arXiv:2304.00771v2 [math.OC] 2 Nov 2023

锚点加速的连续时间分析

Jaewook J. Suh
Seoul National University
jacksuhkr@snu.ac.kr

Jisun Park
首尔国立大学
colleenp0515@snu.ac.kr

ErnestK. Ryu首尔国立大学
ernestryu@snu.ac.kr

摘要

近期，人们在极小极大优化和不动点问题中发现了一种不同于 Nesterov 加速的锚点加速机制，但对其机制的理解尚不深入，远不如对 Nesterov 加速的理解充分。本文分析了锚点加速的连续时间模型。我们针对锚点系数 $\beta(t)$，给出了刻画收敛速率的紧致统一分析，从而深入揭示锚点加速机制及其加速的 $\mathcal{O}(1/k^2)$-收敛速率。最后，我们提出了一种受连续时间分析启发的自适应方法，并通过理论分析和实验验证了其有效性。

1 引言

Nesterov 加速 [51] 是一阶优化理论的基石，但其机制及收敛证明并不直观。深入理解该机制的一种途径是连续时间分析：推导离散时间算法的常微分方程模型，并分析连续时间动力学 [65, 66]。这种方法有助于理解加速动力学，并催生了一系列后续研究 [29, 62, 71]。

近年来，人们发现了一种有别于 Nesterov 加速的新型加速机制。这种锚点加速在极小极大优化和不动点问题中的应用 [35, 54, 75] 一直是研究热点，但与 Nesterov 加速相比，人们对其机制的理解要少得多。为理解 Nesterov 加速而发展出的各种分析技术，包括连续时间分析，仅在非常有限的范围内得到应用 [59]。

贡献。 本文对锚点加速进行了连续时间分析。其连续时间模型为以下微分包含：

$$\dot{X} \in -\mathbb{A}(X) - \beta(t)(X - X_0)$$

其初始条件为 $X(0) = X_0 \in \operatorname{dom} \mathbb{A}$，单调算子为 $\mathbb{A}$，标量值函数为 $\beta(t)$。当 $\beta(t) = \frac{1}{t}$ 时，对应于先前的锚点加速方法 APPM [35],EAG [75], 和 FEG [42]。

我们首先证明，尽管锚点系数 $\beta(t)$在 $t = 0$处趋于无穷，微分包含仍然是适定的。随后，我们给出紧致且统一的分析，以刻画收敛速率如何随锚点系数 $\beta(t)$变化。这是对该锚定动力学首次进行正式且严谨的处理，并揭示了锚点加速机制及其加速$\mathcal{O}(1/k^2)$收敛速率。最后，我们提出一种受连续时间分析启发的自适应方法，并通过理论分析和实验验证其有效性。

1.1 预备知识与符号说明

我们在附录 A 中介绍先前工作的组织结构。这里，我们回顾标准定义并建立符号约定。

第37届神经信息处理系统会议（NeurIPS 2023）。

单调集值算子。 我们遵循 Bauschke 和 Com- bettes [15],Ryu 与 Yin [58] 的标准定义。对于底层空间，考虑具有标准内积$\langle\cdot,\ \cdot\rangle$和范数 $\|\cdot\|$的 R^n。将 $\mathbb{A}$ 的定义域定义为 $\mathrm{dom}\ \mathbb{A}=\{x\in\mathrm{R}^n \mid \mathbb{A}x\neq\emptyset\}$。我们称 $\mathbb{A}$ 是 R^n 上的算子，并记作 $\mathbb{A}\colon \mathrm{R}^n\rightrightarrows\mathrm{R}^n$，如果 $\mathbb{A}$ 将 R^n 中的一个点映射到 R^n 的一个子集。若 $\mathbb{A}\colon \mathrm{R}^n\rightrightarrows\mathrm{R}^n$ 满足以下条件，则称其为单调算子：

$$\langle \mathbb{A}x-\mathbb{A}y, x-y\rangle\geq 0,\qquad \forall x,y\in\mathbb{R}^n,$$

其中该记号表示，对于所有 $u\in\mathbb{A}x$ 和 $v\in\mathbb{A}y$，$\langle u-v,\ x-y\rangle\geq 0$。对于 $\mu\in(0,\infty)$，称$\mathbb{A}\colon \mathrm{R}^n\rightrightarrows\mathrm{R}^n$ 为 μ-强单调，如果

$$\langle \mathbb{A}x-\mathbb{A}y, x-y\rangle\geq \mu\|x-y\|^2,\qquad \forall x,y\in\mathbb{R}^n.$$

用 $\mathrm{Gra}\,\mathbb{A}=\{(x,u)\mid u\in\mathbb{A}x\}$ 表示 $\mathbb{A}$ 的图。若不存在其他单调算子 $\mathbb{B}$ 使得 $\mathrm{Gra}\ \mathbb{A}\subset\mathrm{Gra}\,\mathbb{B}$ 真包含关系成立，则算子 $\mathbb{A}$ 称为极大单调；若不存在其他 μ-强单调算子 $\mathbb{B}$ 使得 $\mathrm{Gra}\ \mathbb{A}\subset\mathrm{Gra}\,\mathbb{B}$ 真包含关系成立，则称其为极大 μ-强单调。

对于 $L\in(0,\infty)$，若单值算子 $\mathbb{T}\colon \mathrm{R}^n\to\mathrm{R}^n$ 满足以下条件，则称其为 L-Lipschitz：

$$\|\mathbb{T}x-\mathbb{T}y\|\leq L\|x-y\|,\qquad \forall x,y\in\mathbb{R}^n.$$

用 $\mathbb{J}_{\mathbb{A}}=(\mathbb{I}+\mathbb{A})^{-1}$ 表示 $\mathbb{A}$ 的预解式，而 $\mathbb{I}\colon \mathrm{R}^n\to\mathrm{R}^n$ 是恒等算子。当 $\mathbb{A}$ 为极大单调时，众所周知，$\mathbb{J}_{\mathbb{A}}$ 是单值的，且 $\mathrm{dom}\,\mathbb{J}_{\mathbb{A}}=\mathrm{R}^n$。

我们称 $x_\star\in\mathrm{R}^n$ 是 $\mathbb{A}$ 的零点，当且仅当 $0\in\mathbb{A}x_\star$。我们称 $y_\star$ 是 $\mathbb{T}$ 的不动点，当且仅当 $\mathbb{T}y_\star=y_\star$。用 $\mathrm{Zer}\ \mathbb{A}$ 表示 $\mathbb{A}$ 的所有零点组成的集合，用 $\mathrm{Fix}\,\mathbb{T}$ 表示 $\mathbb{T}$ 的所有不动点组成的集合。

连续曲线下的单调性。 如果一个算子是单值、连续且作为函数可微的，我们就称它为可微算子。如果可微算子 $\mathbb{A}$ 是单调的，并且 $X\colon [0,\infty)\to\mathrm{R}^n$ 是一条可微曲线，那么对 $h\to 0$ 取极限

$$\frac{1}{h^2}\langle \mathbb{A}(X(t+h))-\mathbb{A}(X(t)), X(t+h)-X(t)\rangle\geq 0$$

得到

$$\left\langle \frac{d}{dt}\mathbb{A}(X(t)), \dot{X}(t)\right\rangle\geq 0. \tag{1}$$

同样，如果 $\mathbb{A}$ 还是 μ-强单调的，那么

$$\left\langle \frac{d}{dt}\mathbb{A}(X(t)), \dot{X}(t)\right\rangle\geq \mu\left\|\dot{X}(t)\right\|^2. \tag{2}$$

2 锚点加速微分包含模型的推导

2.1 锚定常微分方程

设 $\mathbb{A}\colon \mathrm{R}^n\rightrightarrows\mathrm{R}^n$ 是一个极大单调算子，且 $\beta\colon(0,\infty)\to[0,\infty)$ 是一个二阶可微函数。考虑微分包含

$$\dot{X}(t)\in-\mathbb{A}(X(t))-\beta(t)(X(t)-X_0) \tag{3}$$

其初始条件为 $X(0)=X_0\in\mathrm{dom}(\mathbb{A})$。我们将其称为锚定常微分方程。[1] 我们称$X\colon[0,\infty)\to\mathrm{R}^n$ 为一个解，若其绝对连续，且对 $t\in(0,\infty)$ 中几乎处处满足 (3)。

记 S 为 $[0,\infty)$ 的子集，在该子集上 X 满足微分包含。定义

$$\tilde{\mathbb{A}}(X(t))=-\dot{X}(t)-\beta(t)(X(t)-X_0)$$

~ ~对于 $t\in S$。由于当 $t\in S$ 时，$\mathbb{A}(X(t))\in\mathbb{A}(X(t))$，因此我们称 $\mathbb{A}$ 是 $t\in S$的 $\mathbb{A}$ 选择。如果$\|\tilde{\mathbb{A}}(X(t))\|$在 S的所有有界子集上都有界，那么我们可以将 $\tilde{\mathbb{A}}$扩展到 $[0,\infty)$，同时保留某些有利性质。我们在附录 E.1 中讨论这一扩展的技术细节。第 3 节中的陈述均基于这一扩展。

[1]严格来说，这是一个微分包含，而非微分方程，但我们仍将其称为常微分方程。

2.2 离散方法的推导

现在我们证明，以下锚定常微分方程的一个实例

$$\dot{X}(t) = -\mathbb{A}(X(t)) - \frac{1}{t}(X(t) - X_0), \tag{4}$$

其中， $X(0) = X_0$ 是初始条件， $\mathbb{A}$: $\mathrm{R}^n \to \mathrm{R}^n$ 是一个连续算子，它是 APPM [35], EAG [75], 和 FEG [42], 的连续时间模型；这些方法用于求解单调包含和极小极大问题。

考虑算子为 $h\mathbb{A}$ 的 APPM

$$\begin{aligned} x^k &= \mathbb{J}_{h\mathbb{A}} y^{k-1} \\ y^k &= \frac{k}{k+1}(2x^k - y^{k-1}) + \frac{1}{k+1} y^0 \end{aligned} \tag{5}$$

初始条件为 $y^0 = x^0$。假设 $h > 0$ 和 $\mathbb{A}$: $\mathrm{R}^n \to \mathrm{R}^n$ 是连续单调算子。利用由第一行得到的 $y^{k-1} = x^k + h\mathbb{A}x^k$，将 y^k 和 y^{k-1} 代入第二行，我们得到:

$$x^{k+1} + h\mathbb{A}x^{k+1} = \frac{k}{k+1}\left(x^k - h\mathbb{A}x^k\right) + \frac{1}{k+1}x^0.$$

然后重新整理并将等式两边除以 h，得到

$$\frac{x^{k+1} - x^k}{h} = -\mathbb{A}x^{k+1} - \frac{k}{k+1}\mathbb{A}x^k - \frac{1}{h(k+1)}(x^k - x^0).$$

令 $x^0 = X_0$、 $2hk = t$和 $x^k = X(t)$ 分别为相应项，则有 $\frac{k}{k+1} = 1 - \frac{h}{hk+h} = 1 + \mathcal{O}(h)$，因此

$$2\dot{X}(t) + \mathcal{O}(h) = -\mathbb{A}(X(t+2h)) - (1 + \mathcal{O}(h))\,\mathbb{A}(X(t)) - \frac{2}{t + \mathcal{O}(h)}\left(X(t) - X_0\right).$$

取极限 $h \to 0^+$，并将两边除以 2，即得锚定常微分方程 (4)。与 EAG 和 FEG 的对应关系见附录 D.4。

The 以下定理建立了 APPM 与锚定常微分方程之间的严格对应关系
for 一般极大单调算子。

定理 2.1。 设 $\mathbb{A}$ 为一个（可能是集值的）极大单调算子，并假设 零点$\mathbb{A} \neq \emptyset$。设 x^k 为由 *APPM* (5) 生成的序列， X 为微分包含 (3) 的解，其中 $\beta(t) = \frac{1}{t}$。对于所有固定的 $T > 0$,

$$\lim_{h \to 0+} \max_{0 \le k \le \frac{T}{2h}} \left\| x^k - X(2kh) \right\| = 0.$$

证明见附录 D.2。

2.3 $\beta(t) = \frac{\gamma}{t^p}$ 解的存在性

为了进一步理解锚点加速，我们将锚点系数推广到 $\beta(t) = \frac{\gamma}{t^p}$ for$p$， $\gamma > 0$。我们首先证明该解的唯一性和存在性。

定理 2.2。 考虑带有 $\beta(t) = \frac{\gamma}{t^p}$ 的(3) ,即

$$\dot{X}(t) \in -\mathbb{A}(X(t)) - \frac{\gamma}{t^p}(X(t) - X_0). \tag{6}$$

对于 $p, \gamma > 0$，方程(6) 的解唯一存在。

证明见附录 B。

2.4 锚点常微分方程的其他性质

我们陈述微分包含 (3) 的一个正则性引理，我们认为该结果本身也可能具有一定的研究价值。具体而言，我们在多个证明中多次使用了这一结果。

引理 2.3. 设 $X(\cdot)$ 和 $Y(\cdot)$ 分别是微分包含 (3) 的解，其初值和锚点分别为 X_0 和 Y_0。那么对于所有 $t \in [0,\infty)$,

$$\|X(t) - Y(t)\| \le \|X_0 - Y_0\| .$$

证明见附录 B.1。

根据引理 2.3，轨迹的有界性是一个直接推论。具体而言，假设 $X(\cdot)$ 是微分包含 (3) 在初值 X_0 下的解。那么对所有 $X_\star \in \mathrm{Zer}\mathbb{A}$ and$t \in [0,\infty)$,

$$\|X(t) - X_\star\| \le \|X_0 - X_\star\| .$$

这可通过在引理 2.3 中令 $Y_0 = X_\star$ 得到。

3 收敛分析

现在我们分析锚定常微分方程 (3) 中 $\left\|\tilde{\mathbb{A}}(X(t))\right\|^2$ 的收敛速率，其中 $\beta(t) = \frac{\gamma}{t^p}$ and$\gamma,\ p > 0$。结果汇总于表 1。

Case	$p=1,\ \gamma \ge 1$	$p=1,\ \gamma<1$	$p<1$	$p>1$
$\left\|\tilde{\mathbb{A}}(X(t))\right\|^2$	$\mathcal{O}\left(\frac{1}{t^2}\right)$	$\mathcal{O}\left(\frac{1}{t^{2\gamma}}\right)$	$\mathcal{O}\left(\frac{1}{t^{2p}}\right)$	$\mathcal{O}(1)$

表 1：定理 3.1 的收敛速率。

令 β 为式 (3) 的锚定系数函数。定义 $C\colon [0,\infty) \to \mathbb{R}$ 为 $C(t) = e^{\int_v^t \beta(s)\,ds}$，其中 $v \in [0,\infty)$。注意，$C = C_\beta$ 和 C 在相差一个标量因子的意义下是唯一的。我们将 $\mathcal{O}(\beta(t))$称为消失速度，将$\mathcal{O}\left(\frac{1}{C(t)}\right)$ 称为收缩速度，并在下文描述二者之间的权衡。

粗略地说，收缩速度 描述了仅由锚定项使动力系统收缩的速度。考虑 $\dot{X}(t) = -\beta(t)(X(t) - a)$，其中 $a \in \mathbb{R}^n$，这是一个仅含锚点的系统。那么，$X(t) = \frac{C(0)}{C(t)}(X(0) - a) + a$是其解，因此流以速率$\frac{1}{C(t)}$向锚点收缩 a。直观地说，这种收缩行为会带来稳定性和收敛。另一方面，锚点最终必须消失，因为我们的目标是收敛到 $\mathrm{Zer}\mathbb{A}$ 中的一个元素，而不是锚点。因此，消失速度 必须足够快，不能减慢流向 $\mathrm{Zer}\mathbb{A}$ 的收敛速度。

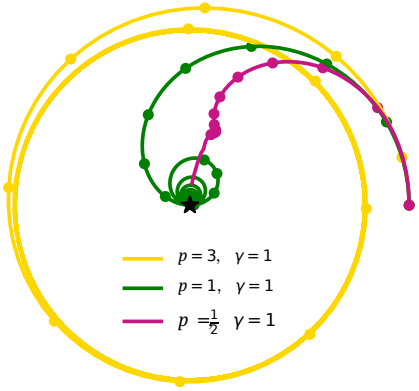


图 1：当 $\mathbb{A} = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$、$\gamma = 1$、$X_0 = (1, 0)$ 且 p 取不同值时，式 (6) 的解的流。流从 $t = 0$ 到$t = 100$。每经过 0.8 个时间单位绘制一个标记，直到 $t = 9.6$。注意，$p = \frac{1}{2}$ 对应流的最后一个标记比 $p = 1$ 对应流的最后一个标记离最优点 $\star$ 更远。

图 1 展示了这一观察结果。考虑 $\mathbb{R}^2$ 上的单调线性算子 $\mathbb{A} = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ 以及满足 $\gamma = 1$ 和 $p > 0$的 $\beta(t) = \frac{\gamma}{t^p}$。注意，如果没有锚点，常微分方程会化简为 $\dot{X} = -\mathbb{A}(X)$，而该方程不会收敛 [31,第 8 章。2]。图 1 显示，当 $p > 1$时，锚点消失得过早，此时流还没有充分收缩，因此无法形成收敛的流。当 $p < 1$时，流确实会收敛，但锚点消失得太晚，减慢了收敛速度。当 $p = 1$时，收敛最快。

下面的定理将这一见解形式化，并给出表 1 中的结果。

定理 3.1。 假设 $\mathbb{A}$ 是一个极大单调算子，且零点$\mathbf{A} \neq \emptyset$。考虑(3)，其中$\beta(t) = \frac{\gamma}{t^p}$。令 $\mathbb{A}(X(t))$ 为 *Section 2.1* 中所述的 $\mathbb{A}(X(t))$ 的选择。那么，

$$\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 = \mathcal{O}\left(\frac{1}{C(t)^2}\right) + \mathcal{O}\left(\beta(t)^2\right) + \mathcal{O}\left(\dot{\beta}(t)\right).$$

注意，

$$C(t) = \begin{cases} t^\gamma & p = 1 \\ e^{\frac{\gamma}{1-p}t^{1-p}} & p \neq 1. \end{cases}$$

我们预计，当各项达到平衡时，定理 3.1 的收敛速率将达到最优。当 $\beta(t) = \frac{1}{t}$，

$$\frac{1}{C(t)^2} = \frac{1}{(e^{\int_1^t \frac{1}{s}ds})^2} = \frac{1}{t^2} = \beta(t)^2 = -\dot{\beta}(t)$$

且这三项均达到平衡。事实上，选择 $\beta(t) = \frac{1}{t}$ 对应于 APPM 或其他加速方法的最优离散时间选择 $\frac{1}{k+2}$。

3.1 定理 3.1 的证明概要

定理 3.1 的证明由引理 3.4 推出，我们将在本节稍后介绍该引理。为推导引理 3.4，我们引入一个守恒律。

命题 3.2. 假设 $\mathbb{A}$ 是 *Lipschitz* 连续且单调的。对于 $t_0 > 0$，定义 $E : (0, \infty) \to \mathbb{R}$ *as*

$$E = \frac{C(t)^2}{2}\left(\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\beta(t)\left\langle\tilde{\mathbb{A}}\left(X(t)\right), X(t) - X_0\right\rangle + \left(\beta(t)^2 + \dot{\beta}(t)\right)\left\|X(t) - X_0\right\|^2\right)$$
$$- \int_{t_0}^t \frac{d}{ds}\left(\frac{C(s)^2\dot{\beta}(s)}{2}\right)\left\|X(s) - X_0\right\|^2 ds + \int_{t_0}^t C(s)^2\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)), \dot{X}(s)\right\rangle ds.$$

那么， E 是一个常函数。

命题 3.2 的证明采用膨胀坐标 $W(t) = C(t)(X(t) - X_0)$，并按照 Suh 等人 [67] 的方法推导其守恒律。详细内容见附录 E.2。

回顾式 (1)，$\left\langle \frac{d}{ds}\tilde{\mathbb{A}}(X(s)),\ \dot{X}(s)\right\rangle \geq 0$，$E$ 的最后一项中的被积函数非负。这促使我们定义

$$V(t) = E - \int_{t_0}^t C(s)^2\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)), \dot{X}(s)\right\rangle ds$$

作为我们的李雅普诺夫函数。

推论 3.3。 设 $\mathbb{A}$ 为极大单调算子 A，且 $\beta(t) = \frac{\gamma}{t^p}$ 满足 $p > 0, \gamma > 0$。设 $\mathbb{A}(X(t))$ 为第 *2.1* 节中 $\mathbb{A}(X(t))$ 的选择。对于 $t_0 \geq 0$，定义 $V : [0, \infty) \to \mathbb{R}$ 为

$$V(t) = \frac{C(t)^2}{2}\left(\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\beta(t)\left\langle\tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle + \left(\beta(t)^2 + \dot{\beta}(t)\right)\left\|X(t) - X_0\right\|^2\right)$$
$$- \int_{t_0}^t \frac{d}{ds}\left(\frac{C(s)^2\dot{\beta}(s)}{2}\right)\left\|X(s) - X_0\right\|^2 ds.$$

