arXiv:2302.02333v1 [cs.GT] 5 Feb 2023

# 量子博弈中的学习

KYRIAKOS LOTIDIS*、PANAYOTIS MERTIKOPOULOS‡ 和 NICHOLAS BAMBOS‡

摘要。本文介绍一类适用于一般量子博弈的学习动力学，称为“跟随量子正则化领导者”（FTQL），这一名称借鉴了有限博弈中的经典 FTRL 模板。我们证明，由此诱导的量子态动力学可分解为两部分：(*i*) 一个经典的交换分量，它支配系统特征值的演化，其方式类似于 FTRL 下混合策略的演化；(*ii*) 一个关于系统特征向量的非交换分量，它没有经典对应物。尽管这一非经典分量带来了复杂性，我们发现，在所有量子博弈中，FTQL 动力学所产生的遗憾至多为常数。此外，我们调整经典稳定性概念，以考虑量子博弈状态空间的非线性几何，并证明在 FTQL 下，只有纯量子均衡才可能稳定且具有吸引性；作为部分逆命题，我们还证明，满足某种“变分稳定性”条件的纯均衡总是具有吸引性。最后，我们证明，在量子极小极大博弈中，FTQL 动力学具有庞加莱复现性，从而推广了近期关于量子复制子动力学的一项结果。

## 1. 引言

量子信息论的兴起，以及由此带来的“量子优势” [4, 36, 50] ，对计算机科学和机器学习产生了深远影响，涉及量子密码学和阴影层析 [1], ，以及量子生成对抗网络（QGANs）和对抗学习 [7, 11, 27]。从宏观层面来看，量子计算的优势源于能够制备二态量子系统的叠加态，这类系统称为量子比特：经典比特无法处于叠加态，因此经典计算机能够执行的计算实际上受限于其二进制字母表和存储结构。鉴于此，量子计算有望大幅加速人工智能算法和模型的发展；谷歌的“Sycamore”54量子比特处理器在不到200 秒的时间内训练了一个自动驾驶汽车模型 [4]。

同样，当这类模型被部署于多智能体场景中——例如 QGANs 或自动驾驶汽车——相较于经典非合作框架，局面会发生巨大变化。其主要原因仍是“量子优势”：由于退相干和纠缠的复杂性——这两个量子概念在经典世界中都没有经典对应物——量子博弈参与者能够相较于“经典”参与者获得明显优势，在均衡时取得原本不可能达到的更高收益 [15, 33]。这同样归因于概率混合在量子世界和经典世界中的运作方式不同：在经典博弈中，混合策略 是

* 斯坦福大学管理科学与工程系。‡格勒诺布尔阿尔卑斯大学、法国国家科学研究中心、法国国家信息与自动化研究所、格勒诺布尔国立综合理工学院、格勒诺布尔信息学实验室，法国格勒诺布尔 38000。
‡斯坦福大学电气工程系与管理科学与工程系。电子邮箱： klotidis@stanford.edu、panayotis.mertikopoulos@imag.fr、bambos@stanford.edu。
2020 数学主题分类。主要分类：91A81、37N40、68Q32；次要分类：68T05、81Q93、91B80。关键词与短语。量子博弈；纳什均衡；正则化学习；渐近稳定性。

各个纯策略的概率凸组合；在量子博弈中，混合状态 则是与各组成状态相关联的量子投影算符 的概率混合。因此，量子混合态可以产生落在经典混合策略凸包之外的收益，从而为能够使用量子技术的参与者带来切实优势——例如，他们能够将行动编码到量子比特寄存器中，再将其提交给“裁判”（决定量子博弈收益的自然机制）。

当然，量子博弈参与者的优势能在多大程度上体现出来，取决于参与者是否真的达到了均衡。Bostanci 和 Watrous 最近的研究 [5] 表明，计算量子博弈的近似纳什均衡这一问题属于 PPAD；因此，根据 Daskalakis 等人的开创性工作 [12, 13], ，该问题必定是这个复杂性类中的完全问题（因为计算量子均衡至少与计算经典均衡一样困难）。因此，鉴于量子博弈的维度相对于每位参与者可用的量子比特数量呈指数增长，除规模最小的博弈外，计算量子博弈的纳什均衡很快就会变得难以处理。鉴于此，采用在线学习范式似乎更为合理：每位玩家都力求将自身的遗憾降至最低，并由此提出以下问题：

> 在量子无遗憾学习方案下，所有均衡结果出现的可能性都一样吗？是否存在一类均衡具有固有的选择偏差——无论是偏向还是排斥某些结果？

**我们的贡献**。首先，为了在量子情境下实现无遗憾，我们基于广受欢迎的有限 $N$人量子博弈“跟随正则化领导者”（FTRL）范式，引入了一种灵活的一般学习模型有限博弈 [41, 42]。我们将由此得到的模型称为“跟随量子正则化领导者”（FTQL）。矩阵乘法／指数权重（MMW）动力学是该模型的一个特例，已广泛应用于量子博弈和矩阵学习 [2, 22–24, 46]，并由此产生量子复制子动力学 [20, 23]。重要的是，正如我们在第 3 节所示，FTQL 的混合态动力学可分解为一个“经典”部分（特征值按照有限博弈中的 FTRL 动力学演化），以及一个“量子”部分，用以刻画系统特征函数的演化（且没有经典对应物）。

就遗憾最小化而言，所有 FTQL 动力学造成的遗憾至多为常数，因此从学习的角度看，它们是颇具吸引力的选择。然而，由于博弈状态空间的非线性几何，推导这些动力学的均衡收敛性质要困难得多。具体来说，与有限博弈（其中纯策略是彼此孤立的极点）不同，量子博弈的纯态构成一个连续的驻点流形（其中每一点都是极点），因此，研究稳定性与收敛问题会变得十分复杂。尽管存在这些拓扑上的复杂性，我们仍证明 FTQL 具有以下基本性质：(*i*) 纳什均衡是驻定的；(*ii*) 内部轨迹的极限和李雅普诺夫稳定态都是纳什均衡；(*iii*) 在 FTQL 下，只有纯量子均衡能够稳定且具有吸引性（排除平凡的驻定状态）；以及 (*iv*) 作为部分逆命题，我们证明，满足某种“变分稳定性”条件的纯态具有吸引性，与所选正则化项无关。因此，我们的结果导出了一种隐含的量子“纯化”原理：在 FTQL 下，混合态本质上是脆弱的，只有纯量子态才能持续具有吸引性。[1]

最后，我们通过进一步考察双人零和 Quantum 博弈，补充了我们的研究结果 games，在这里我们展示了 FTQL 表现出一种称为<样式 id='1'>庞加莱复现</样式>的循环特性*e:*

[1]在经典世界中，上述一系列结果的一个版本有时被称为演化博弈论的“民间定理” [10, 17, 21, 30]。

几乎所有博弈轨迹都会无穷多次任意接近其起点。这一结果是 Jain 等人近期结果的广义推广。 [23], 他们证明了 MMW 动力学具有这一特性。就此而言，我们的结果表明，MMW 的结果并非巧合：尽管存在“量子优势”，完全竞争仍无法通过正则化学习的动力学得到解决。

为简化阐述，全文将重点讨论连续时间演化的模型。这样我们便可避开超参数调优等相关问题，转而突出理论的核心方面。

## 2. 预备知识

W我们首先简要回顾量子博弈论的一些基础知识，并介绍必要的背景，以便理解我们的结果。

**记号**。 给定一个（复）希尔伯特空间 $\mathcal{H}$，我们将使用狄拉克左矢-右矢符号来区分 $\mathcal{H}$ 中的元素 $|\psi\rangle$ 及其伴随 $\langle\psi|$；否则，当上下文已表明所用的基时，我们将使用匕首记号“†”表示 $\psi$ 的厄米转置$\psi^\dagger$。我们还将用 $\mathbb{H}^d$ 表示 $d\times d$ 厄米矩阵的空间，并用 $\mathbb{H}^d_+$ 表示 $\mathbb{H}^d$ 中半正定矩阵构成的锥。最后，给定一个实函数 $f\colon \mathbb{R}\to\mathbb{R}$ 以及一个厄米矩阵 $\mathbf{X}\in\mathbb{H}^d$，其酉特征分解为 $\mathbf{X}=\sum_{\alpha=1}^d x_\alpha \mathbf{u}_\alpha\mathbf{u}_\alpha^\dagger$，我们将用 $f(\mathbf{X})$ 表示矩阵 $f(\mathbf{X})=\sum_{\alpha=1}^d f(x_\alpha)\mathbf{u}_\alpha\mathbf{u}_\alpha^\dagger$（它同样是厄米矩阵）。

**量子博弈。 </style}按照，一个量子博弈** 由以下基本要素构成：

(1) 一个有限的参与者 $i\in\mathcal{N}=\{1,\dots,N\}$集合。(2) 每位玩家 $i\in\mathcal{N}$都可以访问一个复希尔伯特空间 $\mathcal{H}_i\cong\mathbb{C}^{d_i}$，该空间描述玩家可用的（纯态）量子态集合（通常是一个由量子比特构成的离散寄存器）。具体而言，量子态是 $\mathcal{H}_i$ 中的一个单位范数元素 $\psi_i$，因此所有此类状态构成 $\mathcal{H}_i$的单位球面 $\Psi_i:=\{\psi_i\in\mathcal{H}_i:\ \|\psi_i\|=1\}$。我们还将$=\prod_i\Psi_i$记作$\Psi:=\prod_i\Psi_i$，表示由每位玩家独立制备的纯态 $\psi_i\in\Psi_i$组成的所有系综 $\psi=(\psi_1,\dots,\psi_N)$的空间 $i\in\mathcal{N}$。

(3) 参与者的奖励由各自的收益函数 $u_i\colon \Psi\to\mathbb{R}$决定。这些收益函数并非任意设定，而是由一个联合的正算子值测度（POVM）量子测量过程得出，过程如下 [8]：首先，假设存在一个有限的可能测量结果 $\omega\in\Omega$集合，裁判可以从参与者的量子态中观察到这些结果（例如，对玩家制备的量子比特进行测量，判断其为“上”或“下”）。每个此类结果 $\omega\in\Omega$都对应一个半正定算子$\mathbf{P}_\omega\colon \mathcal{H}\to\mathcal{H}$，它作用于参与者各自状态空间的张量积 $\mathcal{H}:=\bigotimes_i\mathcal{H}_i$；我们进一步假设$\sum_{\omega\in\Omega}\mathbf{P}_\omega=\mathbf{I}$，因此，当系统处于状态 $\psi\in\Psi$ 时，观察到 $\omega\in\Omega$的联合概率为

$$P_\omega(\psi)=\langle\psi_1\otimes\cdots\otimes\psi_N|\mathbf{P}_\omega|\psi_1\otimes\cdots\otimes\psi_N\rangle \tag{1}$$

每位玩家的收益 $i\in\mathcal{N}$由该测量过程的结果给出，通过收益可观测量 收益可观测量 $U_i\colon \Omega\to\mathbb{R}$；具体而言，在此语境下，$u_i(\psi)$ 表示玩家在状态 $\psi\in\Psi$ 下的期望收益，即

$$u_i(\psi):=\langle U_i\rangle\equiv\sum\nolimits_\omega P_\omega(\psi)\,U_i(\omega). \tag{2}$$

量子博弈 定义为一个元组 $\mathcal{Q}\equiv\mathcal{Q}(\mathcal{N},\ \Psi,\ u)$，其中包含上述参与者、量子态和收益函数。

**混合态**。 除了纯态外，每位玩家 $i \in \mathcal{N}$ 还可以制备其概率混合，称为混合态。与经典有限博弈中的混合策略不同，这些混合态并非其对应纯态的凸组合；相反，给定一族由 $\alpha_i \in \mathcal{A}_i$ 索引的纯量子态 $\psi_{i\alpha_i} \in \Psi_i$ ，混合态由如下形式的密度矩阵描述：

$$\mathbf{X}_i = \sum_{\alpha_i \in \mathcal{A}_i} x_{i\alpha_i} |\psi_{i\alpha_i}\rangle\langle\psi_{i\alpha_i}| \tag{3}$$

其中， $x_{i\alpha_i} \geq 0$ 是 $\psi_{i\alpha_i}$的混合权重，并且我们假设 $\mathrm{tr}\mathbf{X}_i = 1$ （在此情境下，不假设状态 $\psi_{i\alpha_i}$ 彼此正交）。根据玻恩规则，这意味着，如果每位玩家 $i \in \mathcal{N}$ 按照$\mathbf{X}_i$制备一个混合态，那么在$\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_N)$ 下观察到$\omega \in \Omega$的概率将为 $P_\omega(\mathbf{X}) = \sum_\alpha x_\alpha \langle\psi_\alpha|\mathbf{P}_\omega|\psi_\alpha\rangle$, 其中，采用多重指标记法， $\alpha = (\alpha_1, \dots, \alpha_N)$ 、 $x_\alpha = \prod_i x_{i\alpha_i}$和 $\psi_\alpha = \bigotimes_i \psi_{i\alpha_i}$。 因此，稍微滥用一下记号，玩家 $i \in \mathcal{N}$ 在$\mathbf{X}$ 下的期望收益将为

$$\begin{aligned} u_i(\mathbf{X}) &= \sum_{\omega\in\Omega} \sum_{\alpha\in\mathcal{A}} x_\alpha P_\omega(\psi_\alpha)\, U_i(\omega) = \sum_{\alpha\in\mathcal{A}} x_\alpha u_i(\psi_\alpha) \\ &= \sum_{\alpha_1\in\mathcal{A}_1} \cdots \sum_{\alpha_N\in\mathcal{A}_N} x_{\alpha_1}\cdots x_{\alpha_N}\, u_i(\psi_{\alpha_1}, \dots, \psi_{\alpha_N}). \end{aligned} \tag{4}$$

**与其他类别的博弈相比**。 玩家在混合态下的期望收益表达式（4）让人联想到经典有限博弈的混合扩展，但这种联系十分牵强。从概念角度看，主要区别如下：

**(1)** 纯态构成无限连续体 $\psi \in \Psi$，而非有限个（有限博弈即是如此）。**(2)** 将密度矩阵分解为纯态的方式（3）并不唯一；一般而言，可能存在连续多个（彼此不等价的）纯态族及混合权重，它们都能产生相同的密度矩阵。**(3)** 两个纯态的凸叠加 $\lambda\psi + (1-\lambda)\psi'$ $\psi$ 和 $\psi'$ 可能会在诱导收益中产生形如 $|\psi\rangle\langle\psi'|$ 和 $|\psi'\rangle\langle\psi|$ 的量子干涉项；这些交叉项在有限博弈中没有对应物。

基于上述原因，将量子博弈视为有限博弈的“张量”扩展可能会产生误导。相反，更恰当的做法是将量子博弈理解为一种连续博弈 ，其中每位玩家 $i \in \mathcal{N}$ 控制一个$\mathbf{X}_i$ 矩阵变量，该变量取自“谱单纯形”

$$\boldsymbol{\mathcal{X}}_i = \{\mathbf{X}_i \in \mathbb{H}_+^{d_i} : \mathrm{tr}\,\mathbf{X}_i = 1\} \tag{5}$$

且玩家的收益函数 $u_i\colon \boldsymbol{\mathcal{X}} \equiv \prod_j \boldsymbol{\mathcal{X}}_j \to \mathbb{R}$ 对每位玩家的密度矩阵$\mathbf{X}_j \in \boldsymbol{\mathcal{X}}_j$均为线性的， $j \in \mathcal{N}$。

**纳什均衡**。 在我们的量子情境中，纳什均衡（NE）这一经典解概念刻画了这样的混合量子态$\mathbf{X}^* \in \boldsymbol{\mathcal{X}}$： 它们会抑制单边偏离，其含义是

$$u_i(\mathbf{X}^*) \geq u_i(\mathbf{X}_i; \mathbf{X}^*_{-i}) \quad \text{for all } \mathbf{X}_i \in \boldsymbol{\mathcal{X}}_i,\ i \in \mathcal{N} \tag{NE}$$

其中，我们用 $(\mathbf{X}_i;\ \mathbf{X}_{-i}) = (\mathbf{X}_1, \dots\dots, \mathbf{X}_i, \dots\dots, \mathbf{X}_N)$ 表示玩家 $i$ 相对于所有其他参与者的选择。由于 $\boldsymbol{\mathcal{X}}_i$ 是凸的，且 $u_i$ 关于 $\mathbf{X}_i$ 是线性的，因此根据 Debreu [14] 的开创性定理，纳什均衡的存在性得以证明。

现在，令

$$\mathbf{V}_i(\mathbf{X}) = \nabla_{\mathbf{X}_i^\top} u_i(\mathbf{X}) \tag{6}$$

表示玩家 $i$ 的个体收益梯度，标准论证 [16, 40] 表明，量子博弈的纳什均衡 $\mathcal{Q}$ 恰好是以下变分不等式的解

$$\operatorname{tr}[\mathbf{V}(\mathbf{X}^*)(\mathbf{X} - \mathbf{X}^*)] \leq 0 \quad \text{for all } \mathbf{X} \in \boldsymbol{\mathcal{X}} \tag{VI}$$

其中 $\mathbf{V}(\mathbf{X}) = (\mathbf{V}_1(\mathbf{X}) \dots, \mathbf{V}_N(\mathbf{X}))$。这里需要注意，由于 $u_i$ 关于 $\mathbf{X}_i$ 是线性的，因此 $\mathbf{V}_i(\mathbf{X})$（即 $\mathbf{V}(\mathbf{X})$ 的第 $i$ 个分块）不依赖于$\mathbf{X}_i$，且有

$$u_i(\mathbf{X}_i; \mathbf{X}_{-i}) = \operatorname{tr}[\mathbf{X}_i \mathbf{V}_i(\mathbf{X})] \quad \text{for all } \mathbf{X} \in \boldsymbol{\mathcal{X}}. \tag{7}$$

在下文中，我们将直接运用这些性质。

**遗憾**。 作为纳什均衡概念的补充，动态环境中的一项重要理性要求是最小化玩家的遗憾，即玩家随时间获得的预期累积收益与事后最佳固定状态所能带来的收益之间的差距。形式上，针对博弈轨迹$\mathbf{X}(t) \in \boldsymbol{\mathcal{X}}$， $t \geq 0$，第 $i$ 位玩家的遗憾定义为

$$\operatorname{Reg}_i(T) = \max_{\mathbf{X}_i' \in \boldsymbol{\mathcal{X}}_i} \int_0^T [u_i(\mathbf{X}_i'; \mathbf{X}_{-i}(t)) - u_i(\mathbf{X}(t))]\, dt \tag{8}$$

