

# 病理演化么半群 (PEM) 公理系统

- 作者: GaoZheng
- 日期: 2025-10-21
- 版本: v1.0.0

注: “O3理论/O3元数学理论/主纤维丛版广义非交换李代数(PFB-GNLA)”相关理论参见: [作者 \(GaoZheng\) 网盘分享](#) 或 [作者 \(GaoZheng\) 开源项目](#) 或 [作者 \(GaoZheng\) 主页](#), 欢迎访问!

## 摘要

本文档旨在对病理演化么半群 (Pathological Evolution Monoid, PEM) 构建一个统一且严谨的公理化系统。该系统基于“将病理状态抽象为高维细胞空间中的可测子集, 将病理过程抽象为作用于该子集的非交换算子”这一核心思想, 实现了对拓扑/测度、连续-离散动力学、观测与学习、可治疗性与非交换度量等关键环节的全面重构。通过引入一套“近物理化”的算子分类与正规形 (normal form), 本系统不仅为描述疾病的发生、发展和转归提供了统一的数学语言, 也为构建可计算、可推演、可优化的疾病模型, 特别是在O3/GRL路径积分框架下进行治疗序列优化, 奠定了坚实的理论基石。

## 第一部分: 形式化基础与核心定义

### 定义 0: 记号与基本结构

#### 0.1 细胞状态空间

设  $((C, d))$  为一个带度量的高维细胞状态流形或流形簇。空间中的一个点  $(c \in C)$  完整地编码了细胞的类型、空间位置、表观遗传/转录谱、黏附/受体表型、细胞周期相位等一系列生物学信息。记  $(B(C))$  为其对应的 Borel  $(\sigma)$ -代数,  $(\mu)$  为一个参考测度 (可灵活组合体素体积、细胞计数和细胞类型权重)。

#### 0.2 可用状态族 (PEM的作用空间)

为避免数学上的任意子集带来的病理性问题 (如无限复杂的边界), PEM 的作用空间并非  $C$  的完整幂集  $P(C)$ , 而是其一个受约束的子集, 即可测的、具有有限能量的子族  $\mathcal{G}$ 。

$\mathcal{G} \subseteq B(C)$  满足:  $\mu(S) < \infty$ , 且  $S$  具有有限边界周长/有限变差。

- 数学释义:
  - $\mathcal{G}$  中的每个元素  $S$  代表一个宏观的、可观测的生理或病理状态切片, 如一个肿瘤或一个纤维化区域。
  - “有限测度”和“有限边界”这两个条件保证了病灶的体积、边界长度、连通分支数等关键拓扑-几何量是可定义且可计算的。
- 基本应用业务逻辑:
  - 在一个肝癌的数字模型中, 一个状态  $S_1 \in \mathcal{G}$  可以是“第一象限内所有正常肝细胞”构成的、边界清晰的区域。另一个状态  $S_2 \in \mathcal{G}$  则是“第一象限内所有分化程度为I期的癌细胞”构成的、可通过影像分割算法识别的区域。PEM的目标就是精确描述从  $S_1$  演化到  $S_2$  的过程。

#### 0.3 观测函数与病灶负荷

设  $(\Phi)$  为一个观测泛函簇, 即一组从状态空间映射到非负实数的函数,  $\Phi = \{\varphi : \mathcal{G} \rightarrow \mathbb{R}_{\geq 0}\}$ 。这些函数用于量化病理状态的不同维度。常用者包括:

- 总负荷 (Total Burden):  $B(S) = \mu(S)$
- 连通分支数 (Number of Components):  $N_{\text{comp}}(S)$
- 侵袭度 (Invasiveness / Perimeter):  $P(S) = \mu(\partial S)$  (边界测度)
- 组织功能保真度 (Functional Fidelity):  $F(S) \in$  (由多组学/功能指征回归得到)

这些观测测量将是定义疗效、构建路径代价泛函和量化算子效应的核心。

## 定义 1: PEM 的对象与态射

- 对象 (Objects):** PEM 的对象是作用空间  $\mathcal{S}$  中的状态集合  $S$ 。
- 算子/态射 (Operators/Morphisms):** 算子族  $\mathcal{O}$  是所有保持  $\mathcal{S}$  结构（即可测性与有限能量）的内映射  $\mathcal{O} : \mathcal{S} \rightarrow \mathcal{S}$  的子集。这些算子代表了所有能够导致病理状态拓扑演化的生物学过程。

## 定义 2: PEM 幺半群 (Monoid)

三元组  $((\mathcal{O}, \circ, \mathcal{I}))$  构成一个幺半群:

- $\circ$  是算子的**复合 (Composition)** 运算。
- $\mathcal{I}$  是**恒等算子 (Identity Operator)**, 即单位元。
- 通常情况下, 算子复合**不满足交换律**, 因此 PEM 是一个**非交换幺半群**。

## 第二部分: 结构与可计算性公理

### A1. 可测性与有限能量公理

对于任意算子  $\mathcal{O} \in \mathcal{O}$  和任意状态  $S \in \mathcal{S}$ , 其作用结果  $\mathcal{O}(S)$  必须仍在作用空间  $\mathcal{S}$  内。

$$\mu(\mathcal{O}(S)) < \infty, \quad P(\mathcal{O}(S)) < \infty.$$

### A2. 局域性/支集约束公理

存在一个全局常数  $v_{\max} > 0$  (代表生物组织内信息或物质传播的有效速度上限), 使得对于任意小时间步  $\Delta t$  的推进算子  $\mathcal{O}_{\Delta t}$ , 其影响范围受限。

$$\text{dist}(\mathcal{O}_{\Delta t}(S) \setminus S, S) \leq v_{\max} \Delta t.$$

此公理排除了“瞬时远距离”作用的非物理过程, 确保了演化的生物物理可达性。

### A3. 单调性或次模性公理

- 对于“增殖-外渗类”算子族  $\mathcal{O}^\uparrow$ , 要求满足**单调性**:  $\mathcal{O} \in \mathcal{O}^\uparrow \Rightarrow (S \subseteq T \Rightarrow \mathcal{O}(S) \subseteq \mathcal{O}(T))$ 。
- 对于“竞争-资源受限类”算子, 可放宽为**次模性 (Submodularity)**:

$$\mu(\mathcal{O}(S \cup T)) + \mu(\mathcal{O}(S \cap T)) \leq \mu(\mathcal{O}(S)) + \mu(\mathcal{O}(T))$$

### A4. 并合/拆分相容性公理

$$\mathcal{O}(S \cup T) \subseteq \mathcal{O}(S) \cup \mathcal{O}(T), \quad \mathcal{O}(S \cap T) \subseteq \mathcal{O}(S) \cap \mathcal{O}(T)$$

等号在“线性-无饱和”的理想近似下成立。此公理用于推导系统的可加性与饱和效应。

### A5. 观测的利普希茨连续性公理

对于每个观测函数  $\varphi \in \Phi$ , 存在一个常数  $L_\varphi$ , 使得:

$$|\varphi(\mathcal{O}(S)) - \varphi(S)| \leq L_\varphi d_H(\mathcal{O}(S), S)$$

其中  $d_H$  是状态间的 Hausdorff 距离。这确保了当状态变化很小时, 其观测量不会发生剧烈跳变, 保证了系统的稳健可估计性。

第三部分：基本算子族（规范化命名与类型）

为建立统一的语义，我们将算子进行分类，并给出其类型（扩张/替换/输运/消减/调制）和典型性质。

算子名称	符号	类型	数学释义 / 核心作用	基本应用业务逻辑
恒等算子	$\mathcal{I}$	单位元	$\mathcal{I}(S) = S$ ，代表病理状态无变化。	<b>场景:</b> 疾病的缓解期或稳定期。 <b>逻辑:</b> 系统状态在一段时间内没有显著恶化或好转，此时的演化过程可由 $\mathcal{I}$ 描述。
致病/转化算子	$\mathcal{O}_{\text{carcin}}$	替换型	将 $S_{\text{normal}}$ 的部分替换为 $S_{\text{cancer}}$ 。	<b>场景:</b> 模拟因HBV感染导致的肝细胞癌变。 <b>逻辑:</b> $\mathcal{O}_{\text{carcin}}$ 作用于“长期受病毒影响的肝细胞”子集，输出“早期肝癌细胞”子集，是疾病模型的起点。
炎症/浸润算子	$\mathcal{O}_{\text{inflam}}$	扩张-并合型	$\mathcal{O}_{\text{inflam}}(S) = S \cup S_{\text{infiltrate}}$ 。近似形态学“膨胀”。	<b>场景:</b> 描述肿瘤微环境的形成。 <b>逻辑:</b> 在“肿瘤细胞”子集 $S_{\text{tumor}}$ 上施加 $\mathcal{O}_{\text{inflam}}$ ，会引入免疫细胞等构成的 $S_{\text{infiltrate}}$ ，形成更复杂的病理状态。
纤维化算子	$\mathcal{O}_{\text{fibrosis}}$	替换-收缩型	$S_{\text{functional}} \mapsto S_{\text{fibrotic}}$ 。常降低功能 $F(S)$ 并改变连通性。	<b>场景:</b> 模拟肝硬化过程。 <b>逻辑:</b> 慢性的 $\mathcal{O}_{\text{inflam}}$ 触发 $\mathcal{O}_{\text{fibrosis}}$ ，将受损区域转化为“无功能的纤维组织”，导致肝功能下降。
增殖算子	$\mathcal{O}_{\text{prolif}}$	扩张型	$\mu(\mathcal{O}_{\text{prolif}}(S)) \geq \mu(S)$ ，伴随边界 $P(S)$ 上升。	<b>场景:</b> 肿瘤生长。 <b>逻辑:</b> 描述癌细胞群体的指数级增长，是病灶负荷增加的主要驱动力。
凋亡/坏死算子	$\mathcal{O}_{\text{apop}}$	消减型	$\mu(\mathcal{O}_{\text{apop}}(S)) \leq \mu(S)$ ，可能增加碎裂 ( $N_{\text{comp}}$ 上升)。	<b>场景:</b> 肿瘤内部因缺氧坏死或化疗诱导的细胞死亡。 <b>逻辑:</b> 该算子减小病灶体积，但可能导致病灶内部结构异质性增加。
血管生成算子	$\mathcal{O}_{\text{angio}}$	调制-扩张型	提升资源可达性，放大后续 $\mathcal{O}_{\text{prolif}}$ 的效果。	<b>场景:</b> 肿瘤突破尺寸限制。 <b>逻辑:</b> $\mathcal{O}_{\text{angio}}$ 本身不直接增加肿瘤细胞，但它通过改善微环境，为后续的 $\mathcal{O}_{\text{prolif}}$ 提供“增益”效应。
转移算子	$\mathcal{O}_{\text{meta}}$	输运-播散型	$\mathcal{O}_{\text{meta}}(S) = S \cup \bigcup_i S_{\text{meta},i}$ 。严格增加连通分支数 $N_{\text{comp}}$ 。	<b>场景:</b> 晚期乳腺癌的肺转移。 <b>逻辑:</b> $\mathcal{O}_{\text{meta}}$ 将一个单连通的“原位癌”子集，变为一个包含多个不连通的“远端转移灶”的多连通状态。
耐药获得算子	$\mathcal{O}_{\text{resist}}$	调制-替换型	改变细胞子集对后续治疗算子的敏感度。	<b>场景:</b> 靶向治疗后出现耐药。 <b>逻辑:</b> 在施加治疗算子前， $\mathcal{O}_{\text{resist}}$ 已将部分敏感细胞替换为耐药细胞，导致后续治疗效果大打折扣。
治疗算子	$\mathcal{O}_{\text{therapy}}$	消减/修复/调制	药物/放疗/手术等。一般非可逆。	<b>场景:</b> 靶向治疗肺癌。 <b>逻辑:</b> 对于包含“EGFR突变细胞”的子集 $S_{\text{mutant}}$ ，施加 $\mathcal{O}_{\text{Osimertinib}}$ 会将其转化为“凋亡细胞”子集，使病灶缩小。

第四部分：运算、动力学与非交换性

A6. 复合与结合律公理

算子集合上严格定义了**复合运算**  $\circ$ ，代表病理过程在时间上的相继发生。复合运算满足结合律。

$$(\mathcal{O}_B \circ \mathcal{O}_A)(S) = \mathcal{O}_B(\mathcal{O}_A(S))$$

- 基本应用业务逻辑：**  
一个典型的结肠癌“腺瘤-癌”演化路径可表示为算子序列的复合：

$$S_{\text{cancer}} = (\mathcal{O}_{\text{carcin}} \circ \mathcal{O}_{\text{inflam}} \circ \mathcal{O}_{\text{adenoma}})(S_{\text{normal}})$$

这表示正常组织先演化为腺瘤，后在炎症作用下，最终癌变。

## A7. 非交换性公理与对易子

复合运算。通常**不满足交换律**，这是描述复杂病理过程的核心特征。

$$\mathcal{O}_B \circ \mathcal{O}_A \neq \mathcal{O}_A \circ \mathcal{O}_B$$

为了量化这种顺序敏感性，定义**对易子**  $[\mathcal{O}_A, \mathcal{O}_B]$  和由观测诱导的**非交换度量**  $\Delta_\Phi$ :

$$[\mathcal{O}_A, \mathcal{O}_B] := \mathcal{O}_B \circ \mathcal{O}_A - \mathcal{O}_A \circ \mathcal{O}_B$$

$$\Delta_\Phi(\mathcal{O}_A, \mathcal{O}_B; S) := \sum_{\varphi \in \Phi} w_\varphi |\varphi((\mathcal{O}_B \circ \mathcal{O}_A)(S)) - \varphi((\mathcal{O}_A \circ \mathcal{O}_B)(S))|$$