对于 $t > 0$ 和 $V(0) = \lim_{t\to 0+} V(t)$，则 $V(t) \leq V(0)$ 对 $t \geq 0$ 成立。

一个技术细节是，在定义 V 时排除了所有涉及 $\frac{d}{ds}\tilde{\mathbb{A}}(X(s))$ 的项，这使得 $\mathbb{A}$ 不必是 Lipschitz 连续的。详细内容见附录 E.3。

引理 3.4. 考虑推论 *3.3* 的设定。假设 $\mathrm{Zer}\mathbf{A} \neq \emptyset$。那么对于 $t > 0$ 和 $X_\star \in \mathrm{Zer}\mathbf{A}$，

$$\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 \leq 4\beta(t)^2\left\|X_0 - X_\star\right\|^2 + \frac{4V(0)}{C(t)^2} - 2\left(\beta(t)^2 + \dot{\beta}(t)\right)\left\|X(t) - X_0\right\|^2$$
$$+ \frac{2}{C(t)^2}\int_{t_0}^t \frac{d}{ds}\left(C(s)^2\dot{\beta}(s)\right)\left\|X(s) - X_0\right\|^2 ds. \tag{7}$$

引理 *3.4* 的证明思路。定义

$$\Phi(t) = \left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\beta(t)\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle.$$

˜因此，根据 $\mathbb{A}$ 的单调性和杨氏不等式，

$$\begin{aligned}\Phi(t) &\geq \left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\beta(t)\left\langle \tilde{\mathbb{A}}(X(t)), X_\star - X_0\right\rangle \\ &\geq \left\|\tilde{\mathbb{A}}(X(t))\right\|^2 - 2\left(\left\|\frac{1}{2}\tilde{\mathbb{A}}(X(t))\right\|^2 + \left\|\beta(t)\left(X_\star - X_0\right)\right\|^2\right) \\ &= \frac{1}{2}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 - 2\beta(t)^2\left\|X_0 - X_\star\right\|^2. \end{aligned} \tag{8}$$

根据推论 3.3，$\frac{2V_{(0)}}{C(t)^2} - \frac{2V}{(t)\,C(t)_2} + \Phi(t) \geq \Phi(t)$，其中 $t > 0$。应用式 (8) 并进行整理，即可得到所需结果。详细内容见附录 E.4。 □

定理 *3.1* 的证明概要。剩下需要证明的是，引理 3.4 中最后的积分项为 $\mathcal{O}\left(\frac{1}{C\,t\,()^2}\right) + \mathcal{O}\left(\beta(t)^2\right) + \mathcal{O}\left(\dot{\beta}(t)\right)$。详细内容见附录 E.5。 □

在结束本节之前，我们观察一下，在特殊情形 $\beta(t) = \frac{1}{t}$ 下，我们的分析如何得到简化。在这种情况下，

$$V(t) = \frac{t^2}{2}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + t\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle,$$

这对应于情形 $\gamma = 1$ 下 [59, 第 4] 节的李雅普诺夫函数。由于 $V(0) = 0$，引理 3.4 的结论变为

$$\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 \leq \frac{4}{t^2}\left\|X_0 - X_\star\right\|^2 = \mathcal{O}\left(\frac{1}{t^2}\right),$$

其对应于表 1 中的最佳速率。

3.2 点收敛

APPM 是 Halpern 方法的一种实例 [54, 引理 3.1], 其迭代收敛到 $\mathrm{Zer}\mathbb{A}$ 中最接近 X_0 [34, 73]的元素。锚定常微分方程也表现出这一性质。

定理 3.5。设 $\mathbb{A}$ 是一个极大单调算子，满足 $\mathrm{Zer}\mathbb{A} \neq \emptyset$，且 X 是（3）的解。若 $\lim_{t\to\infty}\left\|\tilde{\mathbb{A}}(X(t))\right\| = 0$ 且 $\lim_{t\to\infty} 1/C(t) = 0$，则当 $t \to \infty$ 时，

$$X(t) \to \operatorname*{argmin}_{z\in\mathrm{Zer}\mathbb{A}}\left\|z - X_0\right\|.$$

证明见附录 E.6。

4 紧性分析

本节通过考察显式示例 $\mathbb{A} = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ 下的动力学，证明表 1 中的收敛速率确实是紧的。在本节中，当算子为线性算子时，我们将 $\mathbb{A}$ 记作 A。

4.1 线性 A的显式解

引理 4.1。设 $A\colon \mathbb{R}^n \to \mathbb{R}^n$ 是一个线性算子，且 $\beta(t) = \frac{\gamma}{t}$。级数

$$X(t) = \sum_{n=0}^{\infty} \frac{(-tA)^n}{\Gamma(n+\gamma+1)}\Gamma(\gamma+1)X_0,$$

其中 Γ 表示伽马函数，是满足 $\mathbb{A} = A$ 的(3) 的解。

注意，当 $\gamma = 0$ 时，这就是矩阵指数的级数定义，且 $X(t) = e^{-tA}$。该解还有一种积分形式，可推广到一般的 $\beta(t)$。

引理 4.2。 假设 $A\colon\ \mathbb{R}^n \to \mathbb{R}^n$ 是一个单调线性算子。那么

$$X(t) = \frac{e^{-tA}}{C(t)} \left(\int_0^t e^{sA} C(s)\beta(s)ds + C(0)I \right) X_0 \tag{9}$$

是(3) 的解，其中 $\mathbb{A} = A$。

详情请参见附录 F.1.1 和附录 F.1.2。

4.2 表 1 中的速率是紧的

首先，我们考虑 $p > 1$ 对于 $\beta(t) = \frac{\gamma}{t^p}$。

定理 4.3。 假设$\lim_{t\to\infty} C^{1}{}_{t()} \neq 0$, 即假设 $\beta(t) \in L^1[t_0, \infty)$ 对某个 $t_0 > 0$成立。那么存在一个算子 $\mathbb{A}$ 使得

$$\lim_{t\to\infty} \left\| \tilde{\mathbb{A}}(X(t)) \right\| \neq 0,$$

其中 X 是(3)的解。

注意，当 $p > 1$时， $\frac{\gamma}{t^p} \in L^1[t_0, \infty)$。定理 4.3 的证明考虑 $\mathbb{A} = 2\pi\xi \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \right)$ for$\xi \in \mathbb{R}$，并使用傅里叶反演公式。详情见附录 F.2。

接下来，我们考虑 $\beta(t) = \frac{\gamma}{t^p}$ for除 $p > 1$以外的情形。

定理 4.4. 设 $A = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \right)$, $\beta(t) = \frac{\gamma}{t^p}$, $0 \leq p \leq 1$和 $\gamma > 0$。设 X 为(9) 给出的解，并且 $X_0 \neq 0$。设

$$r(t) = \begin{cases} t^2 & \textit{for } p = 1, \gamma \geq 1 \\ t^{2\gamma} & \textit{for } p = 1, \gamma < 1 \\ t^{2p} & \textit{for } 0 < p < 1. \end{cases},$$

于是，

$$\lim_{t\to\infty} r(t) \left\| A(X(t)) \right\|^2 \neq 0.$$

证明见附录 F.3。

5 离散化算法

In t本节给出与连续时间收敛速率相匹配的离散时间收敛结果 f Sec第 3 节。

定理 5.1。 设 $\mathbb{A}$ 是一个极大单调算子， $p > 0$，且 $\gamma > 0$。考虑

$$\begin{aligned} x^k &= \mathbb{J}_{\mathbb{A}} y^{k-1} \\ y^k &= \frac{k^p}{k^p + \gamma}(2x^k - y^{k-1}) + \frac{\gamma}{k^p + \gamma} x^0 \end{aligned}$$

对于 $k = 1$、2、.....，其初始条件为 $y^0 = x^0 \in \mathbb{R}^n$。令 $\tilde{\mathbb{A}} x^k = y^{k-1} - x^k$，其中 $k = 1$、2、.....。则该方法具有表 *2* 中所示的收敛速率。

Case	$p = 1$, $\gamma \geq 1$	$p = 1$, $\gamma < 1$	$p < 1$	$p > 1$
$\left\| \tilde{\mathbb{A}}(x^k) \right\|^2$	$\mathcal{O}\left(\frac{1}{k^2}\right)$	$\mathcal{O}\left(\frac{1}{k^{2\gamma}}\right)$	$\mathcal{O}\left(\frac{1}{k^{2p}}\right)$	$\mathcal{O}(1)$

表 2：定理 5.1 的离散时间方法的收敛速率。

注意到，定理 5.1 中的方法在 $\gamma = 1$、 $p = 1$时退化为 APPM。

定理 *5.1* 的证明概要。 总体策略是寻找相应连续时间分析的离散化形式。然而，由于技术上的原因，直接对命题 3.2 的守恒律进行离散化较为困难。因此，我们利用膨胀坐标获得了经过不同缩放但等价的守恒律，然后对其进行离散化。具体的膨胀坐标受 [67], 启发，分别为 $W_1(t) = X(t) - X_0$，适用于 $p > 1$； $W_2(t) = t^p(X(t) - X_0)$，适用于 $0 < p < 1$; $W_3(t) = t(X(t) - X_0)$，适用于 $p = 1$; 以及 $\gamma \geq 1$ 和 $W_4(t) = t^\gamma (X(t) - X_0)$，适用于 $p = 1$、$0 < \gamma < 1$。
在离散时间分析中，主导项的表现是可以预测的，因为它们与连续时间分析中的对应项一致。然而，困难之处在于控制连续时间分析中不存在的高阶项。通过详细分析，我们对这类高阶项进行了界定，并证明它们最终不会影响收敛速率。详情见附录G.3。

□

6 强单调性下的收敛分析

本节分析 μ-强单调 $\mathbb{A}$ 下的锚定常微分方程 (3) 的动力学。当$\beta(t) = \frac{1}{t}$且 $\mathbb{A} = \left(\begin{smallmatrix} \mu & 0 \\ 0 & \mu \end{smallmatrix}\right)$时，引理 4.1 告诉我们 $\mathbb{A}(X(t)) = \frac{1}{t}\left(I - e^{-tA}\right) X_0$，因此$\|\mathbb{A}(X(t))\|^2 = \Theta\left(\frac{1}{t^2}\right)$，这在强单调设定下是一个慢速率。另一方面，我们将看到， $\beta(t) = \frac{2\mu}{e^{2\mu t}-1}$ 是该设定下能带来更快速率的更好选择。

本节的分析同样基于一条守恒律，但我们采用了一个略作修改的版本，以利用强单调性。

命题 6.1。 假设 $\mathbb{A}$ 是单调且利普希茨连续的。设 X 是(3) 的解，并设 $R\colon [0,\infty) \to (0,\infty)$ 是可微函数。对于 $t_0 > 0$，定义 $E\colon (0,\infty) \to \mathbb{R}$ 为

$$E = \frac{C(t)^2 R(t)^2}{2}\left(\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\beta(t)\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0 \right\rangle + \left(\beta(t)^2 + \dot{\beta}(t)\right) \|X(t) - X_0\|^2 \right)$$
$$- \int_{t_0}^{t} \frac{d}{ds}\left(\frac{C(s)^2 R(s)^2 \dot{\beta}(s)}{2} \right) \|X(t) - X_0\|^2 \, ds + \int_{t_0}^{t} C(s)^2 R(s)^2 \left(\left\langle \frac{d}{ds}\tilde{\mathbb{A}}(X(s)), \dot{X}(s) \right\rangle - \frac{\dot{R}(s)}{R(s)} \left\|\dot{X}(s)\right\|^2 \right) ds.$$

于是，对于 $t \in [0,\infty)$， E 是一个常函数。

命题 6.1 推广了命题 3.2，因为后者对应于 $R(t) \equiv 1$的特殊情形。回顾式 (2)，当 $\mathbb{A}$ 是 μ-强单调时，我们有

$$\left\langle \frac{d}{ds}\mathbb{A}(X(t)), \dot{X}(t) \right\rangle - \mu \left\|\dot{X}(t)\right\|^2 \geq 0.$$

这启发我们选择 $R(t) = e^{\mu t}$，因为 $\frac{\dot{R}(s)}{R(s)} = \mu$。根据附录H.2中的计算，选择 $\beta(t) = \frac{2\mu}{e^{2\mu t}-1}$ 会使 $\frac{d}{ds}\left(\frac{C(s)^2 R(s)^2}{2} \right) = 0$。将这些选择代入命题 6.1，并沿用第 3 节的论证，我们便得到以下定理。

定理 6.2。 令 $\mathbb{A}$ 为一个 μ-强极大单调算子，且满足 $\mu > 0$，并假设 $\mathrm{Zer}\mathbb{A} \neq \emptyset$。令 X 为微分包含 (3) 的一个解，且满足 $\beta(t) = \frac{2\mu}{e^{2\mu t}-1}$，即对几乎所有的 t,

$$\dot{X} \in -\mathbb{A}(X) - \frac{2\mu}{e^{2\mu t} - 1}(X - X_0). \tag{10}$$

令 $\mathbb{A}(X(t))$ 为 $\mathbb{A}(X(t))$ 在第 *2.1* 节中的选择。定义 $V\colon [0,\infty) \to \mathbb{R}$ 为

$$V(t) = \frac{(e^{\mu t} - e^{-\mu t})^2}{2} \left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\mu\left(1 - e^{-2\mu t}\right)\left(\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0 \right\rangle - \mu \|X(t) - X_0\|^2 \right).$$

那么 $V(t) \leq V(0)$ 对 $t \geq 0$成立。此外，对于 $X_\star \in \mathrm{Zer}\mathbb{A}$,

$$\|\tilde{\mathbb{A}}(X(t))\|^2 \leq \left(\frac{2\mu}{e^{\mu t} - 1} \right)^2 \|X_0 - X_\star\|^2 = \mathcal{O}\left(\frac{1}{e^{2\mu t}} \right).$$

在附录 H.2.3 中，我们证明式 (10) 是 ParkandRyu [54] 的 OS-PPM 连续时间模型。在附录 C 中，我们证明该解的存在性和唯一性。

由于对于任意 $t_0 \geq 0$，$\beta(t) = \frac{2\mu}{e^{2\mu t}-1} \in L^1[t_0, \infty)$，定理 4.3 表明，当 $\mathbb{A}$ 仅为单调算子时，$\tilde{\mathbf{A}}(X(t)) \nrightarrow 0$。这告诉我们，$\beta(t)$ 的最优选择应取决于 $\mathbf{A}$ 的性质。在下一节中，我们将介绍如何选择 $\beta(t)$，使其适应算子的性质。

7 自适应锚点加速与实验

本节介绍一种选择锚点系数 β 的自适应方法，并从理论和实验两方面表明，这种选择能够使动力学适应算子的性质。当 $\mathbb{A}$ 为单调算子时，该方法实现最优的 $\mathcal{O}(1/k^2)$-收敛速率；当 $\mathbb{A}$ 进一步为 μ-强单调且 Lipschitz 连续时，则实现指数收敛速率。

定理 7.1。设 $\mathbb{A}$ 是 *Lipschitz* 连续且单调的。考虑锚定常微分方程

$$\dot{X} = -\tilde{\mathbb{A}}(X) + \underbrace{\frac{\left\|\tilde{\mathbb{A}}(X)\right\|^2}{2\left\langle \tilde{\mathbb{A}}(X), X - X_0 \right\rangle}}_{=-\beta(t)}(X - X_0) \tag{11}$$

初始条件为 $X(0) = X_0$，且 $\left\|\tilde{\mathbb{A}}(X_0)\right\| \neq 0$。假设解存在，且 $\dot{X}$ 在 $t = 0$处连续。此外，假设 $\beta\colon (0,\infty) \to \mathbb{R}$ 定义良好，即其定义中不会出现除以零的情况。那么，对于 $t > 0$ 和 $X_\star \in \mathrm{Zer}\mathbb{A}$，有 $\beta(t) > 0$，并且

$$\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 \leq 4\beta(t)^2 \left\|X_0 - X_\star\right\|^2$$
$$\beta(t)^2 \leq \frac{1}{t^2}.$$

如果 $\mathbb{A}$ 还具有 μ-强单调性，那么对于 $t > 0$,

$$\beta(t)^2 \leq \left(\frac{\mu/2}{e^{\mu t/2} - 1}\right)^2.$$

证明见附录 I.1。注意，锚点系数 (11) 的选取使得

$$\Phi(t) = \left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\beta(t)\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0 \right\rangle = 0.$$

因此，(8) 的左侧为零，立即得到 $\mathcal{O}\left(\beta(t)^2\right)$ 收敛速率。类似的离散时间结果见下述定理。

定理 7.2. 设 $\mathbb{A}$ 为极大单调算子。设 $x^0 = y^0 \in \mathbb{R}^n$。考虑

$$x^k = \mathbb{J}_{\mathbb{A}} y^{k-1}$$
$$y^k = (1 - \beta_k)(2x^k - y^{k-1}) + \beta_k x^0$$

with

$$\beta_k = \begin{cases} \dfrac{\|\tilde{\mathbf{A}}x^k\|^2}{-\langle \tilde{\mathbf{A}}x^k,\, x^k - x^0\rangle + \|\tilde{\mathbf{A}}x^k\|^2} & \text{if } \|\tilde{\mathbf{A}}x^k\|^2 \neq 0 \\ 0 & \text{if } \|\tilde{\mathbf{A}}x^k\|^2 = 0, \end{cases}$$

对于 $k = 1$、2、.....，其中 $\tilde{\mathbf{A}}x^k = y^{k-1} - x^k$。

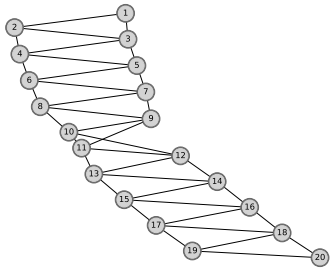
然后 $\beta_k \geq 0$ 和

$$\left\|\tilde{\mathbf{A}}x^{k+1}\right\|^2 \leq \beta_k^2 \left\|x^0 - x^\star\right\|^2$$
$$\beta_k^2 \leq \frac{1}{(k+1)^2}$$

对于 $k = 1$、2、.....以及 $x^\star \in \mathrm{Zer}\mathbf{A}$。

如果 $\mathbb{A}$ 进一步是 μ-强单调且 L-利普希茨连续的，那么对于 $k = 1$，2，.....，

$$\beta_k^2 \leq \left(\frac{\mu/(1+L^2)}{\left(1 + \mu/(1+L^2)\right)^k - 1 + \mu/(1+L^2)}\right)^2.$$



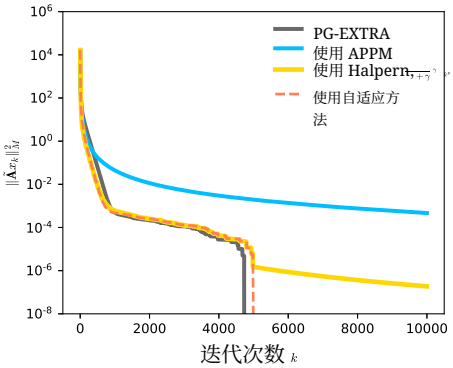


图 2：（左）网络图。（右）平方 M-范数 $\|\tilde{\mathbf{A}}x^k\|_M^2$ 与 k 的对比。Halpern 对应于定理 5.1 中的方法，我们使用 $p = 1.5$ 和 $\gamma = 2.0$。

对于单调设定，速率 $\left\|\tilde{\mathbf{A}}x^{k+1}\right\|^2 \leq \frac{1}{k+1()^2}\left\|x^0 - x^\star\right\|^2$ 与 APPM 的精确最优收敛速率 [54] 相匹配。当取极限 $\mu \to 0$ 时，μ-强单调情形的结果会退化为单调情形的结果。我们在附录 I.3 中给出定理 7.2 的证明和详细内容。

定理 7.2 中的方法是式 (11) 常微分方程的离散时间对应形式。定义 β_k 时分母中的额外项 $\|\tilde{\mathbf{A}}x^k\|^2$在连续时间极限下消失。更多细节见附录 I.2。与连续时间情形类似，离散时间自适应方法的一个关键性质是，对应的 $\Phi(t)$ 始终保持非正。在引理 I.3 的证明中，事实 $\beta_k < 1$ 对于证明这一性质起着关键作用。分母中的额外项 $\|\tilde{\mathbf{A}}x^k\|^2$ 以及事实 $\langle \tilde{\mathbf{A}}x^k, x^k - x^0\rangle < 0$ 当 $\|\tilde{\mathbf{A}}x^k\|^2 \neq 0$ 时，导致 $\beta_k < 1$。

7.1 实验细节

现在，我们展示将定理 7.2 中的方法应用于去中心化压缩感知问题的一个实验，参见 Shi 等人 [63]。我们假设有测量值 $b_i = A_{(i)}x + e_i$，其中 $A_{(i)}$ 是每个本地智能体 i可用的测量矩阵，x 是我们希望恢复的未知共享信号，而 e_i 是测量误差。我们以去中心化方式求解此问题：本地智能体将测量值保密，仅与邻居通信。

与 Shi 等人一样， [63], 我们将问题表述为一个无约束的 ℓ_1-正则化最小二乘问题

$$\underset{x\in\mathbb{R}^d}{\text{minimize}} \quad \frac{1}{n}\sum_{i=1}^n \left\{\frac{1}{2}\|A_{(i)}x - b_i\|^2 + \rho\|x\|_1\right\},$$