若 $\operatorname{Reg}_i(T) = o(T)$，则称玩家 $i$ 具有无遗憾 性质。在本文余下部分，我们将关注不会产生遗憾的学习动态，并考察其相对于博弈纳什均衡的收敛性质。

## 3. 学习动态

**通过量子正则化进行学习**。 在经典有限博弈中，研究最广泛的一类无悔动力学是所谓的“跟随正则化领导者”（FTRL）算法系列 [30, 41, 42]。这一广受欢迎的范式背后的核心思想如下：在每个时刻 $t \geq 0$，每位玩家 $i \in \mathcal{N}$ 都会选择一种混合策略，使其累积收益减去某个正则化惩罚后达到最大。如此一来，表现持续较好的策略往往会优先于表现较差的策略，而“正则化惩罚”则会引入一定程度的探索，以避免陷入停滞。

在量子情形中，混合策略由博弈的混合量子态取代，因此，FTRL 背后的强化机制会导出矩阵值的“跟随量子正则化领导者”（FTQL）动力学

$$\mathbf{X}_i(t) = \arg\max_{\mathbf{X}_i \in \boldsymbol{\mathcal{X}}_i} \left\{ \int_0^t u_i(\mathbf{X}_i; \mathbf{X}_{-i}(s))\, ds - h_i(\mathbf{X}_i) \right\} \tag{9}$$

其中 $h_i\colon \boldsymbol{\mathcal{X}}_i \to \mathbb{R}$ 表示玩家 $i$ 的惩罚函数 ——或正则化项 ——（下文将详细讨论）。如前所述，FTQL 动力学（9）采用积分形式，这并不特别适合我们的分析。为此，首先考虑正则化最佳响应映射，以得到更具体的自治重构

$$\mathbf{Q}_i(\mathbf{Y}_i) := \arg\max_{\mathbf{X}_i \in \boldsymbol{\mathcal{X}}_i} \{\operatorname{tr}[\mathbf{Y}_i \mathbf{X}_i] - h_i(\mathbf{X}_i)\} \tag{10}$$

对所有厄米 $\mathbf{Y}_i \in \boldsymbol{\mathcal{Y}}_i := \mathbb{H}^{d_i}$ 成立。因此，鉴于 (7)，积分动力学 (9) 可以改写为如下微分形式

$$\dot{\mathbf{Y}}_i(t) = \mathbf{V}_i(\mathbf{X}(t)) \qquad \mathbf{X}_i(t) = \mathbf{Q}_i(\mathbf{Y}_i(t)). \tag{FTQL}$$

动力学 (FTQL) 将作为我们分析的基础，因此有必要作一些说明。首先，从解释的角度看，(FTQL) 可视为一个与正则化状态选择方案相结合的梯度跟随过程——分别对应动力学 $Y_i = V_i$ 和映射 $Q_i$: $Y_i \mapsto X_i$。在这方面，作为 $Q_i$ 定义基础的正则化项 $h_i$ 起着至关重要的作用，而对 $h_i$ 的不同选择可能导致截然不同的动力学。为明确起见，以下我们仅假设 $h_i$ 是具有如下形式的迹函数： $h_i(X_i) = \operatorname{tr}[\theta_i(X_i)]$，其中 $\theta_i$: $[0,1] \to \mathbb{R}$ 在 $[0,1]$ 上连续，且 $\inf_{x\in(0,1]} \theta_i''(x) > 0$。我们还称当 $\lim_{x\to 0^+} \theta_i'(x) = -\infty$ 时，$\theta_i$ 是陡峭的；并出于归一化目的，假设 $\theta_i(0) = 0$。

为简便起见，省略玩家指标，以下是一些标准正则化项的示例：

示例 3.1（$L^2$ 正则化）。若 $\theta(x) = x^2/2$，玩家的惩罚函数为弗罗贝尼乌斯范数平方 $h(X) = (1/2)\|X\|_F^2$，此时公式 (10) 给出正交投影算子 $Q(Y) = \mathbf{\Pi}_{\mathcal{X}}(Y) \equiv \arg\min_{X\in\mathcal{X}} \|X - Y\|_F$。这一选择导出（弗罗贝尼乌斯）投影动力学：

$$\dot{\mathbf{Y}} = \mathbf{V}(\mathbf{X}) \qquad \mathbf{X} = \mathbf{\Pi}_{\mathcal{X}}(\mathbf{Y}) \tag{PD}$$

**示例 3.2** （冯·诺伊曼正则化）。另一种标准选择是 $\theta(x) = x\log x$，它导出（负）冯·诺伊曼熵 $h(\mathbf{X}) = \operatorname{tr}[\mathbf{X}\log\mathbf{X}]$。通过标准计算可知，这种正则化项会产生矩阵乘法／指数权重（MMW）动力学

$$\dot{\mathbf{Y}} = \mathbf{V}(\mathbf{X}) \qquad \mathbf{X} = \frac{\exp(\mathbf{Y})}{\operatorname{tr}[\exp(\mathbf{Y})]} \tag{MMW}$$

Tsuda 等人 [46] 和 Kakade 等人 [24]在核学习的背景下引入了这类动力学的离散时间版本；关于量子学习背景下较近的一系列进展，参见 [2, 22, 23]。

例 3.3（Tsallis 正则化）。在上述两者之间进行插值，Tsallis正则化项定义为 $\theta(x) = [q(1-q)]^{-1}(x - x^q)$，其中 $q > 0$ （采用连续性约定（$x - x^q$）/$(1-q) = x\log x$ 当 $q = 1$时）。当 $q \leftarrow 2$时，我们得到投影动力学（PD）；相较之下，选择 $q \leftarrow 1$ 会产生（MMW）；最后，选择 $q \leftarrow 1/2$ 在老虎机在线学习背景下尤其常见，参见 [24, 51] 及其中的参考文献。

**FTQL 的混合态动力学**。在（FTQL）下，参与者混合态的演化 $\mathbf{X}(t)$ 可通过辅助得分矩阵 $\mathbf{Y}(t)$的演化隐式地描述。**另一方面，要得到 $\mathbf{X}(t) = \mathbf{Q}(\mathbf{Y}(t))$ 动力学的显式表达式则困难得多**，因为矩阵微积分的规则无法为张量导数 $\nabla_{\mathbf{Y}}\mathbf{Q}(\mathbf{Y})$ 给出解析表达式，即使 $\mathbf{Q}$本身具有闭式表达式。

为避免这一困难，我们将采用如下形式的$\mathbf{X}$ 的酉特征分解

$$\mathbf{X} = \sum\nolimits_{\alpha=1}^{d} x_\alpha \mathbf{u}_\alpha \mathbf{u}_\alpha^\dagger \tag{11}$$

其中 $x_\alpha \geq 0, \alpha = 1$、……、 $d$ 是$\mathbf{X}$特征值的枚举，$\mathbf{u}_\alpha \in \mathcal{H}$ 是$\mathbf{X}$ 对应于 $x_\alpha$ 的单位范数特征向量，玩家指标同样省略（以简化记号）。一般而言，$\mathbf{X}$ 与$\mathbf{V}$ （因而也与$\mathbf{X}$）不对易，因此$\mathbf{X}$ 的特征值和特征向量将以耦合的方式同步演化；下面的首个结果给出了这种协同演化的显式表达式：

定理 *1.* 设 $X(t) = \sum_{\alpha=1}^{d} x_\alpha(t)u_\alpha(t)u_\alpha^\dagger(t)$ 是 $X(t)$ 按照式 *(11)* 进行的特征分解，并假设 $h$ 是陡峭的。那么，在 *(FTQL)* 下，$\dot{X}$ 的元素 $[\dot{X}]_{\alpha\beta} = u_\alpha^\dagger \dot{X} u_\beta$ 遵循如下量子态动力学：

$$[\dot{\mathbf{X}}]_{\alpha\beta} = \begin{cases} \dfrac{V_{\alpha\alpha}}{\theta''(x_\alpha)} - \dfrac{\sum_\kappa V_{\kappa\kappa}/\theta''(x_\kappa)}{\sum_\kappa \theta''(x_\alpha)/\theta''(x_\kappa)} & \alpha = \beta \\ \dfrac{x_\beta - x_\alpha}{\theta'(x_\beta) - \theta'(x_\alpha)} V_{\alpha\beta} & \alpha \neq \beta \end{cases} \tag{QD}$$

其中 $V_{\alpha\beta} = \mathbf{u}_\alpha^\dagger \mathbf{V}(\mathbf{X})\mathbf{u}_\beta$ 并且我们采用连续性约定 $(x-y)/[\theta'(x) - \theta'(y)] = 1/\theta''(x)$ 当 $y \to x$ 时。

定理 1 的证明基于这样一个事实：如果厄米矩阵 $\mathbf{A}$ 的特征分解为 $\mathbf{A} = \sum_\alpha \lambda_\alpha \mathbf{u}_\alpha \mathbf{u}_\alpha^\dagger$ 遵循动力学 $\dot{\mathbf{A}} = \mathbf{M}$，对特征值方程 $\mathbf{A}\mathbf{u}_\alpha = \lambda_\alpha \mathbf{u}_\alpha$ 求导

$$\dot{\lambda}_\alpha \delta_{\alpha\beta} = (\lambda_\alpha - \lambda_\beta)\mathbf{u}_\alpha^\dagger \dot{\mathbf{u}}_\beta + \mathbf{u}_\alpha^\dagger \mathbf{M} \mathbf{u}_\beta \quad \text{for all } \alpha, \beta. \tag{12}$$

这一恒等式使我们能够通过求解与最大化问题（$\mathbf{X}$ 相关的拉格朗日函数，分析并推导出 10）的表达式。完整证明见附录；这里我们将转而重点讨论一个具有代表性的例子。

**例 3.4** （量子复制子动力学）。 动力学（QD）的一个重要特例可通过 von Neumann 正则化器 $h(\mathbf{X}) = \mathrm{tr}[\mathbf{X}\log\mathbf{X}]$（来自 例 3.2）得到。这给出了量子复制子动力学

$$[\dot{\mathbf{X}}]_{\alpha\beta} = \begin{cases} x_\alpha[V_{\alpha\alpha} - \sum_\kappa x_\kappa V_{\kappa\kappa}] & \text{for } \alpha = \beta \\ \dfrac{x_\beta - x_\alpha}{\log x_\beta - \log x_\alpha} V_{\alpha\beta} & \text{for } \alpha \neq \beta \end{cases} \tag{QRD}$$

分别对应 $\mathbf{X}$ 的对角元素和非对角元素。(QRD) 的对角部分在形式上类似于演化博弈论中的复制子动力学 [39, 45, 48]，并刻画了 $\mathbf{X}(t)$的特征值演化。因此，与其非对角分量结合起来看，(QRD) 给出了 (MMW) 下混合态演化的显式表达式。或者，直接对 (MMW) 应用弗雷歇微分公式，(QRD) 可改写为无基表示法：

$$\dot{\mathbf{X}} = \int_0^1 \mathbf{X}^{1-s}\mathbf{V}(\mathbf{X})\mathbf{X}^s \, ds - \mathrm{tr}[\mathbf{X}\mathbf{V}(\mathbf{X})]\mathbf{X} \tag{13}$$

动力学 (QRD) 和无坐标表达式 (13) 与 Jain 等人的动力学一致 [23]（他们在假设 $\mathbf{X}$ 与 $\dot{\mathbf{X}}$ 可交换的前提下推导出了等价表达式），但与 Hidalgo 的动力学不同 [20]；后者源自一种不同且无关的量子化范式。相关计算见附录 B。 ◊

在接下来的章节中，我们将详细考察 (FTQL) 的经典与量子分量如何相互作用，共同决定玩家的长期行为。

## 4. 遗憾最小化

我们首先分析（FTQL），从其动力学的遗憾最小化性质入手。为提供必要的背景，已知在经典的有限博弈 [25, 32]中，FTRL 动力学所产生的遗憾至多为常数。如下面所示，尽管矩阵值动力学（FTQL）是在由纯态构成的连续统上运行的，但其仍然具备

量子博弈中相同的遗憾最小化保证。形式化地说，我们有以下结果。

**命题 1**。 设 $\mathbf{X}(t) = \mathbf{Q}(\mathbf{Y}(t))$， $t \geq 0$ 是由 (FTQL) 诱导的博弈轨迹。那么，对于所有 $T \geq 0$，我们有

$$\mathrm{Reg}_i(T) \leq |d_i \cdot \theta_i(1/d_i) - \theta_i(1)|. \tag{14}$$

命题 1 的证明基于 [25] 的一般理论，并在附录C中详细给出。这里仅指出，界（14）对博弈维度的依赖很大程度上取决于正则化项的选择：(i) 欧几里得正则化项 $\theta(x) = x^2/2$ 给出一个 $\mathcal{O}(1)$ 界；(ii) 冯·诺伊曼熵导致 $\mathcal{O}(\log d)$ 依赖；最后，(iii) 对于 Tsallis正则化项 $\theta(x) \propto x - x^q$，当所有 $q \in (0, 1)$ 时，其界为 $\mathcal{O}(d^{1-q})$。这一点应与离散时间的在线学习模型形成对比，在那些模型中，二次正则化相较于熵变体和 Tsallis 变体都会导致次优结果 [26, 44, 51]。这种离散到连续的差距源于这样一个事实：(FTQL) 具有一个精确的能量函数，即所谓的 Fenchel 耦合

$$F_i(\mathbf{X}_i, \mathbf{Y}_i) = h_i(\mathbf{X}_i) + h_i^*(\mathbf{Y}_i) - \mathrm{tr}[\mathbf{X}_i \mathbf{Y}_i] \tag{15}$$

其中 $\mathbf{X}_i \in \boldsymbol{\mathcal{X}}_i$、$\mathbf{Y}_i \in \boldsymbol{\mathcal{Y}}_i$，以及

$$h_i^*(\mathbf{Y}_i) = \max_{\mathbf{X}_i \in \boldsymbol{\mathcal{X}}_i} \{\mathrm{tr}[\mathbf{X}_i \mathbf{Y}_i] - h_i(\mathbf{X}_i)\} \tag{16}$$

表示 $h_i$ 的凸共轭。这种原始-对偶耦合的一个版本最早由 [30] 在有限博弈的背景下提出，并具有以下基本性质：

**引理 1.** 设 $\mathbf{X}(t) = \mathbf{Q}(\mathbf{Y}(t))$， $t \geq 0$ 是由 (FTQL) 所诱导的博弈轨迹。则对于所有 $\mathbf{P} \in \boldsymbol{\mathcal{X}}$，我们有

$$\frac{d}{dt} F_i(\mathbf{P}_i, \mathbf{Y}_i(t)) = \mathrm{tr}[\mathbf{V}_i(\mathbf{X}(t)) \, (\mathbf{X}_i(t) - \mathbf{P}_i)]. \tag{17}$$

该引理的重要性（证明见 附录 A）在于，(17) 的右端恰好是遗憾的被积函数，因此，命题 1 可通过将 引理 1应用于定义 (8) 中使玩家 $i$ 的遗憾达到最大值的状态而得出。相关计算留待 附录C。

## 5. 收敛、稳定性与民间定理

鉴于命题 1对遗憾最小化提供了强有力的保证，矩阵值动力学（FTQL）从学习的角度来看是一个极具吸引力的选择。与此同时，即使在经典有限博弈中，遗憾最小化也不足以排除非理性化结果：例如，正如 Viossat 和 Zapechelnyuk 所证明的那样， [47], 采用无遗憾策略时，参与者的博弈经验频率仍可能对（严格）被支配策略赋予正的选择概率，甚至仅对被支配策略赋予正的选择概率。因此，本节余下部分将着重考察（FTQL）相对于博弈纳什均衡的收敛性和稳定性。

稳定性与收敛性的概念。我们的分析需要用到动力系统理论中的一些基本概念，下面将简要介绍。首先，回顾一下，抽象度量空间 $\mathcal{Z}$ 上的流是一个连续映射 $\phi\colon \mathbb{R}\times\mathcal{Z}\to\mathcal{Z}$，满足 (a) $\phi_0(z)=z$；以及 (b) 对所有 $t$、$s\in\mathbb{R}$ 和 $z\in\mathcal{Z}$，均有 $\phi_{t+s}(z)=\phi_t(\phi_s(z))$。简而言之，流通常由适定常微分方程组（ODE）的解轨道生成，例如（FTQL）：在这种解释下，$\phi_t(z)$ 表示从时刻 $t=0$ 位于 $z$ 的常微分方程解在时刻 $t$ 的位置。

基于此，以下不变性与稳定性概念将在我们的分析中发挥关键作用。给定一点 $p\in\mathcal{Z}$，我们称：

(1) 若对所有 $t\in\mathbb{R}$，$\phi_t(p)=p$，则称 $p$ 驻定。(2) 如果对于 $\mathcal{Z}$ 中 $p$ 的每个邻域 $\mathcal{U}$，都存在 $\mathcal{Z}$ 中 $p$ 的某个（更小的）邻域 $\mathcal{U}'$，使得对所有 $t\in\mathbb{R}$，$\phi_t(\mathcal{U}')\subseteq\mathcal{U}$，则称 $p$ 是李雅普诺夫稳定的 —— 或简称稳定的。换言之，如果从足够靠近 $p$ 的位置出发的任何轨道始终保持足够靠近，则称 $p$ 是稳定的。(3) 如果存在一个邻域 $\mathcal{U}$，使得对所有 $z\in\mathcal{U}$，$\lim_{t\to\infty}\phi_t(z)=p$，则称 $p$ 是吸引的。换言之，如果所有邻近轨道都收敛到 $p$，则称 $p$ 是吸引的。(4) 如果 $p$ 既稳定又吸引，则称其渐近稳定。

接下来，我们将寻求精确刻画（FTQL）的稳定状态和／或吸引状态。

**经典情形**。为后续分析做好铺垫，回顾经典有限博弈中的经典学习情形将很有帮助。为具体起见，我们聚焦于标准的复制子动力学情形（参见上文的 例 3.2 和 3.4 ），一般有限博弈的稳定性与收敛图景可以由演化博弈论中所谓的“民间定理”概括，该定理陈述如下 [10, 21]:

(1) 纳什均衡是驻定的。(2) 内部轨道的极限是纳什均衡。(3) 李雅普诺夫稳定状态是纳什均衡。(4) 一个状态渐近稳定，当且仅当它是严格纳什均衡（也就是说，在均衡处，每个玩家都有唯一的最佳回应）。

