}(0)=\frac{(2k)!}{k!}f^{(k)}(0)$$

对于每个 $k\leq m-1$。令 $k=m-1$ 和 $s:=2m+\alpha-\frac{1}{2}$，则 c_{2m} 是一个正倍数 o f

$$f^{(m-1)}(0)=(-1)^{m-1}s(s-1)\cdots(s-m+2)\neq 0.$$

也就是说，$c_{2m}\neq 0$。 □

4 C函数空间的包含关系与定理的证明 5

本节旨在证明定理 5。我们将在下文不同的小节中分别证明这一结果的各个部分。为了证明函数空间的真包含关系，我们会用到可微性。以下关于在原点处取零的 Lipschitz 函数空间 $\mathrm{Lip}_0(\mathbb{R}^d)$ 的一个（标准）结果使这一点成为可能 $\mathbb{R}^d\to\mathbb{R}$：

命题 18。 假设 $\mathcal{F}\subset\mathrm{Lip}_0(\mathbb{R}^d)$ 中的函数在某个开集 $U\subseteq\mathbb{R}^d$上连续可微。那么，$\mathcal{F}$ 中的每个元素在 U上也连续可微。

证明略。固定 $f\in\mathcal{F}$。则存在一个序列 $\{f_n\}_{n=1}^\infty$，其中各项属于 $\mathcal{F}$，且满足 $\|f_n-f\|_{\mathrm{Lip}}\to 0$。特别地，$\{f_n\}_{n=1}^\infty$ 在 $\mathrm{Lip}_0(\mathbb{R}^d)$ 中是柯西序列。考虑不等式

$$\sup_{x\in U}\|\nabla f_m(x)-\nabla f_n(x)\|\leq\|f_m-f_n\|_{\mathrm{Lip}}\qquad\forall\,m,n\in\mathbb{N},$$

这既可以作为命题 10 的推论得到，也可以通过涉及中值定理的直接论证得到。这个不等式说明，$\{\nabla f_n|_U\}_{n=1}^\infty$ 在一致有界函数空间 $U\to\mathbb{R}^d$ 中关于一致范数构成柯西列，也就是说，$\{\nabla f_n|_U\}_{n=1}^\infty$ 一致收敛于某个g：$U\to\mathbb{R}^d$。由于连续函数的一致极限仍然连续，因此 g 是连续的。还需证明 f 在每个 $x\in U$ 处均可微，且 $\nabla f(x)=g(x)$，这是一道标准练习（相关论证示例见 [21] 中的定理 7.17）。 □

接下来，我们推导空间 PL_h 正齐次分段线性函数 $\mathbb{R}^d\to\mathbb{R}$的重要结构性质。（这些性质在文献中基本上已有论述，但为方便起见，我们在此重述。）对于每个 $f\in\mathrm{PL}_\mathrm{h}$，都存在一个有限的镶嵌 $\mathcal{T}$，其 $\mathbb{R}^d$（意 $\mathcal{T}$由覆盖 $\mathbb{R}^d$ 且内部两两不相交的 $\mathbb{R}^d$ 子集组成，并且每个成员

$\mathcal{T}$ 等于其内部的闭包），使得 ∇f 的水平集恰好是 $\mathcal{T}$ 各成员的内部。由于 f 是正齐次的，每个 $C \in \mathcal{T}$ 都在正标量乘法下封闭；此外，对于每个 $C \in \mathcal{T}$，都存在 $v_C \in \mathbb{R}^d$（即 f 在 $\operatorname{int} C$ 上的常梯度），使得 $f(x) = \langle x, v_C \rangle$ 对于所有 $x \in \operatorname{int} C$ 都成立。最后，由于 f 连续，决定 f 的元组 $(v_C)_{C\in\mathcal{T}} \in (\mathbb{R}^d)^{\mathcal{T}}$ 必然满足线性约束

$$\langle x, v_C \rangle = \langle x, v_{C'} \rangle \qquad \forall C, C' \in \mathcal{T},\, C \neq C',\, x \in C \cap C'.$$

给定 $C,\ C' \in \mathcal{T}$ 且满足 $C \neq C'$，由上述约束可知，$C \cap C'$ 包含于超平面 $\operatorname{span}\{v_C - v_{C'}\}^\perp$ 中。由于 $\mathcal{T}$ 中的集合是闭集，且覆盖 $\mathbb{R}^d$ 并且内部互不相交，因此每个 $C \in \mathcal{T}$ 的边界点必然位于某个其他 $C' \in \mathcal{T}$ 中。因此，

$$\bigcup_{C\in\mathcal{T}} \partial C = \bigcup_{\substack{C,C'\in\mathcal{T}\\ C\neq C'}} C \cap C' \subseteq \bigcup_{\substack{C,C'\in\mathcal{T}\\ C\neq C'}} \operatorname{span}\{v_C - v_{C'}\}^\perp.$$

特别地，右侧超平面的并集的集合补集是若干开区域的不交并，且 f 在每个这样的开区域上具有常梯度。

给定一个由 $\mathcal{H}$ 个线性超平面构成的有限集，令 $\mathrm{PL_h}(\mathcal{H})$ 表示 $\mathrm{PL_h}$ 中在由 $\mathcal{H}$ 诱导的每个开区域上具有常梯度的函数所构成的子空间。那么，由上述归约可得，

$$\mathrm{PL_h} = \bigcup_{\mathcal{H}} \mathrm{PL_h}(\mathcal{H}), \tag{8}$$

其中并集取遍所有由线性超平面构成的有限集。有趣的是，这些子空间构成一个格，因为 $\mathrm{PL_h}(\mathcal{H}) \leq \mathrm{PL_h}(\mathcal{H}')$ 等价于 $\mathcal{H} \subseteq \mathcal{H}'$。若令 $\mathcal{T}(\mathcal{H})$ 表示由 $\mathcal{H}$诱导的闭区域镶嵌，则 $\mathrm{PL_h}(\mathcal{H})$ 是有限维的，这是因为存在单射线性映射 $\operatorname{grad} : \mathrm{PL_h}(\mathcal{H}) \to (\mathbb{R}^d)^{\mathcal{T}(\mathcal{H})}$，其定义为：对于任意 $x \in \operatorname{int} C$，$\operatorname{grad} f(C) = \nabla f(x)$。定义一个顶点集为 $\mathcal{T}(\mathcal{H})$ 的图，并在 $C \cap C'$ 的余维为 1 时令 $C \leftrightarrow C'$ 相邻，再令

$$\kappa(f) := \max_{C\leftrightarrow C'} \|\operatorname{grad} f(C) - \operatorname{grad} f(C')\| \qquad \forall f \in \mathrm{PL_h}(\mathcal{H}).$$

由于图是连通的，因此 κ 恰好在子空间 $(\mathbb{R}^d)^* \leq \mathrm{PL_h}(\mathcal{H})$ 上消失，该子空间由线性泛函构成；并且它在商空间 $\mathrm{PL_h}(\mathcal{H})/(\mathbb{R}^d)^*$ 上定义了一个范数。从某种意义上说，$\kappa(f)$量化了关于 f 的二阶导数信息。此外，量 $\kappa(f)$ 与 $\mathcal{H}$ 的选择无关，因为将 f 视为某个 $\mathcal{H}' \supseteq \mathcal{H}$ 对应的 $\mathrm{PL_h}(\mathcal{H}')$ 的成员，得到的 $\kappa(f)$ 值相同。建立了这些预备内容后，我们现在开始证明定理 5。

4.1 定理 5(a) 的证明：平面旋转

对于这一部分 定理 5，只需综合前几节建立的理论即可 .