并应用 PG-EXTRA。我们比较原始 PG-EXTRA 与各种锚定版 PG-EXTRA，其中 β_k 分别对应定理 5.1 和定理 7.2。实验结果见图 2。实验的更多细节见附录 J。

8 结论

本文提出了锚点加速的连续时间模型，即锚定常微分方程 $\dot{X} = -\mathbf{A}(X) - \beta(t)(X - X_0)$。我们刻画了收敛速率与 $\beta(t)$ 的函数关系，从而深入理解锚点加速机制。最后，受连续时间分析的启发，我们提出了一种自适应方法，并通过理论分析和实验验证了其有效性。

此前对 Nesterov 加速连续时间模型的分析启发了多项后续研究，例如基于拉格朗日力学和哈密顿力学的分析 [26, 71, 72], 高分辨率常微分方程模型 [62], 以及连续化框架 [29]。对锚定常微分方程进行类似分析，是未来值得探索的方向。

致谢与资金披露

本研究得到了三星科学技术基金会的资助（项目编号：SSTF-BA2101-02）。感谢 Jaeyeon Kim 提供宝贵的反馈。感谢 Hangjun Cho 就常微分方程的适定性进行有益讨论。我们也感谢匿名审稿人提出建设性意见。

参考文献

[1] J. K. Alcala、Y. T. Chow 和 M. Sunkula。光滑结构化极小极大问题的移动锚点外梯度法。*arXiv:2308.12359*，2023年。[2] V. Apidopoulos、J.-F. Aujol 和 C. Dossal。FISTA 算法的微分包含建模及其在该情形下收敛速率的最优性 $b \leq 3$。*SIAM* 优化杂志，28(1)：551–574，2018年。[3] H. Attouch 和 A. Cabot。具有时变黏性的惯性梯度动力学的渐近稳定化。微分方程杂志，263(9)：5412–5458，2017年。[4] H. Attouch 和 S. C. László。由极大单调算子控制的类牛顿惯性动力学与近端算法。*SIAM* 优化杂志，30(4)：3252–3283，2020年。[5] H. Attouch 和 S. C. László。通过带消失 Tikhonov 正则化的惯性算法进行凸优化：快速收敛至最小范数解。*arXiv:2104.11987*，2021年。[6] H. Attouch 和 J. Peypouquet。由极大单调算子控制的惯性动力学与近端算法的收敛性。数学规划，174(1)：391–432，2019年。[7] H. Attouch、Z. Chbani、J. Peypouquet 和 P. Redont。具有渐近消失黏性的惯性动力学与算法的快速收敛。数学规划，168(1)：123–175，2018年。[8] H. Attouch、Z. Chbani 和 H. Riahi。将用于凸优化的快速惯性动力学与 Tikhonov 正则化相结合。数学分析与应用杂志，457(2)：1065–1094，2018年。[9] H. Attouch、Z. Chbani 和 H. Riahi。亚临界情形下 Nesterov 加速梯度法的收敛速率 $\alpha \leq 3$。*ESAIM*：控制、优化与变分法，25：2，2019年。[10] H. Attouch、Z. Chbani、J. Fadili 和 H. Riahi。带惯性和海森驱动阻尼的一阶优化算法的迭代点收敛性。《优化》，第1–40页，2021年。[11] J.-P. Aubin 和 A. Cellina。微分包含。Springer，1984年。[12] J.-F. Aujol、C. Dossal 和 A. Rondepierre。Nesterov 加速的最优收敛速率。*SIAM* 优化杂志，29(4)：3131–3153，2019年。[13] J.-B. Baillon 和 R. E. Bruck。平均非扩张映射的渐近正则性最优速率。不动点理论与应用，128：27–66，1992年。[14] S. Banach。抽象集合上的运算及其在积分方程中的应用。*Fundamenta Mathematicae*，3(1)：133–181，1922年。[15] H. H. Bauschke 和 P. L. Combettes。*Hilbert* 空间中的凸分析与单调算子理论。Springer International Publishing，第二版，2017年。[16] R. I. Bo¸t 和 E. R. Csetnek。一个与变分不等式相关、具有海森驱动阻尼和惩罚项的二阶动力系统。《优化》，68(7)：1265–1277，2019年。

[17] R. I. Boţ 和 D. A. Hulett。针对加性结构单调包含的消失阻尼二阶分裂动力学。《动力学与微分方程杂志》，2022年。[18] R. I. Boţ、E. R. Csetnek 和 S. C. László。采用可变阻尼的非凸光滑最小化二阶动力学方法。《应用分析》，99(3)：361–378，2020年。[19] R. I. Boţ、E. R. Csetnek 和 S. C. László。具有海森驱动阻尼的二阶动力系统的 Tikhonov 正则化。数学规划，189(1)：151–186，2021年。[20] R. I. Bot、E. R. Csetnek 和 D.-K. Nguyen。连续时间与离散时间下的快速乐观梯度下降—上升法。arXiv:2203.10947，2022年。[21] M. Bravo 和 R. Cominetti。平均非扩张映射的精确收敛速率。《以色列数学杂志》，227：163–188，2018年。[22] A.-L. Cauchy。求解联立方程组的一般方法。《法国科学院每周会议报告》，25：536–538，1847年。[23] R. Cominetti、J. A. Soto 和 J. Vaisman。关于 Krasnosel'skiĭ-Mann 迭代的收敛速率及其与伯努利变量之和的联系。《以色列数学杂志》，199(2)：757–772，2014年。[24] C. Daskalakis、A. Ilyas、V. Syrgkanis 和 H. Zeng。以乐观性训练生成对抗网络。国际学习表征会议，2018年。[25] J. Diakonikolas。用于近乎最优且无参数单调包含以及变分不等式强解的 Halpern迭代。学习理论会议，2020年。[26] J. Diakonikolas 和 M. I. Jordan。广义动量方法：哈密顿视角。SIAM 优化杂志，31(1)：915–944，2021年。[27] Y. Drori。光滑凸最小化的精确信息复杂度。《复杂性杂志》，39：1–16，2017年。[28] J. Eckstein 和 D. P. Bertsekas。关于极大单调算子的 Douglas—Rachford 分裂方法与近端点算法。数学规划，55(1)：293–318，1992年。[29] M. Even、R. Berthier、F. Bach、N. Flammarion、H. Hendrikx、P. Gaillard、L. Massoulié 和 A. Taylor。确定性与随机梯度下降以及八卦算法的连续化加速方法。神经信息处理系统，2021年。[30] George J. Minty。希尔伯特空间中的单调（非线性）算子。《杜克数学杂志》，29(3)：341–346，1962年。[31] I. Goodfellow。NIPS 2016 教程：生成对抗网络。arXiv:1701.00160，2016年。[32] E. Gorbunov、N. Loizou 和 G. Gidel。外梯度法： $O(1/K)$单调变分不等式的末次迭代收敛性及其与余强制性的联系。国际人工智能与统计会议，2022年。[33] G. Gu 和 J. Yang。极大单调包含问题的近端点算法的紧致次线性收敛速率。SIAM 优化杂志，30(3)：1905–1921，2020年。[34] B. Halpern。非扩张映射的不动点。《美国数学会通报》，73(6)：957–961，1967年。[35] D. Kim。极大单调算子的加速近端点方法。数学规划，190(1–2)：57–87，2021年。[36] D. Kim 和 J. A. Fessler。用于光滑凸最小化的优化一阶方法。数学规划，159(1-2)：81–107，2016年。

[37] D. Kim 和 J. A. Fessler。优化一阶方法在降低光滑凸函数梯度方面的效率。《优化理论与应用杂志》，188(1): 192–219，2021年。[38] J. Kim 和 I. Yang。统一适用于凸目标函数和强凸目标函数的 Nesterov 加速梯度方法。《国际机器学习会议》，202，2023年。[39] U. Kohlenbach。关于 F. E. Browder 和 R. Wittmann 定理的定量版本。《数学进展》，226(3):2764–2795，2011年。[40] M. A. Krasnosel'skii。关于逐次逼近方法的两点说明。《数学科学进展》，10(1):123–127，1955年。[41] J. Lee 和 E. K. Ryu。通过锚定加速值迭代。《神经信息处理系统》，2023年。[42] S. Lee 和 D. Kim。用于光滑结构化非凸非凹极小极大问题的快速外梯度方法。《神经信息处理系统》，2021年。[43] L. Leustean。非扩张映射的 Halpern迭代的渐近正则性速率。《通用计算机科学杂志》，13(11):1680–1691，2007年。[44] J. Liang、J. Fadili 和 G. Peyré。非精确非扩张算子的收敛速率。《数学规划》，159(1):403–434，2016年。[45] F. Lieder。关于 Halpern迭代的收敛速率。《优化快报》，15(2): 405–418，2021年。[46] T. Lin 和 M. I. Jordan。单调包含、加速与闭环控制。《运筹学数学研究》，2023年。[47] W. R. Mann。迭代中的平均值方法。《美国数学会会刊》，4(3):506–510，1953年。[48] B. Martinet。通过逐次逼近正则化变分不等式。《法国计算机科学与运筹学杂志》，红色系列，4(3):154–158，1970年。[49] B. Martinet。《优化与极小极大问题求解算法》。博士论文，格勒诺布尔科学与医学大学，1972年。[50] S.-Y. Matsushita。关于 Krasnosel'skiĭ–Mann 迭代的收敛速率。《澳大利亚数学会通报》，96(1):162–170，2017年。[51] Y. Nesterov。一种求解凸规划问题并具有收敛速率$O(1/k^2)$的方法。《苏联科学院通报》，269(3):543–547，1983年。[52] Y. Nesterov。《凸优化导论：基础课程》。Springer，2004年。[53] C. Park、J. Park 和 E. K. Ryu。加速梯度方法的因子-$\sqrt{2}$加速。《应用数学与优化》，2022年。[54] J. Park 和 E. K. Ryu。不动点迭代的精确最优加速复杂度。《国际机器学习会议》，2022年。[55] J. Park 和 E. K. Ryu。约束优化与不动点迭代的加速不可行性检测。《国际机器学习会议》，2023年。[56] L. D. Popov。用于搜索鞍点的 Arrow–Hurwicz方法的改进。《苏联科学院数学札记》，28(5):845–848，1980年。[57] A. Rakhlin 和 K. Sridharan。基于可预测序列的在线学习。《学习理论会议》，2013年。[58] E. K. Ryu 和 W. Yin。《通过单调算子进行大规模凸优化》。剑桥大学出版社，2022年。

–

[59] E. K. Ryu、K. Yuan 和 W. Yin。用于极小极大问题和生成对抗网络的随机梯度方法的乐观性与锚定常微分方程分析。*arXiv:1905.10899*，2019年。[60] S. Sabach 和 S. Shtern。一种求解凸双层优化问题的一阶方法。*SIAM* 优化杂志，27(2):640–660，2017年。[61] A. Salim、L. Condat、D. Kovalev 和 P. Richtárik。一种在仿射约束下进行强凸最小化的最优算法。*AISTATS*，2022年。[62] B. Shi、S. S. Du、M. I. Jordan 和 W. J. Su。通过高分辨率微分方程理解加速现象。数学规划，2021年。[63] W. Shi、Q. Ling、G. Wu 和 W. Yin。一种去中心化复合优化的近端梯度算法。*IEEE* 信号处理汇刊，63(22):6013–6023，2015年。[64] M. V. Solodov 和 B. F. Svaiter。一种利用极大单调算子扩张的混合近似外梯度–近端点算法。集值分析，7(4):323–345，1999年。[65] W. Su、S. Boyd 和 E. J. Candès。一种用于建模 Nesterov 加速梯度法的微分方程：理论与见解。神经信息处理系统，2014年。[66] W. Su、S. Boyd 和 E. J. Candès。一种用于建模 Nesterov 加速梯度法的微分方程：理论与见解。机器学习研究杂志，17(153):1–43，2016年。[67] J. J. Suh、G. Roh 和 E. K. Ryu。通过扩张坐标系中的守恒律进行加速梯度法的连续时间分析。国际机器学习会议，2022年。[68] A. Taylor 和 Y. Drori。一种用于光滑强凸最小化的最优梯度法。数学规划，2022年。[69] Q. Tran-Dinh 和 Y. Luo。用于单调包含的 Halpern 型加速与分裂算法。*arXiv:2110.08150*，2021年。[70] B. Van Scoy、R. A. Freeman 和 K. M. Lynch。已知最快的用于最小化强凸函数的全局收敛一阶方法。*IEEE* 控制系统快报，2(1):49–54，2018年。[71] A. Wibisono、A. C. Wilson 和 M. I. Jordan。优化中加速方法的变分视角。美国国家科学院院刊，113(47):E7351–E7358，2016年。[72] A. C. Wilson、B. Recht 和 M. I. Jordan。优化中加速方法的 Lyapunov 分析。机器学习研究杂志，22(113):1–34，2021年。[73] R. Wittmann。非扩张映射不动点的逼近。数学档案，58(5):486–491，1992年。[74] T. Wu、K. Yuan、Q. Ling、W. Yin 和 A. H. Sayed。具有异步性和延迟的去中心化共识优化。*IEEE* 网络信号与信息处理汇刊，4(2):293–307，2018年。[75] T. Yoon 和 E. K. Ryu。具有 $\mathcal{O}(1/k^2)$ 梯度范数平方收敛速率的光滑凸-凹极小极大问题加速算法。国际机器学习会议，2021年。[76] T. Yoon 和 E. K. Ryu。加速极小极大算法齐步前进。*arXiv:2205.11093*，2022年。

前期工作

离散情形下光滑凸函数的加速。关于光滑凸函数加速的研究成果颇丰。Nesterov [51] 提出了加速梯度法（AGM），其降低函数值的 $\mathcal{O}(1/k^2)$ 收敛速度快于梯度下降法 [22] 的$\mathcal{O}(1/k)$速度。优化梯度法（OGM） [36] 将 AGM 的收敛速度提高了一个常数因子，并已证明其具有最优性 [27]。对于光滑强凸情形，强凸加速梯度法 [52]实现了加速收敛，后续研究还探索了进一步的改进 [53, 61, 68, 70]。最近，研究者提出了 OGM-G [37] ，这是一种用于光滑凸最小化、以降低梯度模平方为目标的加速方法。

连续时间情形下光滑凸函数的加速。对 Nesterov 加速的连续时间分析也已得到深入研究。Su 等人 [65, 66] 提出了加速梯度法常微分方程模型 $\ddot{X}(t) + \frac{r}{t}\dot{X} + \nabla f(X) = 0$，为 $r \geq 3$提供了$f(X(t)) - f_\star \in \mathcal{O}\left(1/t^2\right)$ 收敛速度。Attouch 等人 [7] 改进了 $r > 3$ 的界常数，并证明了轨迹的收敛性。Attouch 等人 [9] 实现了 $0 < r < 3$的 $\mathcal{O}\left(1/t^{-2r/3}\right)$ 收敛速度。Apidopoulos 等人 [2] 将其结果推广到具有不可微凸函数的微分包含。此外，人们还研究了加速梯度法常微分方程的多种变体 [3, 5, 8, 10, 12, 16, 18, 19]。此外，Attouch 和 László[4], 、Attouch 和 Peypouquet [6],、Bo̧t 和 Hulett [17]还研究了其在单调包含问题中的应用。

受上述加速方法的连续时间分析启发，用于分析常微分方程的工具得到了进一步发展。Wibisono 等人 [71], Wilson 等人 [72] 和 Kim 和 Yang [38] 采用拉格朗日力学，先后提出了第一、第二及统一的 Bregman 拉格朗日量，以统一分析一类广义常微分方程，其中后者还分析了强凸加速梯度法。研究者提出了利用哈密顿力学 [26] 和扩张坐标系 [67] 系统地构造李雅普诺夫函数的方法，并通过膨胀坐标框架分析了 OGM-G。此后还发展出多种不同形式的连续时间模型，例如高分辨率常微分方程 [62] 和连续化框架 [29] 。

另一方面，另一种称为锚点加速 的加速方式最近受到关注。正如 Yoon 和 Ryu [76]所指出的，近期发现的许多适用于极小极大优化和不动点问题的加速方法都基于锚点加速。最近，研究者还提出了锚点加速在检测约束优化问题 [55], 的不可行性，以及加速动态规划和强化学习 [41] 中的值迭代等实际应用。

不动点问题。 关于不动点问题的研究历史可追溯至巴拿赫的工作 [14], ，该工作证明了采用压缩算子的皮卡迭代具有收敛性。随后，人们提出了克拉斯诺谢尔斯基-曼迭代（KM） [40, 47] ，这是皮卡迭代的一种推广。Martinet [49]证明了采用一般非扩张算子的KM迭代具有收敛性。Wittmann [73]则证明了Halpern迭代 [34],在参数选择范围较广时仍具有收敛性。

不动点残差的平方范数 $\|y_k - \mathbb{T}y_k\|^2$ 是不动点问题中常见的误差度量。研究表明，KM迭代具有 $\mathcal{O}(1/k)$速率 [21, 23, 44] 和 $o(1/k)$-速率 [13, 50]。对于Halpern迭代，Leustean [43], 首先确立了 $\mathcal{O}(1/(\log k)^2)$-速率，随后Kohlenbach [39]将其改进为 $\mathcal{O}(1/k)$速率。Sabach和Shtern [60] 首次实现了加速的 $\mathcal{O}(1/k^2)$速率，而Lieder [45]将该常数降低了16倍。

众所周知，求解不动点问题与求解单调包含问题[28, 30, 54]是等价的。近端点方法（PPM） [48] 在 $\left\|\tilde{\mathbb{A}}x_k\right\|^2$ [33]意义下达到 $\mathcal{O}(1/k)$速率。加速近端点方法（APPM） [35] 将速率提升至加速的 $\mathcal{O}\left(1/k^2\right)$速率。Park 和 Ryu [54] 证明，APPM 是该问题的精确最优方法，并提出了一种针对 μ强单调算子的精确最优方法 OS-PPM，其达到 $\mathcal{O}\left(1/e^{4\mu k}\right)$ 速率。最优方法 APPM 和 OS-PPM 均基于锚点加速 [76]。

极小极大问题。形式为 $\min_x \max_y \mathbf{L}(x,\ y)$ 的极小极大优化问题近年来受到机器学习领域的关注。常见的理论设定之一是光滑凸-凹设定，并以梯度范数平方作为误差度量。对于 $\|\partial\mathbf{L}(x,\ y)\|^2$，经典外梯度法（EG） [64] 和 OG [24, 56, 57] 被证明可达到$\mathcal{O}(1/k)$速率 [32]。随机梯度下降上升法（SGDA） [59] 通过引入“锚点”这一术语，达到了 $p > 1/2$ 的 $\mathcal{O}\left(1/k^{2-2p}\right)$ 速率。通过引入一种无参数的 Halpern 型方法，Diakonikolas [25] 达到了 $\mathcal{O}(\log k/k^2)$。最近，EAG [75] 首次借助锚点加速达到加速速率 $\mathcal{O}(1/k^2)$，随后出现了 FEG [42], 锚定的 Popov 方案 [69] 和移动锚点方法 [1]。快速 ODGA [20] 也达到了加速的 $o(1/k^2)$速率。对于 $\partial\mathbf{L}$ 进一步满足条件数为 κ的强单调性时，SM-EAG+ [76] 达到了加速的 $\mathcal{O}\left(1/e^{2k\kappa}\right)$ 速率。

然而，据我们所知，目前关于锚点加速的连续时间分析还不充分。Bot 等人 [20], Lin 和 Jordan [46], 研究了单调包含问题的加速连续时间分析，但没有考虑锚点加速。Ryu 等人 [59] 研究了锚点加速的连续时间分析，但仅限于少数情形 $\dot{X}(t) = -\mathbb{A}(X(t)) - \frac{\gamma}{t}(X - X_0)$，适用于 $\gamma \geq 1$。本文针对广义锚点系数下的锚点加速，给出统一的连续时间分析。

B 定理 2.2 的证明

B.1 唯一性的证明

唯一性的证明可直接由引理 2.3 得出，我们先证明该引理。

引理 *2.3* 的证明。 对于 $t=0$，结论显然成立，因此我们可以假设 $t>0$。由 $\mathbb{A}$ 的单调性和杨氏不等式，可得下列不等式。

$$\begin{aligned}\frac{d}{dt}\|X(t)-Y(t)\|^2 &= 2\left\langle \dot{X}(t)-\dot{Y}(t), X(t)-Y(t)\right\rangle \\ &= -2\left\langle \tilde{\mathbb{A}}(X(t))-\tilde{\mathbb{A}}(Y(t))+\beta(t)(X(t)-Y(t))-\beta(t)(X_0-Y_0), X(t)-Y(t)\right\rangle \\ &\le -2\beta(t)\|X(t)-Y(t)\|^2+2\beta(t)\langle X_0-Y_0, X(t)-Y(t)\rangle \\ &\le -2\beta(t)\|X(t)-Y(t)\|^2+\beta(t)\left(\|X_0-Y_0\|^2+\|X(t)-Y(t)\|^2\right) \\ &= -\beta(t)\|X(t)-Y(t)\|^2+\beta(t)\|X_0-Y_0\|^2.\end{aligned}$$