除了一些技术性细节外，这些性质也适用于经典有限博弈中的整类跟随正则化领导者动力学，参见 [17, 30], 及其中的参考文献。然而，量子博弈中玩家状态空间的非线性几何，对（FTQL）的非交换矩阵值情形中哪些性质能够保留，施加了严格的结构性限制。我们将在下文探讨这一问题。

**谱单纯形中的正则化学习**。快速考察量子复制子动力学 (QRD) 可以发现以下结构性质：$\mathbf{X}$ 在 (QRD) 中初始为零的特征值将始终为零；同样，初始为正的特征值也将始终为正。形式上，这意味着核$\ker(\mathbf{X})$ 在 $\mathbf{X}$ 在 (QRD) 的作用下保持不变；因此，鉴于厄米矩阵的线性张成是其核的正交补，$\operatorname{im}(\mathbf{X})$ 也具有同样的性质。

密度矩阵的核——或等价地，其像——在（QRD）下保持不变，这与混合策略组合的支撑集在标准复制子动力学下保持不变互为量子类比。在有限博弈中，这种不变性的一个直接推论是所有 纯策略组合都是驻定的（作为单纯形的零维面）。这一性质也适用于（QRD）

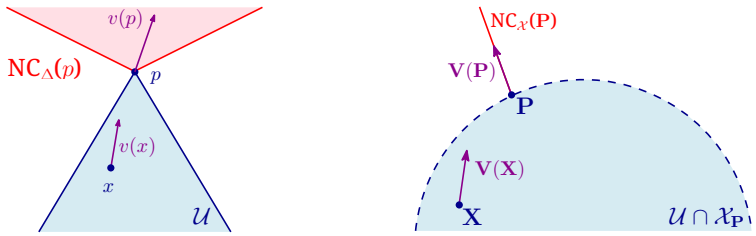


**图 1：** 经典情形与量子情形（分别为左图和右图）之间根本性的几何差异。在经典有限博弈中，单纯形在纯策略 $p$ 处的法锥 $\mathrm{NC}_\Delta(p)$具有非空的拓扑内部，因此纯态纳什均衡通常是稳健的：如果 $v(p)$ 在 $p$ 处垂直于 $\Delta$，那么经过微小扰动后，它在 $\Delta$ 处仍将保持垂直。另一方面，在量子博弈中，谱单纯形在纯态$\mathbf{P}$ 处的法锥 $\mathrm{NC}_{\boldsymbol{\mathcal{X}}}(\mathbf{P})$是一条射线，因此纯态纳什均衡无法具有这种意义上的稳健性。我们还注意到纯态的几何形态有所不同：在单纯形中，纯策略组合是孤立的极点；在谱单纯形中，纯态构成一个连续流形。

事实上，也适用于所研究的整个混合态动力学类：形式上，在（QD）下，所有纯量子态都是驻定的。

尽管如此，量子情形与经典情形之间的主要定性差异在于，量子博弈中存在连续统个纯态，即由秩为 1 的密度矩阵构成的整个流形 $\Psi$ （球面的乘积）。相比之下，在有限博弈中，纯态是由玩家的纯策略张成的单纯形 $\Delta$ 的顶点，因此数量有限且彼此孤立。因此，在经典有限博弈中，纯策略组合可以渐近稳定；而在量子博弈中，由于每个纯态周围都有其他不变态，则不可能。

第二个主要区别在于，在有限博弈中，严格纳什均衡具有鲁棒性：对博弈收益进行微小扰动，不会改变该博弈的严格均衡。在量子博弈中，这种鲁棒性不复存在：事实上，纳什均衡的变分刻画（VI）意味着 $\mathbf{V}(\mathbf{X}^*)$ 必须属于法锥$\boldsymbol{\mathcal{X}}$ 在 $\mathbf{X}^*$处的法锥；然而，谱单纯形在秩为 1 的矩阵处的法锥，其拓扑内部为空，因此所要求的隶属性质不可能具有鲁棒性（图示见图 1）。特别是，对量子博弈收益的任何一般性扰动，无论多么微小，都可能导致相关均衡在纯态流形 $\Psi$上发生微小偏移。

**一致性与变分稳定性。** 鉴于上述内容，我们已经可以得出关于量子情境的两个主要结论：

(1) 渐近稳定性的任何概念都必须包含一致性的概念：如果某个状态不在线性张成的动力学初始状态中，就无法达到该状态。(2) 鲁棒性的任何概念同样必须纳入变分稳定性的概念：对平衡态的微小扰动应当趋于使其恢复。

我们将这两个想法形式化如下：

**定义 1.** 固定一个状态 $\mathbf{P} \in \boldsymbol{\mathcal{X}}$，并令

$$\boldsymbol{\mathcal{X}}_{\mathbf{P}} \coloneqq \{\mathbf{X} \in \boldsymbol{\mathcal{X}} : \ker(\mathbf{X}) \leq \ker(\mathbf{P})\} \tag{18}$$

表示 $\mathbf{P}$ 在$\boldsymbol{\mathcal{X}}$中的一致性域，即线性张成包含$\mathbf{P}$的线性张成的混合态集合。那么，给定一个流 $\boldsymbol{\chi}$: $\mathbb{R} \times \boldsymbol{\mathcal{X}} \to \boldsymbol{\mathcal{X}}$，我们将这样表述：

(1) 若存在一个邻域 $\mathbf{P}$，使得对所有 $\mathbf{X} \in \boldsymbol{\mathcal{U}} \cap \boldsymbol{\mathcal{X}}_{\mathbf{P}}$，都有 $\lim_{t\to\infty} \boldsymbol{\chi}_t(\mathbf{X}) = \mathbf{P}$，则称$\mathbf{P}$ 具有一致吸引性。换言之，若$\mathbf{P}$ 能吸引所有邻近的一致初始化，则称其具有一致吸引性。

(2) 若$\mathbf{P}$是李雅普诺夫稳定且具有一致吸引性，则称其一致渐近稳定。

**定义 2**。 若存在$\mathbf{X}^*$ 在$\boldsymbol{\mathcal{X}}$ 中的一个邻域$\boldsymbol{\mathcal{U}}$，使得

$$\operatorname{tr}[\mathbf{V}(\mathbf{X})(\mathbf{X} - \mathbf{X}^*)] < 0 \quad \text{for all } \mathbf{X} \in \boldsymbol{\mathcal{U}} \backslash \{\mathbf{X}^*\}. \tag{VS}$$

直观而言，定义 1恰好刻画了我们上文讨论的可达性条件：$\boldsymbol{\mathcal{X}}_{\mathbf{P}}$ 是一个稠密凸集，由 $\mathbf{X}$ 中所有包含$\mathbf{P}$ 的面的相对内部组成（包括$\boldsymbol{\mathcal{X}}$ 本身）。至于定义 2，变分稳定性应被视为一种均衡精炼，其思想源自奠定有限博弈“民间定理”基础的经典概念——演化稳定性 [28, 29]; 关于连续博弈背景下的（VS），详见 [31]。

备注。 在定义 1中，李雅普诺夫稳定性的概念并未附带“一致性”这一限定。正如我们在 附录 D中所讨论的，这是因为对于李雅普诺夫稳定性，这两个概念最终是一致的，因此无需加以区分。 ◊

**量子“民间定理”**。 现在我们可以陈述本节的主要结果。为简化表述，以下假设 (FTQL) 使用陡峭正则化函数；并且根据 定理 1，量子态动力学 (QD) 指的是由 (FTQL) 在 $\boldsymbol{\mathcal{X}}$ 上诱导的流。

**定理 2**。 设 $\mathcal{Q} \equiv \mathcal{Q}(\mathcal{N}, \Psi, u)$ 是一个量子博弈，固定某个状态 $\mathbf{X}^* \in \boldsymbol{\mathcal{X}}$，并设$\mathbf{X}(t) = \mathbf{Q}(\mathbf{Y}(t))$ 是由 (FTQL) 在使用陡峭正则化函数时诱导的博弈轨迹。那么：

*(1)* 若 $X^*$ 是纳什均衡，则它是 *(QD)* 的静止点。*(2)* 若 $X(t) \to X^*$ 当 $t \to \infty$ 时，$X^*$ 是 $\mathcal{Q}$ 的纳什均衡。*(3)* 若 $X^*$ 在 *(QD)* 下稳定，则它是 $\mathcal{Q}$ 的纳什均衡。*(4)* 若 $X^*$ 一致渐近稳定，则它是纯态。*(5)* 若 $X^*$ 满足 *(VS)*，则它一致渐近稳定。

在讨论定理 2的证明之前，有必要先作几点说明。或许最重要的一点涉及该定理中关于渐近稳定性的部分（可以说，这也是经典民间定理最显著的要点）。在这里，尽管博弈状态空间的几何结构排除了标准意义上的渐近稳定性，FTQL仍实现了次优的结果：根据定义，一致渐近稳定的状态会吸引附近除零测集之外的所有初始条件，而定理 2 表明，只有纯态才具有这一性质。这一选择结果对量子博弈有重要意义，因为它表明，正则化学习实质上会将一个初始混合态“坍缩”为某个特定的纯态——尽管任何混合态都可以由无穷多种纯态组合制备而成。

另一方面，变分稳定态也（持续地）渐近稳定这一含义，为（FTQL）提供了一个相关的收敛判据，并表明其对参与者信念和预测的变化具有内在的鲁棒性。具体而言，

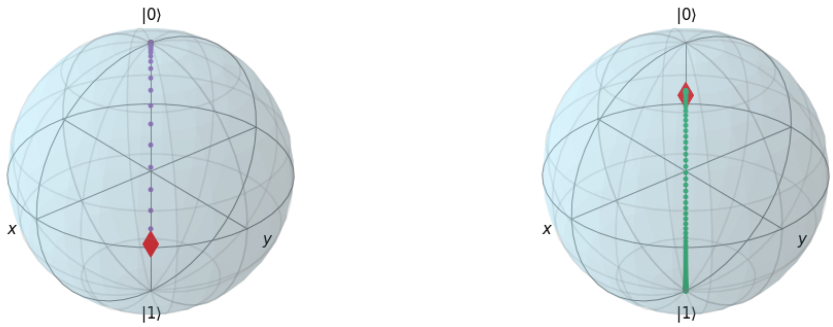


**图 2：** 量子反协调博弈中向变分稳定均衡的收敛。每位玩家的轨迹都表示在其各自的布洛赫球上；详见 附录 F 。

由于（VS）仅涉及底层博弈的基本要素，这类状态在 所有 FTQL 动力学下都具有吸引性，意味着它们可以被视为普适吸引子——而且由于只有纯态才具有这一性质，我们也可以间接推断，变分稳定态 自然也 是纯态。

**证明概要**。 我们以简要概述最后两部分的证明来结束本节，相关内容见 定理 2 （完整证明见 附录 D）。首先，只有纯态才能成为稳定吸引子的这一推论，取决于以下体积保持原理：

命题 *2*。设 $W \subseteq Y$ 是 *(FTQL)* 的一个初始条件开集，并令 $W_t = \psi_t(W)$，$t \geq 0$ 表示 $W$ 在 *(FTQL)* 的流 $\psi\colon \mathbb{R} \times Y \to Y$ 下的演化。那么，对所有 $t \geq 0$，$\mathit{vol}(W_t) = \mathit{vol}(W)$。

命题 2 在我们工作的多个方面都至关重要。首先，这一体积保持性质适用于 所有 量子博弈和 所有 FTQL 动力学，从而将 Jain 等人针对双人零和博弈中的 ( [23] （MMW）/（QRD）的最新分析推广至更一般的情形。这种一般性源于每位玩家的收益函数关于其自身密度矩阵分别呈线性，因此，各自的梯度场 $\mathbf{V}_i$ 不依赖于 $\mathbf{X}_i$； 命题 2 随后可由刘维尔公式推出。

在这方面， 命题 2可视为有限 博弈中跟随正则化领导者动力学普遍体积保持性质的一种量子类比 [9, 17]。然而，与经典世界一样，值得注意的是， 命题 2 讨论的是 FTQL 在 $\mathcal{Y}$ 上的流，而非在 $\mathcal{X}$ 上的诱导流；特别是，$\mathbf{Q}(\mathcal{W}_t)$ 仍可能在 $\mathcal{X}$ 的边界处收敛为一个状态 $\mathcal{W}_t$逃逸至无穷远时。正因如此，尽管命题 2 相对容易排除收敛到满秩均衡，但要排除低秩均衡则需要更精细的论证，其中一致性的概念发挥着重要作用（且在经典世界中没有对应物）。

最后，定理 2 的最后一部分借鉴了李雅普诺夫直接法中的能量论证。与 (FTQL) 的原始-对偶性质相一致，一个自然的能量函数选择是 (15) 中的 Fenchel 耦合 $E(t) = F(\mathrm{X}^*, \mathrm{Y}(t))$，而它又通过引理 1 与 (VS) 相联系。将这两个要素结合起来，便可直接得到 $\lim_{t\to\infty} E(t)$ 存在；随后，通过俘获论证可以证明，存在一个时间序列 $t_n \to \infty$使得 $\mathrm{X}(t_n) \to \mathrm{X}^*$，由此我们可以得出 $E(t) \to 0$ ，最终得到我们的结论。

## 6. 零和博弈中的庞加莱复现

在最后一节中，我们旨在研究 (FTQL) 在零和博弈中的长期行为。零和博弈是一类在应用中经常出现的博弈，也可能是研究最为广泛的量子博弈家族。

在经典有限博弈中，众所周知，复制子动力学——更一般地说，所有 FTRL 动力学——都会产生围绕内部均衡循环的轨迹，这种行为称为庞加莱复现 [32, 35]。形式上，在我们的设定中，如果对于几乎所有初始条件$\mathbf{X}(0) = \mathbf{Q}(\mathbf{Y}(0)) \in \boldsymbol{\mathcal{X}}$（按照勒贝格意义），都存在一个递增的时间序列$t_n \to \infty$ 使得$\mathbf{X}(t_n) \to \mathbf{X}(0)$ 当 $n \to \infty$时成立，则称动力学（FTQL）具有庞加莱复现性；也就是说，（FTQL）的几乎所有轨迹都会无限多次地回到无限接近其起点的位置。

按照我们的记法，双人零和量子博弈 $\mathcal{Q} \equiv \mathcal{Q}(\mathcal{N},\ \Psi,\ u)$ 对应于以下情形：$\mathcal{N} = \{1, 2\}$ 且 $u_1 = -u_2 = -\mathcal{L}$，其中某个极小极大收益函数 $\mathcal{L}$：$\boldsymbol{\mathcal{X}}_1 \times \boldsymbol{\mathcal{X}}_2 \to \mathbb{R}$。在此背景下，Jain 等人 [23] 最近证明，矩阵乘法权重／指数权重动力学（MMW）在所有存在满秩均衡的双人零和博弈中都具有庞加莱复现性。如下文所示，这一性质可以推广到所研究的正则化学习动力学整个类别。

*定理 3.* 设 $\mathcal{Q} \equiv \mathcal{Q}(\mathcal{N}, \Psi, u)$ 是一个存在满秩均衡 $X^* \in \mathcal{X}$的相对内部的双人零和博弈。那么，*(FTQL)* 的几乎每条解轨道 $X(t) = Q(Y(t))$ 都是庞加莱复现的：存在一个递增的时间序列 $t_n \to \infty$，使得 $\lim_{n\to\infty} X(t_n) = X(0)$。

定理 3 的证明包含三个主要要点：首先是命题 2，它表明 (FTQL) 在 $\boldsymbol{\mathcal{Y}}$ 上诱导的流是保体积的。其次，在零和博弈中，起到局部能量函数作用的 Fenchel 耦合（15）变成了一个运动常数（这也是满秩假设发挥作用之处）。第三，除去某种不会影响 $\boldsymbol{\mathcal{X}}$ 上动力学轨迹的商过程后，Fenchel 耦合的等值集是有界的；这是证明中最具挑战性的部分，也是各种相关过程的非交换性使问题复杂得多的地方。具备这些基本要点后，应用庞加莱定理即可得到我们的结果。该定理指出，具有有界轨迹的保体积流是复现的。相关细节见附录 E。

## 7. 总结

当量子计算模型部署于多智能体场景中——从自动驾驶汽车到量子生成对抗网络——参与者之间的交互格局相较于经典交互发生了巨大变化。量子情境下的博弈论学习仍处于起步阶段，因此目前尚不清楚有限理性的量子博弈参与者会呈现何种表现。在这方面，对 FTQL 动力学的研究带来了以下重要洞见：量子态空间的几何结构使“学习陷阱”（稳态）大为增多，而这些陷阱在经典对应物中并不存在；尽管如此，在 FTQL 下能够稳定且具有吸引性的状态，只有该博弈的纯量子均衡。深入理解量子博弈论学习的局限，是一个尤其值得探索的研究方向，并可能对多智能体量子计算系统的部署产生深远影响。

## 附录 A. 正则化最佳响应与 Fenchel 耦合

在本附录中，我们介绍正则化最佳响应映射$\mathbf{Q}$ 和 Fenchel 耦合的基本性质。为简化记号，全文省略玩家指标；形式上，这意味着，例如，$\mathbf{X}$ 表示策略组合$\mathbf{X} = (\mathbf{X}_i)_{i\in\mathcal{N}}$；整体镜像映射$\mathbf{Q}\colon \mathcal{Y} \to \mathcal{X}$表示乘积映射$\mathbf{Q} = \prod_i \mathbf{Q}_i$；定义在$\mathcal{X}$ 上的聚合正则项 $h$ 表示总和 $h = \sum_i h_i$，等等。根据上下文很容易推断出相应的替换，因此无需担心混淆。

我们还假设 V 配备了某种抽象范数 $\|\cdot\|$，该范数不一定由量子态空间 $\mathcal{H}$的希尔伯特结构诱导。此外，出于技术原因，将 $h$视为一个扩展实值函数 $h$：定义在厄米矩阵 $d\times d$构成的整个空间 V $\equiv \mathbb{H}^d$上，即对所有满足 X $\in$ V $\setminus$ X 的 X，赋值 $h$(X) $= +\infty$。按照 Rockafellar [37], 的做法，这使我们能够将 $h$在 X $\in$X 处的次微分定义为

$$\partial h(\mathbf{X}) \coloneqq \{\mathbf{Y} \in \mathcal{Y} : h(\mathbf{X}') \geq h(\mathbf{X}) + \operatorname{tr}[\mathbf{Y}(\mathbf{X}' - \mathbf{X})] \text{ for all } \mathbf{X}' \in \mathcal{V}\}, \tag{A.1}$$