定理 5(a) 的证明。回顾一下，$h\colon\ \mathbb{R}^2 \to \mathbb{R}\times\mathbb{C} \cong \mathbb{R}^3$ 由以下方式定义：$h(0) = 0$ 并且

$$h(x) = \|x\| \cdot \Big(\cos\tfrac{\pi}{2r},\ (\tfrac{x_1+\mathrm{i}x_2}{\|x\|})^r \sin\tfrac{\pi}{2r} \Big)$$

对于 $x \neq 0$。那么对于 $x \in S^1$,

$$h_1(x) = \cos\tfrac{\pi}{2r}, \qquad h_2(x) = \operatorname{Re}[(x_1+\mathrm{i}x_2)^r] \cdot \sin\tfrac{\pi}{2r}, \qquad h_3(x) = \operatorname{Im}[(x_1+\mathrm{i}x_2)^r] \cdot \sin\tfrac{\pi}{2r}.$$

值得注意的是，这些分别与次数为 0 的 G-不变量多项式函数 $S^1 \to \mathbb{R}$ r 和 r 相一致。因此，对于每个 i，我们有 $h_i = \operatorname{ext}(h_i|_{S^1}) \in \operatorname{ext}\mathcal{P}^G$，于是

$$\operatorname{span}\{h_i\}_{i=1}^n \leq \overline{\operatorname{ext}\mathcal{P}^G}.$$

F此外，由于一个空间是有限维的，而另一个不是，因此这种包含关系是严格的。接下来，任取 $p \in \operatorname{ext}\mathcal{P}^G$。由定理 14，存在一个多项式 q 使得

$$p(x) = \int_{S^1} q(y)\, \langle\!\langle [x], [y] \rangle\!\rangle\, d\omega(y)$$

对于所有 $x \in S^1$，因此定理 13 给出 $p \in$ 张成空间 $\mathcal{M}$。由此可得

$$\overline{\operatorname{ext} \mathcal{P}^G} \leq \overline{\operatorname{span} \mathcal{M}}.$$

此外，这种包含关系是严格的，因为左侧的每个元素在 $\mathbb{R}^2 \setminus \{0\}$ 上都是连续可微的（命题 18），而非平凡的最大滤波器则不是（见命题 11）。

最后，每个最大滤波器都是 G-不变量、正齐次且分段线性的。由于这些性质在线性组合下保持不变，因此可得

$$\operatorname{span} \mathcal{M} \leq \mathrm{PL}_{\mathrm{h}}^G.$$

为证明等式，我们将证明任意 $f \in \mathrm{PL}_{\mathrm{h}}^G$ 都可表示为最大滤波器的线性组合。取一个 G-不变量超平面集合 $\mathcal{H}$（在此情形下即为直线），满足 $|\mathcal{H}| \geq 2$，使得$f \in \mathrm{PL}_{\mathrm{h}}(\mathcal{H})^G$。于是，只需证明 $\mathrm{PL}_{\mathrm{h}}(\mathcal{H})^G$ 由最大滤波器张成。为此，令$U \subseteq S^1$ 表示包含在这些超平面中的所有单位向量构成的集合。我们声称，$\mathrm{PL}_{\mathrm{h}}(\mathcal{H})$ 中的每个函数都唯一地由一个函数 $U \to \mathbb{R}$ 经锥形延拓得到。事实上，若按逆时针顺序排列 U 中的点，则该函数在相邻点所张成的锥包上是线性的。（此处可以取锥包，因为 $\mathcal{H}$ 的大小至少为 2，因此每个扇形的角度都严格小于 π。）于是，$\mathrm{PL}_{\mathrm{h}}(\mathcal{H})$ 中的 G-不变量也相应地对应于 G-不变量函数$U \to \mathbb{R}$，因此 $\mathrm{PL}_{\mathrm{h}}(\mathcal{H})^G$ 的维数为 $n := |U|/|G|$。接下来，取集合 S，其中包含 U 的 G-轨道代表元。于是 $|S| = |U|/|G| = n$。将每个 $s \in S$ 逆时针旋转 $\pi/|G|$ 弧度，得到模板 $z_s \in \mathbb{R}^2$。于是，最大滤波器 $\{\langle\!\langle [\cdot], [z_s] \rangle\!\rangle\}_{s \in S}$ 属于 $\mathrm{PL}_{\mathrm{h}}(\mathcal{H})^G$。此外，由于这些函数的奇异集（除原点外）两两不相交，它们线性无关。因此，它们张成 $\mathrm{PL}_{\mathrm{h}}(\mathcal{H})^G$ 的一个 n 维子空间，也就是说，张成整个 $\mathrm{PL}_{\mathrm{h}}(\mathcal{H})^G$。

□

4.2 定理 5(b)：实相位恢复

对于定理 5的这一部分，前几节的内容已涵盖所有要点，唯独尚未证明最后的包含关系 $\operatorname{span} \mathcal{M} \lneq \mathrm{PL}_{\mathrm{h}}^G$是严格的。本节将逐步建立必要的理论，以证明这一最后的结论，然后写出定理 5(b) 的证明。

我们主要关注 G-不变量函数 PL_{h}，即 $\mathrm{PL}_{\mathrm{h}}^G$。就我们的目的而言，这个函数空间最重要的性质体现在以下引理中：

引理 19。 取任意 $d > 2$ 并令 $G := \{\pm I\} \leq \mathrm{O}(d)$。给定 $f \in \mathrm{PL}_{\mathrm{h}}^G$，记为

$$K(f) := \{\, x \in \mathbb{R}^d : f \text{ is not differentiable at } x \,\}.$$

若 $K(f)$ 不是有限个超平面的并，则 $f \notin$ 张成空间 $\mathcal{M}$。

Ex引理 20。 取任意 $d > 2$ 并令 $G:= \{\pm I\} \leq \mathrm{O}(d)$。考虑由以下条件定义的 $f \in \mathrm{PL}_{\mathrm{h}}^G$

$$f(x) = \begin{cases} \min\limits_{i \in [d]} |x_i| & \text{if } \operatorname{sign}(x_1) = \cdots = \operatorname{sign}(x_d) \\ 0 & \text{else.} \end{cases}$$