### 基本应用业务逻辑:

- 场景:** 炎症与癌变的关系。
- 路径1 (炎症驱动癌变):** 先有慢性炎症，再诱发癌变。这是临床常见的“炎-癌转化”。
$$S_{\text{path1}} = (\mathcal{O}_{\text{carcin}} \circ \mathcal{O}_{\text{inflam}})(S_{\text{normal}})$$
- 路径2 (癌变引发炎症):** 先有肿瘤形成，再诱导免疫炎症反应。
$$S_{\text{path2}} = (\mathcal{O}_{\text{inflam}} \circ \mathcal{O}_{\text{carcin}})(S_{\text{normal}})$$
- 结论:**  $S_{\text{path1}}$  和  $S_{\text{path2}}$  是两个微环境、治疗策略和预后都截然不同的病理状态。PEM 的非交换性精确地捕捉了这种**路径依赖性 (Path-dependency)**。

## A8. 离散时间动力学与路径代价

给定一个治疗演化策略序列  $\mathcal{O} = (\mathcal{O}_{t_1}, \dots, \mathcal{O}_{t_n})$ ，状态的演化路径为  $S_{t_{k+1}} = \mathcal{O}_{t_{k+1}}(S_{t_k})$ 。定义路径的**拉格朗日-代价泛函**:

$$\mathcal{L}(S_t, \mathcal{O}_t) = \alpha B(S_t) + \beta P(S_t) + \gamma(1 - F(S_t)) + \lambda \text{Tox}(\mathcal{O}_t)$$

路径的**总作用量**为  $\mathcal{A} = \sum_t \mathcal{L}$ 。在 GRL 视角下，以  $\exp(-\tau \mathcal{A})$  加权采样最优序列，这与 O3 理论的路径积分语义完全对应。

## A9. 连续时间半群与生成元

若存在强连续半群  $(T_t)_{t \geq 0} \subset \mathbb{O}$  满足  $T_{t+s} = T_t \circ T_s$ ，则可定义其**生成元**  $\mathcal{G}$ :

$$\mathcal{G}(S) := \lim_{t \downarrow 0} \frac{T_t(S) - S}{t}$$

$\mathcal{G}$  可被看作是“病理流”的矢量场。治疗对应于引入控制项，形成受控动力学系统  $\mathcal{G}_u = \mathcal{G} + \sum_k u_k \mathcal{G}_k^{(\text{therapy})}$ 。

# 第五部分：治疗、可逆性与系统分析

## A10. 部分可逆与伪逆公理

治疗算子  $\mathcal{O}_{\text{therapy}}$  通常不存在严格的数学逆。我们定义其在观测空间  $\Phi$  下的 **Moore-Penrose 型伪逆**  $\mathcal{O}_{\text{th}}^\dagger$ ，其目标是找到一个操作，使得治疗后的状态与目标状态  $S_{\text{target}}$  在观测上最接近:

$$\mathcal{O}_{\text{th}}^\dagger \text{ minimizes } \sum_{\varphi \in \Phi} w_\varphi |\varphi(\mathcal{O}_{\text{th}}(S)) - \varphi(S_{\text{target}})|$$

例如，纤维化是否可逆转，就等价于在一个功能保真度  $F(\cdot)$  的约束下的伪逆可达性问题。

## A11. 正规形与重写系统公理

为比较不同算子序列的等效性，引入一套基于近似对易关系的**基本重写规则**，任何有限长度的算子序列都可被化简到一个**PEM正规形**。

- $\mathcal{O}_{\text{inflam}} \circ \mathcal{O}_{\text{prolif}} \Rightarrow \mathcal{O}_{\text{prolif}}^{\dagger} \circ \mathcal{O}_{\text{inflam}}$  (炎症放大增殖)
- $\mathcal{O}_{\text{resist}} \circ \mathcal{O}_{\text{therapy}} \Rightarrow \mathcal{O}_{\text{therapy}} \circ \mathcal{O}_{\text{resist}}$  (耐药更新治疗敏感度)
- $\mathcal{O}_{\text{meta}} \circ \mathcal{O}_{\text{apop}} \Rightarrow \mathcal{O}_{\text{apop}}^{\oplus} \circ \mathcal{O}_{\text{meta}}$  (转移灶内并行凋亡)