其中，$\mathcal{Y} \equiv \mathcal{Y}$ 在此扮演 $\mathcal{V}$的代数对偶的角色。于是， $h$ 的次可微域定义为

$$\operatorname{dom}\partial h \coloneqq \{\mathbf{X} \in \operatorname{dom} h : \partial h \neq \varnothing\} \tag{A.2}$$

$h$ 的凸共轭 由下式给出

$$h^*(\mathbf{Y}) = \max_{\mathbf{X}\in\mathcal{X}} \{\operatorname{tr}[\mathbf{Y}\mathbf{X}] - h(\mathbf{X})\} \quad \text{for all } \mathbf{Y} \in \mathcal{Y}. \tag{A.3}$$

由于 $h$ 是形如 $h$(X) $= \operatorname{tr}[\theta(\mathbf{X})]$ 的迹函数，且在 $[0,1]$,上强凸，因此可以直接得出，$h$ 也关于 $\|\cdot\|$ 是 $K$-强凸的，其中 $K > 0$为某个正数。于是我们得到以下基本结果：

**引理 A.1.** 设 $h$ 是如上所述定义在 $\mathcal{X}$ 上的 $K$-强凸正则化项，且$\mathbf{Q}\colon \mathcal{Y} \to \mathcal{X}$ 是其诱导的镜像映射。则：

*(1)* $Q$ 在 $Y$ 上是单值的：具体而言，对于所有 $X \in X$、$Y \in Y$，均有 $X = Q(Y) \iff Y \in \partial h(X)$。*(2)* $h$ 的近端域 $X_h \coloneqq \operatorname{im} Q$ 满足 $\operatorname{ri} X \subseteq X_h \subseteq X$。*(3)* $Q$ 是 $(1/K)$-利普希茨连续的，且 $Q = \nabla h^*$。

文献中的多个地方都可以找到这些性质的变体——例如参见 [31]及其中的参考文献——不过为求完整，我们在此简要列出几条线索。

证明。对于 Q 的第一项性质，注意到对于所有 Y $\in$Y，式 (10) 中的最大值都能取到，因为 $h$ 下半连续（l.s.c.）且强凸。此外，根据关于驻点的费马定理，当且仅当 Y $-\partial h$(X) $\ni 0$ 时，X 是式 (10) 的解；也就是说，当且仅当 Y $\in \partial h$(X)。上述结论还表明，$\mathrm{X}_h = \operatorname{dom}\partial h$；由于 ri X $\subseteq$ dom $\partial h \subseteq$X [37,第 26], 章，故得出第二个结论。最后，关于第三个结论，表达式 Q $= \nabla h^*$ 是 Danskin 定理 [43], 的直接推论，而 Q 的 Lipschitz 连续性则由标准的变分分析结果得出， [例如参见文献 38，定理 12.60(b)]。■

Q 的下一些性质更为特殊，并且与谱单纯形的结构密切相关：

**引理 A.2.** 采用引理 *A.1* 中的记号和假设引理 *A.1*，我们有：

(1) $\partial h(\mathbf{X}) = \theta'(\mathbf{X}) + \mathrm{NC}(\mathbf{X})$，其中$\mathrm{NC}(\mathbf{X})$ 表示$\mathcal{X}$ 在$\mathbf{X}$处的法锥；特别地，对于所有$\mathbf{X} \in \operatorname{ri}\mathcal{X}$，有 $\partial h(\mathbf{X}) = \{\theta'(\mathbf{X}) + \lambda\mathbf{I} : \lambda \in \mathbb{R}\}$。

(2) $\mathbf{Y}$ 与$\mathbf{Q}(\mathbf{Y})$ 对所有$\mathbf{Y}\ \in\boldsymbol{\mathcal{Y}}$均可交换。(3) 如果$\mathrm{eig}(\mathbf{Y}) = \{\mathbf{Y}_\alpha\}_{\alpha=1}^d$ 和$\mathrm{eig}(\mathbf{X})$ $= \{x_\alpha\}_{\alpha=1}^d$ 分别表示$\mathbf{Y}\ \in\boldsymbol{\mathcal{Y}}$和$\mathbf{X}\ =\mathbf{Q}(\mathbf{Y})\ \in\boldsymbol{\mathcal{X}}$的特征值，则当$\mathbf{Y}_\alpha - \mathbf{Y}_\beta \to -\infty$ 对于某个$\beta \neq \alpha$时，有 $x_\alpha \to 0$ 。

证明。对于第一个断言，令 $\delta_{\boldsymbol{\mathcal{X}}}$表示 X 的（凸）指示函数，即对于所有 X $\in$X， $\delta_{\boldsymbol{\mathcal{X}}}$(X) $=0$，并且对于所有 X $\in$V $\setminus$X， $\delta_{\boldsymbol{\mathcal{X}}}$(X) $=\infty$。根据标准凸分析论证， [37],我们有 $\partial\delta_{\boldsymbol{\mathcal{X}}}$(X) $=$NC(X)，对于所有 X $\in$X，因此

$$\partial h(\mathbf{X}) = \partial[\mathrm{tr}[\theta(\mathbf{X})] + \delta_{\boldsymbol{\mathcal{X}}}(\mathbf{X})] = \theta'(\mathbf{X}) + \mathrm{NC}(\mathbf{X}), \tag{A.4}$$

我们的断言得证。

对于第二个断言，令 $\mathbf{Y}\ =\mathbf{U}\boldsymbol{\Psi}\mathbf{U}^\dagger$ 为 $\mathbf{Y}$ 的特征分解。于是，(10) 的目标函数可改写为

$$\begin{aligned}\mathrm{tr}[\mathbf{YX}] - h(\mathbf{X}) = \mathrm{tr}[\mathbf{U}\boldsymbol{\Psi}\mathbf{U}^\dagger\mathbf{X}] - \mathrm{tr}[\theta(\mathbf{X})] &= \mathrm{tr}[\boldsymbol{\Psi}\mathbf{U}^\dagger\mathbf{X}\mathbf{U}] - \mathrm{tr}[\mathbf{U}^\dagger\theta(\mathbf{X})\mathbf{U}] \\ &= \mathrm{tr}[\boldsymbol{\Psi}\mathbf{U}^\dagger\mathbf{X}\mathbf{U}] - \mathrm{tr}[\theta(\mathbf{U}^\dagger\mathbf{X}\mathbf{U})]\end{aligned} \tag{A.5}$$

因为$\mathbf{X}$ 和 $\mathbf{U}^\dagger\mathbf{X}\mathbf{U}$ 具有相同的特征值，而 $\mathrm{tr}[\theta(\cdot)]$ 只取决于其自变量的特征值。由于$\boldsymbol{\mathcal{X}}$ 在任意酉矩阵的共轭作用下保持不变（即对于每个酉矩阵 $\mathbf{U}$，$\mathbf{U}^\dagger\boldsymbol{\mathcal{X}}\mathbf{U}\ =\boldsymbol{\mathcal{X}}$），因此

$$\max_{\mathbf{X}\in\boldsymbol{\mathcal{X}}}\{\mathrm{tr}[\mathbf{YX}] - h(\mathbf{X})\} = \max_{\mathbf{X}\in\boldsymbol{\mathcal{X}}}\{\mathrm{tr}[\boldsymbol{\Psi}\mathbf{X} - \theta(\mathbf{X})]\} = \max_{x\in\Delta_d}\left\{\sum_{\alpha=1}^d[\mathbf{Y}_\alpha x_\alpha - \theta(x_\alpha)]\right\} \tag{A.6}$$

其中，按显然的记号， $x_\alpha$ 和$\mathbf{Y}_\alpha$ 分别表示 $\mathbf{X}$ 和 $\mathbf{Y}$ 的特征值。(10) 的最大值在 $\mathbf{Y}$ 和 $\mathbf{X}$ 具有共同的特征向量基时取得，这反过来意味着二者可交换。由于 $\theta$ 根据假设是强凸的，(10) 存在唯一解，因此我们的断言得证。

最后，对于我们的第三个论断，设 $\mathbf{X}_n = \mathbf{Q}(\mathbf{Y}_n)$，其中 $_n \in\boldsymbol{\mathcal{Y}}$ 是某个序列，并分别用 $x_{\alpha,n}$ 和 $\mathbf{Y}_{\alpha,}$ $_n$ 表示 $\mathbf{X}_n$ 和 $\mathbf{Y}_n$ 的特征值。我们要证明，若存在某个 $\beta$ 使得 $\mathbf{Y}_{\alpha,n} -\mathbf{Y}_{\beta,n} \to -\infty$，则 $x_{\alpha,n} \to 0$ 。

现在，必要时取一个子序列，我们不妨假设存在某个 $\varepsilon > 0$ 使得 $x_{\alpha,n} \geq \varepsilon > 0$ 对所有 $n$ 均成立。然后，根据 10 所定义的关系 $\mathbf{Q}(\mathbf{Y})$，可得

$$\sum_{\alpha=1}^d[\mathbf{Y}_{\alpha,n}x_{\alpha,n} - \theta(x_{\alpha,n})] \geq \sum_{\alpha=1}^d[\mathbf{Y}_{\alpha,n}x'_{\alpha,n} - \theta(x'_{\alpha,n})] \tag{A.7}$$

对所有 $x' \in \Delta_d$ 均成立。因此，取 $x'_n = x_n + \varepsilon(e_\beta - e_\alpha)$，即可得到

$$\varepsilon(\mathbf{Y}_{\alpha,n} - \mathbf{Y}_{\beta,n}) \geq \sum_\kappa[\theta(x_{\kappa,n}) - \theta(x'_{\kappa,n})] \geq \min h - \max h \tag{A.8}$$

这与我们最初的假设 $\mathbf{Y}_{\alpha,n} -\mathbf{Y}_{\beta,n} \to -\infty$ 矛盾。由于 $\boldsymbol{\mathcal{X}}$ 是紧集，上述结果表明，对于 $x_n$ 的任意极限点 $x^*$， $x^*_\alpha = 0$ 均成立，证毕。■

有了这些准备，下面我们继续讨论 Fenchel 耦合的基本性质——之所以这样命名，是因为它包含了 Fenchel 不等式的所有项，参见 [31]。我们需要的两个主要性质体现在以下命题中

**命题 A.1**。 沿用 引理 *A.1* 中的记号和假设，我们有：

*(1)* $F(P, Y) \geq (K/2)\ \|Q(Y) - P\|^2$ 对所有 $P\ \in X$、$Y\ \in Y$ 均成立。

*(2)* $\nabla_{\mathbf{Y}^\top}F(P, Y)\ = Q(Y)\ - P$ 对所有 $P\ \in X$、$Y\ \in Y$ 均成立。

证明。 对于我们的第一个论断，令 $\mathbf{X} = \mathbf{Q}(\mathbf{Y})$ 因此 $h^*(\mathbf{Y}) = \operatorname{tr}[\mathbf{Y}\mathbf{X}] - h(\mathbf{X})$ 根据定义（A.3），有 $h^*$。由 $h$ 的强凸性可得

$$h(\mathbf{X}) + t\operatorname{tr}[\mathbf{Y}(\mathbf{P} - \mathbf{X})] \leq h(\mathbf{X} + t(\mathbf{P} - \mathbf{X})) \leq th(\mathbf{P}) + (1-t)h(\mathbf{X}) - \tfrac{1}{2}Kt(1-t)\|\mathbf{X} - \mathbf{P}\|^2, \tag{A.9}$$

从而得到以下界

$$\tfrac{1}{2}K(1-t)\|\mathbf{X} - \mathbf{P}\|^2 \leq h(\mathbf{P}) - h(\mathbf{X}) - \operatorname{tr}[\mathbf{Y}(\mathbf{P} - \mathbf{X})] = F(\mathbf{P}, \mathbf{Y}) \tag{A.10}$$

对所有 $t \in (0, 1]$。令 $t \to 0^+$ 代入（A.10），即可得出我们的断言。

至于我们的第二个结论，引理 A.1 可直接推出

$$\nabla F(\mathbf{P}, \mathbf{Y}) = \nabla_{\mathbf{Y}^\top} h^*(\mathbf{Y}) - \nabla_{\mathbf{Y}^\top} \operatorname{tr}[\mathbf{Y}\mathbf{P}] = \mathbf{Q}(\mathbf{Y}) - \mathbf{P} \tag{A.11}$$

如上所述。 ∎

现在我们可以证明关于引理 1在（FTQL）动力学下 Fenchel 耦合演化的结论；为方便起见，我们在下文重述该结果。

**引理 1**。设$\mathbf{X}(t) = \mathbf{Q}(\mathbf{Y}(t))$， $t \geq 0$是由(FTQL)诱导的博弈轨迹。那么，对于所有$\mathbf{P} \in \mathcal{X}$，都有

$$\frac{d}{dt} F_i(\mathbf{P}_i, \mathbf{Y}_i(t)) = \operatorname{tr}[\mathbf{V}_i(\mathbf{X}(t))\,(\mathbf{X}_i(t) - \mathbf{P}_i)]. \tag{17}$$

证明。根据定义，我们有

$$\begin{aligned}\frac{d}{dt} F_i(\mathbf{P}_i, \mathbf{Y}_i(t)) &= \frac{d}{dt}[h_i(\mathbf{P}_i) + h_i^*(\mathbf{Y}_i(t)) - \operatorname{tr}[\mathbf{Y}_i(t)\mathbf{P}_i]] \\ &= \operatorname{tr}[\dot{\mathbf{Y}}_i(t)\,(\mathbf{X}_i(t) - \mathbf{P}_i)] = \operatorname{tr}[\mathbf{V}_i(\mathbf{X}(t))\,(\mathbf{X}(t) - \mathbf{P})],\end{aligned} \tag{A.12}$$

其中，在最后一行，我们用到了命题 A.1。 ∎

我们还需要的最后一个结果，是一种受 [31] 启发的“互易性”性质，它表明，由 $F$ 在 $\mathcal{X}$ 上诱导的收敛拓扑与通常的拓扑相容。在线性多胞体中，这一性质对于可分解正则项相对容易证明；但在矩阵情形下，问题要复杂得多。

**命题 A.2**。设$\mathbf{Y}_n$, $n = 1, 2, \ldots$, 是$\mathcal{Y}$中的一个序列。那么$\mathbf{X}_n = \mathbf{Q}(\mathbf{Y}_n)$ 收敛到$\mathbf{P} \in \mathcal{X}$当且仅当li$\mathrm{m}_{n\to\infty} F(\mathbf{P}, \mathbf{Y}_n) = 0$。

证明。 我们首先证明直接蕴含“ $\Longrightarrow$”。事实上，假设$\mathbf{X}_n \to \mathbf{P}$，根据 $F$ 的定义可得

$$\begin{aligned}F(\mathbf{P}, \mathbf{X}_n) &= h(\mathbf{P}) + h^*(\mathbf{Y}_n) - \operatorname{tr}[\mathbf{Y}_n\mathbf{P}] \\ &= h(\mathbf{P}) - h(\mathbf{X}_n) + \operatorname{tr}[\mathbf{Y}_n\,(\mathbf{X}_n - \mathbf{P})] \\ &= h(\mathbf{P}) - h(\mathbf{X}_n) + \operatorname{tr}[\theta'(\mathbf{X}_n)(\mathbf{X}_n - \mathbf{P})]\end{aligned} \tag{A.13}$$

其中，在最后一个等式中，我们使用了引理 A.2 的第一部分。由于 $\mathrm{X}_n \to \mathrm{P}$，项$h(\mathrm{P}) - h(\mathrm{X}_n)$ 消失，当 $n \to \infty$ 时，因此只需证明第二项也同样如此。为此，若令 $\mathrm{X}_n = \sum_\alpha x_{\alpha,n}\mathrm{u}_{\alpha,n}\mathrm{u}_{\alpha,n}^\dagger$ 为 $\mathrm{X}_n$ 的特征分解，并设$p_{\alpha,n} = \mathrm{u}_{\alpha,n}^\dagger \mathrm{P}\mathrm{u}_{\alpha,n}$ ， 且 $p_\alpha = \lim_{n\to\infty} p_{\alpha,n}$ （该极限存在，因为根据假设，$\mathrm{X}_n$ 收敛于 P

），我们便可直接得到

$$\begin{aligned}\operatorname{tr}[\theta'(\mathbf{X}_n)(\mathbf{X}_n-\mathbf{P})] &= \sum_{\alpha=1}^{d}\theta'(x_{\alpha,n})(x_{\alpha,n}-p_{\alpha,n}) \\ &= \sum_{\alpha\in\operatorname{supp}(p)}\theta'(x_{\alpha,n})(x_{\alpha,n}-p_{\alpha,n}) + \sum_{\alpha:p_\alpha=0}\theta'(x_{\alpha,n})(x_{\alpha,n}-p_{\alpha,n}) \quad \text{(A.14)}\end{aligned}$$

由于 $\mathbf{X}_n$ 收敛于 $\mathbf{P}$，上面的第一项和趋于零：这是因为 $\lim_{n\to\infty} x_{\alpha,n} > 0$，因此对于所有 $\alpha \in \operatorname{supp}(p)$，$\theta'(x_{\alpha,n})$ 都有有限极限。至于第二项，由于 $x_{\alpha,n} - p_{\alpha,n} \le x_{\alpha,n}$，只需证明 $z\theta'(z) \to 0$ 当 $z \to 0^+$ 时成立；我们将在下面的引理 A.3 中证明这一点，其中 $f = \theta'$。

最后，关于逆向蕴含“$\Longleftarrow$”，我们的断言由 命题 A.1 的第一部分立即得出。 ■

**引理 A.3**。设 $f\colon (0,\ 1] \to \mathbb{R}$ 可微，并满足 *(i)* $\lim_{x\to 0^+} f(x) = -\infty$、$\inf_{x\in(0,1]} f'(x)$ *(ii)* $> 0$，以及 *(iii)* $\int_0^1 |f(x)|\,dx < \infty$。则，$\lim_{x\to 0^+} xf(x) = 0$。

证明。为方便起见，令 $g = -f$。于是，我们很容易得到 (i) $\lim_{x\to 0^+} g(x) = +\infty$、(ii) $\inf_{x\in(0,1]} g'(x) < 0$，以及 (iii) $\int_0^1 |g(x)|\,dx < \infty$。由于 $\inf_{x\in(0,1]} g'(x) < 0$，我们得到 $g$ 严格递减。现在，$\lim_{x\to 0^+} g(x) = +\infty$ 意味着 $g$ 在 0 附近严格为正，即对于某个 $c > 0$，$g(x) > 0$ 在 $(0, c)$ 中。令 $\varepsilon \in (0, c)$。由于 $g$ 在 $(0, c)$ 中递减且为正，我们有：

$$0 \le \varepsilon g(\varepsilon) \le \int_0^{\varepsilon} g(x)\,dx \quad \text{(A.15)}$$