参见图 1 了解示意图。于是 $K(f)$ 是 $\mathbb{R}^d$的非负正交象限和非正正交象限的非空子集，即它不是有限个超平面的并。由引理 19，可得$f \in \mathrm{PL}_{\mathrm{h}}^G \setminus$张成空间 $\mathcal{M}$。

引理 19 的证明。首先，我们证明一个中间结论，它是逆否命题的一个特例：对于每个非零的 $f \in$张成空间 $\mathcal{M}$，$K(f)$ 是有限个超平面的并。取任意非零的 $f \in$张成空间 $\mathcal{M}$，并考虑任意一种分解 $f = \sum_{i=1}^n c_i |\langle \cdot, y_i \rangle|$ 为最大滤波器 $\{|\langle \cdot, y_i \rangle|\}_{i=1}^n$，这些滤波器作为函数线性无关。我们断言，$K(f) = \bigcup_{i=1}^n$ 张成空间$\{y_i\}^\perp$。为方便记号，我们用 U 表示右侧的集合。$\subseteq$方向可由以下事实推出：f 中任何不可微点都是某个最大滤波器 $|\langle \cdot, y_i \rangle|$的不可微点。对于$\supseteq$方向，任取一个 $x \in U$。令 $\delta_i := \operatorname{sign}\langle x, y_i \rangle \in \{-1, 0, 1\}$。选取一个一般位置的 $v \in \mathbb{R}^d$。于是

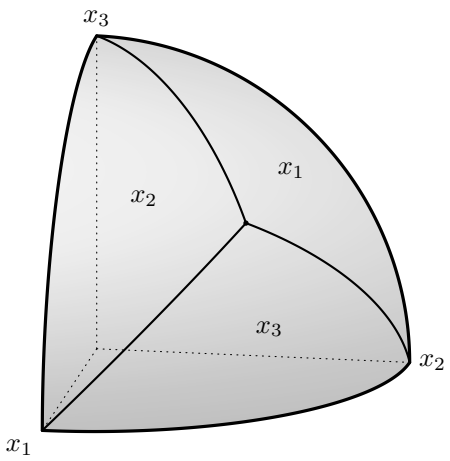

图1：例子 20 在 $d = 3$ 的特殊情形下的示意图。此处展示的是$f(x)$ 在 x 位于单位球面的第一卦限内时的取值。对于 x 位于大多数其他卦限内的情形，我们有$f(x) = 0$；不过，与第一卦限相对的卦限内的函数值由f 是偶函数这一事实决定。使 $x \in \mathbb{R}^3$ 处的 $f(x)$ 不可微的点集，由实心黑色曲线上各点的所有标量倍数组成。

对于所有绝对值充分小的 $t \in \mathbb{R}$，都有

$$\begin{aligned} f(x+tv) = \sum_{i=1}^{n} c_i |\langle x+tv, y_i\rangle| &= \sum_{\substack{i\in[n]\\ y_i \not\perp x}} c_i |\langle x+tv, \delta_i y_i\rangle| + \sum_{\substack{i\in[n]\\ y_i \perp x}} c_i |\langle x+tv, y_i\rangle| \\ &= \sum_{\substack{i\in[n]\\ y_i \not\perp x}} c_i \langle x+tv, \delta_i y_i\rangle + \sum_{\substack{i\in[n]\\ y_i \perp x}} c_i |\langle tv, y_i\rangle| \\ &= f(x) + t \sum_{\substack{i\in[n]\\ y_i \not\perp x}} c_i \langle v, \delta_i y_i\rangle + |t| \sum_{\substack{i\in[n]\\ y_i \perp x}} c_i |\langle v, y_i\rangle|. \end{aligned}$$

由于 $x \in U$，必然存在至少一个 i 使得 $y_i \perp x$。此外，$|t|$的系数非零，因为 v 处于一般位置。因此，f 在 x 沿着 v方向不可微，于是$x \in K(f)$，如所声称。

现在我们可以证明引理了。任取一个 $f \in \mathrm{PL}_{\mathrm{h}}^{G}$，使得 $K(f)$ 不是有限个超平面的并。由于 $f \in \mathrm{PL}_{\mathrm{h}}$，分层 (8) 表明存在一个由超平面组成的集合 $\mathcal{H}$，使得 $f \in \mathrm{PL}_{\mathrm{h}}(\mathcal{H})$。对于每个 $H \in \mathcal{H}$，选取一个单位向量 $z_H \in H^{\perp}$，并令 $T \leq \mathrm{PL}_{\mathrm{h}}(\mathcal{H})$ 表示由所得的最大滤波器 $\{|\langle \cdot, z_H\rangle|\}_{H\in\mathcal{H}}$张成的空间。注意，对于每个线性泛函 $\ell \in (\mathbb{R}^d)^*$，函数 $f+\ell \in \mathrm{PL}_{\mathrm{h}}(\mathcal{H})$ 非零，且 $K(f+\ell) = K(f)$。与此同时，根据我们的中间结论，T 的每个非零元素在整个线性超平面上都不可微。由于$f + (\mathbb{R}^d)^*$ 避开了 T，且由于 κ 在 $\mathrm{PL}_{\mathrm{h}}(\mathcal{H})/(\mathbb{R}^d)^*$ 上定义了一个范数，因此可得

$$c := \inf_{g\in T} \kappa(g-f) > 0.$$

接下来，我们证明对于每个最大滤波器的线性组合 $\sum_{i=1}^{n} c_i g_i$ $\{g_i\}_{i=1}^{n}$，

$$\Big\| \sum_{i=1}^{n} c_i g_i - f \Big\|_{\mathrm{Lip}} \geq \frac{c}{2}, \tag{9}$$

因此，f 不属于最大滤波器的张成空间的 Lipschitz 闭包。令 $I \subseteq [n]$ 表示满足 $g_i \in T$ 的指标，并设 $g := \sum_{i\in I} c_i g_i$ 和 $h := \sum_{i\in[n]\setminus I} c_i g_i$。那么 $g - f \in \mathrm{PL}_{\mathrm{h}}(\mathcal{H})$ 且 $\kappa(g-f) \geq c$。此外，由于对每个 $i \in [n] \setminus I$ 都有 $g_i \notin T$，h 中不可微的线性超平面均不与 $\mathcal{H}$ 中的任何线性超平面重合。因此，我们可以

从 $\mathcal{T}(\mathcal{H})$ 中相邻室的内部选取 $x, y \in \mathbb{R}^d$，使得

$$\|\nabla(g-f)(x) - \nabla(g-f)(y)\| = \kappa(g-f) \qquad \text{and} \qquad \nabla h(x) = \nabla h(y).$$