## 第六部分：观测-学习-校准闭环

### A12. 探测算子与数据生成公理

引入**探测算子**  $M: \mathcal{S} \rightarrow \mathcal{Y}$ ，它将真实病理状态映射到可观测的数据空间  $\mathcal{Y}$ （如影像、病理切片、ctDNA读数）。 $M$  必须满足统计可识别性，即能够从数据中无偏地估计出我们关心的观测量  $\varphi(S)$ 。

### A13. 算子同定公理

给定时间序列观测数据  $\{M(S_{t_k})\}$ ，我们可以在一个参数化的算子类别  $\mathcal{O}(\Theta)$  内，通过求解优化问题来**同定 (identify)** 出最符合数据的个体化算子参数：

$$\min_{\theta \in \Theta} \sum_k \|\hat{\varphi}(M(S_{t_{k+1}})) - \hat{\varphi}(M(\mathcal{O}_{\theta}(S_{t_k})))\|_2^2$$

这个过程最终得到一个“**个体化的PEM模型**”。

## 第七部分：关键推论与定量指标

### T1. 转移的拓扑判据

任何导致状态的连通分支数  $N_{\text{comp}}$  期望值严格增加的演化，必然涉及转移算子  $\mathcal{O}_{\text{meta}}$  的作用。因此，任何旨在控制转移的治疗序列，必须优先抑制  $\mathcal{O}_{\text{meta}}$  的活性。

### T2. 炎症-癌变顺序效应的可观测性

如果非交换度量  $\Delta_{\Phi}(\mathcal{O}_{\text{inflam}}, \mathcal{O}_{\text{carcin}}; S_0)$  在  $\{B, P, F\}$  等观测指标下显著为正，则“炎→癌”与“癌→炎”这两条路径演化出的终末状态，在影像组学或病理组织学上是可区分的。这直接提示了针对不同路径的个体化治疗排序策略。

### T3. 疗法组合的“近对易窗口”

在某些参数域  $\mathcal{U}$  内（例如，低肿瘤负荷、低炎症水平、低耐药比例），不同的治疗算子可能近似满足交换律 ( $[\mathcal{O}_{\text{therapy}}^{(a)}, \mathcal{O}_{\text{therapy}}^{(b)}] \approx 0$ )。在此窗口内，“并行给药”与“交替给药”在最终疗效上是近似等价的；而一旦超出此窗口，则会出现强烈的顺序效应。

### 可计算的定量指标

- 非交换得分**:  $\text{NC}(\mathcal{O}_A, \mathcal{O}_B; S) = \frac{\Delta_{\Phi}(\mathcal{O}_A, \mathcal{O}_B; S)}{1 + \sum_{\varphi \in \Phi} \varphi(S)}$
- 拓扑风险**:  $\text{TopoRisk}(S) = \alpha_1 N_{\text{comp}}(S) + \alpha_2 P(S)$
- 伪逆可达度**:  $\text{Reach}(\mathcal{O}_{\text{th}}^{\dagger}; S \rightarrow S^*) = \exp(-\min_{\mathcal{O} \in \mathcal{U}} \mathcal{A}(\mathcal{O}; S \rightarrow S^*))$

## 第八部分：极简工作流（从定义到可计算）

- 构建空间**: 基于临床数据（如影像分割、空间转录组）构建状态空间  $\mathcal{S}$  和观测函数集  $\Phi$ 。
- 选择算子**: 从规范算子库中挑选与病理过程相关的算子子集，构建参数化算子族  $\mathcal{O}(\Theta)$ ，并施加 A1–A5 约束。

- 3. **个体化同定:** 使用 A13, 通过拟合患者的纵向数据, 同定出个体化的算子参数, 并估计关键算子对的非交换得分 NC。
  - 4. **优化与比较:** 在给定的治疗预算和约束下, 通过 GRL/最优控制等方法求解  $\min_{\mathcal{O}} \mathcal{A}$ , 找到最优治疗序列。使用正规形 (A11) 对不同的候选方案进行等价类比较。
  - 5. **验证与迭代:** 利用 T1-T3 等推论和拓扑风险等指标进行体内外实验设计与自治性检查, 迭代优化模型。
- 

**许可声明 (License)**

Copyright (C) 2025 GaoZheng

本文档采用[知识共享-署名-非商业性使用-禁止演绎 4.0 国际许可协议 \(CC BY-NC-ND 4.0\)](#)进行许可。