现在我们将证明 $\lim_{\varepsilon\to 0^+} \int_0^{\varepsilon} g(x)\,dx = 0$。为此，我们可以将 $\int_0^{\varepsilon} g(x)\,dx$ 写为：

$$\int_0^{\varepsilon} g(x)\,dx = \int \mathbb{1}_{(0,\varepsilon)}\, g(x)\,dx \quad \text{(A.16)}$$

其中 $\mathbb{1}_A(x) = 1$ 当 $x \in A$ 时，否则为 0。设 $\{\varepsilon_n\}_{n\in\mathbb{N}}$ 是一个正实数序列，且当 $n \to \infty$ 时 $\varepsilon_n \to 0$。则有：

$$\mathbb{1}_{(0,\varepsilon_n)}\, g(x) \xrightarrow{n\to\infty} 0 \quad \text{for all } x \in (0, 1] \quad \text{(A.17)}$$

因为对于固定的 $x \in (0, 1]$，，我们有 $\varepsilon_n < x$ 对所有充分大的 $n$ 都成立，这意味着 $\mathbb{1}_{(0,\varepsilon_n)} g(x) = 0$。此外，对于所有 $x \in (0, 1]$，，都有：

$$|\mathbb{1}_{(0,\varepsilon_n)}\, g(x)| \le |g(x)| \quad \text{(A.18)}$$

其中 $\int_0^1 |g(x)|\,dx < \infty$。因此，根据控制收敛定理 [18]，，我们得到：

$$\lim_{n\to\infty} \int_0^1 \mathbb{1}_{(0,\varepsilon_n)}\, g(x)\,dx = 0 \quad \text{(A.19)}$$

并且，由于 $\{\varepsilon_n\}_{n\in\mathbb{N}}$ 是任意的，我们得出 极限$_{\varepsilon\to 0^+} \int_0^{\varepsilon} g(x)dx = 0$。因此，将其与（A.15）结合，我们得到 极限$_{x\to 0^+} xg(x) = 0$，即 极限$_{x\to 0^+} xf(x) = 0$。 ■

## Appendix B. General properties of the dynamics

定理 *1.* 设 $X(t) = \sum_{\alpha=1}^{d} x_\alpha(t) u_\alpha(t) u_\alpha^\dagger(t)$ 是 $X(t)$ 根据 *(11)* 的特征分解，并假设 $h$ 是陡峭的。那么，在 *(FTQL)* 下，$\dot{X}$ 的条目 $[\dot{X}]_{\alpha\beta} = u_\alpha^\dagger \dot{X} u_\beta$

遵循量子态动力学：

$$[\dot{\mathbf{X}}]_{\alpha\beta} = \begin{cases} \dfrac{V_{\alpha\alpha}}{\theta''(x_\alpha)} - \dfrac{\sum_\kappa V_{\kappa\kappa}/\theta''(x_\kappa)}{\sum_\kappa \theta''(x_\alpha)/\theta''(x_\kappa)} & \alpha = \beta \\ \dfrac{x_\beta - x_\alpha}{\theta'(x_\beta) - \theta'(x_\alpha)} V_{\alpha\beta} & \alpha \neq \beta \end{cases} \tag{QD}$$

其中 $V_{\alpha\beta} = \mathbf{u}_\alpha^\dagger \mathbf{V}(\mathbf{X})\mathbf{u}^\beta$，并且我们采用连续性约定 $(x-y)/[\theta'(x)-\theta'(y)] = 1/\theta''(x)$ 当 $y \to x$ 时。

证明。首先，根据 第 3 节 中描述的动力学，$\mathbf{X}(t)$ 是通过正则化最佳响应得到的，即最大化问题的一个解：

$$\mathbf{Q}(\mathbf{Y}) = \arg\max_{\mathbf{X}' \in \boldsymbol{\mathcal{X}}} \{\mathrm{tr}[\mathbf{Y}\mathbf{X}'] - \mathrm{tr}[\theta(\mathbf{X}')]\}. \tag{B.1}$$

令 $\mathbf{X}(t) = \sum_{\alpha=1}^d x_\alpha(t)\, \mathbf{u}_\alpha(t)\mathbf{u}_\alpha^\dagger(t)$ 为 $\mathbf{X}(t)$ 的特征分解，由于 $h$ 是陡峭的，我们立即得到 $\mathbf{X}(t) \in \mathcal{X}$ 的相对内部，这意味着对所有 $\alpha$，$x_\alpha(t) > 0$。由于 $\inf_{x\in(0,1]} \theta''(x) > 0$，函数 $x \mapsto \theta(x)$ 是严格凸的，因此 $\mathbf{X} \mapsto \mathrm{tr}[\theta(\mathbf{X})]$，也是严格凸的；参见 [6, 定理 2.10]。因此，(B.1) 在 $\mathcal{X}$ 中有唯一解。

根据KKT条件，与半正定约束 $x_\alpha(t) \geq 0$（$\alpha = 1$、.....、$d$）对应的对偶变量在最优解中均为零，因为如前所述，这些不等式是严格的。因此，只需考虑“约化”拉格朗日函数：

$$\mathcal{L}(\mathbf{X}; \lambda) = \mathrm{tr}[\mathbf{Y}\mathbf{X}] - \mathrm{tr}[\theta(\mathbf{X})] - \lambda(\mathrm{tr}[\mathbf{X}] - 1) \tag{B.2}$$

其中，$\lambda \in \mathbb{R}$ 是与约束 $\mathrm{tr}[\mathbf{X}] = 1$对应的对偶变量。然后，对$\mathbf{X}^\top$求$\mathcal{L}$ 的导数，解需要满足 $\nabla_{\mathbf{X}^\top}\mathcal{L}(\mathbf{X}; \lambda) = \mathbf{Y} - \theta'(\mathbf{X}) - \lambda\mathbf{I} = 0$，即：

$$\theta'(\mathbf{X}) = \mathbf{Y} - \lambda\mathbf{I} \tag{B.3}$$

这里我们用到了 $\nabla_{\mathbf{X}^\top}\mathrm{tr}[\theta(\mathbf{X})] = \theta'(\mathbf{X})$，见 [34]。对 B.3 式关于 $t$ 求导，并利用 $\dot{\mathbf{Y}} = \mathbf{V}$，可得：

$$\begin{aligned} \frac{d}{dt}\theta'(\mathbf{X}) &= \dot{\mathbf{Y}} - \dot{\lambda}\mathbf{I} \\ &= \mathbf{V} - \dot{\lambda}\mathbf{I} \end{aligned} \tag{B.4}$$

将 $\theta'(\mathbf{X})$ 写成与$\mathbf{X}$相同的特征基，即 $\theta'(\mathbf{X}) = \sum_{\alpha=1}^d \theta'(x_\alpha)\mathbf{u}_\alpha\mathbf{u}_\alpha^\dagger$，则上式变为：

$$\sum_{k=1}^d \theta''(x_k)\dot{x}_k\mathbf{u}_k\mathbf{u}_k^\dagger + \sum_{k=1}^d \theta'(x_k)\dot{\mathbf{u}}_k\mathbf{u}_k^\dagger + \sum_{k=1}^d \theta'(x_k)\mathbf{u}_k\dot{\mathbf{u}}_k^\dagger = \mathbf{V} - \dot{\lambda}\mathbf{I} \tag{B.5}$$

在（$\mathbf{u}_\alpha^\dagger$ 的左侧和$\mathbf{u}^\beta$ 的右侧应用（B.5），可得：

$$\sum_{k=1}^d \theta''(x_k)\dot{x}_k\mathbf{u}_\alpha^\dagger\mathbf{u}_k\mathbf{u}_k^\dagger\mathbf{u}_\beta + \sum_{k=1}^d \theta'(x_k)\mathbf{u}_\alpha^\dagger\dot{\mathbf{u}}_k\mathbf{u}_k^\dagger\mathbf{u}_\beta + \sum_{k=1}^d \theta'(x_k)\mathbf{u}_\alpha^\dagger\mathbf{u}_k\dot{\mathbf{u}}_k^\dagger\mathbf{u}_\beta = \mathbf{u}_\alpha^\dagger\mathbf{V}\mathbf{u}_\beta - \dot{\lambda}\mathbf{u}_\alpha^\dagger\mathbf{u}_\beta \tag{B.6}$$

利用 $\mathbf{u}_k^\dagger\,\mathbf{u}_\ell = \delta_{k\ell}$，上述关系变为：

$$\theta''(x_\alpha)\dot{x}_\alpha\delta_{\alpha\beta} + \theta'(x_\beta)\mathbf{u}_\alpha^\dagger\dot{\mathbf{u}}_\beta + \theta'(x_\alpha)\dot{\mathbf{u}}_\alpha^\dagger\mathbf{u}_\beta = \mathbf{V}_{\alpha\beta} - \dot{\lambda}\delta_{\alpha\beta} \tag{B.7}$$

现在，我们注意到，由于$\mathbf{u}_\alpha^\dagger\mathbf{u}_\beta = \delta_{\alpha\beta}$，对其关于 $t$ 求导，得到：

$$\dot{\mathbf{u}}_\alpha^\dagger\mathbf{u}_\beta + \mathbf{u}_\alpha^\dagger\dot{\mathbf{u}}_\beta = 0 \tag{B.8}$$

因此，（B.7）变为：

$$\theta''(x_\alpha)\dot{x}_\alpha\delta_{\alpha\beta} + (\theta'(x_\beta) - \theta'(x_\alpha))\mathbf{u}_\alpha^\dagger\dot{\mathbf{u}}_\beta = \mathbf{V}_{\alpha\beta} - \dot{\lambda}\delta_{\alpha\beta} \tag{B.9}$$

有了上述方程，我们继续证明的最后几步。按照与之前相同的步骤，$\mathbf{X}$ 的 $\alpha\beta$ 项可写为：

$$[\dot{\mathbf{X}}]_{\alpha\beta} = \dot{x}_\alpha\delta_{\alpha\beta} + (x_\beta - x_\alpha)\mathbf{u}_\alpha^\dagger\dot{\mathbf{u}}_\beta \tag{B.10}$$

**(1)** 当 $\alpha = \beta$ 时，方程 **(B.10)** 给出：

$$[\dot{\mathbf{X}}]_{\alpha\alpha} = \dot{x}_\alpha \tag{B.11}$$

并且，方程 **(B.9)** 变为：

$$\theta''(x_\alpha)\dot{x}_\alpha = \mathbf{V}_{\alpha\alpha} - \dot{\lambda} \tag{B.12}$$

由于 $x_\alpha > 0$，根据关于 $\theta$ 的假设，我们有 $\theta''(x_\alpha) > 0$，因此：

$$\dot{x}_\alpha = \frac{\mathbf{V}_{\alpha\alpha}}{\theta''(x_\alpha)} - \frac{\dot{\lambda}}{\theta''(x_\alpha)} \tag{B.13}$$

对 $\beta = 1, \ldots, d$ 求和，并利用 $\sum_{\beta=1}^d \dot{x}_\beta = 0$ **(**因为$\sum_{\beta=1}^d x_\beta = 1$**)** 这一事实，我们得到：

$$\dot{\lambda} = \frac{\sum_{\beta=1}^d V_{\beta\beta}/\theta''(x_\beta)}{\sum_{\beta=1}^d 1/\theta''(x_\beta)} \tag{B.14}$$

因此，将（B.14）与（B.11）和（B.13）结合起来，可得：

$$[\dot{\mathbf{X}}]_{\alpha\alpha} = \frac{V_{\alpha\alpha}}{\theta''(x_\alpha)} - \frac{\sum_\beta V_{\beta\beta}/\theta''(x_\beta)}{\sum_\beta \theta''(x_\alpha)/\theta''(x_\beta)} \tag{B.15}$$

（2）对于 $\alpha = \beta$，方程（B.10）给出：

$$[\dot{\mathbf{X}}]_{\alpha\beta} = (x_\beta - x_\alpha)\mathbf{u}_\alpha^\dagger\dot{\mathbf{u}}_\beta \tag{B.16}$$

并且，方程（B.9）变为：

$$(\theta'(x_\beta) - \theta'(x_\alpha))\mathbf{u}_\alpha^\dagger\dot{\mathbf{u}}_\beta = \mathbf{V}_{\alpha\beta} \tag{B.17}$$

将 $\dot{\mathbf{u}}_\beta$ 表示在基 $\mathbf{u}_1, \ldots, \mathbf{u}_d$ 下，可写成 $\dot{\mathbf{u}}_\beta = \sum_{k=1}^d B_{\beta k}\mathbf{u}_k$，其中系数 $B_{\beta k}$'s 尚待确定。因此，（B.17）可写为：

$$(\theta'(x_\beta) - \theta'(x_\alpha))\sum_{k=1}^d B_{\beta k}\mathbf{u}_\alpha^\dagger\mathbf{u}_k = \mathbf{V}_{\alpha\beta} \tag{B.18}$$

并且，由于 $\mathbf{u}_k^\dagger\,\mathbf{u}_\ell = \delta_{k\ell}$，我们很容易得到：

$$(\theta'(x_\beta) - \theta'(x_\alpha))B_{\beta\alpha} = \mathbf{V}_{\alpha\beta} \tag{B.19}$$

或者，等价地：

$$B_{\beta\alpha} = \frac{\mathbf{V}_{\alpha\beta}}{\theta'(x_\beta) - \theta'(x_\alpha)} \tag{B.20}$$

现在，很容易看出$\mathbf{u}_\alpha^\dagger\dot{\mathbf{u}}_\beta = B_{\beta\alpha}$。因此，结合式（B.16）和式（B.20），我们得出：

$$[\dot{\mathbf{X}}]_{\alpha\beta} = \frac{x_\beta - x_\alpha}{\theta'(x_\beta) - \theta'(x_\alpha)}\mathbf{V}_{\alpha\beta} \tag{B.21}$$

证明到此结束。 ■

**命题 B.1**。 量子复制子动力学（*QRD*）可以写成以下形式：

$$\dot{\mathbf{X}} = \int_0^1 \mathbf{X}^{1-s}\mathbf{V}(\mathbf{X})\mathbf{X}^s \, ds - \operatorname{tr}[\mathbf{X}\mathbf{V}(\mathbf{X})]\mathbf{X} \tag{B.22}$$

证明。 对于 von Neumann 正则化器，我们得到复制子动力学：

$$\dot{\mathbf{Y}} = \mathbf{V}(\mathbf{X}) \qquad \mathbf{X} = \frac{\exp(\mathbf{Y})}{\operatorname{tr}[\exp(\mathbf{Y})]} \tag{B.23}$$

因此，对 $t$ 求导并使用链式法则，我们得到：

$$\begin{aligned}
\dot{\mathbf{X}} &= \frac{1}{\operatorname{tr}[\exp(\mathbf{Y})]}\frac{d}{dt}\exp(\mathbf{Y}) + \exp(\mathbf{Y})\frac{d}{dt}\frac{1}{\operatorname{tr}[\exp(\mathbf{Y})]} \\
&= \frac{1}{\operatorname{tr}[\exp(\mathbf{Y})]}\frac{d}{dt}\exp(\mathbf{Y}) - \frac{\exp(\mathbf{Y})}{\operatorname{tr}[\exp(\mathbf{Y})]^2}\frac{d}{dt}\operatorname{tr}[\exp(\mathbf{Y})] \\
&= \frac{1}{\operatorname{tr}[\exp(\mathbf{Y})]}\frac{d}{dt}\exp(\mathbf{Y}) - \frac{\exp(\mathbf{Y})}{\operatorname{tr}[\exp(\mathbf{Y})]^2}\operatorname{tr}\left[\frac{d}{dt}\exp(\mathbf{Y})\right]
\end{aligned} \tag{B.24}$$

根据 Fréchet 导数 [49],有：

$$\begin{aligned}
\frac{d}{dt}\exp(\mathbf{Y}) &= \int_0^1 e^{(1-s)\mathbf{Y}}\dot{\mathbf{Y}}e^{s\mathbf{Y}}\, ds \\
&= \int_0^1 e^{(1-s)\mathbf{Y}}\mathbf{V}(\mathbf{X})e^{s\mathbf{Y}}\, ds \\
&= \operatorname{tr}[\exp(\mathbf{Y})]\int_0^1 \mathbf{X}^{1-s}\mathbf{V}(\mathbf{X})\mathbf{X}^s\, ds
\end{aligned} \tag{B.25}$$

并在 (B.25) 中使用 (B.24)，可得：

$$\begin{aligned}
\dot{\mathbf{X}} &= \frac{1}{\operatorname{tr}[\exp(\mathbf{Y})]}\operatorname{tr}[\exp(\mathbf{Y})]\int_0^1 \mathbf{X}^{1-s}\mathbf{V}(\mathbf{X})\mathbf{X}^s\, ds \\
&\qquad - \frac{\exp(\mathbf{Y})}{\operatorname{tr}[\exp(\mathbf{Y})]^2}\operatorname{tr}\left[\operatorname{tr}[\exp(\mathbf{Y})]\int_0^1 \mathbf{X}^{1-s}\mathbf{V}(\mathbf{X})\mathbf{X}^s\, ds\right] \\
&= \int_0^1 \mathbf{X}^{1-s}\mathbf{V}(\mathbf{X})\mathbf{X}^s\, ds - \frac{\exp\mathbf{Y}}{\operatorname{tr}[\exp(\mathbf{Y})]}\operatorname{tr}\left[\int_0^1 \mathbf{X}^{1-s}\mathbf{V}(\mathbf{X})\mathbf{X}^s\, ds\right] \\
&= \int_0^1 \mathbf{X}^{1-s}\mathbf{V}(\mathbf{X})\mathbf{X}^s\, ds - \frac{\exp\mathbf{Y}}{\operatorname{tr}[\exp(\mathbf{Y})]}\int_0^1 \operatorname{tr}[\mathbf{X}^{1-s}\mathbf{V}(\mathbf{X})\mathbf{X}^s]\, ds \\
&= \int_0^1 \mathbf{X}^{1-s}\mathbf{V}(\mathbf{X})\mathbf{X}^s\, ds - \frac{\exp\mathbf{Y}}{\operatorname{tr}[\exp(\mathbf{Y})]}\int_0^1 \operatorname{tr}[\mathbf{X}\mathbf{V}(\mathbf{X})]\, ds \\
&= \int_0^1 \mathbf{X}^{1-s}\mathbf{V}(\mathbf{X})\mathbf{X}^s\, ds - \operatorname{tr}[\mathbf{X}\mathbf{V}(\mathbf{X})]\mathbf{X}
\end{aligned} \tag{B.26}$$