现在定义 $C(t)=e^{\int_v^t \beta(s)ds}$，其中 $v>0$ 为某个值，则可得 $\dot{C}(t)=C(t)\beta(t)$。将 $-\beta(t)\|X(t)-Y(t)\|^2$ 移至左边，并将两边乘以 $C(t)$，可得

$$\begin{aligned}\frac{d}{dt}\left(C(t)\|X(t)-Y(t)\|^2\right) &= C(t)\frac{d}{dt}\|X(t)-Y(t)\|^2+C(t)\beta(t)\|X(t)-Y(t)\|^2 \\ &\le C(t)\beta(t)\|X_0-Y_0\|^2=\frac{d}{dt}C(t)\|X_0-Y_0\|^2\end{aligned}$$

将两边从 $\epsilon>0$ 到 t 积分，可得

$$C(t)\|X(t)-Y(t)\|^2-C(\epsilon)\|X(\epsilon)-Y(\epsilon)\|^2\le C(t)\|X_0-Y_0\|^2-C(\epsilon)\|X_0-Y_0\|^2.$$

由于 C 非负且单调不减，$\lim_{\epsilon\to 0+}C(\epsilon)$ 存在。对两边取极限 $\epsilon\to 0+$，可得

$$C(t)\|X(t)-Y(t)\|^2\le C(t)\|X_0-Y_0\|^2.$$

最后，两边除以 $C(t)$，得到

$$\|X(t)-Y(t)\|^2\le\|X_0-Y_0\|^2.$$

□

广义情形 (3) 的唯一性证明同样可以完成。

定理 B.1 (解的唯一性)。 若(3)的解存在，则该解唯一。

证明。 设 X_1, X_2 是具有相同初值 X_0的 (3) 的解。由引理 2.3，得

$$\|X_1(t)-X_2(t)\|\le\|X_0-X_0\|=0$$

因此，对于所有 $t\in[0,\infty)$，$X_1(t)=X_2(t)$，由此得到所需结果。 □

由于 (6) 是 (3) 的特例，定理 2.2 的唯一性证明由定理 B.1 得出。

B.2 存在性证明

由于 $\frac{\gamma}{t^p}$ at $t=0$ 处存在奇异性，存在性证明需要进行繁琐的工作，因此我们通过若干小节给出证明。在开始之前，我们先简要概述证明思路。该证明基本上基于 [65] and [11] 中给出的证明。

我们首先证明 $\mathbb{A}$ is Lipschitz 的情形。当 $\mathbb{A}$ is Lipschitz 时，微分包含变为常微分方程，因此我们可以采用 [65] X $X\ t\ =-\mathbb{A}\ X\ t\ -\frac{\gamma}{t^p}\ X_0-$ 中类似的论证。我们考虑一系列相对 $X\ t$ 于 为 Lipschitz 的常微分方程，用以逼近 ( ) (( )) (( ))。根据常微分方程经典理论，这些近似常微分方程存在解；通过选取解的适当子序列，我们便可得到真正的解。

由于 Lipschitz $\mathbb{A}$ 的解已经得到，我们可以采用 [11]中类似的论证。首先考虑解 X_λ 及其 Yosida逼近 $\mathbb{A}_\lambda=\frac{1}{\lambda}(\mathbb{I}-(\mathbb{I}+\lambda\mathbb{A})^{-1})$，该逼近是 $\mathbb{A}$ 的一个近似，并且具有利普希茨连续性。随后，我们可以得到一个子序列 X_{λ_n}，其收敛于原微分包含。

B.2.1 Lipschitz 存在性证明 A

由于我们将使用利普希茨连续函数逼近 A，因此首先考虑具有利普希茨连续 A的常微分方程。**定理 B.2**。 （*Lipschitz*解的存在性 A）

假设$A:\ \tilde{\mathbb{R}}^n \to \mathbb{R}^n$ 是 *L-Lipschitz* 连续函数。考虑具有初值条件~ ~$X(0) = X_0 \in dom(A)$的微分方程，

$$\dot{\tilde{X}}(t) = -\tilde{A}(\tilde{X}(t)) - \frac{\gamma}{t^p}(\tilde{X}(t) - X_0). \tag{12}$$

其中 γ、 $p > 0$。那么存在唯一的解 $\tilde{X} \in \mathcal{C}^1([0,\infty), \mathbb{R}^n)$，对于所有 $t \in (0,\infty)$，都满足 (12)。此外，对于由 $\tilde{X}(0) = \lim_{t\to 0+} {}^{X_{(t)}}{}_t^{-X_0}$ 定义的. ——— . ~$\tilde{X}(0)$，以下结论成立

$$\dot{\tilde{X}}(0) = \begin{cases} -\tilde{A}(X_0) & \text{if } 0 < p < 1 \\ -\frac{1}{\gamma+1}\tilde{A}(X_0) & \text{if } p = 1 \\ 0 & \text{if } p > 1. \end{cases}$$

~ ~证明需要做一些准备。我们将考虑近似解 X_δ，得到一个收敛到解 X 的序列。~ ~因此，我们先定义 X_δ。从现在起，我们将把 A 称为 L-Lipschitz 单调函数。

定义 B.3。令 $0 < \delta < 1$。考虑

$$\dot{\tilde{X}}_\delta(t) = \begin{cases} -\tilde{A}(\tilde{X}_\delta(t)) - \frac{\gamma}{\delta^p}(\tilde{X}_\delta(t) - X_0) & 0 \le t < \delta \\ -\tilde{A}(\tilde{X}_\delta(t)) - \frac{\gamma}{t^p}(\tilde{X}_\delta(t) - X_0) & t > \delta \end{cases} \tag{13}$$

由于上述右侧关于 $\tilde{X}_\delta$ 是 $\left(L + \frac{\gamma}{\delta^p}\right)$-Lipschitz 的，根据经典常微分方程理论，该解唯一存在。~将该解定义为 X_δ。那么，对于满足 $\delta_m < 1$ 且 $\lim_{m\to\infty}\delta_m = 0$ 的正数列 $\{\delta_m\}_{m\in\mathbb{N}}$，考虑序列$\left\{\tilde{X}_{\delta_m}\right\}_{m\in\mathbb{N}}$。

开始之前，我们先证明一个有用的引理，后续在算子为利普希茨连续的情形下会频繁用到它。

~**引理 B.4**。设 $A\colon \mathbb{R}^n \to \mathbb{R}^n$ 是一个利普希茨连续函数。设 $\beta\colon D \to [0,\infty)$ 是一个满足$D \subset [0,\infty)$的连续函数。考虑微分方程

$$\dot{\tilde{X}} = -\tilde{A}(\tilde{X}) - \beta(t)(\tilde{X} - X_0),$$

~其中 $X_0 \in \mathbb{R}^n$。设 $X\colon D \to \mathbb{R}^n$ 是满足上述方程的可微曲线，其中 $t \in D$。那么，对于所有满足 $[a,b] \subset D$ 的 $0 \le a < b$， X 以及复合函数 $A \circ X\colon [a,b] \to \mathbb{R}^n$ 都是利普希茨连续的。

. ~此外，若 β（t）几乎对所有 $t \in [a,b]$都有定义且有界，则 X 关于 $[a,b]$是利普希茨连续的。因此，若 β 是~二阶可微的，则 X 是利普希茨连续的。

. ~ ~ ~ ~由于 X、 $A \circ X$ 和 X 是利普希茨连续函数，因此它们都是绝对连续函数。

~ ~ ~ ~ ~证明。我们先证明 X 是利普希茨连续的。由于 A、 X、 β 在 $[a,b]$ 上连续，因此可知 $A(X(t)) + \beta(t)(X(t) - X_0)$ 在 $[a,b]$上连续，于是

$$M_1 = \max_{t\in[a,b]} \left\|\dot{X}(t)\right\| = \max_{t\in[a,b]} \left\|\tilde{A}(\tilde{X}(t)) + \beta(t)(\tilde{X}(t) - X_0)\right\|$$

~存在。由于其导数被 $M_1 < \infty$所界定，因此 X 是 M_1-Lipschitz 连续的。作为两个 Lipschitz~ ~连续函数的复合函数， $A \circ X$ 是 Lipschitz 连续的。

. ~首先注意，如果 β二阶可微，那么由于 β连续，它在 $[a,b]$ 上有界。现在，如果 $\beta(t)$在 $t \in [a,b]$上有界，即对于某个 $M_2 > 0$，有 $\left|\dot{\beta}(t)\right| \le M_2$，那么对于 $M_3 = \max_{t\in[a,b]} |\beta(t)|$，我们有

$$\left\|\frac{d}{dt}\left(\beta(t)(\tilde{X}(t) - X_0)\right)\right\| = \left\|\dot{\beta}(t)(\tilde{X}(t) - X_0)\right\| + \left\|\beta(t)\dot{\tilde{X}}(t)\right\| \le M_2 M_1 |b-a| + M_3 M_1$$

~因此， $\beta(t)(X(t) - X_0)$是（$M_2 M_1 |b-a| + M_3 M_1$）-Lipschitz 连续的。因此，作为 Lipschitz 连续~函数的和，X 是 Lipschitz 连续的。 □

我们将证明，对于每个 $T > 0$，导数集合 $\left\{\dot{\tilde{X}}_\delta \mid \delta \in (0,\ 1)\right\}$ 在 $[0,T]$ 上一致有界，并且 $\left\{\tilde{X}_{\delta_m}\right\}_{m\in}$ ℕ在 $[0,T]$上一致收敛。我们首先证明导数的有界性。为此，我们先证明一个有用的引理。

引理 B.5。*Fo* $_{[a,b]\to\mathbb{R}}$ ˜r $0<a<b$，假设 $X:\ [a,b]\to\mathbb{R}^n$ 满足常微分方程 *(12)*，其中 $t\in[a,b]$。定义 $U:$ *as*

$$\tilde{U}_1(t)=\left\|\dot{\tilde{X}}(t)\right\|^2+\frac{\gamma p}{t^{p+1}}\left\|\tilde{X}(t)-X_0\right\|^2$$
$$\tilde{U}_2(t)=\frac{1}{t^{p-1}}\left\|\dot{\tilde{X}}(t)\right\|^2+\frac{\gamma p}{t^{2p}}\left\|\tilde{X}(t)-X_0\right\|^2$$

˜那么，如果 $0<p\le 1$，U_1 是非递增的；并且对于 $p>1$，U_2 是非递增的。

˜ ˜证明。 根据引理 B.4，我们可以验证 U_1 和 U_2 在 $[a,b]$上绝对连续。因此，只需验证导数几乎处处非正。根据引理 B.4，我们可以对 (12) 两边求导。因此，几乎处处的 t 都有

$$\ddot{\tilde{X}}(t)=-\frac{d}{dt}\tilde{A}(\tilde{X}(t))+\frac{\gamma p}{t^{p+1}}(\tilde{X}(t)-X_0)-\frac{\gamma}{t^p}\dot{\tilde{X}}(t).$$

回顾 $\tilde{A}$ 的单调性及 (1)，我们知道几乎处处 t 有 $\left\langle \dot{\tilde{X}}(t),\ \frac{d}{dt}\tilde{A}(\tilde{X}(t))\right\rangle \ge 0$ 。因此，几乎处处的 t，

(i) $0<p\le 1$

$$\begin{aligned}
\dot{\tilde{U}}_1(t)&=2\left\langle \dot{\tilde{X}}(t),\ddot{\tilde{X}}(t)\right\rangle+\frac{2\gamma p}{t^{p+1}}\left\langle \dot{\tilde{X}}(t),\tilde{X}(t)-X_0\right\rangle-\frac{\gamma p(p+1)}{t^{p+2}}\left\|\tilde{X}(t)-X_0\right\|^2\\
&=-2\left\langle \dot{\tilde{X}}(t),\frac{d}{dt}\tilde{A}(\tilde{X}(t))\right\rangle-\frac{2\gamma}{t^p}\left\|\dot{\tilde{X}}(t)\right\|^2+\frac{2\gamma p}{t^{p+1}}\left\langle \dot{\tilde{X}}(t),\tilde{X}(t)-X_0\right\rangle\\
&\quad+\frac{2\gamma p}{t^{p+1}}\left\langle \dot{\tilde{X}}(t),\tilde{X}(t)-X_0\right\rangle-\frac{2\gamma p^2}{t^{p+2}}\left\|\tilde{X}(t)-X_0\right\|^2-\frac{\gamma p(1-p)}{t^{p+2}}\left\|\tilde{X}(t)-X_0\right\|^2\\
&=-2\left\langle \dot{\tilde{X}}(t),\frac{d}{dt}\tilde{A}(\tilde{X}(t))\right\rangle-\frac{2\gamma}{t^p}\left\|\dot{\tilde{X}}(t)-\frac{p}{t}(\tilde{X}-X_0)\right\|^2-\frac{\gamma p(1-p)}{t^{p+2}}\left\|\tilde{X}(t)-X_0\right\|^2\\
&\le 0.
\end{aligned}$$

(ii) $p>1$

$$\begin{aligned}
\dot{\tilde{U}}_2(t)&=-\frac{p-1}{t^p}\left\|\dot{\tilde{X}}(t)\right\|^2+\frac{2}{t^{p-1}}2\left\langle \dot{\tilde{X}}(t),\ddot{\tilde{X}}(t)\right\rangle+\frac{2\gamma p}{t^{2p}}\left\langle \dot{\tilde{X}}(t),\tilde{X}-X_0\right\rangle-\frac{2\gamma p}{t^{2p+1}}\left\|\tilde{X}(t)-X_0\right\|^2\\
&=-\frac{p-1}{t^p}\left\|\dot{\tilde{X}}(t)\right\|^2-\frac{2}{t^{p-1}}\left\langle \dot{\tilde{X}}(t),\frac{d}{dt}\tilde{A}(\tilde{X}(t))\right\rangle\\
&\quad-\frac{2\gamma}{t^{2p-1}}\left\|\dot{\tilde{X}}(t)\right\|^2+\frac{2\gamma p}{t^{2p}}\left\langle \dot{\tilde{X}}(t),\tilde{X}-X_0\right\rangle+\frac{2\gamma p}{t^{2p}}\left\langle \dot{\tilde{X}}(t),\tilde{X}-X_0\right\rangle-\frac{2\gamma p}{t^{2p+1}}\left\|\tilde{X}(t)-X_0\right\|^2\\
&=-\frac{p-1}{t^p}\left\|\dot{\tilde{X}}(t)\right\|^2-\frac{2}{t^{p-1}}\left\langle \dot{\tilde{X}}(t),\frac{d}{dt}\tilde{A}(\tilde{X}(t))\right\rangle-\frac{2\gamma}{t^{2p-1}}\left\|\dot{\tilde{X}}(t)-\frac{p}{t}(\tilde{X}-X_0)\right\|^2\\
&\le 0.
\end{aligned}$$

□

.˜现在我们可以证明导数的一致有界性了 X_δ。

引理 B.6。 取 $T>0$。对于所有 $t\in[0,T]$ 和 $\delta\in(0、1)$，以下不等式成立。

$$\left\|\dot{\tilde{X}}_\delta(t)\right\|\le\tilde{M}_{dot}(T)=\begin{cases}\sqrt{1+\gamma p}\left\|\tilde{A}(X_0)\right\| & 0<p\le 1\\ \sqrt{T^{p-1}\left(\frac{1}{2\gamma}+2p\left(2L^2+2\gamma+1\right)\right)}\left\|\tilde{A}(X_0)\right\| & p\ge 1.\end{cases}$$

证明。 我们通过考虑两种情况来证明第一个结论。

(1) $t \in [0, \delta]$

根据引理 B.4，我们知道 $\dot{\tilde{X}}_\delta$ 是绝对连续的，因此 $\left\|\dot{\tilde{X}}_\delta\right\|^2$ 也是绝对连续的。对 (13) 求导，对于几乎所有 $t \in (0, \delta)$，都有

$$\ddot{\tilde{X}}_\delta = -\frac{d}{dt}\tilde{A}(\tilde{X}_\delta(t)) - \frac{\gamma}{\delta^p}\dot{\tilde{X}}_\delta.$$

现在，对于几乎所有 $t \in (0, \delta)$,

$$\frac{d}{dt}\left\|\dot{\tilde{X}}_\delta\right\|^2 = 2\left\langle \dot{\tilde{X}}_\delta, \ddot{\tilde{X}}_\delta \right\rangle = -2\left\langle \dot{\tilde{X}}_\delta, \frac{d}{dt}\tilde{A}(\tilde{X}_\delta(t)) \right\rangle - \frac{2\gamma}{\delta^p}\left\|\dot{\tilde{X}}_\delta\right\|^2 \le -\frac{2\gamma}{\delta^p}\left\|\dot{\tilde{X}}_\delta\right\|^2 \le 0.$$

(i) $0 < p \le 1$

From above we know $\left\|\dot{\tilde{X}}_\delta\right\|$ is nonincreasing in $t \in [0, \delta]$, we have $\left\|\dot{\tilde{X}}_\delta(t)\right\| \le \left\|\dot{\tilde{X}}_\delta(0)\right\| = \left\|\tilde{A}(X_0)\right\|$.

(ii) $p > 1$

对 $\frac{d}{dt}\left\|\dot{\tilde{X}}_\delta\right\| \le -\frac{2\gamma}{\delta^p}\left\|\dot{\tilde{X}}_\delta\right\|$ 积分，得到

$$\begin{aligned}\left\|\dot{\tilde{X}}_\delta(t)\right\|^2 - \left\|\dot{\tilde{X}}_\delta(0)\right\|^2 &\le -\frac{2\gamma}{\delta^p}\int_0^t \left\|\dot{\tilde{X}}_\delta(s)\right\|^2 ds \\ &\le -\frac{2\gamma}{\delta^p}\int_0^t \left\|\dot{\tilde{X}}_\delta(t)\right\|^2 ds = -\frac{2\gamma t}{\delta^p}\left\|\dot{\tilde{X}}_\delta(t)\right\|^2.\end{aligned}$$

按 $\left\|\dot{\tilde{X}}_\delta(t)\right\|$ 整理，得到

$$\left\|\dot{\tilde{X}}_\delta(t)\right\| \le \frac{\left\|\dot{\tilde{X}}_\delta(0)\right\|}{\sqrt{1+\frac{2\gamma t}{\delta^p}}} = \sqrt{\frac{\delta^p}{\delta^p + 2\gamma t}}\left\|\tilde{A}(X_0)\right\| \le \sqrt{\frac{\delta^p}{2\gamma t}}\left\|\tilde{A}(X_0)\right\|.$$

(2) $t \ge \delta$

注意到 $t > \delta$ (12) 成立，我们可以应用引理 B.5。

(i) $0 < p \le 1$

由 (i) 可知 $\left\|\dot{\tilde{X}}_\delta(t)\right\| \le \left\|\tilde{A}(X_0)\right\|$ 对 $t \in [0, \delta]$成立。因此 $\left\|\dot{\tilde{X}}_\delta(\delta)\right\| \le \left\|\tilde{A}(X_0)\right\|$，我们看到

$$\left\|\tilde{X}_\delta(\delta) - X_0\right\| = \left\|\int_0^\delta \dot{\tilde{X}}_\delta(s)\, ds\right\| \le \int_0^\delta \left\|\dot{\tilde{X}}_\delta(s)\right\| ds \le \delta\left\|\tilde{A}(X_0)\right\|.$$

由引理 B.5 可知，$\tilde{U}_1(t) = \left\|\dot{\tilde{X}}_\delta(t)\right\|^2 + \frac{\gamma p}{t^{p+1}}\left\|\tilde{X}_\delta(t) - X_0\right\|^2$ 是非递增的。因此 $\left\|\dot{\tilde{X}}_\delta(t)\right\|^2 \le U_1(\delta)$ 对 $t \ge \delta$成立。由于 $\delta < 1$，我们有

$$\begin{aligned}\left\|\dot{\tilde{X}}_\delta(t)\right\| \le \sqrt{\tilde{U}_1(\delta)} &= \sqrt{\left\|\dot{\tilde{X}}_\delta(\delta)\right\|^2 + \frac{\gamma p}{\delta^{p+1}}\left\|\tilde{X}_\delta(\delta) - X_0\right\|^2} \\ &\le \sqrt{\left\|\tilde{A}(X_0)\right\|^2 + \gamma p \delta^{1-p}\left\|\tilde{A}(X_0)\right\|^2} \le \sqrt{1+\gamma p}\left\|\tilde{A}(X_0)\right\|.\end{aligned}$$

(ii) $p > 1$

由 (i) 可知，对于 $t \in [0, \delta]$，有 $\left\|\dot{\tilde{X}}_\delta(t)\right\| \le \sqrt{\frac{\delta^p}{2\gamma t}}\left\|\tilde{A}(X_0)\right\|$。因此，$\left\|\dot{\tilde{X}}_\delta(\delta)\right\| \le \sqrt{\frac{\delta^{p-1}}{2\gamma}}\left\|\tilde{A}(X_0)\right\|$，我们看到

$$\left\|\tilde{X}_\delta(\delta) - X_0\right\| \le \int_0^\delta \left\|\dot{\tilde{X}}_\delta(s)\right\| ds \le \int_0^\delta \sqrt{\frac{\delta^p}{2\gamma s}}\left\|\tilde{A}(X_0)\right\| ds = \sqrt{\frac{2\delta^{p+1}}{\gamma}}\left\|\tilde{A}(X_0)\right\|.$$