由此得到

$$\begin{aligned} c \le \kappa(g-f) = \|\nabla(g-f)(x) - \nabla(g-f)(y)\| &= \|\nabla h(x) - \nabla h(y) + \nabla(g-f)(x) - \nabla(g-f)(y)\| \\ &\le \|\nabla(h+g-f)(x)\| + \|\nabla(h+g-f)(y)\| \\ &\le 2\|h+g-f\|_{\mathrm{Lip}}, \end{aligned}$$

整理后即得所声称的界（9）。 □

定理 5(b) 的证明。 回忆一下，$h\colon \mathbb{R}^d \to \mathbb{R}^{d\times d}_{\mathrm{sym}} \cong \mathbb{R}^n$ 由算子平方根定义为

$$h(x) = \sqrt{xx^\top}.$$

于是，对每个 $i \in [n]$，都存在 $j, k \in [d]$，使得对每个 $x \in S^{d-1}$，

$$h_i(x) = x_j x_k,$$

这给出了一个次数为 2 的 G-不变多项式函数 $S^{d-1} \to \mathbb{R}$。因此，

$$\operatorname{span}\{h_i\}_{i=1}^n \le \overline{\operatorname{ext} \mathcal{P}^G},$$

且该包含关系是严格的，因为一个空间是有限维的，而另一个空间不是 .

接下来，$\operatorname{ext}\mathcal{P}$ 的每个成员 ^{G} 都属于张成空间 $\mathcal{M}$，根据定理16 和13，因此

$$\overline{\operatorname{ext} \mathcal{P}^G} \le \overline{\operatorname{span} \mathcal{M}}.$$

此外，包含关系是严格的，因为左侧的每个成员在 $\mathbb{R}^d \setminus \{0\}$ 上连续可微，根据命题18；而非平凡的最大滤波器并非如此（见命题11）。

最后，$\mathcal{M} \subseteq \mathrm{PL}_{\mathrm{h}}^G$ 蕴含

$$\overline{\operatorname{span} \mathcal{M}} \le \overline{\mathrm{PL}_{\mathrm{h}}^G},$$

并且，根据引理19 和例20，包含关系是严格的。 □

4.3 定理5（c）的证明：反射群

定理 5 的这一部分在性质上与前几部分截然不同，因为我们的分析并未通过第3节中发展的思路进行。相反，它依赖于这一显著特征：张成空间 $\mathcal{M}$ 是有限维的。

定理 5(c) 的证明。 回忆一下，$h\colon \mathbb{R}^d \to \mathbb{R}^d$ 定义为

$$h(x) \in [x] \cap W.$$

考虑映射 $L\colon \mathbb{R}^d \to \mathrm{Lip_h}(\mathbb{R}^d)$，其定义为

$$L(y)(x) = \langle h(x), y\rangle \qquad (x \in \mathbb{R}^d).$$

那么 L 是线性的，并且由于对每个 $i \in [d]$ 都有 $h_i = L(e_i)$，可知 $\operatorname{im} L =$ 张成空间 $\{h_i\}_{i=1}^d$。接下来，对每个 $z \in \mathbb{R}^d$，我们在 [19] 中应用公式 (2)，得到

$$\langle\!\langle [\cdot], [z] \rangle\!\rangle = \langle h(\cdot), h(z)\rangle = L(h(z)).$$

由于 $\operatorname{im} h = W$ 张成 $\mathbb{R}^d$，因此

$$\operatorname{span} \mathcal{M} = \operatorname{span}\{L(h(z)) : z \in \mathbb{R}^d\} = L(\operatorname{span} W) = \operatorname{im} L = \operatorname{span}\{h_i\}_{i=1}^d.$$

由于该子空间是有限维的，因此它等于其在 $\mathrm{Lip_0}(\mathbb{R}^d)$ 中的闭包。由于该子空间中的每个函数都是 G-不变量、正齐次且分段线性的，我们还得到

$$\overline{\operatorname{span} \mathcal{M}} \le \overline{\mathrm{PL}_{\mathrm{h}}^G}.$$

这种包含关系是严格的，因为 $\mathrm{PL}_{\mathrm{h}}^{G}$ 是无限维的；这一点可以通过 $\mathrm{PL}_{\mathrm{h}}^{G}$ 中的无限序列 $\{f_n\}_{n=1}^{\infty}$ 来体现，其中每个 f_n 在某个点 x_n 处不可微，而在该点处f_1、.....、 f_{n-1}均可微。（例如，先定义 $f_n|_W$，使其不可微点位于一个新的超平面上，再将其作 G-不变扩张。）

最后，假设没有非零向量被所有 G 固定，即 G 是本质的。我们证明 $\operatorname{ext}\mathcal{P}^G \cap \operatorname{span}\mathcal{M} = \{0\}$。根据命题 18，$\operatorname{ext}\ \mathcal{P}^G$ 的每个成员在 $\mathbb{R}^d \setminus \{0\}$ 上都是连续可微的。与此同时，每个非零 $f \in \operatorname{span}\mathcal{M}$ 都可以看作是 $\langle \cdot, y\rangle|_W$ 的 G-不变量扩张，其中某个非零 $y \in \mathbb{R}^d$ 存在。我们将利用这一表示证明 f 在$\mathbb{R}^d \setminus \{0\}$ 中某点不可微。由于作为反射群的 G 是本质的，因此 W 是形如 $\{x \in \mathbb{R}^d : \langle x, \alpha_i\rangle \geq 0\}$ （其中 $i \in [d]$）的半空间的交，其中 $\{\alpha_i\}_{i=1}^{d}$ 构成 $\mathbb{R}^d$的一个基。由于 y 非零，存在 $i \in [d]$ 使得 $\langle y, \alpha_i\rangle \neq 0$。我们声称 f 在W 的面 $W \cap \alpha_i^{\perp}$ 上不可微。为此，令 $s_i \in G$ 表示关于 $\alpha_i^{\perp}$ 的反射。那么对于每个 $x \in \operatorname{int}(W)$，我们有

$$f(s_i x) = f(x) = \langle x, y\rangle = \langle s_i^2 x, y\rangle = \langle s_i x, s_i y\rangle.$$

由于我们选择了 $i \in [d]$ 使得 $\langle y,\ \alpha_i\rangle \neq 0$，因此

$$\nabla f(s_i x) = s_i y = y - 2\frac{\langle y, \alpha_i\rangle}{\|\alpha_i\|^2}\alpha_i \neq y = \nabla f(x).$$

由于 f在 $\operatorname{int}(W)$ 和 $\operatorname{int}(s_i W)$ 上的梯度不同，因此结论成立。 □

5 数值结果

在前几节中，我们证明了，至少对于某些有限子群 $G \leq \mathrm{O}(d)$，通过将适当的最大滤波器组与线性映射进行后复合，可以得到轨道空间 $\mathbb{R}^d/G$ 到欧几里得空间的近乎最优双 Lipschitz 嵌入。本节将介绍数值实验，以展示这一现象似乎在多大程度上具有普遍性。我们考虑多种情形，其中有