如所断言。 ■

## Appendix C. Regret minimization

本附录旨在证明，FTQL的动力学所产生的遗憾至多为常数。为方便起见，我们在下文重述相关结果。

**命题 1.** 设 $\mathbf{X}(t)=\mathbf{Q}(\mathbf{Y}(t))$， $t\geq 0$是由 (FTQL)所诱导的博弈轨迹。那么，对所有 $T\geq 0$，均有

$$\mathrm{Reg}_i(T)\leq |d_i\cdot\theta_i(1/d_i)-\theta_i(1)|. \tag{14}$$

我们的证明基本遵循 [25] 中的分析；证明的特殊之处在于推导遗憾界的精确值（14），但为求完整，我们给出整个证明。

证明。 令 $\mathbf{P}_i\in\boldsymbol{\mathcal{X}}$ 为事后最佳固定行动，即:

$$\mathbf{P}_i\in\arg\max_{\mathbf{X}_i'\in\boldsymbol{\mathcal{X}}_i}\int_0^T u_i(\mathbf{X}_i';\mathbf{X}_{-i}(t))\,dt \tag{C.1}$$

并定义函数 $E_i(t)$: $=F_i(\mathbf{P}_i,\mathbf{Y}_i(t))$ 对 $t\geq 0$成立。根据 引理 1，我们有:

$$\frac{d}{dt}E_i(t)=\mathrm{tr}[\mathbf{V}_i(\mathbf{X}(t))(\mathbf{X}_i(t)-\mathbf{P}_i)]=u_i(\mathbf{X}(t))-u_i(\mathbf{P}_i;\mathbf{X}_{-i}(t)) \tag{C.2}$$

对时间积分，并结合（8）与（C.1），得到:

$$\begin{aligned}\mathrm{Reg}_i(T)&=E_i(0)-E_i(T)\\&=h_i^*(\mathbf{Y}_i(0))-\mathrm{tr}[\mathbf{P}_i\mathbf{Y}_i(0)]-h_i^*(\mathbf{Y}_i(T))+\mathrm{tr}[\mathbf{P}_i\mathbf{Y}_i(T)]\\&\leq h_i^*(\mathbf{Y}_i(0))-\mathrm{tr}[\mathbf{P}_i\mathbf{Y}_i(0)]+h_i(\mathbf{P}_i)\\&=h_i(\mathbf{P}_i)-h_i(\mathbf{Q}_i(\mathbf{Y}_i(0)))\\&=\mathrm{tr}[\theta_i(\mathbf{P}_i)]-\mathrm{tr}[\theta_i(\mathbf{Q}_i(\mathbf{Y}_i(0)))]\\&\leq\max_{\mathbf{X}_i'\in\boldsymbol{\mathcal{X}}_i}\mathrm{tr}[\theta_i(\mathbf{X}_i')]-\min_{\mathbf{X}_i'\in\boldsymbol{\mathcal{X}}_i}\mathrm{tr}[\theta_i(\mathbf{X}_i')]\end{aligned} \tag{C.3}$$

其中我们用到了这一事实:

$$h_i^*(\mathbf{Y}_i)=\max_{\mathbf{X}_i'\in\boldsymbol{\mathcal{X}}_i}\{\mathrm{tr}[\mathbf{X}_i'\mathbf{Y}_i]-h_i(\mathbf{X}_i')\}\geq\mathrm{tr}[\mathbf{P}_i\mathbf{Y}_i]-h_i(\mathbf{P}_i) \tag{C.4}$$

当且仅当 $\mathbf{X}_i'=\mathbf{Q}_i(\mathbf{Y}_i)$时成立。精确界（14）可通过注意到以下事实得出： $h$ 取值的最大差异出现在谱单纯形的重心与任意秩为 1的纯态密度矩阵之间，因此可得界

$$\max h_i-\min h_i=|d_i\cdot\theta_i(1/d_i)-\theta_i(1)|, \tag{C.5}$$

由此即可得出我们的断言。 ■

## Appendix D. The folk theorem

在本附录中，我们的目标是证明 定理 2。为做好必要准备，我们先给出一系列辅助引理。第一个引理与定理 2的验证部分有关，并讨论动力学在其捕获区域附近的行为。

引理 *D.1*。取定某个 $X^*\in X$，并假设对于 $X^*$ 在 $X$ 中的每个邻域 $U$，都存在 *(FTQL)* 的一个解轨迹 $X(t)=Q(Y(t))$，使得对所有 $t\geq 0$，均有 $X(t)\in U$。那么，$X^*$是一个纳什均衡。

证明。我们用反证法。事实上，若 $\mathrm{X}^*$ 不是纳什均衡，则根据纳什均衡的变分刻画（VI），存在某个偏离 $\mathrm{P}\in\mathrm{X}$，使得 $\mathrm{tr}[\mathrm{V}(\mathrm{X}^*)(\mathrm{P}-\mathrm{X}^*)]>0$。于是，根据连续性，存在 $\mathrm{X}^*$ 在 X 中的一个充分小的紧邻域 C，使得对所有 $\mathrm{X}\in\mathrm{C}$，都有 $\mathrm{tr}[\mathrm{V}(\mathrm{X})(\mathrm{P}-\mathrm{X})]>0$，进而可得 $c:=\inf_{\mathbf{X}\in\boldsymbol{C}}\mathrm{tr}[\mathrm{V}(\mathrm{X})(\mathrm{P}-\mathrm{X})]>0$。现在，根据假设，存在一条解轨道

$\mathbf{X}(t) = \mathbf{Q}(\mathbf{Y}(t))$ of (FTQL) such that $\mathbf{X}(t) \in \mathcal{U}$ for all $t \geq 0$ so, by 引理 1, we readilyobtain

$$\frac{d}{dt} F(\mathbf{P}, \mathbf{Y}(t)) = \operatorname{tr}[\mathbf{V}(\mathbf{X}(t))\,(\mathbf{X}(t) - \mathbf{P})] \leq -c < 0 \quad \text{for all } t \geq 0. \tag{D.1}$$

简单积分即可得出 $F(\mathbf{P}, \mathbf{Y}(t)) \to -\infty$ 当 $t \to \infty$，这与结论矛盾，从而证明了我们的论断。■

我们还需要第二个中间结果：一个关于一致渐近稳定状态的一致收敛引理：

**引理 D.2.** 假设 $\mathbf{X}^* \in$ 相对内部$\mathcal{X}$ 在 (FTQL)下是一致渐近稳定的。那么，对每个充分小的紧邻域$\mathcal{U}$，若其为 $\mathbf{X}^*$在 相对内部 $\mathcal{X}$中的邻域，则有

$$\lim_{t \to \infty} \sup_{\mathbf{X} \in \mathcal{U}} \|\boldsymbol{\chi}_t(\mathbf{X}) - \mathbf{X}^*\| = 0. \tag{D.2}$$

证明。首先，由于 $\mathrm{X}^* \in$相对内部 X，我们有 $\ker(\mathrm{X}^*) = \{0\}$，这意味着 $\mathrm{X}\,\mathrm{X}_* =$相对内部 X。因此，对于 ri X 中的任意开集 $O$，均有 $O \cap \mathrm{X}\,\mathrm{X}_* = O$。现在，根据定义 1 中一致渐近稳定性的定义，令 $\mathrm{U}_0$为 $\mathrm{X}^*$的吸引域，并令 $\mathrm{U} \subseteq \mathrm{U}_0$ 为 $\mathrm{X}^*$的一个紧邻域。为导出矛盾，假设 $\sup_{\mathbf{X} \in \mathcal{U}} \|\boldsymbol{\chi}_t(\mathrm{X}) - \mathrm{X}^*\| \not\to 0$。这意味着存在 $\varepsilon > 0$、一个序列 $\{t_n\}_{n \in \mathbb{N}}$，使得当 $n \to \infty$时 $t_n \to \infty$，且对 $n \in \mathbb{N}$，有 $\mathrm{X}_n \in$U，从而

$$\|\boldsymbol{\chi}_{t_n}(\mathbf{X}_n) - \mathbf{X}^*\| \geq \varepsilon \quad \text{for all } n \in \mathbb{N} \tag{D.3}$$

由于$\mathcal{U}$ 是紧集，我们可以假设（必要时通过取子序列）$\mathbf{X}_n$ 收敛于某个极限点$\mathbf{X}_\infty \in \mathcal{U}$。现在，由于$\mathbf{X}^*$ 是李雅普诺夫稳定的，存在$\mathcal{U}'$ 的一个邻域$\mathcal{U}'$，使得当$\mathbf{X} \in \mathcal{U}'$时，轨迹 $\boldsymbol{\chi}_t($ $\mathbf{X})$始终保持在距$\mathbf{X}^*$的 $\varepsilon/2$距离范围内，或者等价地：

$$\|\boldsymbol{\chi}_t(\mathbf{X}) - \mathbf{X}^*\| < \frac{\varepsilon}{2} \quad \text{for all } t \geq 0 \text{ and } \mathbf{X} \in \mathcal{U}' \tag{D.4}$$

定义命中时间 $\tau := \inf\{t \geq 0 : \boldsymbol{\chi}_t(\mathbf{X}_\infty) \in \mathcal{U}'\}$，即从$\mathbf{X}_\infty$出发时，轨迹首次进入$\mathcal{U}'$的时间。容易看出 $\tau < \infty$，因为$\mathbf{X}_\infty \in \mathcal{U} \subseteq \mathcal{U}_0$，因此，根据$\mathbf{X}^*$的渐近稳定性，当 $t \to \infty$ 时，$\|\boldsymbol{\chi}_t(\mathbf{X}_\infty) - \mathbf{X}^*\| \to 0$。

由 $\boldsymbol{\chi}$ 的连续性，我们很容易得到，存在 $\mathbf{X}_\infty$ 的一个邻域 $\mathcal{D}$，使得$\boldsymbol{\chi}_\tau(\mathcal{D}) \subseteq \mathcal{U}''$，其中 $\mathcal{U}''$ 是 $\mathbf{X}^*$ 的一个邻域，且对于所有 $t \geq 0$ 和 $\mathbf{X} \in \mathcal{U}''$，都有 $\|\boldsymbol{\chi}_t(\mathbf{X}) - \mathbf{X}^*\| < \varepsilon$。

由于 $\mathbf{X}_n \to \mathbf{X}_\infty$，我们得出，对于所有充分大的 $n$，$\mathbf{X}_n \in \mathcal{D}$，这进而意味着 $\boldsymbol{\chi}_\tau(\mathbf{X}_n) \in \mathcal{U}''$。此外，根据序列 $\{t_n\}_{n \in \mathbb{N}}$的定义，我们有$t_n \to \infty$，由此可知，对于充分大的 $n$，$t_n > \tau$，因为 $\tau < \infty$，正如前文所述。

因此，对于所有 $n$ 充分大的，我们有 $\|\boldsymbol{\chi}_{t_n}(\mathbf{X}_n) - \mathbf{X}^*\| < \varepsilon$，这与 (D.3) 矛盾。■

在证明 定理 2 之前的最后一步，是比较由作用于 $\psi$ 上的 $\mathcal{X}$ 的 $\mathbf{Q}$ 所诱导的李雅普诺夫稳定性概念，以及相对于作用于 $\boldsymbol{\chi}$ 上的 $\mathcal{X}$ 的共轭流所对应的概念。形式化地说，我们有以下定义。

**定义3。** 我们称 $\mathbf{X}^* \in \mathcal{X}$ 为：

(1) $\boldsymbol{\chi}$-稳定：若对于所有$\mathcal{U}\mathbf{X}^* \in \mathcal{X}$的邻域，都存在一个$\mathcal{U}'\mathbf{X}^* \in \mathcal{X}$的邻域，使得对于所有 $\mathbf{X} \in \mathcal{U}'$ 以及所有 $t \geq 0$，$\boldsymbol{\chi}_t(\mathbf{X}) \in \mathcal{U}$。(2) $\psi$-稳定：若对于所有$\mathcal{U}\mathbf{X}^* \in \mathcal{X}$的邻域，都存在一个$\mathcal{U}'\mathbf{X}^* \in \mathcal{X}$的邻域，使得对于所有 $\mathbf{X} \in \mathcal{U}' \cap \operatorname{im} \mathbf{Q}$ 以及所有 $t \geq 0$，$\boldsymbol{\chi}_t(\mathbf{X}) \in \mathcal{U}$。

**引理 D.3。** 点 $\mathbf{X}^* \in \mathcal{X}$ 是 $\boldsymbol{\chi}$-稳定的，当且仅当它是 $\psi$-稳定的。

证明。“仅当”部分显然，因此我们着重证明“当”部分。为此，假设 $X^*$ 是 $\psi$-稳定的，但不是 $\chi$-稳定的。那么，存在 $X^*$ 在 X 中的一个邻域 U，使得对于 $X^*$的所有邻域 $U'$，都存在 $X \in U'$，使得 $\chi_t(X) \notin \mathcal{U}$，其中某个 $t \geq 0$成立。现在，取 $X^*$的一个邻域 $U_0$，使得 $cl(U_0) \subseteq U$。于是，根据 $X^*$的 $\psi$-稳定性，存在 $X^*$的一个邻域 $U_0'$，使得对于所有 $X \in U_0' \cap im\ Q$ 以及所有$t \geq 0$，都有 $\chi_t(X) \in U_0$。但由于 $X^*$ 不是 $\chi$-稳定的，存在 $X \in U_0' \cap (im\ Q)^c$，使得命中时间 $\tau = \inf\{t \geq 0 : \chi_t(X) \notin U\}$ 有限。此外，取 $\{X_n\}_{n\in\mathbb{N}} \subseteq U_0' \cap im\ Q$ 中一个收敛到 X 的序列（这样的序列存在，因为 $cl(im\ Q) = X$）。由于 $X^*$ 是 $\psi$-稳定的，对于所有 $n$，都有 $\chi_\tau(X_n) \in U_0$；因此，利用 $\chi$的连续性，我们得到：

$$\chi_\tau(\mathbf{X}) = \lim_{n\to\infty} \chi_\tau(\mathbf{X}_n) \tag{D.5}$$

这意味着 $\chi_\tau(\mathbf{X}) \in cl(\mathcal{U}_0) \subseteq \mathcal{U}$，因为它是$\mathcal{U}_0$的一个极限点。但是，由于根据假设，$\chi_\tau(\mathbf{X}) \notin \mathcal{U}$，我们得到一个矛盾，从而完成证明。■

多亏了上述命题，我们无需区分“李雅普诺夫稳定”和“一致李雅普诺夫稳定”状态，因为这两个概念是等价的。因此，既然这一问题已经解决，我们便可以证明量子版的“民间定理”，为方便起见，下面重新陈述该定理：

**定理 2.** 设 $\mathcal{Q} \equiv \mathcal{Q}(\mathcal{N}, \Psi, u)$ 是一个量子博弈，固定一个状态 $\mathbf{X}^* \in \boldsymbol{\mathcal{X}}$，并设$\mathbf{X}(t) = \mathbf{Q}(\mathbf{Y}(t))$ 是由 (FTQL) 与陡峭正则化函数诱导的博弈轨迹。那么：
*(1)* 若 $X^*$ 是纳什均衡，则它是 *(QD)* 的静止点。*(2)* 若 $X(t) \to X^*$ 当 $t \to \infty$ 时，$X^*$ 是 $\mathcal{Q}$ 的纳什均衡。*(3)* 若 $X^*$ 在 *(QD)* 下稳定，则它是 $\mathcal{Q}$ 的纳什均衡。*(4)* 若 $X^*$ 一致渐近稳定，则它是纯态。*(5)* 若 $X^*$ 满足 *(VS)*，则它一致渐近稳定。

我们在下文证明该定理的各个部分。

第 1 部分的证明。假设 $X^*$ 是 $\mathcal{Q}$ 的一个纳什均衡。若 $rank(X^*) = 1$，则无需证明，因此只需在 $rank(X^*) > 1$ 的情形下确立我们的主张。在这种情形下，如有必要，可将动力学限制在包含 $X^*$ 的、X 的面 $X^* = \{X \in X : \ker X \geq \ker X^*\}$ 的相对内部；于是，不失一般性，我们可以假设 $X^*$ 实际上具有满秩。基于此，固定某个 $Y \in Q^{-1}(X^*)$，并考虑从时刻$t = 0$的 Y 出发的轨迹 $Y(t) = Y + tV(X^*)$。由于根据纳什均衡的变分刻画 (VI)，有 $V(X^*) \in NC(X^*)$，由引理 A.1 可知，对所有 $t \geq 0$，$Q(Y(t)) = Q(Y)$。进而，这表明对所有 $t = 0$，有$\dot{Y}(t) = V(X^*) = V(Q(Y(t)))$，即 $Y(t)$ 是 (FTQL) 的一条解轨道。根据皮卡-林德洛夫定理，这意味着 $Y(t)$ 是其唯一解；因此，鉴于对所有 $t \geq 0$ 都有 $X(t) = Q(Y(t)) = X^*$，我们的主张得证。■

第 *2* 和第 *3* 部分的证明。收敛到 $\mathbf{X}^*$ 以及李雅普诺夫稳定性均意味着满足 引理 D.1的假设，因此根据该引理的结论，我们的断言立即成立。■

第 *4* 部分的证明。本部分的证明分为三个步骤。首先，我们将在 $ri\boldsymbol{\mathcal{X}}$ 上构造一个测度 $\mu$，使其在流 $\chi$下保持不变，即对于所有可测集 $A$ 和 $t \geq 0$，都有 $\mu(A) = \mu(\chi_t(A))$，其中 $\chi_t(A)$表示 $A$ 经过时间 $t$后的像。接下来，我们将证明，若$\mathbf{X}^* \in ri\,\boldsymbol{\mathcal{X}}$，则它不可能是一致渐近稳定的。最后，我们将得出结论