˜应用 (13)，并回顾 A 是 L-Lipschitz 的这一事实，我们有

$$\begin{aligned}\frac{\gamma^2}{\delta^{2p}}\left\|X_\delta(\delta)-X_0\right\|^2 &= \left\|\tilde{A}(X_\delta(\delta))+\dot{X}_\delta(\delta)\right\|^2 = \left\|\tilde{A}(X_\delta(\delta))-\tilde{A}(X_0)+\tilde{A}(X_0)+\dot{X}_\delta(\delta)\right\|^2\\ &\le 2\left\|\tilde{A}(X_\delta(\delta))-\tilde{A}(X_0)\right\|^2+2\left\|\tilde{A}(X_0)+\dot{X}_\delta(\delta)\right\|^2\\ &\le 2L^2\left\|\tilde{X}_\delta(\delta)-X_0\right\|^2+4\left\|\tilde{A}(X_0)\right\|^2+4\left\|\dot{X}_\delta(\delta)\right\|^2\\ &\le \left(\frac{4}{\gamma}L^2\delta^{p+1}+4+\frac{2\delta^{p-1}}{\gamma}\right)\left\|\tilde{A}(X_0)\right\|^2.\end{aligned}$$

由引理 B.5 可知，$\tilde{U}_2(t)=\frac{1}{t^{p-1}}\left\|\dot{\tilde{X}}(t)\right\|^2+\frac{\gamma p}{t^{2p}}\left\|\tilde{X}(t)-X_0\right\|^2$ 是非递增的。因此，对于 $\frac{\left\|\dot{\tilde{X}}_\delta(t)\right\|^2}{t^{p-1}}\le$ ˜ $U_2(\delta)$，$t\ge\delta$。由于 $\delta<1$，我们有

$$\begin{aligned}\frac{\left\|\dot{\tilde{X}}_\delta(t)\right\|}{\sqrt{t^{p-1}}} &\le \sqrt{\tilde{U}_2(\delta)}=\sqrt{\frac{\left\|\dot{\tilde{X}}_\delta(\delta)\right\|^2}{\delta^{p-1}}+\frac{\gamma p}{\delta^{2p}}\left\|\tilde{X}_\delta(\delta)-X_0\right\|^2}\\ &\le \sqrt{\frac{\left\|\tilde{A}(X_0)\right\|^2}{2\gamma}+p\left(4L^2\delta^{p+1}+4\gamma+2\delta^{p-1}\right)\left\|\tilde{A}(X_0)\right\|^2}\le\sqrt{\frac{1}{2\gamma}+2p\left(2L^2+2\gamma+1\right)}\left\|\tilde{A}(X_0)\right\|.\end{aligned}$$

因此，对于 $t\in[0,T]$，我们有

$$\left\|\dot{\tilde{X}}_\delta(t)\right\|\le\sqrt{T^{p-1}\left(\frac{1}{2\gamma}+2p\left(2L^2+2\gamma+1\right)\right)}\left\|\tilde{A}(X_0)\right\|$$

由 (1) 和 (2)，可得所需结果。 □

现在我们证明，对于任意 $T>0$，序列 $\left\{\tilde{X}_{\delta_m}\right\}_{m\in}\mathbb{N}$ 在 $[0,T]$ 上一致收敛。只需证明以下命题。

命题 B.7. 对于 T >0，序列 $\left\{\tilde{X}_{\delta_m}\right\}_{m\in}\mathbb{N}$ 关于 $[0,T]$ 上的上确界范数是柯西序列。

证明。取 $\epsilon>0$。我们要证明，存在 $N\in\mathbf{N}$，使得若 $n,m>N$，则对于所有 $t\in[0,\infty)$，都有 $\left\|\tilde{X}_{\delta_n}(t)-\tilde{X}_{\delta_m}(t)\right\|\le\epsilon$。定义 $d_{\delta,\nu}(t)=\frac{1}{2}\left\|\tilde{X}_\delta(t)-\tilde{X}_\nu(t)\right\|^2$。不失一般性，我们可以假设 $\delta\ge\nu$。利用引理 B.6 中定义的 $M_{dot}(T)$，我们将证明 $d_{\delta,\nu}(t)\le 2\delta^2\tilde{M}_{dot}(T)^2$。

首先考虑情形 $0\le t\le\delta$。根据引理 B.6，我们有

$$\begin{aligned}\left\|\tilde{X}_\delta(t)-\tilde{X}_\nu(t)\right\| &\le \int_0^t\left\|\dot{\tilde{X}}_\delta(s)-\dot{\tilde{X}}_\nu(s)\right\|ds\le\int_0^t\left(\left\|\dot{\tilde{X}}_\delta(s)\right\|+\left\|\dot{\tilde{X}}_\nu(s)\right\|\right)ds\\ &\le\int_0^t 2\tilde{M}_{dot}(T)ds=2t\tilde{M}_{dot}(T)\end{aligned}$$

因此，对于 $t\in[0,\delta]$，我们有

$$d_{\delta,\nu}(t)\le 2t^2\tilde{M}_{dot}(T)^2\le 2\delta^2\tilde{M}_{dot}(T)^2.$$

˜现在考虑情形 $t \geq \delta$。根据 A 的单调性，我们有

$$\begin{aligned}
\frac{d}{dt}\frac{1}{2}\left\|\tilde{X}_\delta(t)-\tilde{X}_\nu(t)\right\|^2 &= \left\langle \dot{X}_\delta(t)-\dot{X}_\nu(t), \tilde{X}_\delta(t)-\tilde{X}_\nu(t)\right\rangle \\
&= \left\langle -\left(\tilde{A}(\tilde{X}_\delta(t))-\tilde{A}(\tilde{X}_\nu(t))\right)-\frac{\gamma}{t^p}(X_\delta(t)-X_0)+\frac{\gamma}{t^p}(X_\nu(t)-X_0), X_\delta(t)-X_\nu(t)\right\rangle \\
&\leq \left\langle -\frac{\gamma}{t^p}(X_\delta(t)-X_0)+\frac{\gamma}{t^p}(X_\nu(t)-X_0), X_\delta(t)-X_\nu(t)\right\rangle \\
&= -\frac{\gamma}{t^p}\left\|X_\delta(t)-X_\nu(t)\right\|^2 \\
&\leq 0.
\end{aligned}$$

因此我们有$d_{\delta,\ \nu}(t) \leq d_{\delta,\ \nu}(\delta)$，其中 $t \geq \delta$。

现在结合两种情况，我们得到

$$d_{\delta,\nu}(t) \leq 2\delta^2 \tilde{M}_{dot}(T)^2 \tag{14}$$

由于 $\lim_{k\to\infty}\delta_n = 0$，存在 $N \in \mathbb{N}$，使得 $m > n > N$ 蕴含 $d_{\delta_n,\ \delta_m} \leq \frac{\epsilon^2}{2}$，证毕。 □

现在我们可以证明定理 B.2 了。

证明。(1) 解的存在性。由命题 B.7 可知，$\left\{\tilde{X}_{\delta_m}\right\}_{m\in\mathbb{N}}$ 在 $[0,T]$ 上一致收敛，对每个 $T >$ 都成立。将极限记为 $\tilde{X}$，˜，即定义 $X\colon [0,\infty) \to \mathbb{R}^n$ 为

$$\lim_{m\to\infty}\tilde{X}_{\delta_m}(t) = \tilde{X}(t).$$

˜ ˜ ˜ck $X(0) = X_0$ 容易得出，因为 $X_{\delta_m}(0) = X_0$ 对所有 $m \in \mathbb{N}$ 都成立。

我们可以检剩下要证明 X 满足 (12)。

取 $t > 0$。我们希望证明

$$\lim_{h\to 0}\left\|\frac{\tilde{X}(t+h)-\tilde{X}(t)}{h}+\tilde{A}(\tilde{X}(t))+\frac{\gamma}{t^p}\left(\tilde{X}(t)-X_0\right)\right\| = 0.$$

考虑 h, δ, T，使得 $0 < |h| < t, 0 < \delta <$ min $\{1, t-|h|\}$ 且 $T > t+|h|$。那么 $(t-|h|, t+|h|) \in [0,T]$且 $\dot{\tilde{X}}_\delta(t) = -\tilde{A}(\tilde{X}_\delta(t)) - \frac{\gamma}{t^p}\left(\tilde{X}_\delta(t) - X_0\right)$。考虑不等式

$$\begin{aligned}
&\left\|\frac{\tilde{X}(t+h)-\tilde{X}(t)}{h}+\tilde{A}(\tilde{X}(t))+\frac{\gamma}{t^p}\left(\tilde{X}(t)-X_0\right)\right\| \\
&\leq \left\|\frac{\tilde{X}(t+h)-\tilde{X}(t)}{h}-\frac{\tilde{X}_\delta(t+h)-\tilde{X}_\delta(t)}{h}\right\| + \left\|\frac{\tilde{X}_\delta(t+h)-\tilde{X}_\delta(t)}{h}-\dot{\tilde{X}}_\delta(t)\right\| + \left\|\dot{\tilde{X}}_\delta(t)+\tilde{A}(\tilde{X}(t))+\frac{\gamma}{t^p}\left(\tilde{X}(t)-X_0\right)\right\|.
\end{aligned}$$

˜ ˜现在证明，当 $h \to 0$ 时，右侧趋于零。证明的关键在于，X_δ 一致地收敛到 X，并且 ¨ ˜X_δ在 $[t-|h|, T]$ 上一致有界。

由（14）可得

$$\begin{aligned}
&\left\|\frac{\tilde{X}(t+h)-\tilde{X}(t)}{h}-\frac{\tilde{X}_\delta(t+h)-\tilde{X}_\delta(t)}{h}\right\| \\
&\leq \frac{1}{|h|}\left(\left\|\tilde{X}(t+h)-\tilde{X}_\delta(t+h)\right\| + \left\|\tilde{X}(t)-\tilde{X}_\delta(t)\right\|\right) \leq 4\frac{\delta}{|h|}\tilde{M}_{dot}(T).
\end{aligned}$$

˜同样由（14）可知，且由于 A 是 L-利普希茨连续的，

$$\begin{aligned}
\left\|\dot{\tilde{X}}_\delta(t)+\tilde{A}(\tilde{X}(t))+\frac{\gamma}{t^p}\left(\tilde{X}(t)-X_0\right)\right\| &= \left\|-\left(\tilde{A}(\tilde{X}_\delta(t))+\frac{\gamma}{t^p}\left(\tilde{X}_\delta(t)-X_0\right)\right)+\tilde{A}(\tilde{X}(t))+\frac{\gamma}{t^p}\left(\tilde{X}(t)-X_0\right)\right\| \\
&\leq L\left\|\tilde{X}_\delta(t)-\tilde{X}(t)\right\| + \frac{\gamma}{t^p}\left\|\tilde{X}_\delta(t)-\tilde{X}(t)\right\| \leq 2\delta\left(L+\frac{\gamma}{t^p}\right)\tilde{M}_{dot}(T).
\end{aligned}$$

˙现在由引理 B.4 可知， $X_\delta(t)$是绝对连续的，因此

$$\left\|\frac{\tilde{X}_\delta(t+h)-\tilde{X}_\delta(t)}{h}-\dot{\tilde{X}}_\delta(t)\right\| = \left\|\frac{\int_t^{t+h}\left(\dot{\tilde{X}}_\delta(s)-\dot{\tilde{X}}_\delta(t)\right)ds}{h}\right\|$$
$$= \frac{1}{|h|}\left\|\int_t^{t+h}\int_t^s \ddot{\tilde{X}}_\delta(u)\,du\,ds\right\| \le \frac{1}{|h|}\left|\int_t^{t+h}\int_t^s \left\|\ddot{\tilde{X}}_\delta(u)\right\|\,du\,ds\right|$$

注意，对于几乎所有的 $u\in[t-|h|,t+|h|]\subset[0,T]$，根据引理 B.6，我们有

$$\left\|\ddot{\tilde{X}}_\delta(u)\right\| = \left\|-\frac{d}{du}\tilde{A}(\tilde{X}_\delta(u))+\frac{p\gamma}{u^{p+1}}\left(\tilde{X}_\delta(u)-X_0\right)-\frac{\gamma}{u^p}\dot{X}_\delta(u)\right\|$$
$$= \left\|\frac{d}{du}\tilde{A}(\tilde{X}_\delta(u))\right\|+\left\|\frac{p\gamma}{u^{p+1}}\int_0^u\dot{\tilde{X}}_\delta(v)dv\right\|+\left\|\frac{\gamma}{u^p}\dot{X}_\delta(u)\right\|$$
$$\le L\tilde{M}_{dot}(T)+\frac{p\gamma}{(t-|h|)^{p+1}}\int_0^T\tilde{M}_{dot}(T)dv+\frac{\gamma}{(t-|h|)^p}\tilde{M}_{dot}(T)$$
$$= \left(L+\frac{p\gamma T}{(t-|h|)^{p+1}}+\frac{\gamma}{(t-|h|)^p}\right)\tilde{M}_{dot}(T)=:M.$$

˜ ˜ ˜注意， M与 h或 δ无关。在得到该不等式时，我们用到了这样一个事实：$A(X_\delta(\cdot))$关于 $[0,T]$是 $LM_{dot}(T)$-Lipschitz 连续的。现在

$$\frac{1}{|h|}\left|\int_t^{t+h}\int_t^s\left\|\ddot{\tilde{X}}_\delta(u)\right\|\,du\,ds\right| \le \frac{1}{|h|}\left|\int_t^{t+h}\int_t^s M\,du\,ds\right| = \frac{1}{|h|}M\left|\int_t^{t+h}(s-t)\,ds\right| = \frac{|h|}{2}M.$$

现在考虑 $\delta=h^2$，取足够小的 h，使其满足 $|h|<1$ 和 $|h|+h^2<t$。此时条件 $|h|<t$、$0<\delta<\min\{1,t-|h|\}$ 均成立，上述论证有效。综合以上结果，我们有

$$\left\|\frac{\tilde{X}(t+h)-\tilde{X}(t)}{h}+\tilde{A}(\tilde{X}(t))+\frac{\gamma}{t^p}\left(\tilde{X}(t)-X_0\right)\right\|$$
$$\le 2\sqrt{2}|h|\tilde{M}_{dot}(T)+\sqrt{2}|h|^2\left(L+\frac{\gamma}{t^p}\right)\tilde{M}_{dot}(T)+\frac{|h|}{2}M=\mathcal{O}\left(|h|\right).$$

这就推出了所需结果。

˜.(2) $X(t)$ 在 $t=0$处的值与连续性。

定义 $C(t)$如下

$$C(t)=\begin{cases}t^\gamma & p=1\\ e^{\frac{\gamma}{1-p}t^{1-p}} & p>0,p\neq1.\end{cases}$$

那么我们看到，对于 $t>0$

$$\frac{d}{dt}\left(C(t)\left(\tilde{X}(t)-X_0\right)\right)=C(t)\left(\dot{\tilde{X}}(t)+\frac{\gamma}{t^p}(\tilde{X}(t)-X_0)\right)=-C(t)\tilde{A}(\tilde{X}(t)).$$

对两边从 $\epsilon>0$ 积分到 t，可得

$$C(t)(\tilde{X}(t)-X_0)-C(\epsilon)(\tilde{X}(\epsilon)-X_0)=-\int_\epsilon^t C(s)\tilde{A}(\tilde{X}(s))\,ds.$$

由于 C 是单调不减函数且下界为 0，因此 $\lim_{\epsilon\to0+}C(\epsilon)$ 存在，取极限 $\epsilon\to0+$，得

$$C(t)(\tilde{X}(t)-X_0)=-\int_0^t C(s)\tilde{A}(\tilde{X}(s))\,ds.$$

两边同时除以 $tC(t)$，令变量代换 $v = s/t$，得

$$\frac{\tilde{X}(t) - X_0}{t} = -\int_0^t \frac{C(s)}{C(t)} \tilde{A}(\tilde{X}(s)) \frac{ds}{t} = -\int_0^1 \frac{C(tv)}{C(t)} \tilde{A}(\tilde{X}(tv))\, dv.$$

注意

$$\frac{C(tv)}{C(t)} = \begin{cases} v^\gamma & p = 1 \\ e^{\frac{\gamma}{1-p} t^{1-p} \left(v^{1-p} - 1\right)} & p \neq 1, p > 0. \end{cases}$$

注意，对于 $\frac{\gamma}{1-p} t^{1-p} \left(v^{1-p} - 1\right) \leq 0$，由于 $\frac{\gamma}{1-p}$ 和 $\left(v^{1-p} - 1\right)$ 符号相反，因此 $0 < p < 1$ 或 $p > 1$。因此 $e^{\frac{\gamma}{1-p} t^{1-p} \left(v^{1-p} - 1\right)} \leq 1$。此外，对于 $v \in [0, 1]$，$\tilde{A}(\tilde{X}(tv))$ 有界，因为根据引理 B.4，$A(\tilde{X}(\cdot))$ 连续。

因此，我们可以应用控制收敛定理并取极限 $t \to 0+$。由于

$$\lim_{t \to 0+} \frac{C(tv)}{C(t)} = \begin{cases} 1 & 0 < p < 1 \\ v^\gamma & p = 1 \\ 0 & p > 1 \end{cases}$$

对于 $v \neq 0$，我们有

$$\dot{\tilde{X}}(0) = \lim_{t \to 0+} \frac{\tilde{X}(t) - X_0}{t} = -\int_0^1 \lim_{t \to 0+} \left(\frac{C(tv)}{C(t)} \tilde{A}(\tilde{X}(tv)) \right) dv = \begin{cases} -\tilde{A}(X_0) & 0 < p < 1 \\ -\frac{1}{\gamma+1} \tilde{A}(X_0) & p = 1 \\ 0 & p > 1. \end{cases} \tag{15}$$

现在我们来验证 $X(t)$ 在 $t = 0$处连续。

(i) $0 < p \leq 1$对于 $t > 0$,
我们知道

$$\dot{\tilde{X}}(t) = -\tilde{A}(\tilde{X}(t)) - \frac{\gamma}{t^p} (\tilde{X}(t) - X_0).$$

观察

$$\lim_{t \to 0+} \frac{\gamma}{t^p} (\tilde{X}(t) - X_0) = \lim_{t \to 0+} \left(\gamma t^{1-p} \cdot \frac{\tilde{X}(t) - X_0}{t} \right) = \begin{cases} 0 & 0 < p < 1 \\ -\frac{\gamma}{\gamma+1} \tilde{A}(X_0) & p = 1. \end{cases} \tag{16}$$

现在根据常微分方程 $\dot{\tilde{X}}(t) = -A(X(t)) - \frac{\gamma}{t^p}(X(t) - X_0)$，取极限 $t \to 0+$，我们得到

$$\lim_{t \to 0+} \dot{\tilde{X}}(t) = -\tilde{A}(X_0) - \lim_{t \to 0+} \left(\frac{\gamma}{t^p} (\tilde{X}(t) - X_0) \right) = \begin{cases} -\tilde{A}(X_0) & 0 < p < 1 \\ -\frac{1}{\gamma+1} \tilde{A}(X_0) & p = 1. \end{cases}$$

因此，$X(t)$ 在 $t = 0$处连续。

(ii) $p > 1$
对于 $p > 1$，我们不知道 $\lim_{t \to 0} \frac{\gamma}{t^p}(X(t) - X_0)$ 的值。因此我们先求这个极限。
让我们回到

$$C(t)(\tilde{X}(t) - X_0) = -\int_0^t C(s) \tilde{A}(\tilde{X}(s))\, ds.$$

回顾 $\dot{C}(t) = \frac{\gamma}{t^p} C(t)$。对右侧进行分部积分，可得

$$\int_0^t C(s) \tilde{A}(\tilde{X}(s))\, ds = \frac{t^p}{\gamma} C(t) \tilde{A}(\tilde{X}(t)) - \int_0^t p \frac{s^{p-1}}{\gamma} C(s) \tilde{A}(\tilde{X}(s))\, ds - \int_0^t \frac{s^p}{\gamma} C(s) \frac{d}{ds} \tilde{A}(\tilde{X}(s))\, ds.$$

根据引理 B.4，我们知道 $A(X(s))$ 几乎处处可微。现在，将等式两边同时除以$t^p C(t)\, \gamma$。于是，对于 $s = tv$，我们有

$$\begin{aligned} \frac{\gamma}{t^p} (\tilde{X}(t) - X_0) &= -\tilde{A}(\tilde{X}(t)) + \int_0^t \frac{p s^{p-1}}{t^p} \frac{C(s)}{C(t)} \tilde{A}(\tilde{X}(s))\, ds + \int_0^t \frac{s^p}{t^p} \frac{C(s)}{C(t)} \left(\frac{d}{ds} \tilde{A}(\tilde{X}(s)) \right) ds \\ &= -\tilde{A}(\tilde{X}(t)) + \int_0^1 p v^{p-1} \frac{C(tv)}{C(t)} \tilde{A}(\tilde{X}(s))\, dv + \int_0^1 v^p \frac{C(tv)}{C(t)} \left(\frac{d}{ds} \tilde{A}(\tilde{X}(s)) \right) t\, dv. \end{aligned}$$

根据命题 B.7，且由于 X 满足 (12)，对于 $t>0$，当 $s\in(0,\ t]$ 时，我们有

$$\dot{\tilde{X}}(s)=\tilde{A}(\tilde{X}(s))+\frac{\gamma}{t^p}\left(\tilde{X}(s)-X_0\right)=\lim_{m\to\infty}\left(\tilde{A}(\tilde{X}_{\delta_m}(s))+\frac{\gamma}{t^p}\left(\tilde{X}_{\delta_m}(s)-X_0\right)\right)=\lim_{m\to\infty}\dot{\tilde{X}}_{\delta_m}(s).$$