- 一个实希尔伯特空间 V（通常为 $\mathbb{R}^d$），
- 一个群 $G \leq \mathrm{O}(V)$,
- 一个训练集 $\mathcal{X} = \{x_i\}_{i=1}^{N}$ 取自 V 中的不同 G-轨道，并且
- 目标维数 n,

我们的目标是找到一个 G-不变映射 $f\colon V \to \mathbb{R}^n$，使经验失真 $\operatorname{dist}_{\mathcal{X}}(f^{\downarrow})$最小，其定义为

$$\operatorname{dist}_{\mathcal{X}}(f^{\downarrow}) := \frac{\beta_{\mathcal{X}}(f^{\downarrow})}{\alpha_{\mathcal{X}}(f^{\downarrow})}, \quad \alpha_{\mathcal{X}}(f^{\downarrow}) := \min_{\substack{i,j\in[N]\\ i\neq j}} \frac{\|f(x_i) - f(x_j)\|}{d_{V/G}([x_i],[x_j])}, \quad \beta_{\mathcal{X}}(f^{\downarrow}) := \max_{\substack{i,j\in[N]\\ i\neq j}} \frac{\|f(x_i) - f(x_j)\|}{d_{V/G}([x_i],[x_j])}.$$

我们考虑两种类型的（近似） G-不变映射：训练映射 （映射属于某个指定的参数化族，且针对经验失真对参数进行优化）和未训练映射 （映射由显式公式给出，不过在某些情况下，映射的定义中带有一定的随机性）。对于这两种情况，我们都会报告某个测试集 $\mathcal{Y}$（取自 V）上的经验失真 $\operatorname{dist}_{\mathcal{Y}}(f^{\downarrow})$; 该测试集与训练集 $\mathcal{X}$相互独立。

我们测试了多种用于训练（近似） G-不变映射的架构:

- **训练的最大滤波器组（MF）**。学习最大滤波器组的单位范数模板。
- **训练后的线性层** $\circ$ **随机最大滤波器组（LRMF）**。随机生成一个最大滤波器组 Φ，其模板包含标准高斯分布元素，然后训练一个线性映射 L，使得到的映射为 $L \circ \Phi$。
- **训练好的线性层** $\circ$ **最大滤波器组 (LMF)**。同时训练最大滤波器组 Φ 和线性层 L，使得所得映射为 $L \circ \Phi$。

表1：群不变多项式映射

V	G	$p(x)$	原生目标空间	n
$\mathbb{R}^d$	$\{\pm I\}$	$xx^\top$	$\mathbb{R}^{d\times d}_{\text{sym}}$	$\binom{d+1}{2}$
$\mathbb{R}^d$	S_d	$\{\sum_{j=1}^d x_j^k\}_{k=1}^d$	$\mathbb{R}^d$	d
$\ell^2(\mathbb{Z}_d)$	C_d	$\{\hat{x}(a)\hat{x}(b)\hat{x}(-a-b)\}_{a,b\in\mathbb{Z}_d}$	$\mathbb{C}^{d^2}$	$2d^2$
$\mathbb{C}$	C_r	x^r	$\mathbb{C}$	2
$\mathbb{C}^d$	S^1	xx^*	$\mathbb{C}^{d\times d}_{\text{sa}}$	d^2
$(\mathbb{R}^d)^k$	$\mathrm{O}(d)$	$\{\langle x_i, x_j\rangle\}_{i,\ j\in[k]}$	$\mathbb{R}^{k\times k}_{\text{sym}}$	$\binom{k+1}{2}$

- **训练好的 ReLU 神经网络 (ReLU)**。训练两个线性映射 $W\colon\ V\to\mathbb{R}^m$ 和 $L\colon\ \mathbb{R}^m\to\mathbb{R}^n$，使得所得映射为 $L\circ\mathrm{ReLU}\circ W$。值得注意的是，该映射不是 G-不变量。我们使用训练集 $\mathcal{X}$的 G-不变增强$\{gx_i\}_{i\in[n],\ g\in G}$ 来训练该映射，不过当G 为无限时，我们会在增强过程中随机抽取 G 。

我们还考 虑以下未经训练的 G-不变量 映射：

- **随机最大滤波器组 (RMF)**。绘制一个随机最大滤波器组，其模板包含标准高斯分布的元素。

- G**-不变多项式 (Poly)**。这是一个特定的 G-不变多项式，由 n映射的实坐标 p 给出，详见表1。（在 $\ell^2(\mathbb{Z}_d)$的情况下，我们用 $\hat{}x$ 表示 x的离散傅里叶变换。）

- **的齐次延拓** G**-不变多项式 (HPoly)**。对于表G中描述的相应1-不变多项式，我们将其限制在单位球面上，并取其齐次延拓，从而得到一个正齐次映射。

在每个实验中，我们将最终嵌入维数固定为 $n=3\dim(V/G)$，当$3d$ 且 $V=\mathbb{R}^d$ 为有限值时，该维数恰好为 G 。训练集由各元素相互独立且服从标准高斯分布的向量构成。我们使用不同的初始化进行训练，共训练 10 次。对于 RMF，我们记录了在 2000 个随机最大滤波器组中，测试集上的最小经验失真。我们报告使用相同测试集所取得的最佳经验失真；该测试集包含 2000 个样本，并根据相同分布独立生成。结果见表2 。每一行中，我们将失真最小的训练映射加粗，因为我们旨在寻找一种通用方法来训练低失真嵌入。在某些情况下，我们已知 HPoly 是最优的，因此将这些情况下的失真值设为斜体。对于$\mathbb{R}^2/C_r$这种形式的情况，我们指出，HPoly 未能达到最优失真，这是因为我们在表1中选取了直接的不变量，而例1 中的最优不变量则是某种非常特殊的多项式选择的齐次延拓。表2中省略了一个失真值，因为我们的数据增强方法代价过高，因此未考虑对 S_5使用 ReLU。在所有设置中，训练得到的最佳G不变量映射都是 LMF。在 $\mathbb{R}^3/\{\pm I\}$的情况下，LMF 的表现略逊于近似的G不变量映射 ReLU，这或许是因为符号不变性与 ReLU 激活函数十分契合。

接下来，我们单独讨论以下情形：$\mathbb{R}^2/\{\pm I\}$、$\mathbb{C}^2/S^1$和$(\mathbb{R}^2)^2/\mathrm{O}(2)$。对于第一种情形，已知最优畸变为 $\sqrt{2}$；此外，根据定理3（b），LMF 可以逼近这一数值。表3也印证了这一点。请记住，最大滤波器组所能达到的下确界失真略大一些：$\sqrt{\pi/(\pi-2)}\approx 1.659$。图2展示了此情形下 MF 与 LMF 的差异。显然，LMF 中的最终线性层使冰淇淋锥变得更加扁平。对于另外两种情形，即 $\mathbb{C}^2/S^1$和$(\mathbb{R}^2)^2/\mathrm{O}(2)$，我们也知道最优畸变为 $\sqrt{2}$ [10, 7]，但目前还没有关于 LMF 在这种设定下的理论。不过，从表4 和5来看，LMF 似乎也能在这些情形下逼近最优畸变！（后一种情形下经验失真低于 $\sqrt{2}$ ，这是因为经验失真衡量的是测试集而非整个轨道空间的畸变。）显然，这两个轨道空间值得进一步研究，也有助于推进问题6。