若$\mathbf{X}^* \in \mathcal{X}$是一致渐近稳定的，则通过一个简单的约化论证可知它是纯态。

步骤1。根据 Q 的定义，我们有 $\mathrm{Y} = \nabla_{\mathbf{X}^\top}\theta(\mathrm{X}) + c\cdot\mathrm{I}$，即 $\mathrm{Q}^{-1}(\mathrm{X}) = \{\nabla_{\mathbf{X}^\top}\theta(\mathrm{X}) + c\cdot\mathrm{I} : c \in \mathbb{R}\}$。令 $\mathrm{Y}_0 \coloneqq \mathrm{Y}/\sim$ 为由等价关系 $\sim$诱导的商空间，其中 $\mathrm{Y} \sim \mathrm{Y}'$ 当且仅当对于某个 $c \in \mathbb{R}$，$\mathrm{Y} - \mathrm{Y}' = c\cdot\mathrm{I}$；并令 $q : \mathrm{Y} \to \mathrm{Y}_0$ 为相应的商映射，定义为 $q(\mathrm{Y}) = \{\mathrm{Y} + c\cdot\mathrm{I} : c \in \mathbb{R}\}$。因此，通过下沉到商空间 $\mathrm{Y}_0$，镜像映射 Q 可分解为经过 $\mathrm{Y}_0$ 的映射 $\mathrm{Q}_0 : \mathrm{Y}_0 \to \mathrm{X}$，其定义为 $\mathrm{Q}_0 \coloneqq \mathrm{Q}\circ q$。现在，令 $\lambda$表示 $\mathrm{Y}_0$, 上的勒贝格测度，并令 $\mu = \mathrm{Q}_0\#\lambda$ 表示 $\lambda$ 经 $\mathrm{Q}_0$, 到 X 的推前测度，即对于 X 中所有可测集 $A$， $\mu(A) = \lambda(\mathrm{Q}_0^{-1}(A))$。最后，根据命题 E.1，我们立即得到，对于 X 中所有可测集 $A \subseteq$ 以及所有 $t \geq 0$， $\mu(\boldsymbol{\chi}_t(A)) = \mu(A)$。

**步骤 2.** 为导出矛盾，假设 $\mathbf{X}^* \in$ 相对内部$\mathcal{X}$ 是一致渐近稳定的，并设 $\mathcal{U}$ 是 $\mathbf{X}^*$ 的一个足够小的紧邻域，如引理 D.2 所述。于是，有 $\lim_{t\to\infty}\sup_{\mathbf{X}\in\mathcal{U}}\|\boldsymbol{\chi}_t(\mathbf{X}) - \mathbf{X}^*\| = 0$，这意味着：

$$\lim_{t\to\infty}\mu(\boldsymbol{\chi}_t(\mathcal{U})) = \mu(\{\mathbf{X}^*\}) = 0 < \mu(\mathcal{U}). \tag{D.6}$$

然而，根据证明的步骤 *1*，我们有 $\lim_{t\to\infty}\mu(\boldsymbol{\chi}_t(\mathcal{U})) = \mu(\mathcal{U})$，因为 $\mu(\boldsymbol{\chi}_t(\mathcal{U})) = \mu(\mathcal{U})$ 对所有 $t \geq 0$。因此，将其与 (D.6 结合，我们便得到矛盾。

**步骤3**。 最后，假设 $\mathbf{X}^*$ 是一致渐近稳定的，且 $r \coloneqq \mathrm{rank}(\mathbf{X}^*) > 1$。注意，若 $\mathbf{X}^*$ 是满秩的，则它不可能是一致渐近稳定的，如证明的步骤 *2* 所示。令 $\mathcal{A}_{\mathbf{X}^*} \coloneqq \{\mathbf{X} \in \mathcal{X} : \ker(\mathbf{X}) \geq \ker(\mathbf{X}^*)\}$，这是凸的。事实上，设$\mathbf{X}$， $\mathbf{X}' \in \mathcal{A}_{\mathbf{X}^*}$ 且 $\lambda \in (0,1)$。对于 $y \in \ker(\mathbf{X}^*)$，我们有：

$$\begin{aligned}[\lambda\mathbf{X} + (1-\lambda)\mathbf{X}']y &= \lambda\mathbf{X}y + (1-\lambda)\mathbf{X}'y\\ &= 0\end{aligned} \tag{D.7}$$

也就是说， $y \in \ker(\lambda\mathbf{X} + (1-\lambda)\mathbf{X}')$，因此，$\ker(\lambda\mathbf{X} + (1-\lambda)\mathbf{X}') \geq \ker(\mathbf{X}^*)$。

因此，考虑动力学在$\mathcal{A}_{\mathbf{X}^*}$上的限制，证明的步骤 *2* 表明，$\mathrm{ri}\mathcal{A}_{\mathbf{X}^*}$ 中的点在诱导拓扑下不可能是一致渐近稳定的。因此，$\mathbf{X}^*$ 在$\mathcal{A}_{\mathbf{X}^*}$上的限制（记为$\mathbf{X}^*|_{\mathcal{A}_{\mathbf{X}^*}}$）不可能在$\mathcal{A}_{\mathbf{X}^*}$上一致渐近稳定，因为它在$\mathcal{A}_{\mathbf{X}^*}$上是满秩的。所以，剩下需要证明的是$\mathbf{X}^*$ 不可能在$\mathcal{X}$上一致渐近稳定。

反设结论不成立。那么，根据定义 1，存在$\mathcal{U}$属于 $\mathbf{X}^*$ 中 $\mathcal{X}$，使得 $\lim_{t\to\infty}\boldsymbol{\chi}_t(\mathbf{X}) = \mathbf{X}^*$ 对所有 $\mathbf{X} \in \mathcal{U} \cap \mathcal{X}_{\mathbf{X}^*}$均成立。但由于 $\mathcal{U}$ 是 $\mathbf{X}^*$ 在 $\mathcal{X}$中的一个邻域，

$$(\mathcal{U}\cap\mathcal{X}_{\mathbf{X}^*})\cap\mathcal{A}_{\mathbf{X}^*} = \mathcal{U}\cap\{\mathbf{X}'\in\mathcal{X} : \ker(\mathbf{X}') = \ker(\mathbf{X}^*)\} \supsetneq \{\mathbf{X}^*\} \tag{D.8}$$

因此，$\mathcal{U} \cap \{\mathbf{X}' \in \mathcal{X} : \ker(\mathbf{X}') = \ker(\mathbf{X}^*)\}$ 在 $\mathcal{A}_{\mathbf{X}^*}$ 上的限制是$\mathbf{X}^*$ 在$\mathcal{A}_{\mathbf{X}^*}$中的一个邻域。因此，根据我们先前关于$\mathcal{A}_{\mathbf{X}^*}$上诱导动力学的论证，存在$\mathbf{X}_0 \in \mathcal{U} \cap \{\mathbf{X}' \in \mathcal{X} : \ker(\mathbf{X}') = \ker(\mathbf{X}^*)\}$，使得 $\boldsymbol{\chi}_t(\mathbf{X}_0) \not\to \mathbf{X}^*$，如 $t \to \infty$。这与我们的原始假设矛盾，因此证明完毕。■

第 5 部分的证明。假设 $\mathrm{X}^*$ 是（局部）变分稳定的，即，对于所有 $\mathrm{X} \in \mathrm{U}$，都有 $\mathrm{tr}[\mathrm{V}(\mathrm{X})(\mathrm{X} - \mathrm{X}^*)] < 0$，其中 U 是 $\mathrm{X}^*$的一个邻域。那么，根据能量函数 $E(t) \coloneqq F(\mathrm{X}^*, \mathrm{Y}(t))$应用李雅普诺夫直接法，并利用引理 D.3，可得 $\mathrm{X}^*$ 是稳定的。因此，存在 $\mathrm{X}^*$的邻域 $\mathrm{U}'$，使得若 $\mathrm{X}(0) \in \mathrm{U}'$，则对所有 $t \geq 0$，$\mathrm{X}(t) \in \mathrm{U}$。这意味着，若 $\mathrm{X}(0) \in \mathrm{U}'$，则对所有 $t \geq 0$，都有 $\mathrm{tr}[\mathrm{V}(\mathrm{X}(t))(\mathrm{X}(t) - \mathrm{X}^*)] < 0$。接下来，我们将证明，若 $\mathrm{X}(0) \in \mathrm{U}'$，则 $\mathrm{X}(t) \to \mathrm{X}^*$。

现在，令 $\mathbf{X}(0) \in \mathcal{U}'$。由于 $\mathbf{X}(t) \in \mathcal{U}$ 对所有 $t \geq 0$均成立，且利用 引理 1，可得：

$$\frac{d}{dt}E(t) = \mathrm{tr}[\mathbf{V}(\mathbf{X}(t))(\mathbf{X}(t) - \mathbf{X}^*)] < 0 \tag{D.9}$$

其中，最后一个不等式由 (VS) 得出。因此， $E(t)$ 在 $\mathbf{X}(0) \in \mathcal{U}$时递减。

现在，我们将证明$\mathbf{X}^*$ 是该流的一个 $\omega$-极限，即存在一个序列 $\{t_n\}_{n\in\mathbb{N}}$，使得 $\mathbf{X}(t_n) \to \mathbf{X}^*$ 当 $n \to \infty$ 时成立。反证法，假设这样的序列不存在。那么， $\|\mathbf{X}(t) - \mathbf{X}^*\|$ 与零之间的距离有正下界，这意味着存在 $c > 0$，使得 $\mathrm{tr}[\mathbf{V}(\mathbf{X}(t))(\mathbf{X}(t) - \mathbf{X}^*)] < -c$，即 $\frac{d}{dt}E(t) < -c$。对时间积分，得到

$$E(t) - E(0) < -ct$$

这意味着当 $t \to \infty$ 时， $E(t) \to -\infty$ 。这与 Fenchel 耦合的非负性相矛盾，因此我们得出结论：存在一个序列 $\{t_n\}_{n\in\mathbb{N}}$ ，使得$\mathbf{X}(t_n) \to \mathbf{X}^*$ 当 $n \to \infty$ 时成立。

最后，根据命题 A.2，我们得到 $F(\mathrm{X}^*, \mathrm{Y}(t_n)) \to 0$ as $n \to \infty$，即 $E(t_n) \to 0$。因此，由于 $E(t)$ 是递减的，且 $E(t_n) \to 0$，我们得出 $E(t) \to 0$ as $t \to \infty$，即 $F(\mathrm{X}^*, \mathrm{Y}(t)) \to 0$。再次利用命题 A.2，我们得出 $\mathrm{X}(t) \to \mathrm{X}^*$，即 $\mathrm{X}^*$ 一致渐近稳定。■

## 附录 E. 庞加莱复现

本附录旨在证明关于具有满秩均衡的极小极大博弈的复现结果。为便于参考，我们在下方重述该结果：

定理 *3.* 设 $\mathcal{Q} \equiv \mathcal{Q}(\mathcal{N}, \Psi, u)$ 是一个具有满秩均衡 $X^* \in$ 相对内部 $X$ 的双人零和博弈。那么，*(FTQL)* 的几乎每条解轨道 $X(t) = Q(Y(t))$ 都是庞加莱复现的：存在一个递增的时间序列 $t_n \to \infty$ ，使得 $\lim_{n\to\infty} X(t_n) = X(0)$。

正如我们在论文正文中简要讨论的那样，我们的证明策略将包括以下步骤：

（1）首先，我们需要将 V 的对偶空间 $\mathrm{Y} \equiv \mathbb{H}^d$ 压缩到具有恒定迹的子空间 Z。原因在于，即使 $h$ 是陡峭的，镜像映射 $\mathrm{Q} : \mathrm{Y} \to \mathrm{X}$ 也不是单射：对于所有 $\lambda \in \mathbb{R}$，都有 $\mathrm{Q}(\mathrm{Y} + \lambda \mathrm{I}) = \mathrm{Q}(\mathrm{Y})$，这意味着 im Q 中任意点的逆像始终包含实线的一个副本。进行这种压缩旨在“商去”这一冗余，从而为之后的体积比较提供便利。（2）第二步是证明动力学（FTQL）在商空间 Z 上仍然定义良好，并且是不可压缩的，也就是说，一个初始条件集合的体积会随时间保持不变。正文中的命题 2 对此进行了刻画。（3）证明的下一步是说明 Z 上的诱导动力学具有一个运动常数。这将由相对于该博弈内点均衡的 Fenchel 耦合给出（这也是需要满秩假设的地方）。（4）最后，也是证明中最具挑战性的部分，是证明 Fenchel 耦合的等值集在 Z 中有界（相比之下，它们在 Y 中并不有界）。

这些步骤完成后，定理 3 将由庞加莱定理推出。接下来，我们将这些步骤分别归纳为一系列引理和命题。为简化

记号，除非绝对有必要，否则我们将省略玩家指标。

**步骤1：下降到商空间**。为消除上述冗余，考虑变换后的矩阵变量

$$\mathbf{Z} = \mathbf{Y} - (1/d)\operatorname{tr}[\mathbf{Y}]\mathbf{I} \tag{E.1}$$

so $\operatorname{tr}\mathbf{Z} = 0$ 对所有 $\mathbf{Y} \in \mathcal{Y}$。形式上，这一变换可通过以下方式表示 p
$\mathbf{\Pi}\colon \mathcal{Y} \to \mathcal{Z}$，其中

$$\mathcal{Z} = \{\mathbf{Z} \in \mathcal{Y} : \operatorname{tr}\mathbf{Z} = 0\} \tag{E.2}$$

并且 $\mathbf{\Pi}\colon \mathbf{Y} \mapsto \mathbf{\Pi}(\mathbf{Y}) = \mathbf{Z}$ 由上文的（E.1）定义。因此，$\mathbf{Z}$ 可视为等价关系的一个代表元 $\mathbf{Y} \sim \mathbf{Y} + \lambda\mathbf{I}$，$\lambda \in \mathbb{R}$在此关系下$\mathcal{Z}$ 可与商空间$\mathcal{Y}/\sim$等同。

根据构造以及$\mathbf{Q}$ (引理 A.1 的性质（尤其是），可得 $\mathbf{Q}(\mathbf{Y}) = \mathbf{Q}(\mathbf{\Pi}(\mathbf{Y})) = \mathbf{Q}(\mathbf{Z})$ 对所有 $\mathbf{Y} \in \mathcal{Y}$成立。因此，由（FTQL）可直接得到

$$\begin{aligned}\dot{\mathbf{Z}} = \frac{d}{dt}[\mathbf{Y} - (1/d)\operatorname{tr}[\mathbf{Y}]\mathbf{I}] &= \dot{\mathbf{Y}} - (1/d)\operatorname{tr}[\dot{\mathbf{Y}}]\mathbf{I} \\ &= \mathbf{V}(\mathbf{Q}(\mathbf{Y})) - (1/d)\operatorname{tr}[\mathbf{V}(\mathbf{Q}(\mathbf{Y}))]\mathbf{I} \\ &= \mathbf{V}(\mathbf{Q}(\mathbf{Z})) - (1/d)\operatorname{tr}[\mathbf{V}(\mathbf{Q}(\mathbf{Z}))]\mathbf{I}\end{aligned} \tag{E.3}$$

因此我们可以将 FTQL) 改写为关于 $\mathbf{Z}$ 的形式：

$$\dot{\mathbf{Z}} = \mathbf{V}(\mathbf{X}) - (1/d)\operatorname{tr}[\mathbf{V}(\mathbf{X})]\mathbf{I} \qquad \mathbf{X} = \mathbf{Q}(\mathbf{Z}). \tag{FTQL$_{\mathcal{Z}}$}$$

由于 $\operatorname{tr}[\dot{Z}] = \operatorname{tr}[\dot{V}(X)] - (1/d)\operatorname{tr}[\dot{V}(X)]\operatorname{tr} I = 0$，因此（FTQL$_{\mathcal{Z}}$）的任何无迹初始条件都将保持无迹——从而对于所有 $n \geq 0$ 始终保持在 X 中。由此我们得出，（FTQL$_{\mathcal{Z}}$）是定义在 Z 上的适定动力系统，其诱导流 $\zeta\colon \mathbb{R} \times Z \to Z$ 按如下意义与 $\psi$ 共轭：对于所有 $t \geq 0$，$\zeta_t \circ \mathbf{\Pi} = \mathbf{\Pi} \circ \psi_t$。

**步骤 2：体积保持**。我们的第二个步骤是证明，动力学（FTQL$_{\mathcal{Z}}$）是保体积的。形式上，我们建立以下 命题 2 的直接类比：

命题 *E.1.* 设 $W \subseteq Z$ 是（*FTQL*$_{\mathcal{Z}}$）的一个初始条件开集，并令 $W_t = \zeta_t(W)$，$t \geq 0$ 表示在（*FTQL*$_{\mathcal{Z}}$）的流 $\zeta\colon \mathbb{R} \times Z \to Z$ 下 $W$ 的演化。那么，对于所有 $t \geq 0$，$\mathrm{vol}(W_t) = \mathrm{vol}(W)$。

备注。 如上文（以及命题 2中）所述，体积形式$\mathrm{vol}(\cdot)$ 表示$\mathcal{Z}$上的普通勒贝格测度。

命题命题 *E.1*的证明。 这里的关键观察是，动力学（FTQL$_{\mathcal{Z}}$）是不可压缩的——也就是说，是无散的。事实上，由于$\mathbf{V}_i$ 不依赖于$\mathbf{X}_i$ （这里无法省略玩家指标），我们立即得到$\nabla_{\mathbf{Z}_i^\top} \mathbf{V}_i(\mathbf{Q}(\mathbf{Y})) = 0$ 对所有 $i \in \mathcal{N}$均成立。这立即意味着向量场

$$\mathbf{W}_i(\mathbf{Z}) = \mathbf{V}_i(\mathbf{Q}(\mathbf{Z})) - (1/d_i)\operatorname{tr}[\mathbf{V}_i(\mathbf{Q}(\mathbf{Z}))]\mathbf{I} \tag{E.4}$$

具有$\operatorname{div}_{\mathbf{Z}}(\mathbf{W}(\mathbf{Z})) = 0$，因此（FTQL$_{\mathcal{Z}}$）是不可压缩的。我们的论断由刘维尔公式 [3]立即得出。
■

命题 2 的证明完全相同，因此我们省略其证明（反正后文也用不到）。

**步骤3：Fenchel 耦合作为运动常数。** 我们现在继续证明，动力学（FTQL$_{\mathcal{Z}}$）存在一个运动常数，即相对于该博弈的满秩均衡的 Fenchel 耦合。事实上，令 $E(t) = F(\mathbf{X}^*、\mathbf{Z}(t))$，引理 1 即可得出：

$$\begin{aligned}\frac{dE}{dt} &= \mathrm{tr}[\mathbf{V}_1(\mathbf{X})(\mathbf{X}_1 - \mathbf{X}_1^*)] + \mathrm{tr}[\mathbf{V}_2(\mathbf{X})(\mathbf{X}_2 - \mathbf{X}_2^*)] \\ &= u_1(\mathbf{X}_1, \mathbf{X}_2) - u_1(\mathbf{X}_1^*, \mathbf{X}_2) + u_2(\mathbf{X}_1, \mathbf{X}_2) - u_2(\mathbf{X}_1, \mathbf{X}_2^*) \\ &= -\mathcal{L}(\mathbf{X}_1, \mathbf{X}_2) + \mathcal{L}(\mathbf{X}_1^*, \mathbf{X}_2) + \mathcal{L}(\mathbf{X}_1, \mathbf{X}_2) - \mathcal{L}(\mathbf{X}_1, \mathbf{X}_2^*) = \mathcal{L}(\mathbf{X}_1^*, \mathbf{X}_2) - \mathcal{L}(\mathbf{X}_1, \mathbf{X}_2^*) = 0\end{aligned} \tag{E.5}$$