由引理 B.6 可知，对于 $s\in[0,t]$，$\left\|\dot{\tilde{X}}_{\delta_m}(s)\right\|\le\tilde{M}_{dot}(t)$；取极限 $m\to\infty$，可得对于 $s\in[0,t]$，$\left\|\dot{\tilde{X}}(s)\right\|\le\tilde{M}_{dot}(t)$。因此，$X(s)$ 关于 $s\in[0,t]$ 变为 $M_{dot}(t)$-Lipschitz 连续。于是，$(A\circ X)(s)$ 关于 $s\in[0,t]$ 变为 $L\tilde{M}_{dot}(t)$-Lipschitz 连续，且对于几乎所有 $s\in[0,t]$，有 $\left\|\frac{d}{ds}\tilde{A}(\tilde{X}(s))\right\|\le L\tilde{M}_{dot}(t)$。此外，由于 C 是非负的单调不减函数，$C(tv)/C(t)$ 关于 $v\in[0,1]$ 有界。

因此，我们可以再次应用控制收敛定理。回想一下，当 $p>1$ 时，$\lim_{t\to0+}\frac{C_{(tv)}}{C(t)}=0$，取极限 $t\to0+$，得到

$$\lim_{t\to0+}\frac{\gamma}{t^p}(\tilde{X}(t)-X_0)=-\tilde{A}(X_0)+0=-\tilde{A}(X_0).\tag{17}$$

最后，我们得到

$$\lim_{t\to0+}\dot{\tilde{X}}(t)=-\tilde{A}(X_0)-\lim_{t\to0+}\left(\frac{\gamma}{t^p}(\tilde{X}(t)-X_0)\right)=-\tilde{A}(X_0)-(-\tilde{A}(X_0))=0=\dot{\tilde{X}}(0).$$

□

在转向原包含关系之前，我们先证明几个重要推论，它们给出了 $\left\|\dot{\tilde{X}}(t)\right\|$ 和 $\left\|\tilde{A}(\tilde{X}(t))\right\|$ 在 $[0,T]$ 上的一致界。请注意，下面的推论与引理 B.6 的主要区别在于，对于情形 $p>1$，不再依赖 Lipschitz 常数 L，这将在下一节中起到关键作用。

推论 B.8. 设 X 为 (12) 的解。那么，对于 $T>0$，以下不等式对 $t\in[0,T]$ 成立。

$$\left\|\dot{\tilde{X}}(t)\right\|\le\begin{cases}\left\|\tilde{A}(X_0)\right\| & 0<p<1\\ \frac{1}{\sqrt{\gamma+1}}\left\|\tilde{A}(X_0)\right\| & p=1\\ \sqrt{\frac{p}{\gamma}T^{p-1}}\left\|\tilde{A}(X_0)\right\| & p>1.\end{cases}$$

证明。(i) $0<p\le1$

由于 X 是 (12) 的解，根据引理 B.5，我们知道

$$\tilde{U}_1(t)=\left\|\dot{\tilde{X}}(t)\right\|^2+\frac{\gamma p}{t^{p+1}}\left\|\tilde{X}(t)-X_0\right\|^2$$

是一个非增函数。由 (15)、(16) 以及 $X(t)$ 在 $t=0$处的连续性可得

$$\lim_{\epsilon\to0+}\tilde{U}_1(\epsilon)=\begin{cases}\left\|\tilde{A}(X_0)\right\|^2 & 0<p<1\\ \frac{1}{\gamma+1}\left\|\tilde{A}(X_0)\right\|^2 & p=1.\end{cases}\tag{18}$$

因此，对于 $t>0$，$\left\|\dot{\tilde{X}}(t)\right\|^2\le\tilde{U}_1(t)\le\lim_{\epsilon\to0+}\tilde{U}_1(\epsilon)$，从而得到所需结果。

(ii) $p>1$

As $\tilde{X}$ 是 (12) 的解，根据引理 B.5，我们知道

$$\tilde{U}_2(t)=\frac{1}{t^{p-1}}\left\|\dot{\tilde{X}}(t)\right\|^2+\frac{\gamma p}{t^{2p}}\left\|\tilde{X}(t)-X_0\right\|^2$$

是一个非增函数。然而，由于我们不知道 $\lim_{t\to0+}\frac{1}{t^{p-1}}\left\|\dot{\tilde{X}}(t)\right\|^2$ 的值，因此先计算它。为此，我们考虑

$$\tilde{U}_1(t)=\left\|\dot{\tilde{X}}(t)\right\|^2+\frac{\gamma p}{t^{p+1}}\left\|\tilde{X}(t)-X_0\right\|^2.$$

在引理 B.5 的证明中，我们观察到它的导数变为

$$\begin{aligned}\dot{\tilde{U}}_1(t) &= -2\left\langle \dot{\tilde{X}}(t), \frac{d}{dt}\tilde{A}(\tilde{X}(t))\right\rangle - \frac{2\gamma}{t^p}\left\|\dot{\tilde{X}}(t) - \frac{p}{t}(\tilde{X}(t) - X_0)\right\|^2 - \frac{\gamma p(1-p)}{t^{p+2}}\left\|\tilde{X}(t) - X_0\right\|^2 \\ &\leq \frac{\gamma p(p-1)}{t^{p+2}}\left\|\tilde{X}(t) - X_0\right\|^2.\end{aligned}$$

对于 $\epsilon > 0$，从 ϵ 到 t 积分可得

$$\tilde{U}_1(t) \leq \tilde{U}_1(\epsilon) + \int_\epsilon^t \frac{\gamma p(p-1)}{s^{p+2}}\left\|\tilde{X}(s) - X_0\right\|^2 ds.$$

我们考虑取极限 $\epsilon \to 0+$。由 (17) 和 (15) 可知，$\lim_{t\to 0+} \frac{\tilde{X}-X_0}{t^p} = -\tilde{A}_{(X_0)\gamma}$，且 $\left\|\dot{\tilde{X}}(0)\right\| = 0$，因此可得

$$\lim_{\epsilon\to 0+}\tilde{U}_1(\epsilon) = 0 + \lim_{\epsilon\to 0+}\frac{\gamma p}{\epsilon^{2p}}\left\|\tilde{X}(\epsilon) - X_0\right\|^2 \cdot \epsilon^{p-1} = \frac{p}{\gamma}\left\|\tilde{A}(X_0)\right\|^2 \lim_{\epsilon\to 0+}\epsilon^{p-1} = 0.$$

此外，由于 $\lim_{t\to 0+} \frac{\|\tilde{X}_{(t)} - X_0\|^2}{t^{2p}} = \frac{\|\tilde{A}(X_0)\|^2}{\gamma^2}$，存在 $\delta > 0$，使得当 $0 < s < \delta$ 时，有 $\frac{\|\tilde{X}(t) - X_0\|^2}{s^{2p}} \leq \frac{2\|\tilde{A}(X_0)\|^2}{\gamma^2}$。回顾 $p > 1$，对于 $0 < \epsilon < \delta$，有

$$\begin{aligned}\int_0^\epsilon \frac{\gamma p(p-1)}{s^{p+2}}\left\|\tilde{X}(s) - X_0\right\|^2 ds &\leq \int_0^\epsilon \frac{\gamma p(p-1)}{s^{2-p}}\frac{2\left\|\tilde{A}(X_0)\right\|^2}{\gamma^2} ds \\ &= \frac{2p(p-1)\left\|\tilde{A}(X_0)\right\|^2}{\gamma}\left[\frac{1}{p-1}s^{p-1}\right]_0^\epsilon = \frac{2p\left\|\tilde{A}(X_0)\right\|^2}{\gamma}\epsilon^{p-1},\end{aligned}$$

因此

$$\int_0^t \frac{\gamma p(p-1)}{s^{p+2}}\left\|\tilde{X}(s) - X_0\right\|^2 ds \leq \frac{2p\left\|\tilde{A}(X_0)\right\|^2}{\gamma}\epsilon^{p-1} + \int_\epsilon^t \frac{\gamma p(p-1)}{s^{p+2}}\left\|\tilde{X}(s) - X_0\right\|^2 ds < \infty.$$

因此，当 $\epsilon \to 0+$时，该积分是良定义的。取极限 $\epsilon \to 0+$，我们有

$$\tilde{U}_1(t) \leq \int_0^t \frac{\gamma p(p-1)}{s^{p+2}}\left\|\tilde{X}(s) - X_0\right\|^2 ds.$$

将 $\frac{\gamma p}{t^{p+1}}\left\|\tilde{X}(t) - X_0\right\|^2$ 移至右侧，我们得到关于 $\left\|\dot{\tilde{X}}(t)\right\|^2$的一个不等式，

$$\left\|\dot{\tilde{X}}(t)\right\|^2 = \tilde{U}_1(t) - \frac{\gamma p}{t^{p+1}}\left\|\tilde{X}(t) - X_0\right\|^2 \leq \int_0^t \frac{\gamma p(p-1)}{s^{p+2}}\left\|\tilde{X}(s) - X_0\right\|^2 ds - \frac{\gamma p}{t^{p+1}}\left\|\tilde{X}(t) - X_0\right\|^2.$$

注意，根据洛必达法则，我们有

$$\lim_{t\to 0+}\frac{\int_0^t \frac{\gamma p(p-1)}{s^{p+2}}\left\|\tilde{X}(s) - X_0\right\|^2 ds}{t^{p-1}} = \lim_{t\to 0+}\frac{\frac{\gamma p(p-1)}{t^{p+2}}\left\|\tilde{X}(t) - X_0\right\|^2}{(p-1)t^{p-2}} = \lim_{t\to 0+}\gamma p\frac{\left\|\tilde{X}(t) - X_0\right\|^2}{t^{2p}} = \frac{p}{\gamma}\left\|\tilde{A}(X_0)\right\|^2.$$

现在将 t^{p-1} 除以前一个不等式并取极限，我们得出

$$\begin{aligned}\lim_{t\to 0+}\frac{1}{t^{p-1}}\left\|\dot{\tilde{X}}\right\|^2 &\leq \lim_{t\to 0+}\frac{1}{t^{p-1}}\int_0^t \frac{\gamma p(p-1)}{s^{p+2}}\left\|\tilde{X} - X_0\right\|^2 ds - \lim_{t\to 0+}\frac{\gamma p}{t^{2p}}\left\|\tilde{X} - X_0\right\|^2 \\ &= \frac{p}{\gamma}\left\|\tilde{A}(X_0)\right\|^2 - \frac{p}{\gamma}\left\|\tilde{A}(X_0)\right\|^2 \\ &= 0.\end{aligned}$$

因此，我们有 $\lim_{t\to 0+} \frac{1}{t^{p-1}} \left\|\dot{\tilde{X}}\right\|^2 = 0$.所以,

$$\lim_{\epsilon\to 0+} \tilde{U}_2(\epsilon) = \lim_{\epsilon\to 0+} \left(\frac{1}{\epsilon^{p-1}} \left\|\dot{\tilde{X}}(\epsilon)\right\|^2 + \frac{\gamma p}{\epsilon^{2p}} \left\|\tilde{X}(\epsilon) - X_0\right\|^2 \right) = \frac{p}{\gamma} \left\|\tilde{A}(X_0)\right\|^2.$$

由 $\frac{\left\|\dot{\tilde{X}}\right\|^2}{t^{p-1}} \le \lim_{\epsilon\to 0+} \tilde{U}_2(\epsilon) = \frac{p}{\gamma} \left\|\tilde{A}(X_0)\right\|^2$，我们得出所需结果。

□

推论 B。 **9.** 令 X 为 (12) 的解。那么对于 $T > 0$，以下不等式对 $t \in [0, T$ 成立].

$$\left\|\tilde{A}(\tilde{X}(t))\right\| \le \begin{cases} \sqrt{(\gamma+1)\left(1 + \frac{T^{1-p}}{p}\right)} \left\|\tilde{A}(X_0)\right\| & 0 < p < 1 \\ \left\|\tilde{A}(X_0)\right\| & p = 1 \\ \sqrt{\frac{p(\gamma+1)}{\gamma}\left(T^{p-1} + \frac{1}{p}\right)} \left\|\tilde{A}(X_0)\right\| & p > 1. \end{cases}$$

证明。 首先注意到,

$$\begin{aligned} \left\|\tilde{A}(\tilde{X}(t))\right\|^2 &= \left\|\dot{\tilde{X}}(t) + \frac{\gamma}{t^p}\left(\tilde{X}(t) - X_0\right)\right\|^2 \\ &= \left\|\dot{\tilde{X}}(t)\right\|^2 + 2\gamma \left\langle \dot{\tilde{X}}(t), \frac{\tilde{X}(t) - X_0}{t^p} \right\rangle + \frac{\gamma^2}{t^{2p}} \left\|\tilde{X}(t) - X_0\right\|^2 \\ &\le \left\|\dot{\tilde{X}}(t)\right\|^2 + \gamma \left(\left\|\dot{\tilde{X}}(t)\right\|^2 + \frac{\left\|\tilde{X}(t) - X_0\right\|^2}{t^{2p}} \right) + \frac{\gamma^2}{t^{2p}} \left\|\tilde{X}(t) - X_0\right\|^2 \\ &= (\gamma+1)\left(\left\|\dot{\tilde{X}}(t)\right\|^2 + \frac{\gamma}{t^{2p}} \left\|\tilde{X}(t) - X_0\right\|^2 \right) \end{aligned}$$

该不等式由杨氏不等式得出。现在我们考虑各种情况。

(i) $p = 1$ 对于这种情况，右侧各项恰好变为 $U_1(t)$。因此由 (18)

$$\left\|\tilde{A}(\tilde{X}(t))\right\|^2 \le (\gamma+1)\tilde{U}_1(t) \le (\gamma+1) \lim_{\epsilon\to 0+} \tilde{U}_1(\epsilon) = \left\|\tilde{A}(X_0)\right\|^2.$$

(ii) $0 < p < 1$

回顾推论 B.8 的证明，我们知道 $\tilde{U}_1(t) \le \lim_{\epsilon\to 0+} \tilde{U}_1(\epsilon) = \left\|\tilde{A}(X_0)\right\|^2$。因此我们有 $\frac{\gamma p}{t^{p+1}} \left\|\tilde{X}(t) - X_0\right\|^2 \le \tilde{U}_1(t) \le \left\|\tilde{A}(X_0)\right\|^2$，应用推论 B.8 可得

$$\begin{aligned} \left\|\tilde{A}(\tilde{X}(t))\right\|^2 &\le (\gamma+1)\left(\left\|\dot{\tilde{X}}(t)\right\|^2 + \frac{\gamma p}{t^{p+1}} \left\|\tilde{X}(t) - X_0\right\|^2 \cdot \frac{t^{1-p}}{p} \right) \\ &\le (\gamma+1) \left\|\tilde{A}(X_0)\right\|^2 \left(1 + \frac{T^{1-p}}{p}\right). \end{aligned}$$

(iii) $p > 1$

回顾推论 B.8 的证明，我们知道 $\tilde{U}_2(t) \le \lim_{\epsilon\to 0+} \tilde{U}_2(\epsilon) = \frac{p}{\gamma} \left\|\tilde{A}(X_0)\right\|^2$。因此我们有

$\frac{\gamma p}{t^{2p}}\left\|\tilde{X}(t)-X_0\right\|^2 \le \tilde{U}_2(t) \le \frac{p}{\gamma}\left\|\tilde{A}(X_0)\right\|^2$，应用推论 B.8 可得

$$\begin{aligned}\left\|\tilde{A}(\tilde{X}(t))\right\|^2 &\le (\gamma+1)\left(\left\|\dot{\tilde{X}}(t)\right\|^2+\frac{\gamma p}{t^{2p}}\left\|\tilde{X}(t)-X_0\right\|^2\cdot\frac{1}{p}\right)\\ &\le \frac{p(\gamma+1)}{\gamma}\left(T^{p-1}+\frac{1}{p}\right)\left\|\tilde{A}(X_0)\right\|^2.\end{aligned}$$

□

B.2.2 一般 $\mathbb{A}$的存在性证明

现在我们转向原包含关系。如前所述，我们将用一个称为 Yosida逼近的 Lipschitz 函数 $\mathbb{A}_\lambda$ 来逼近 $\mathbb{A}$。我们将定义并陈述关于 $\mathbb{A}_\lambda$ 的一些事实，并将其作为引理使用，而不加证明。感兴趣的读者可参见 [11, 第 3.1 章，定理 2]。

引理 B.10. 定义 $\mathbb{A}_\lambda:\ \mathbb{R}^n\to\mathbb{R}^n$ 为

$$\mathbb{A}_\lambda=\frac{1}{\lambda}\left(\mathbb{I}-\mathbb{J}_{\lambda\mathbb{A}}\right)=\frac{1}{\lambda}\left(\mathbb{I}-(\mathbb{I}+\lambda\mathbb{A})^{-1}\right)$$

这就是所谓的 $\mathbb{A}$ 的 Yosida逼近。以下结论成立。

(i) $\mathbb{A}_\lambda$ 是 $\frac{1}{\lambda}$-利普希茨连续且极大单调的。*(ii)* $\lim_{\lambda\to0+}\mathbb{A}_\lambda x=m(\mathbb{A}(x))$。这里，$m(\mathbb{A}(x))$ 定义为 $m(\mathbb{A}(x))=\Pi_{\mathbb{A}(x)}(0)$，即 $\mathbb{A}(x)$ 中范数最小的元素。*(iii)* $\|\mathbb{A}_\lambda(x)\|\le\|m(\mathbb{A}(x))\|$。*(iv)* $\forall x\in R^n,\ \mathbb{A}_\lambda(x)\in\mathbb{A}(\mathbb{J}_{\lambda\mathbb{A}}x)$。

现在我们可以陈述用于证明定理 2.2 存在性的命题

命题 B.11。对于 *Yosida*逼近 $\mathbb{A}_\lambda$，考虑常微分方程

$$\dot{X}_\lambda(t)=-\mathbb{A}_\lambda(X_\lambda(t))-\frac{\gamma}{t^p}(X_\lambda(t)-X_0) \tag{19}$$

带有初值条件 $X_\lambda(0)=X_0\in dom(\mathbb{A})$。根据定理 *B.2*，解唯一存在，将该解记为 X_λ。现在对于 $T>0$ 以及满足 $\lim_{n\to\infty}\lambda_n=0$ 的正数列 $\{\lambda_n\}_{n\in\mathbb{N}}$，定义解序列为$\mathcal{F}_T=\{X_{\lambda_n}:\ [0,T]\to\mathbb{R}^n\mid m\in\mathbb{N}\}$。于是存在一个子序列 $\{\lambda_{n_k}\}_{k\in\mathbb{N}}$，使得 $X_{\lambda_{n_k}}$ 在 $[0,T]$ 上一致地收敛于 (6) 的解。

由推论 B.8、推论 B.9 和引理 B.10 (iii)，以下引理显然成立。

引理 B.12。设 X_λ 为 (19) 的解。则对于所有 $T>0,\ t\in[0,T]$ 均满足以下结论。

$$\left\|\dot{X}_\lambda(t)\right\|\le M_{dot}(T)=\begin{cases}\|m(\mathbb{A}(X_0))\| & 0<p<1\\ \frac{1}{\sqrt{\gamma+1}}\|m(\mathbb{A}(X_0))\| & p=1\\ \sqrt{\frac{p}{\gamma}T^{p-1}}\|m(\mathbb{A}(X_0))\| & p>1\end{cases}$$

and

$$\|\mathbb{A}_\lambda(X_\lambda(t))\|\le M_{\mathbb{A}}(T)=\begin{cases}\sqrt{(\gamma+1)\left(1+\frac{T^{1-p}}{p}\right)}\|m(\mathbb{A}(X_0))\| & 0<p<1\\ \|m(\mathbb{A}(X_0))\| & p=1\\ \sqrt{\frac{p(\gamma+1)}{\gamma}\left(T^{p-1}+\frac{1}{p}\right)}\|m(\mathbb{A}(X_0))\| & p>1.\end{cases}$$

~ ~证明。在推论 B.8 和推论 B.9 中分别将 A 替换为 $\mathbb{A}_\lambda$，将 X 替换为 X_λ。应用引理 B.10 (iii) 中的 $\|\mathbb{A}_\lambda(X_0)\|\le\|m(\mathbb{A}(X_0))\|$，即可得证。 □

在得出收敛序列的存在性时，我们将利用 [11, 第 0.3 章定理 4]中的以下引理。为方便起见，我们在此重述该引理。

引理 B.13。 考虑一个从区间 I 到巴拿赫空间X 的绝对连续函数序列 $x_k(\cdot)$，满足

(i) $\forall t \in I$, $\{x_k(t)\}_{k\in\mathbb{N}}$ 是 X的相对紧子集

(ii) 存在一个正函数 $c(\cdot) \in L^1(I, [0,\infty))$，使得对几乎所有的 $t \in I$, $\|\dot{x}_k(t)\| \le c(t)$