表2：不同轨道空间的最小经验失真

	MF	LRMF	LMF	ReLU RMF	Poly		HPoly
$\mathbb{R}^3/\{\pm I\} \to \mathbb{R}^9$	1.85	2.56	1.76	**1.73**	2.40	17.70	*1.41*
$\mathbb{R}^5/S_5 \to \mathbb{R}^{15}$	3.06	**1.00**	**1.00**	—	5.70	4.70e3	46.3
$\ell^2(\mathbb{Z}_5)/C_5 \to \mathbb{R}^{15}$	4.72	3.57	**2.88**	19.0	5.58	5.18e2	132
$\mathbb{C}/C_3 \to \mathbb{R}^6$	2.93	2.08	**1.60**	2.58e3	3.05	2.42e4	3.00
$\mathbb{C}/C_4 \to \mathbb{R}^6$	4.16	2.08	**1.54**	16.20	4.43	3.89e6	4.00
$\mathbb{C}^2/S^1 \to \mathbb{R}^9$	2.21	1.84	**1.58**	9.71	2.59	22.37	*1.41*
$(\mathbb{R}^3)^3/\mathrm{O}(3) \to \mathbb{R}^{18}$	2.80	1.65	**1.41**	2.35e3	4.79	10.79	3.78

表3： $\mathbb{R}^2/\{\pm I\}$的最小经验失真

	MF	LRMF	LMF	ReLU RMF	Poly	HPoly	
$\mathbb{R}^2/\{\pm I\} \to \mathbb{R}^8$	1.84	2.01	1.66	**1.64**	1.92	163	*1.41*
$\mathbb{R}^2/\{\pm I\} \to \mathbb{R}^{16}$	1.67	1.54	**1.46**	1.63	1.78	163	*1.41*
$\mathbb{R}^2/\{\pm I\} \to \mathbb{R}^{32}$	1.66	1.49	**1.44**	1.58	1.80	163	*1.41*
$\mathbb{R}^2/\{\pm I\} \to \mathbb{R}^{256}$	1.66	**1.42**	**1.42**	1.51	1.76	163	*1.41*

表4： $\mathbb{C}^2/S^1$的最小经验失真

	MF	LRMF	LMF	ReLU RMF	Poly	HPoly	
$\mathbb{C}^2/S^1 \to \mathbb{R}^8$	2.20	1.81	**1.64**	15.26	2.52	16.89	*1.41*
$\mathbb{C}^2/S^1 \to \mathbb{R}^{16}$	2.16	1.66	**1.47**	7.19	2.46	16.89	*1.41*
$\mathbb{C}^2/S^1 \to \mathbb{R}^{32}$	2.16	1.52	**1.44**	5.61	2.30	16.89	*1.41*
$\mathbb{C}^2/S^1 \to \mathbb{R}^{256}$	2.16	**1.42**	**1.42**	3.29	2.18	16.89	*1.41*

表5：$(\mathbb{R}^2)^2/\mathrm{O}(2)$的最小经验失真

	MF	LRMF	LMF	ReLU RMF	Poly	HPoly	
$(\mathbb{R}^2)^2/\mathrm{O}(2) \to \mathbb{R}^4$	3.23	2.15	**1.41**	6.32e4	3.47	47.63	8.75
$(\mathbb{R}^2)^2/\mathrm{O}(2) \to \mathbb{R}^8$	2.55	1.46	**1.40**	5.23e4	2.91	47.63	8.75
$(\mathbb{R}^2)^2/\mathrm{O}(2) \to \mathbb{R}^{16}$	2.43	**1.40**	**1.40**	1.13e5	2.77	47.63	8.75

表6：连通的美国国会选区的最小失真

	MF	LRMF	LMF	ReLU RMF	
$(\mathbb{R}^2)^{50}/(\mathrm{O}(2) \times C_{50}) \to \mathbb{R}^{100}$	3.73	1.81	**1.50**	10.51	5.39

表7：二维形状结构数据集的最小失真

	MF	LRMF	LMF	ReLU RMF	
$(\mathbb{R}^2)^{100}/(\mathrm{O}(2) \times C_{100}) \to \mathbb{R}^{200}$	4.62	2.16	**2.00**	491	6.25

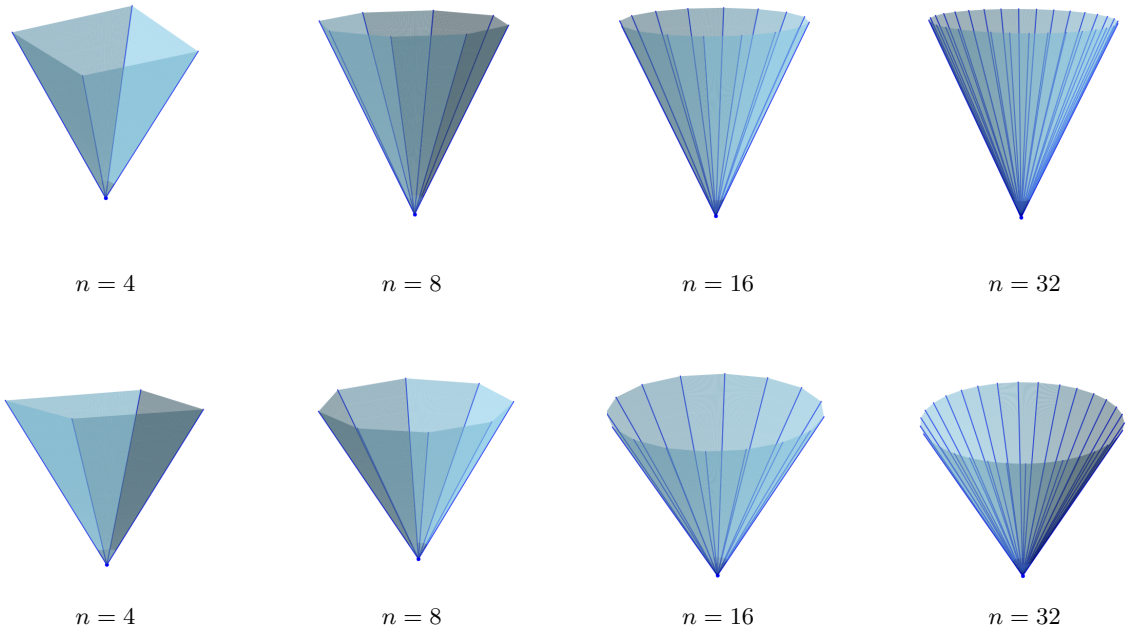

图2：借助主成分分析，将 $\mathbb{R}^2/\{\pm I\}$ 嵌入到维数逐渐增加的空间中进行可视化。MF嵌入位于上方，LMF嵌入位于下方。两种嵌入均使用最优最大滤波器组，即单位根。