其中在最后一行，我们使用了$\mathbf{X}^*$ 是满秩均衡这一假设；另请参阅 [22, 23]中的相关讨论。

**步骤 4：轨迹有界性。** 接下来只需证明 (FTQL$_{\mathcal{Z}}$) 的轨迹有界。以下技术引理是证明这一点的关键：

**引理 E.1。** 设$\mathbf{Z}_n$, $n = 1, 2, \ldots$ 是$\mathcal{Z}$ 中的一个序列，使得对于某个内部$\mathbf{P} \in \mathcal{X}$, $F(\mathbf{P}, \mathbf{Z}_n)$ 保持不变。那么$\mathbf{Z}_n$ 的特征值$\{z_{\alpha,n}\}_{\alpha=1,\ldots,d}$ 有界，即$\sup_{\alpha,n}|z_{\alpha,n}| < \infty$。

证明。 我们用反证法。事实上，假设 $F(\mathbf{P}, \mathbf{Z}_n)$为常数，但$\sup_{\alpha,n} z_{\alpha,n} = \infty$。于是，令 $z_n^+ = \max_\alpha z_{\alpha,n}$ 和 $z_n^- = \min_\alpha z_{\alpha,n}$，由$\mathbf{Z}_n \in \mathcal{Z}$（因此$\mathrm{tr}\mathbf{Z}_n = 0$）这一事实，我们也有$\limsup_{n\to\infty}(z_n^+ - z_n^-) = \infty$。然后，必要时通过重新标记指标并取更细的子序列，不失一般性地假设以下性质同时成立：

(1) $\mathbf{X}_n = Q(\mathbf{Z}_n)$ 收敛到某个 $\mathbf{X} \in \mathcal{X}$（由 $\mathcal{X}$ 的紧致性可知）。(2) $\mathbf{X}_n = \sum_\alpha x_{\alpha,n}\mathbf{u}_{\alpha,n}\mathbf{u}_{\alpha,n}^\dagger$ 的特征分解中的所有元素也都收敛（由酉群 $U(d)$ 和单位单纯形的紧致性可知）。(3) $z_{d,n} = z_n^+$ 和 $z_{1,n} = z_n^-$ 对所有 $n = 1, 2, \ldots$ 均成立。(4) 指标集 $\{1, \ldots, d\}$ 可划分为两个非空集合 $\mathcal{S}^+$ 和 $\mathcal{S}^-$，使得 a) 对所有 $\alpha \in \mathcal{S}^+$，$\sup_n|z_n^+ - z_{\alpha,n}| < \infty$；且 b) 对所有 $\alpha \in \mathcal{S}^-$，$z_n^+ - z_{\alpha,n} \to \infty$。

由于 $\mathbf{X}_n$ 和 $\mathbf{Z}_n$ 可交换（根据 引理 A.2），因此得到：

$$\begin{aligned}\mathrm{tr}[\mathbf{Z}_n \cdot (\mathbf{P} - \mathbf{X}_n)] &= \sum_{\alpha=1}^d z_{\alpha,n}(p_{\alpha,n} - x_{\alpha,n}) \\ &= \sum_{\alpha=1}^d (z_{\alpha,n} - z_n^+)(p_{\alpha,n} - x_{\alpha,n}) \\ &= \sum_{\alpha\in\mathcal{S}^+} (z_{\alpha,n} - z_n^+)(p_{\alpha,n} - x_{\alpha,n}) + \sum_{\alpha\in\mathcal{S}^-} (z_{\alpha,n} - z_n^+)(p_{\alpha,n} - x_{\alpha,n})\end{aligned} \tag{E.6}$$

其中令 $p_{\alpha,n} = \mathbf{u}_{\alpha,n}^{\mathrm{diag}}\mathbf{P}\mathbf{u}_{\alpha,n}$，并且第二行用到了 $\mathrm{tr}\mathbf{P} = \mathrm{tr}\,\mathbf{X}_n = 1$ 这一事实。现在，根据 引理 A.2的第三部分，对于所有 $\alpha \in \mathcal{S}^-$，有 $x_{\alpha,n} \to 0$，因此对所有 $\alpha \in \mathcal{S}^-$，$\lim_{n\to\infty}(p_{\alpha,n} - x_{\alpha,n}) > 0$（回顾一下，$p_{\alpha,n}$ 收敛，且已假设$\mathbf{P}$ 为满秩）。因此，我们得到

$$\sum_{\alpha\in\mathcal{S}^-} (z_{\alpha,n} - z_n^+)(p_{\alpha,n} - x_{\alpha,n}) \to -\infty \quad \text{as } n \to \infty. \tag{E.7}$$

因此，由于对 $\alpha \in \mathcal{S}^+$ 求和中涉及的所有量都有界，我们可以得出 $\mathrm{tr}[\mathbf{Z}_n \cdot (\mathbf{P} - \mathbf{X}_n)] \to -\infty$ as $n \to \infty$。然而，根据 Fenchel 耦合的定义，

我们有

$$F(\mathbf{P}, \mathbf{Z}_n) = h(\mathbf{P}) + h^*(\mathbf{Z}_n) - \operatorname{tr}[\mathbf{Z}_n \mathbf{P}] = h(\mathbf{P}) - h(\mathbf{X}_n) - \operatorname{tr}[\mathbf{Z}_n \cdot (\mathbf{P} - \mathbf{X}_n)], \tag{E.8}$$

这进而推出 $\lim_{n\to\infty} F(\mathbf{P}, \mathbf{Z}_n) = -\infty$，矛盾。 ■

**步骤 5：综合上述结果。** 至此，我们终于可以证明 定理 3。

定理 3 的证明。由引理 E.1 以及 $F(\mathrm{X}^*, \mathrm{Z}(t))$ 是 (FTQL$_{\mathcal{Z}}$) 下的运动常数这一事实，可知 (FTQL$_{\mathcal{Z}}$) 的轨迹有界。因此，由命题 E.1 和庞加莱回归定理可知，(FTQL$_{\mathcal{Z}}$) 的动力学是复现的；也就是说，对于（勒贝格）几乎所有初始条件 $\mathrm{Z} \in \mathcal{Z}$，所诱导的轨迹 $\mathrm{Z}(t) \equiv \zeta_t(\mathrm{Z})$ 会无限多次返回到任意接近 Z 的位置。更准确地说，这意味着对于 Z 在 Z 中的每个邻域 W，都存在某个有限返回时间 $\tau_{\mathcal{W}}$，使得 $\zeta_{\tau_{\mathcal{W}}}(\mathrm{Z}) \in \mathrm{W}$。因此，取以 Z 为中心、半径为 $1/n$ 的递缩球网 $\mathrm{W}_n$，并令 $t_n = \tau_{\mathcal{W}_n}$，可得 $\zeta(t_n) \to \mathrm{Z}$，当 $n \to \infty$ 时成立。接着，将上述构造应用于某个满足 $\mathrm{Z} \in \mathrm{Q}^{-1}(\mathrm{X}(0)) \cap \mathcal{Z}$ 的点，并注意到 $\mathrm{X}(t_n) = \chi_{t_n}(\mathrm{X}(0)) = \mathrm{Q}(\zeta_{t_n}(\mathrm{Z})) \to \mathrm{Q}(\mathrm{Z}) = \mathrm{X}(0)$，即可得到所述结论。■

## 附录 F。数值结果

在本附录中，我们介绍模拟设置。我们考虑一个简单的双人零和博弈，其行玩家和列玩家的收益矩阵分别为 $P_1 = \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix}$ 和 $P_2 = -P_1$，并在量子情形下表示。将各行记为（$\alpha_1$，$\alpha_2$），各列记为（$\beta_1$，$\beta_2$），可知行动组合（$\alpha_1$，$\beta_2$）是一个严格纳什均衡。对于初始条件分别为（0.2，0.8）和（0.8，0.2）的行玩家和列玩家，我们在该双人量子博弈上运行（MMW），并生成玩家状态截至时间 $t = 100$ 的轨迹。我们通过每位玩家的布洛赫球将其可视化。请注意，图中的红色菱形“♦”表示玩家的初始混合态。正如我们所观察到的，轨迹会收敛到布洛赫球的边界，即纯态。为便于复现我们的结果，我们使用了以下公开代码：[23, https://openreview.net/forum?id=FFZYhY2z3j]

## 参考文献

[1] Aaronson, S. 量子态的阴影层析。*SIAM*计算杂志，49(5):STOC18– 368–STOC18–394，2020年1月。[2] Aaronson, S., Chen, X., Hazan, E., Kale, S., and Nayak, A. 量子态的在线学习。载于*NeurIPS '18*：第*32*届神经信息处理系统国际会议论文集，2018年。[3] Arnold, V. I. 经典力学的数学方法。数学研究生教材第60卷。Springer，纽约州纽约，第二版，1989年。[4] Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., Biswas, R., Boixo, S., Brandao, F. G., Buell, D. A., Burkett, B., Chen, Y., Chen, Z., Chiaro, B., Collins, R., Courtney, W., Dunsworth, A., Farhi, E., Foxen, B., Fowler, A., Gidney, C., Giustina, M., Graff, R., Guerin, K., Habegger, S., Harrigan, M. P., Hartmann, M. J., Ho, A., Hoffmann, M. R., Huang, T., Humble, T. S., Isakov, S. V., Jeffrey, E., Jiang, Z., Kafri, D., Kechedzhi, K., Kelly, J., Klimov, P. V., Knysh, S., Korotkov, A. N., Kostritsa, F., Landhuis, D., Lindmark, M., Lucero, E., Lyakh, D., Mandra, S., McClean, J. R., McEwen, M., Megrant, A., Mi, X., Michielsen, K., Mohseni, M., Mutus, J., Naaman, O., Neeley, M., Neill, C., Niu, M. Y., Ostby, E., Petukhov, A., Platt, J. C., Quintana, C., Rieffel, E. G., Roushan, P., Rubin, N. C., Sank, D., Satzinger, K. J., Smelyanskiy, V., Sung, K. J., Trevithick, M. D., Vainsencher, A., Villalonga,

B., White, T., Yao, Z. J., Yeh, P., Zalcman, A., Neven, H., and Martinis, J. M. 使用可编程超导处理器实现量子霸权。《自然》，574(7779):505–510，2019年。[5] Bostanci, J. and Watrous, J. 量子博弈论及量子纳什均衡近似计算的复杂性。Quantum，6:882，2022年。[6] Carlen, E. A. 迹不等式与量子熵：入门课程。2009年。[7] Chakrabarti, S., Huang, Y., Li, T., Feizi, S., and Wu, X. 量子Wasserstein生成对抗网络。载于NeurIPS ’19：第33届神经信息处理系统国际会议论文集，2019年。[8] Chuang, I. and Nielsen, M. 量子计算与量子信息。剑桥大学出版社，第二版，2010年。[9] Coucheney, P., Gaujal, B., and Mertikopoulos, P. 博弈中的惩罚调节动力学与稳健学习过程。《运筹学数学》，40(3):611–633，2015年8月。[10] Cressman, R. 演化动力学与扩展式博弈。麻省理工学院出版社，2003年。[11] Dallaire-Demers, P.-L. and Killoran, N. 量子生成对抗网络。《物理评论A》，98(1):012324，2018年。[12] Daskalakis, C., Goldberg, P. W., and Papadimitriou, C. H. 计算纳什均衡的复杂性。载于STOC ’06：第38届美国计算机协会计算理论年度研讨会论文集，2006年。[13] Daskalakis, C., Goldberg, P. W., and Papadimitriou, C. H. 计算纳什均衡的复杂性。《美国计算机协会通讯》，52(2):89–97，2009年。[14] Debreu, G. 社会均衡存在定理。《美国国家科学院院刊》，38(10):886–893，1952年10月。[15] Eisert, J., Wilkens, M., and Lewenstein, M. 量子博弈与量子策略。Physical Review Letters，83(15):3077–3080，1999年10月。[16] Facchinei, F. and Pang, J.-S. 有限维变分不等式与互补问题。施普林格运筹学丛书。Springer，2003年。[17] Flokas, L., Vlatakis-Gkaragkounis, E. V., Lianeas, T., Mertikopoulos, P., and Piliouras, G. 无悔学习与混合纳什均衡：它们并不混合。载于NeurIPS ’20：第34届神经信息处理系统国际会议论文集，2020年。[18] Folland, G. B. 实分析。威利-跨科学出版社，第二版，1999年。[19] Gutoski, G. and Watrous, J. 迈向量子博弈的一般理论。载于STOC ’07：第39届美国计算机协会计算理论年度研讨会论文集，2007年。[20] Hidalgo, E. G. 量子复制子动力学。《物理学A：统计力学及其应用》，369(2):393–407，2006年9月。[21] Hofbauer, J. and Sigmund, K. 演化博弈动力学。《美国数学会通报》，40(4):479–519，2003年7月。[22] Jain, R. and Watrous, J. 非交互零和量子博弈的并行近似。载于CCC ’09：2009年电气电子工程师学会国际计算复杂性会议论文集，2009年。[23] Jain, R., Piliouras, G., and Sim, R. 量子博弈中的矩阵乘法权重更新：守恒律与复现。载于NeurIPS ’22：第36届神经信息处理系统国际会议论文集，2022年。[24] Kakade, S. M., Shalev-Shwartz, S., and Tewari, A. 矩阵学习的正则化技术。《机器学习研究杂志》，13:1865–1890，2012年。[25] Kwon, J. and Mertikopoulos, P. 在线优化的连续时间方法。《动力学与博弈杂志》，4(2):125–148，2017年4月。[26] Lattimore, T. and Szepesvári, C. 多臂老虎机算法。剑桥大学出版社，英国剑桥，2020年。[27] Lloyd, S. and Weedbrook, C. 量子生成对抗学习。Physical Review Letters，121(4): 040502–040507，2018年7月。[28] Maynard Smith, J. 演化与博弈论。剑桥大学出版社，剑桥，1982年。[29] Maynard Smith, J. and Price, G. R. 动物冲突的逻辑。《自然》，246:15–18，1973年11月。[30] Mertikopoulos, P. and Sandholm, W. H. 通过强化与正则化实现博弈中的学习。《运筹学数学》，41(4):1297–1324，2016年11月。

[31] Mertikopoulos, P. 和 Zhou, Z. 在连续动作集和未知收益函数的博弈中学习。Mathematical Programming, 173(1-2):465–507，2019年1月。[32] Mertikopoulos, P.、Papadimitriou, C. H. 和 Piliouras, G. 对抗正则化学习中的周期。载于 SODA '18：第29届年度ACM-SIAM离散算法研讨会论文集，2018年。[33] Meyer, D. A. 量子策略。Physical Review Letters，82(5):1052–1055，1999年2月。[34] Petersen, K. B. 和 Pedersen, M. S. 矩阵手册，2012年11月。[35] Piliouras, G. 和 Shamma, J. S. 混沌中的优化：通过庞加莱复现对复杂极限集进行凸松弛。载于 SODA '14：第 25 届年度 ACM-SIAM 离散算法研讨会论文集，2014 年。[36] Preskill, J. NISQ 时代及之后的量子计算。Quantum，2:79，2018年8月。[37] Rockafellar, R. T. 凸分析。普林斯顿大学出版社，普林斯顿，纽约州，1970年。[38] Rockafellar, R. T. 和 Wets, R. J. B. 变分分析，数学综合研究丛书第317卷。Springer-Verlag，柏林，1998年。[39] Sandholm, W. H. 群体博弈与演化动力学。麻省理工学院出版社，剑桥，马萨诸塞州，2010年。[40] Scutari, G.、Facchinei, F.、Palomar, D. P. 和 Pang, J.-S. 多用户通信系统中的凸优化、博弈论与变分不等式理论。IEEE Signal Process. Mag.，27(3): 35–49，2010年5月。[41] Shalev-Shwartz, S. 在线学习与在线凸优化。Foundations and Trends in Machine Learning，4(2):107–194，2011年。[42] Shalev-Shwartz, S. 和 Singer, Y. 凸重复博弈与Fenchel对偶。载于 NIPS' 06：第19届神经信息处理系统年会论文集，第1265–1272页。麻省理工学院出版社，2006年。[43] Shapiro, A.、Dentcheva, D. 和 Ruszczyński, A. 随机规划讲义。工业与应用数学学会，费城，宾夕法尼亚州，2009年。[44] Slivkins, A. 多臂老虎机导论。Foundations and Trends in Machine Learning，12 (1-2):1–286，2019年11月。[45] Taylor, P. D. 和 Jonker, L. B. 演化稳定策略与博弈动力学。Mathematical Biosciences，40(1-2):145–156，1978年。[46] Tsuda, K.、Rätsch, G. 和 Warmuth, M. K. 用于在线Bregman投影的矩阵指数梯度更新。Journal of Machine Learning Research，6:995–1018，2005年。[47] Viossat, Y. 和 Zapechelnyuk, A. 无悔动力学与虚拟博弈。Journal of Economic Theory，148(2):825–842，2013年3月。[48] Weibull, J. W. 演化博弈论。麻省理工学院出版社，剑桥，马萨诸塞州，1995年。[49] Wilcox, R. M. 量子物理中的指数算符与参数微分。Journal of Mathematical Physics，8(4):962–982，1967年。[50] Zhong, H.-S.、Wang, H.、Deng, Y.-H.、Chen, M.-C.、Peng, L.-C.、Luo, Y.-H.、Qin, J.、Wu, D.、Ding, X.、Hu, Y.、Hu, P.、Yang, X.-Y.、Zhang, W.-J.、Li, H.、Li, Y.、Jiang, X.、Gan, L.、Yang, G.、You, L.、Wang, Z.、Li, L.、Liu, N.-L.、Lu, C.-Y. 和 Pan, J.-W. 利用光子实现量子计算优势。Science，370(6523):1460–1463，2020年12月。[51] Zimmert, J. 和 Seldin, Y. Tsallis-INF：随机与对抗老虎机的最优算法。Journal of Machine Learning Research，22(28):1–49，2021年。