那么存在一个子序列（仍记为） $x_k(\cdot)$，它收敛到一个从 I 到 X 的绝对连续函数 $x(\cdot)$，其含义是

(i) $x_k(\cdot)$ 在 I的紧子集上一致收敛到 $x(\cdot)$

(ii) $\dot{x}_k(\cdot)$ 弱地收敛到 $\dot{x}(\cdot)$ 于 $L^1(I, X)$

证明。 参见 [11, 第 0.3 章，定理 4]。 □

根据引理 B.12，我们可以立即验证所有导数的范数 X_{λ_m} 均由 $M_{dot}(T)$ 有界。因此，条件 (ii) 在 $M_{dot}(T)$ 下成立。对于条件 (i)，我们证明 $\mathcal{F}_T$ 在 $C([0,T], \mathbb{R}^n)$ 中收敛。

引理 B.14. $\mathcal{F}_T$ 是 $C([0,T], \mathbb{R}^n)$ 中的收敛序列。换言之，$\forall \epsilon > 0$，存在 $N > 0$，使得 $n, m > N$ 蕴含 上确界 $_{t\in[0,T]} \|X_{\lambda_n}(t) - X_{\lambda_m}(t)\| < \epsilon$

证明. 我们将证明 X_{λ_n} 是 $C([0,T], \mathbb{R}^n)$ 中的柯西序列。令 $\nu, \lambda > 0$。由 (19) 可知，对于 $t \in (0, T]$

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}\|X_\nu(t) - X_\lambda(t)\|^2 &= \left\langle \dot{X}_\nu(t) - \dot{X}_\lambda(t), X_\nu(t) - X_\lambda(t)\right\rangle \\ &= -\langle \mathbb{A}_\nu(X_\nu(t)) - \mathbb{A}_\lambda(X_\lambda(t)),\, X_\nu(t) - X_\lambda(t)\rangle - \frac{\gamma}{t^p}\|X_\nu(t) - X_\lambda(t)\|^2 \\ &\le -\langle \mathbb{A}_\nu(X_\nu(t)) - \mathbb{A}_\lambda(X_\lambda(t)),\, X_\nu(t) - X_\lambda(t)\rangle .\end{aligned}$$

根据预解式的定义 我们知道 $\mathbb{I} - \mathbb{J}_{\lambda\mathbb{A}} = \lambda\mathbb{A}_\lambda$。并且由引理 B.10 (iv) 可知 $\mathbb{A}_\lambda(X_\lambda(t)) \in \mathbb{A}(\mathbb{J}_{\lambda\mathbb{A}}(X_\lambda(t)))$). 因此，由单调不等式 可知

$$\begin{aligned}&-\langle \mathbb{A}_\nu(X_\nu(t)) - \mathbb{A}_\lambda(X_\lambda(t)),\, X_\nu(t) - X_\lambda(t)\rangle \\ &\quad= -\langle \mathbb{A}_\nu(X_\nu(t)) - \mathbb{A}_\lambda(X_\lambda(t)),\, \nu\mathbb{A}_\nu(X_\nu(t)) - \lambda\mathbb{A}_\lambda(X_\lambda(t))\rangle \\ &\qquad-\langle \mathbb{A}_\nu(X_\nu(t)) - \mathbb{A}_\lambda(X_\lambda(t)),\, \mathbb{J}_{\nu\mathbb{A}}(X_\nu(t)) - \mathbb{J}_{\lambda\mathbb{A}}(X_\lambda(t))\rangle \\ &\quad\le -\langle \mathbb{A}_\nu(X_\nu(t)) - \mathbb{A}_\lambda(X_\lambda(t)),\, \nu\mathbb{A}_\nu X_\nu(t) - \lambda\mathbb{A}_\lambda(X_\lambda(t))\rangle \\ &\quad= (\nu+\lambda)\langle \mathbb{A}_\nu(X_\nu(t)), \mathbb{A}_\lambda(X_\lambda(t))\rangle - \left(\nu\|\mathbb{A}_\nu X_\nu(t)\|^2 + \lambda\|\mathbb{A}_\lambda X_\lambda(t)\|^2\right).\end{aligned}$$

根据杨氏不等式

$$\begin{aligned}&(\nu+\lambda)\langle \mathbb{A}_\nu(X_\nu(t)), \mathbb{A}_\lambda(X_\lambda(t))\rangle - \left(\nu\|\mathbb{A}_\nu X_\nu(t)\|^2 + \lambda\|\mathbb{A}_\lambda X_\lambda(t)\|^2\right) \\ &\le \nu\left(\|\mathbb{A}_\nu(X_\nu(t))\|^2 + \frac{1}{4}\|\mathbb{A}_\lambda(X_\lambda(t))\|^2\right) + \lambda\left(\|\mathbb{A}_\lambda(X_\lambda(t))\|^2 + \frac{1}{4}\|\mathbb{A}_\nu(X_\nu(t))\|^2\right) \\ &\quad-\left(\nu\|\mathbb{A}_\nu X_\nu(t)\|^2 + \lambda\|\mathbb{A}_\lambda X_\lambda(t)\|^2\right) \\ &= \frac{1}{4}\left(\nu\|\mathbb{A}_\lambda(X_\lambda(t))\|^2 + \lambda\|\mathbb{A}_\nu(X_\nu(t))\|^2\right).\end{aligned}$$

现在应用引理 B.12，可得

$$\frac{d}{dt}\frac{1}{2}\|X_\nu(t) - X_\lambda(t)\|^2 \le \frac{1}{4}(\nu+\lambda)M_{\mathbb{A}}(T)^2.$$

然后对两边从 0 到 t 积分，可得

$$\|X_\nu(t) - X_\lambda(t)\|^2 \le \frac{t}{2}(\nu+\lambda)M_{\mathbb{A}}(T)^2. \tag{20}$$

现在取 $\epsilon > 0$。那么存在 $N > 0$，使得对于 $n > N$， $\lambda_n < \frac{\epsilon}{TM_{\mathbb{A}}(T)2}$ 成立。然后对于 $t \in [0,T]$， n， $m > N$，我们有

$$\|X_{\lambda_n}(t) - X_{\lambda_m}(t)\|^2 \leq \frac{t}{2}(\lambda_n + \lambda_m)M_{\mathbb{A}}(T)^2 < \epsilon.$$

因此，我们得到了所需的结果。 □

最后，我们可以证明命题 B.11，而它蕴含主定理。

命题 *B.11* 的证明。取 $T > 0$。由引理 B.14 可知， X_{λ_n} 在 $[0,T]$ 上一致收敛。将极限记为 X，即定义 $X\colon [0,T] \to \mathbf{R}^n$ 为 $X(t) = \lim_{n\to\infty} X_{\lambda_n}(t)$。由于对所有 $n \in \mathbf{N}$， $X_{\lambda_n}(0) = X_0$，我们看到 X 满足初始条件。接下来只需证明 X 几乎处处满足 (6)。

回想一下，由引理 B.12 可知， $\{X_{\lambda_n}\}$ 在 $L^\infty([0,T],\mathbf{R}^n)$ 中有界。因此可以应用引理 B.13，存在一个子序列 $\{X_{\lambda_{n_k}}\}$，它在 $L^1([0,T],\mathbf{R}^n)$ 中弱收敛至 $\dot{X}$。此外，我们有 $\dot{X} \in L^\infty([0,T],\mathbf{R}^n)$，因此 $\{\dot{X}_{\lambda_{n_k}}\}$ 在 $L^2([0,T],\mathbf{R}^n)$ 中也弱收敛至 $\dot{X}$。

对于 $\lambda > 0$，定义 $f_\lambda\colon [0,T] \to \mathbf{R}^n$ 为

$$f_\lambda(t) = \begin{cases} \frac{\gamma}{t^p}(X_\lambda(t) - X_0) & \text{if } t > 0 \\ 0 & \text{if } t = 0. \end{cases}$$

那么对于 $f\colon [0,T] \to \mathbf{R}^n$ 定义为 $f(t) = \frac{\gamma}{t^p}(X(t) - X_0)$，其中 $t > 0$ 且 $f(0) = 0$，我们有 $\lim_{k\to\infty} f_{\lambda_{n_k}}(t) = f(t)$。由于当 $t \in (0,T]$ 时，$\|f_\lambda(t)\| = \left\|\dot{X}_\lambda(t) + \mathbb{A}_{\lambda_n}(X_\lambda)(t)\right\| \leq M_{dot}(T) + M_{\mathbb{A}}(T)$，由引理 B.12，我们有

$$\|f_\lambda(t) - f(t)\|^2 \leq (\|f_\lambda(t)\| + \|f(t)\|)^2 \leq 4(M_{dot}(T) + M_{\mathbb{A}}(T))^2.$$

因此，由控制收敛定理，我们有

$$\lim_{k\to\infty} \int_0^T \left\|f_{\lambda_{n_k}}(t) - f(t)\right\|^2 dt = \int_0^T \lim_{k\to\infty} \left\|f_{\lambda_{n_k}}(t) - f(t)\right\|^2 dt = 0,$$

我们得出，$f_{\lambda_{n_k}}$ 在 $L^2([0,T],\mathbf{R}^n)$ 中强收敛到 f。

现在考虑 $F_\lambda\colon [0,T] \to \mathbf{R}^n$，定义为

$$F_\lambda(t) = \begin{cases} \dot{X}_\lambda(t) + \frac{\gamma}{t^p}(X_\lambda(t) - X_0) & \text{if } t > 0 \\ -\mathbb{A}_\lambda(X_0) & \text{if } t = 0 \end{cases}$$

注意，由于 X_λ 是常微分方程 (19) 的解，我们有

$$F_\lambda(t) = -\mathbb{A}_\lambda(X_\lambda(t)).$$

于是对于 $F\colon [0,T] \to \mathbf{R}^n$ 定义为

$$F(t) = \begin{cases} \dot{X}(t) + \frac{\gamma}{t^p}(X(t) - X_0) & \text{if } t > 0 \\ -m(\mathbb{A}(X_0)) & \text{if } t = 0, \end{cases}$$

我们看到 $\{F_{\lambda_{n_k}}\}_{k\in\mathbb{N}}$ 在 $L^2([0,T],\ \mathbf{R}^n)$ 中 F 弱地收敛。

另一方面，根据引理 B.10 (iv) 以及 $-F_{\lambda_{n_k}}(t) = \mathbb{A}_{\lambda_{n_k}}(X_{\lambda_{n_k}}(t))$ 这一事实，我们看到

$$-F_{\lambda_{n_k}}(t) \in \mathbb{A}(\mathbb{J}_{\lambda_{n_k}\mathbb{A}}(X_{\lambda_{n_k}}(t))).$$

注意，根据 $\mathbb{A}_{\lambda_n}$ 的定义以及引理 B.12，我们看到

$$\|X_{\lambda_n}(t) - \mathbb{J}_{\lambda_n\mathbb{A}}(X_{\lambda_n}(t))\| = \lambda_n \|\mathbb{A}_{\lambda_n}(X_{\lambda_n}(t))\| \leq \lambda_n M_{\mathbb{A}}(T).$$

由于 X_{λ_n} 在 $\mathcal{C}([0,T],\ \mathbf{R}^n)$ 中收敛于 X，对上述不等式取 $n \to \infty$，我们看到 $\mathbb{J}_{\lambda_n\mathbb{A}}(X_{\lambda_n})$ 也在$\mathcal{C}([0,T],\ \mathbf{R}^n)$ 中收敛于 X。

现在对于 $\mathcal{A}\colon\ L^2([0,T],\ \mathbf{R}^n) \to L^2([0,T],\ \mathbf{R}^n)$ 定义为 $(\mathcal{A}(x))(t) = \mathbb{A}(x(t))$，几乎处处成立。根据 [11, 第3.1章，命题 4]， $\mathcal{A}$ 是极大单调的，因为 $\mathbb{A}$ 是极大单调的。由于 $-F_{\lambda_k}$ 在 $L^2([0,T],\ \mathbf{R}^n)$ 中弱地收敛到 $-F$

并且 $\mathbb{J}_{\lambda_k\mathbb{A}}(X_{\lambda_k})$ 在 $L^2([0,T],\ \mathbb{R}^n)$ 中强收敛到 X，根据 [11, 第3.1章，命题 2]，我们有 $-F \in \mathcal{A}(X)$ 属于 $L^2([0,T],\ \mathbb{R}^n)$。因此，对于几乎所有 $t \in (0,\ T]$，我们有

$$-\left(\dot{X}(t) + \frac{\gamma}{t^p}(X(t) - X_0)\right) \in \mathbb{A}(X(t)).$$

将结果按 X 重新整理后，对于几乎所有的 $t \in (0, T]$，以下结论成立：

$$\dot{X}(t) \in -\mathbb{A}(X(t)) - \frac{\gamma}{t^p}(X(t) - X_0).$$

由于 $T > 0$ 是任意的，我们得出 X 对几乎所有的 $t \in (0, \infty)$ 都满足上述包含关系。 □

C 解(10)的存在性与唯一性证明

唯一性由定理 B.1 得出，因此我们只需证明存在性。存在性证明需要以下条件：

(i) 引理 B.5 中所示的包含 $\left\|\dot{\tilde{X}}\right\|^2$ 的非增函数 $\tilde{U}(t)$。(ii) 引理 B.6 中所示的 $X_\delta(t)$ 关于 $t \in [0,T]$ 的一致有界性。(iii) Lipschitz 情形的存在性证明中所示的 $\tilde{X}(0) = \lim_{t\to 0+} {}^{X_{(t)}-X_0}_{t} = \lim_{t\to 0+} \tilde{X}(t)$。(iv) 引理 B.12 中所示的 $\left\|\dot{X}_\lambda(t)\right\|$ 和 $\|\mathbb{A}_\lambda(X_\lambda(t))\|$ 关于 $t \in [0,T]$ 的一致有界性

现在我们说明，这些步骤也可以应用于 $\beta(t) = {}^{2\mu}_{e^{2\mu t}-1}$。

(i) 非增函数 $\tilde{U}(t)$，其中包含 $\left\|\dot{\tilde{X}}\right\|^2$

From

$$\dot{\tilde{X}}(t) = -\tilde{A}(\tilde{X}(t)) - \frac{2\mu}{e^{2\mu t}-1}(\tilde{X}(t) - X_0)$$

对于几乎所有的 t，我们有

$$\ddot{\tilde{X}}(t) = -\frac{d}{dt}\tilde{A}(\tilde{X}(t)) + \left(\frac{2\mu}{e^{2\mu t}-1}\right)^2 e^{2\mu t}(\tilde{X}(t) - X_0) - \frac{2\mu}{e^{2\mu t}-1}\dot{\tilde{X}}(t)$$

定义

$$\tilde{U}(t) = e^{-2\mu t}\left\|\dot{\tilde{X}}(t)\right\|^2 + \left(\frac{2\mu}{e^{2\mu t}-1}\right)^2 \left\|\tilde{X}(t) - X_0\right\|^2$$

因此，对于几乎所有 $t > 0$,

$$\begin{aligned}
\dot{\tilde{U}}(t) &= 2e^{-2\mu t}\left\langle \dot{\tilde{X}}(t), \ddot{\tilde{X}}(t)\right\rangle - 2\mu e^{-2\mu t}\left\|\dot{\tilde{X}}(t)\right\|^2 \\
&\quad + 2\left(\frac{2\mu}{e^{2\mu t}-1}\right)^2 \left\langle \dot{\tilde{X}}(t), \tilde{X}(t) - X_0\right\rangle - \left(\frac{2\mu}{e^{2\mu t}-1}\right)^3 2e^{2\mu t}\left\|\tilde{X}(t) - X_0\right\|^2 \\
&= -2e^{-2\mu t}\left\langle \dot{\tilde{X}}(t), \frac{d}{dt}\tilde{A}(\tilde{X}(t))\right\rangle + 2\left(\frac{2\mu}{e^{2\mu t}-1}\right)^2 \left\langle \dot{\tilde{X}}(t), \tilde{X}(t) - X_0\right\rangle - 2e^{-2\mu t}\frac{2\mu}{e^{2\mu t}-1}\left\|\dot{\tilde{X}}(t)\right\|^2 \\
&\quad - 2\mu e^{-2\mu t}\left\|\dot{\tilde{X}}(t)\right\|^2 + 2\left(\frac{2\mu}{e^{2\mu t}-1}\right)^2 \left\langle \dot{\tilde{X}}(t), \tilde{X}(t) - X_0\right\rangle - \left(\frac{2\mu}{e^{2\mu t}-1}\right)^3 2e^{2\mu t}\left\|\tilde{X}(t) - X_0\right\|^2 \\
&= -2\mu e^{-2\mu t}\left\|\dot{\tilde{X}}(t)\right\|^2 - 2e^{-2\mu t}\left\langle \dot{\tilde{X}}(t), \frac{d}{dt}\tilde{A}(\tilde{X}(t))\right\rangle - \frac{2\mu e^{-2\mu t}}{e^{2\mu t}-1}\left\|\dot{\tilde{X}}(t) - \frac{2\mu e^{2\mu t}}{e^{2\mu t}-1}\left(\tilde{X}(t) - X_0\right)\right\|^2 \\
&\le 0.
\end{aligned}$$

(ii) $X_\delta(t)$**的一致有界性**

根据式 (13)，我们定义 $X_\delta(t)$ 作为.....的解

$$\dot{\tilde{X}}_\delta(t) = \begin{cases} -\tilde{A}(\tilde{X}_\delta(t)) - \frac{2\mu}{e^{2\mu\delta}-1}(\tilde{X}_\delta(t) - X_0) & 0 \le t \le \delta \\ -\tilde{A}(\tilde{X}_\delta(t)) - \frac{2\mu}{e^{2\mu t}-1}(\tilde{X}_\delta(t) - X_0) & t \ge \delta \end{cases}$$

再次采用引理 B.6 中相同的论证，我们有 $\left\|\tilde{X}_\delta(\delta) - X_0\right\| \le \delta \left\|\tilde{A}(X_0)\right\|$ 且 $\left\|\dot{\tilde{X}}_\delta(\delta)\right\| \le \left\|\tilde{A}(X_0)\right\|$。现在对于$t \in [0,T]$

$$\begin{aligned} \left\|\dot{\tilde{X}}_\delta(t)\right\| \le \sqrt{e^{2\mu t}\tilde{U}(t)} &\le \sqrt{e^{2\mu t}\tilde{U}(\delta)} = e^{\mu t}\sqrt{e^{-2\mu\delta}\left\|\dot{\tilde{X}}_\delta(\delta)\right\|^2 + \left(\frac{2\mu}{e^{2\mu\delta}-1}\right)^2 \left\|\tilde{X}_\delta(\delta) - X_0\right\|^2} \\ &\le e^{\mu t}\sqrt{e^{-2\mu\delta}\left\|\tilde{A}(X_0)\right\|^2 + \left(\frac{2\mu\delta}{e^{2\mu\delta}-1}\right)^2 \left\|\tilde{A}(X_0)\right\|^2} \\ &\le e^{\mu T}\sqrt{2}\left\|\tilde{A}(X_0)\right\|. \end{aligned}$$

(iii) $\dot{\tilde{X}}(0) =$极限$_{t\to 0+}\frac{\tilde{X}(t)-X_0}{t} =$ 极限$_{t\to 0+}\dot{\tilde{X}}(t)$

定义 $C(t) := 1 - e^{-2\mu t}$。那么，对于 $\beta(t) = \frac{2\mu}{e^{2\mu t}-1}$，$\dot{C}(t) = C(t)\beta(t)$。并且由于

$$\lim_{t\to 0+}\frac{C(tv)}{C(t)} = \lim_{t\to 0+}\frac{1-e^{-2\mu tv}}{1-e^{-2\mu t}} = v,$$

用推导（15）时相同的论证，我们得到

$$\dot{\tilde{X}}(0) = \lim_{t\to 0+}\frac{\tilde{X}(t) - X_0}{t} = -\int_0^1 \lim_{t\to 0+}\left(\frac{C(tv)}{C(t)}\tilde{A}(\tilde{X}(tv))\right) dv = -\frac{1}{2}\tilde{A}(X_0)$$

现在由常微分方程 $\dot{\tilde{X}}(t) = -\tilde{A}(\tilde{X}(t)) - \frac{2\mu}{e^{2\mu t}-1}(\tilde{X}(t) - X_0)$，对两边取 $t \to 0+$ 的极限，得到

$$\lim_{t\to 0+}\dot{\tilde{X}}(t) = -\tilde{A}(\tilde{X}(0)) - \lim_{t\to 0+}\frac{2\mu t}{e^{2\mu t}-1}\frac{\tilde{X}(t) - X_0}{t} = -\frac{1}{2}\tilde{A}(X_0)$$

因此，$\lim_{t\to 0+}\frac{\tilde{X}(t)-X_0}{t} = \lim_{t\to 0+}\dot{\tilde{X}}(t)$。

(iv) $t \in [0,T]$**时** $\left\|\dot{X}_\lambda(t)\right\|$**和** $\|\mathbb{A}_\lambda(t)\|$ **的一致有界性**

回忆一下，$U(t) = e^{-2\mu t}\left\|\dot{X}_\lambda(t)\right\|^2 + \left(\frac{2\mu}{e^{2\mu t}-1}\right)^2 \|X_\lambda(t) - X_0\|^2$ 是非递增的。因此，由（iii）可得

$$\begin{aligned} e^{-2\mu t}\left\|\dot{X}_\lambda(t)\right\|^2 + \left(\frac{2\mu}{e^{2\mu t}-1}\right)^2 \|X_\lambda(t) - X_0\|^2 &\le \lim_{t\to 0+}\left(e^{-2\mu t}\left\|\dot{X}_\lambda(t)\right\|^2 + \left(\frac{2\mu}{e^{2\mu t}-1}\right)^2 \|X_\lambda(t) - X_0\|^2\right) \\ &= \frac{1}{4}\|\mathbb{A}_\lambda(X_0)\|^2 + \frac{1}{4}\|\mathbb{A}_\lambda(X_0)\|^2 = \frac{1}{2}\|\mathbb{A}_\lambda(X_0)\|^2 \end{aligned}$$

因此，由引理 B.10 (iii) 可得

$$e^{-2\mu t}\left\|\dot{X}_\lambda(t)\right\|^2 \le \frac{1}{2}\|\mathbb{A}_\lambda(X_0)\|^2 \le \frac{1}{2}\|m(\mathbb{A}(X_0))\|^2 \quad \Longrightarrow \quad \left\|\dot{X}_\lambda(t)\right\| \le \frac{e^{\mu T}}{\sqrt{2}}\|m(\mathbb{A}(X_0))\|, \tag{21}$$

根据杨氏不等式

$$\begin{aligned}\|\mathbb{A}_\lambda(X_\lambda(t))\|^2 &\le \left\|\dot{X}_\lambda(t)+\frac{2\mu}{e^{2\mu t}-1}(X_\lambda(t)-X_0)\right\|^2\\ &= 2\left(\left\|\dot{X}_\lambda(t)\right\|^2+\left(\frac{2\mu}{e^{2\mu t}-1}\right)^2\|X_\lambda(t)-X_0\|^2\right)\\ &\le 2\left(\frac{e^{2\mu T}}{2}\|\mathbb{A}_\lambda(X_0)\|^2+\frac{1}{2}\|\mathbb{A}_\lambda(X_0)\|^2\right)=(e^{2\mu T}+1)\|\mathbb{A}_\lambda(X_0)\|^2\le(e^{2\mu T}+1)\|m(\mathbb{A}(X_0))\|^2.\end{aligned}$$

因此

$$\|\mathbb{A}_\lambda(X_\lambda(t))\|\le\sqrt{e^{2\mu T}+1}\,\|m(\mathbb{A}(X_0))\|. \tag{22}$$

D 锚定常微分方程(4)的推导中省略的证明

D.1 定理 2.1 证明的准备工作

我们首先以引理的形式给出轨迹的有界性。正如引理 2.3 后的讨论所述，轨迹的有界性是引理 2.3 的直接推论。然而，为了处理更一般一些的情形，我们给出一个证明，其论证与引理 2.3 的证明类似。请注意，以下引理的证明论证对于满足定理 7.1 中假设的微分方程 (11) 的解同样成立。

引理 D.1. （解的有界性）假设 $X(\cdot)$ 是微分包含(3)的解。那么对于所有 $X_\star\in\mathrm{Zer}\mathbb{A}$,$t\in[0,\infty)$，以下结论成立。

$$\|X(t)-X_\star\|\le\|X_0-X_\star\|.$$

And so, $\|X(t)-X_0\|\le 2\|X_0-X_\star\|$.