实际上，我们可能不需要对整个轨道空间进行双利普希茨嵌入，而只需要对数据所在的轨道空间子集 $\mathcal{S}$ 进行嵌入。这意味着我们可以利用训练集（假设该训练集能够代表 $\mathcal{S}$）来学习双利普希茨嵌入 $\mathcal{S} \to \mathbb{R}^n$。下文中，我们将两个不同的群不变训练集以较低畸变嵌入到欧几里得空间中。这里我们不评估泛化能力，因此不考虑与训练集不同的测试集。在两个实验中，我们都考察平面上的形状数据。我们将每个样本表示为一个以原点为中心、具有 k个顶点的多边形，并将其视为轨道空间的一个元素 $(\mathbb{R}^2)^k/(\mathrm{O}(2) \times C_k)$。这里，O(2)作用于作为平面子集的形状，而 C_k 则通过循环置换顶点来作用。

我们的第一个数据集由第119届美国国会选区组成。我们剔除了17个不连通的国会选区，剩下427个选区需要进行嵌入。对每个选区的边界进行采样，取 $k = 50$ 个均匀间隔的点，并将其嵌入到 $\mathbb{R}^n$ 中，采用 $n = 100$。表6 报告了每种类型的模型在独立训练10次后达到的最小畸变（对于 RMF，则是在2000个随机最大滤波器组中取最小值）。图3展示了效果最佳的 LMF 嵌入的前两个主成分。显然，横轴反映了选区的凸程度，纵轴则表示偏心率。红色选区曾被新闻媒体批评为不公正划分选区 [9, 18]。这些选区通常位于左下方，这表明我们的嵌入在几何上具有相关性。

接下来，我们嵌入二维形状结构数据集 [12]。该数据集包含来自70个不同类别的形状。每个形状采样 $k = 100$ 个点，因此我们将嵌入维数加倍至 $n = 200$。畸变结果见表7 。再次由 LMF 达到最小畸变。与之前一样，我们在图4中将这个形状空间可视化。这里，横轴似乎表示偏心率的某种测度，较长的形状位于左侧，较紧凑的形状位于右侧。与此同时，纵轴似乎表示凸性的一种测度，但这只是近似的，“玻璃杯”类别并不符合这一趋势，便可见一斑。我们选取了几个形状类别作为示例。在这些簇中，视觉上离群的形状在形状空间中也被嵌入为离群点，例如与典型玻璃杯形状迥异的玻璃杯，或拓扑结构平凡的马克杯。

这两项实验都展示了 LMF 方法在群不变数据上训练低失真映射的实用性。这些结果表明，群不变机器学习可以从训练 LMF特征映射中受益。

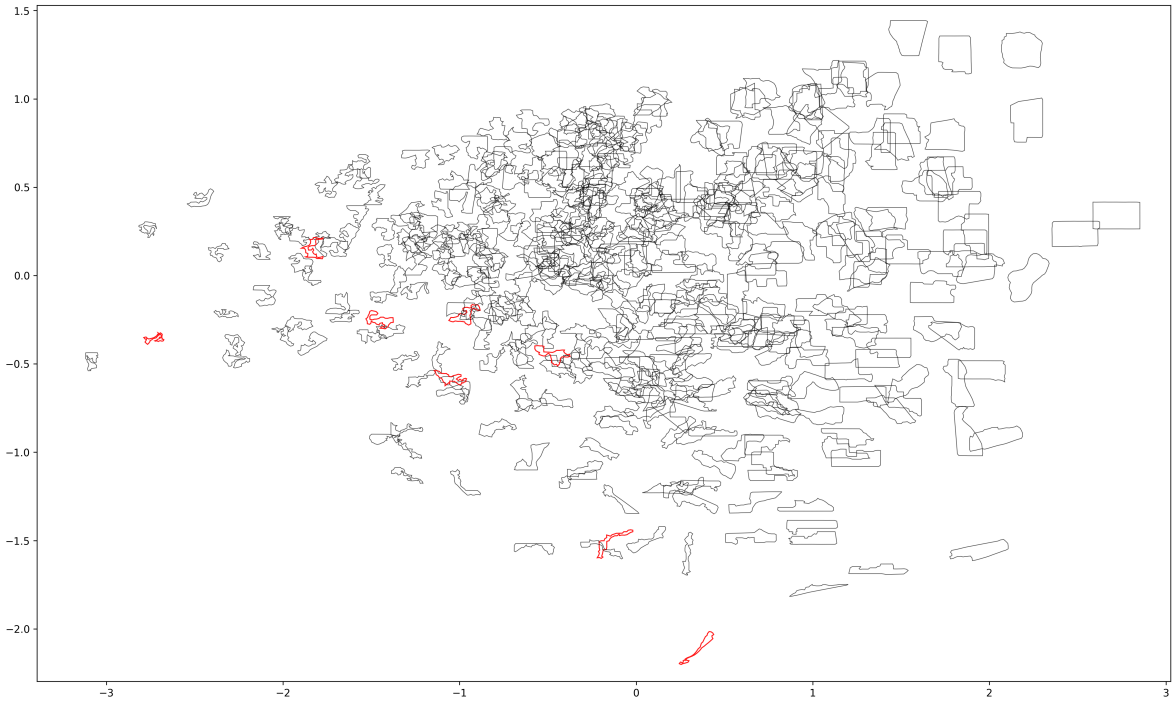

图3：使用 LMF 嵌入连通的美国国会选区，然后使用主成分分析进行可视化。

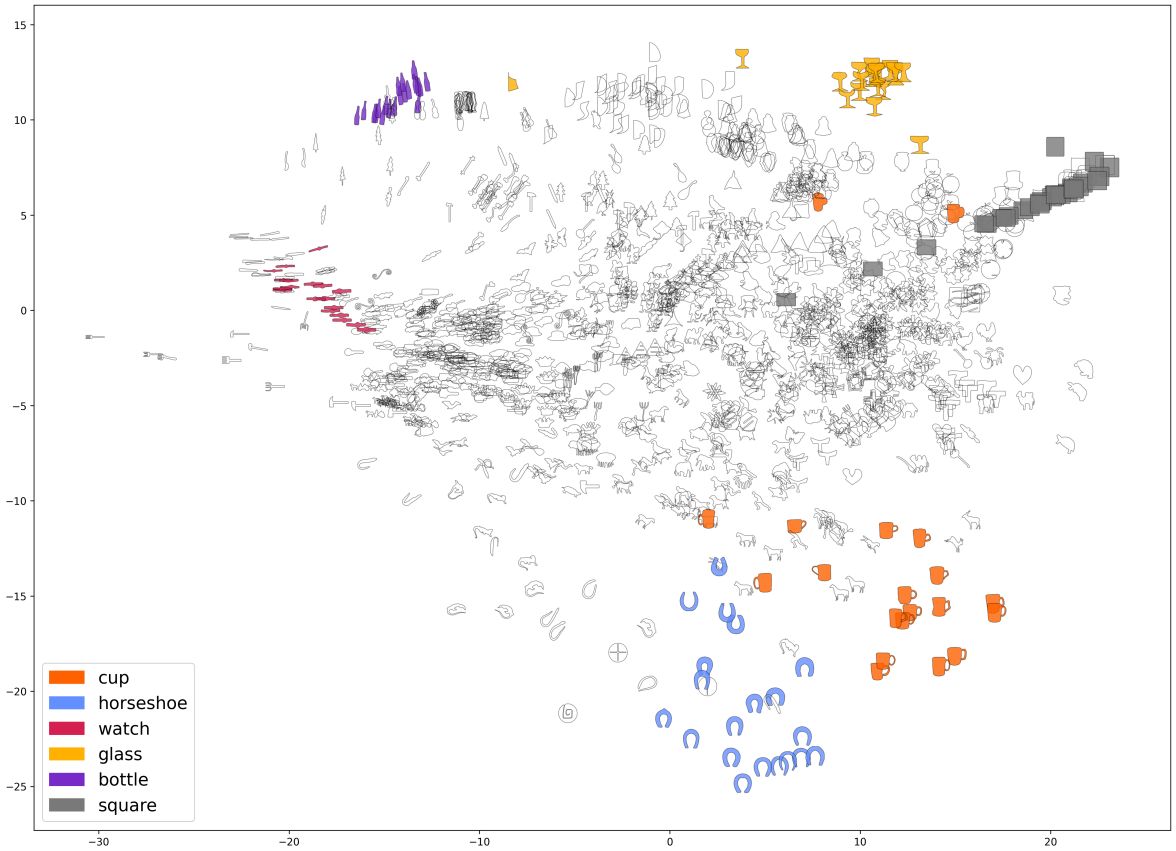

图4：使用 LMF 嵌入二维形状结构数据集，然后使用主成分分析进行可视化。

致谢

JWI 得到了西蒙斯基金会的资助，以及美国国家科学基金会 DMS 2220301 项目的资助。DGM 得到了美国国家科学基金会 DMS 2220304 项目的资助，并感谢 OpenAI 慷慨提供 ChatGPT Pro 版的免费使用权限；该版本建议在定理 16. 的证明中使用罗德里格公式。

参考文献

[1] T. Amir, T. Bendory, N. Dym, D. Edidin, 《紧群上的广义相位恢复问题的稳定性》，应用与计算调和分析 82 (2026) 101838。[2] R. Balan, N. Haghani, M. Singh, 《置换不变表示及其在图深度学习中的应用》，应用与计算调和分析 79 (2025) 101798。[3] R. Balan, E. Tsoukanis, 《利用余轨道的 G-不变表示：双 Lipschitz 性质》，arXiv:2308.11784 (2023)。[4] R. Balan, E. Tsoukanis, M. Wellershoff, 《基于排序的嵌入的稳定性》，arXiv:2410.05446 (2024)。[5] A. S. Bandeira, B. Blum-Smith, J. Kileel, J. Niles-Weed, A. Perry, A. S. Wein, 《群作用下的估计：从不变量恢复轨道》，应用与计算调和分析 66 (2023) 236–319。[6] T. Bendory, N. Boumal, C. Ma, Z. Zhao, A. Singer, 《双谱反演及其在多参考对齐中的应用》，IEEE 信号处理汇刊 66 (2017) 1037–1050[7] B. Blum-Smith, H. Derksen, D. G. Mixon, Y. Qaddura, B. Vose, 《轨道空间欧几里得畸变的估计》，arXiv:2506.04425 (2025)。[8] B. Blum-Smith 和 S. Villar, 《机器学习与不变量理论》，美国数学会通报 70 (2023) 1205–1213。[9] K. Buchholz, 《2020 年人口普查促使重新划分不公正选区》，Statistica，2020 年 4 月 1 日，www.statista.com/chart/21313/most-gerrymandered-districts-us/。[10] J. Cahill, J. W. Iverson, D. G. Mixon, 《迈向双 Lipschitz 不变量理论》，应用与计算调和分析 72 (2024) 101669。[11] J. Cahill, J. W. Iverson, D. G. Mixon, D. Packer, 《群不变最大值滤波》，计算数学基础 25 (2025) 1047–1084。[12] A. Carlier, K. Leonard, S. Hahmann, G. Morin, M. Collins, 《二维形状结构数据集》，https://2dshapesstructure.github.io/[13]H. Derksen, 《欧几里得群作用于数据的双 Lipschitz 商嵌入》，arXiv:2409.06829 (2024)。[14] N. Dym 和 S. Gortler, 《通用几何学习的低维不变嵌入》，计算数学基础 25 (2025) 375–415。[15] N. Dym, M. Wellershoff, E. Tsoukanis, D. Levy, R. Balan, 《基于排序的置换不变嵌入的定量界》，arXiv:2510.22186 (2025)。[16] S. Eriksson-Bique, 《有界曲率流形与轨形的定量双利普希茨嵌入》，几何与拓扑 22 (2018) 1961–2026。[17] D. Freeman, T. Oikhberg, B. Pineau, M. A. Taylor, 《函数空间中的稳定相位恢复》，数学年刊 390 (2024) 1–43。[18] D. Meyers, 《2020 年代最严重的国会选区杰利蝾螈划分》，The Fulcrum，2024 年 12 月 13 日，thefulcrum.us/electoral-reforms/worst-gerrymandered-districts。

[19] D. G. Mixon, D. Packer, 《反射群的最大滤波》, 《Adv. Comput. Math.》49 (2023) 82。[20] D. G. Mixon, Y. Qaddura, 《最大滤波的单射性、稳定性与正定性》, 《Constr. Approx.》(2025) 1–38。[21] W. Rudin, 《数学分析原理》, 第 3 版, 麦格劳-希尔, 1976。[22] E. M. Stein, G. Weiss, 《欧几里得空间上的傅里叶分析导论》, 普林斯顿大学出版社, 1971。[23] N. Weaver, 《利普希茨代数》, 第 2 版, 世界科技出版公司, 新加坡, 2018。[24] Y. Xia, Z. Xu, Z. Xu, 《相位恢复中的稳定性：条件数与最优向量集的刻画》, 《Math. Comput.》(2024)。