证明。对于 $t=0$，这是显然的，因此我们可以假设 $t>0$。

Take $X_\star\in\mathrm{Zer}\mathbb{A}$。由 $\mathbb{A}$ 的单调性和杨氏不等式，我们得到以下不等式 y.

$$\begin{aligned}\frac{d}{dt}\|X(t)-X_\star\|^2=2\left\langle\dot{X}(t),X(t)-X_\star\right\rangle&=-2\left\langle\tilde{\mathbb{A}}(X(t))+\beta(t)(X(t)-X_0),X(t)-X_\star\right\rangle\\ &=-2\left\langle\tilde{\mathbb{A}}(X(t))+\beta(t)(X(t)-X_\star)-\beta(t)(X_0-X_\star),X(t)-X_\star\right\rangle\\ &\le-2\beta(t)\|X(t)-X_\star\|^2+2\beta(t)\langle X_0-X_\star,X(t)-X_\star\rangle\\ &\le-2\beta(t)\|X(t)-X_\star\|^2+\beta(t)\left(\|X_0-X_\star\|^2+\|X(t)-X_\star\|^2\right)\\ &=-\beta(t)\|X(t)-X_\star\|^2+\beta(t)\|X_0-X_\star\|^2.\end{aligned}$$

现在再次定义 $C(t)=e^{\int_v^t\beta_{(s)}ds}$，其中 $v>0$ 为某个值，则我们看到 $\dot{C}(t)=C(t)\beta(t)$。将 $-\beta(t)\|X(t)-X_\star\|^2$ 移至等式左侧，并将两边乘以 $C(t)$，得到

$$\begin{aligned}\frac{d}{dt}\left(C(t)\|X(t)-X_\star\|^2\right)&=C(t)\frac{d}{dt}\|X(t)-X_\star\|^2+C(t)\beta(t)\|X(t)-X_\star\|^2\\ &\le C(t)\beta(t)\|X_0-X_\star\|^2=\frac{d}{dt}C(t)\|X_0-X_\star\|^2.\end{aligned}$$

对两边从 $\epsilon>0$ 到 t 积分，得到

$$C(t)\|X(t)-X_\star\|^2-C(\epsilon)\|X(\epsilon)-X_\star\|^2\le C(t)\|X_0-X_\star\|^2-C(\epsilon)\|X_0-X_\star\|^2.$$

当 $\beta>0$ 时，我们有 C 非负且单调不减，$\lim_{\epsilon\to0+}C(\epsilon)$ 存在。对两边取极限 $\epsilon\to0+$，并将两边除以 $C(t)$，我们得出

$$\|X(t)-X_\star\|^2\le\|X_0-X_\star\|^2.$$

后一个结论可直接由三角不等式得出，

$$\|X(t)-X_0\|\le\|X(t)-X_\star\|+\|X_\star-X_0\|\le2\|X_0-X_\star\|.$$

□

以下引理表明，APPM 是 Halpern 方法的一个实例。由归纳法可立即得出这一结论，但鉴于其重要性，我们将其列为引理。

引理 D.2. 考虑如下定义的方法

$$\begin{aligned} x^k &= \mathbb{J}_{\mathbb{A}} y^{k-1} \\ y^k &= (1-\beta_k)\left(2x^k - y^{k-1}\right) + \beta_k x^0, \end{aligned}$$

对于 $k = 1, 2, \ldots,$ 初始条件为 $y^0 = x^0$。那么，对于反射预解算子 $\mathrm{R}_{\mathbb{A}}$，定义为 $\mathrm{R}_{\mathbb{A}} = 2\mathbb{J}_{\mathbb{A}} - \mathbb{I}$，上述方法等价于

$$\tilde{y}^k = \beta_k \tilde{y}^0 + (1-\beta_k)\,\mathbb{T}\tilde{y}^{k-1}$$

当 $\mathbb{T} = \mathrm{R}_{\mathbb{A}}, \tilde{y}^0 = y^0$ 时。此处等价是指 $\tilde{y}^k = y^k$ 对 $k = 0, 1, \ldots$ 均成立。

证明。 归纳证明。根据假设， $\tilde{y}^0 = y^0$ ，该命题对 $k = 0$ 成立。现在假设 $y^k = \tilde{y}^k$ 对 $k \in \mathrm{N}$ 成立，那么

$$\begin{aligned} \tilde{y}^{k+1} = (1-\beta_{k+1})\mathbb{R}_{\mathbb{A}}\tilde{y}^k + \beta_{k+1}\tilde{y}^0 &= (1-\beta_{k+1})\left(2\mathbb{J}_{\mathbb{A}} - \mathbb{I}\right)\tilde{y}^k + \beta_{k+1}\tilde{y}^0 \\ &= (1-\beta_{k+1})\left(2\mathbb{J}_{\mathbb{A}} - \mathbb{I}\right)y^k + \beta_{k+1}y^0 = (1-\beta_{k+1})\left(2x^{k+1} - y^k\right) + \beta_{k+1}y^0 = y^{k+1}. \end{aligned}$$

因此 $\tilde{y}^{k+1} = y^{k+1}$，该命题对 $k+1$ 成立。由归纳法可得所需结果。 □

D.2 定理 2.1 的证明

设 $\mathbb{A}\colon \mathrm{R}^n \to \mathrm{R}^n$ 是一个极大单调算子，并且 $h,\ \lambda,\ \delta > 0$。同样，记 $\mathbb{A}_\lambda = \frac{1}{\lambda}(\mathbb{I} - \mathbb{A}_{\lambda\mathbb{A}}) = \frac{1}{\lambda}(\mathbb{I} - (\mathbb{I} + \lambda\mathbb{A})^{-1})$。由于证明中会出现多种项，我们先整理这些项和记号。

- X： 微分包含的解,

$$\dot{X} \in -\mathbb{A}(X) - \frac{1}{t}(X - X_0).$$

- X_λ： 微分方程的解,

$$\dot{X}_\lambda = -\mathbb{A}_\lambda(X_\lambda) - \frac{1}{t}(X_\lambda - X_0).$$

- $X_{\lambda,\ \delta}$： 近似微分方程的解,

$$\dot{X}_{\lambda,\delta}(t) = \begin{cases} -\mathbb{A}_\lambda(X_{\lambda,\delta})(t) - \frac{1}{\delta}(X_{\lambda,\delta}(t) - X_0) & 0 \le t < \delta \\ -\mathbb{A}_\lambda(X_{\lambda,\delta})(t) - \frac{1}{t}(X_{\lambda,\delta}(t) - X_0) & t \ge \delta. \end{cases} \tag{23}$$

- $X^k_{\lambda,\ \delta}$： 通过对常微分方程 (23) 进行欧拉离散化得到的序列,

$$X^{k+1}_{\lambda,\delta} = \begin{cases} X^k_\delta - \left(2h\mathbb{A}_\lambda(X^k_{\lambda,\delta}) + \frac{2h}{\delta}(X^k_{\lambda,\delta} - X_0)\right) & 0 \le k < \frac{\delta}{2h} \\ X^k_\delta - \left(2h\mathbb{A}_\lambda(X^k_{\lambda,\delta}) + \frac{1}{k}(X^k_{\lambda,\delta} - X_0)\right) & k \ge \frac{\delta}{2h}. \end{cases}$$

- $x^k_{h,\ \lambda}$： 由算子为 $h\mathbb{A}_\lambda$ 的 APPM 得到的序列，即

$$\begin{aligned} x^k_{h,\lambda} &= \mathbb{J}_{h\mathbb{A}_\lambda} y^{k-1}_{h,\lambda} \\ y^k_{h,\lambda} &= \frac{k}{k+1}(2x^k_{h,\lambda} - y^{k-1}_{h,\lambda}) + \frac{1}{k+1}X_0. \end{aligned}$$

- x^k_h： 由算子 $h\mathbb{A}$对应的 APPM 得到的序列，即

$$\begin{aligned} x^k_h &= \mathbb{J}_{h\mathbb{A}} y^{k-1}_h \\ y^k_h &= \frac{k}{k+1}(2x^k_h - y^{k-1}_h) + \frac{1}{k+1}X_0. \end{aligned}$$

我们要证明，对于固定的 $T>0$,

$$\lim_{h\to 0+}\sup_{0\le k<\frac{T}{2h}}\left\|x_h^k-X(2hk)\right\|=0.$$

等价地，对于固定的 $T>0$，我们可以证明：对于满足 $\{h_n\}_{n\in\mathbb{N}}$ 且收敛于 0 的每个 $h_n>0$，

$$\lim_{n\to\infty}\sup_{0\le k<\frac{T}{2h_n}}\left\|x_{h_n}^k-X(2h_nk)\right\|=0.$$

我们将通过考虑不等式来证明这一点

$$\begin{aligned}\left\|x_h^k-X(2hk)\right\|&\le\|X(2hk)-X_\lambda(2hk)\|+\|X_\lambda(2hk)-X_{\lambda,\delta}(2hk)\|\\&\quad+\left\|X_{\lambda,\delta}(2hk)-X_{\lambda,\delta}^k\right\|+\left\|X_{\lambda,\delta}^k-x_\lambda^k\right\|+\left\|x_{h,\lambda}^k-x_h^k\right\|=:S\left(h,\lambda,\delta,k\right).\end{aligned}\tag{24}$$

我们的目标是证明：对于满足 $\{h_n\}_{n\in\mathbb{N}}$ 且收敛于 0 的每个 $h_n>0$，存在一个序列 $\{(\delta_n,\ \lambda_n)\}_{n\in\mathbb{N}}$，使得 $\lim_{n\to\infty}\sup_{0\le k<\frac{T}{2h_n}}S(h_n,\ \lambda_n,\ \delta_n,\ k)=0$，从而

$$\lim_{n\to\infty}\sup_{0\le k<\frac{T}{2h_n}}\left\|x_{h_n}^k-X(2h_nk)\right\|\le\lim_{n\to\infty}\sup_{0\le k<\frac{T}{2h_n}}S\left(h_n,\lambda_n,\delta_n,k\right)=0.$$

为阐明我们的目标，我们需要根据 $\{h_n\}_{n\in\mathbb{N}}$ 找到合适的 $\{(\delta_n,\lambda_n)\}_{n\in\mathbb{N}}$。由于 $\{h_n\}_{n\in\mathbb{N}}$ 已确定，$\left\|x_{h_n}^k-X(2h_nk)\right\|$ 也随之固定，不会因选择 $\{(\delta_n,\lambda_n)\}_{n\in\mathbb{N}}$ 而改变。但是，如果我们找到一个 $\{(\delta_n,\lambda_n)\}_{n\in\mathbb{N}}$，使得 $S(h_n,\lambda_n,\delta_n,k)$ 很小，那么由于（24）对 δ、λ 的任意选择都成立，适当地选择 $\{(\delta_n,\lambda_n)\}_{n\in\mathbb{N}}$ 就能保证 $\left\|x_{h_n}^k-X(2h_nk)\right\|$ 很小。因此，为了找到这样的 $\{(\delta_n,\lambda_n)\}_{n\in\mathbb{N}}$，我们将考察 S 中的每一项，以找出所需的条件。

引理 D.3。对于 $h,\lambda,\delta>0$，以下结论成立。

(i) $\|X(2hk)-X_\lambda(2hk)\|=\mathcal{O}\left(\sqrt{\lambda}\right)$

(ii) $\|X_\lambda(2hk)-X_{\lambda,\delta}(2hk)\|=\mathcal{O}\left(\delta\right)$

(iii) $\left\|x_{h,\lambda}^k-x_h^k\right\|=O(\lambda)$.

对于 $L_{\lambda,\ \delta}=$最大值 $\left\{\frac{1}{\lambda},\ \frac{\sqrt{2}}{\lambda}\|m(\mathbb{A}(X_0))\|,\ \frac{1}{\delta},\ \frac{4\sqrt{2}}{\delta}\|m(\mathbb{A}(X_0))\|\right\}$,

(iv) $\left\|X_{\lambda,\delta}\left(2hk\right)-X_{\lambda,\delta}^k\right\|=\mathcal{O}\left(he^{2L_{\lambda,\delta}T}\right)$.

此外，如果 $0<\frac{h}{\lambda}\prec\frac{1}{2}$，

$$\begin{aligned}\text{(v)}\ \left\|X_{\lambda,\delta}^k-x_\lambda^k\right\|&=\mathcal{O}\left(h\right)+\mathcal{O}\left(h^2L_{\lambda,\delta}e^{2L_{\lambda,\delta}T}\right)\\&\quad+3^{\frac{T}{\lambda}}\left(\mathcal{O}\left(\tfrac{h}{\lambda}\right)+\mathcal{O}\left(\tfrac{h}{\lambda}e^{2L_{\lambda,\delta}T}\right)+\mathcal{O}(h)+\mathcal{O}\left(\tfrac{h^2}{\lambda}e^{2L_{\lambda,\delta}T}\right)+\mathcal{O}\left(\tfrac{e^{2\delta L_{\lambda,\delta}}}{L_{\lambda,\delta}}\right)+\mathcal{O}\left(\delta\right)\right).\end{aligned}$$

我们在下一小节证明这个引理；这里假设引理成立，并证明定理 2.1。假设上述引理成立。计算过程虽然繁杂，但策略很简单：平衡各项 h，δ，λ 趋于零的速度，使上述各项趋于零。上述引理促使我们将序列取为

$$\begin{aligned}\delta_n&=\min\left\{\frac{h_n}{2},\frac{8MT}{\log_3\left(\frac{1}{h_n}\right)}\right\}\\\lambda_n&=\max\left\{\delta_n,\frac{\sqrt{2}\left\|m(\mathbb{A}(X_0))\right\|\delta_n}{M},\frac{2T}{\log_3\left(\frac{1}{\delta_n}\right)}\right\}.\end{aligned}$$

其中 $M=$最大值 $\left\{1,4\sqrt{2}\|m(\mathbb{A}(X_0))\|\right\}$。当 $\lim_{n\to\infty}h_n\to 0$ 且 $h_n>0$ 时，我们可以轻易验证 $\lim_{n\to\infty}\delta_n=0$ 和 $\lim_{n\to\infty}\lambda_n=0$。因此，情形 (i)、(ii)、(iii) 趋于零点。

现在根据 λ_n 的定义，我们有:

$$\lambda_n \geq \max\left\{\frac{\sqrt{2}\left\|m(\mathbb{A}(X_0))\right\|\delta_n}{M}, \delta_n\right\} \quad \Longrightarrow \quad \frac{M}{\delta_n} \geq \max\left\{\frac{\sqrt{2}\left\|m(\mathbb{A}(X_0))\right\|}{\lambda_n}, \frac{1}{\lambda_n}\right\} \quad \Longrightarrow \quad L_{\lambda_n,\delta_n} = \frac{M}{\delta_n}.$$

Thus

$$\begin{aligned}
e^{2L_{\lambda,\delta}T} \leq e^{\frac{2MT}{\delta_n}} \leq 3^{\frac{2MT}{\delta_n}} \leq 3^{\frac{1}{4}\log_3\left(\frac{1}{h_n}\right)} &= \frac{1}{h_n^{1/4}} \\
3^{\frac{T}{\lambda_n}} \leq 3^{\frac{MT}{\delta_n}} \leq 3^{MT\frac{\log_3\left(\frac{1}{h_n}\right)}{8MT}} &= \frac{1}{h_n^{1/8}} \\
\frac{e^{2\delta_n L_{\lambda_n,\delta_n}}}{L_{\lambda_n,\delta_n}} = e^{2M}\frac{\delta_n}{M} &= \mathcal{O}\left(\delta_n\right).
\end{aligned}$$

因此，当 $\lim_{n\to\infty} h_n \to 0$ 时，

$$\begin{aligned}
h_n e^{2L_{\lambda_n,\delta_n}T} &\leq h_n^{3/4} \to 0 \\
h_n^2 L_{\lambda_n,\delta_n} e^{2L_{\lambda_n,\delta_n}T} = \frac{h_n^2}{2T}(2L_{\lambda_n,\delta_n}T)e^{2L_{\lambda_n,\delta_n}T} &\leq \frac{h_n^2}{2T}e^{4L_{\lambda_n,\delta_n}T} \leq \frac{h_n^{3/2}}{2T} \to 0 \\
3^{\frac{T}{\lambda_n}}h_n &= h_n^{7/8} \to 0 \\
3^{\frac{T}{\lambda_n}}\frac{h_n}{\lambda_n} = 3^{\frac{T}{\lambda_n}}\frac{T}{\lambda_n}\frac{h_n}{T} \leq 3^{\frac{2T}{\lambda_n}}\frac{h_n}{T} &\leq \frac{h_n^{3/4}}{T} \to 0 \\
3^{\frac{T}{\lambda_n}}\frac{h_n}{\lambda_n}e^{2L_{\lambda_n,\delta_n}T} &\leq \frac{h_n^{3/4}}{T}e^{2L_{\lambda_n,\delta_n}T} \leq \frac{h_n^{1/2}}{T} \to 0 \\
3^{\frac{T}{\lambda_n}}\frac{h_n^2}{\lambda_n}e^{2L_{\lambda_n,\delta_n}T} &\leq \frac{h_n^{3/2}}{T} \to 0 \\
3^{\frac{T}{\lambda_n}}\delta_n \leq 3^{T\frac{\log_3(\delta_n)}{2T}}\delta_n &= \delta_n^{1/2} \to 0.
\end{aligned}$$

由于 $\lim_{n\to\infty} h_n = 0$，不失一般性，我们可以假设 $h_n < 1$。由于 $h_n > 0$，我们有 λ_n 定义良好，并满足引理 D.3 的条件。因此，情形（iv）和（v）的项也趋于零。因此，我们有 $\lim_{n\to\infty}\sup_{0\leq k<\frac{T}{2h_n}} S(h_n,\ \lambda_n,\ \delta_n,\ k) = 0$。由于 $\{h_n\}_{n\in\mathbb{N}}$ 是任意的，我们得到所需结果。

D.2.1 引理 D.3 的情形（i）、（ii）、（iii）、（iv）的证明

由于情形（v）的证明需要大量工作，我们将在另一个小节中给出，而这里我们给出从（i）到（iv）各情形的证明。

(i) $\|X(2hk) - X_\lambda(2hk)\| = \mathcal{O}\left(\sqrt{\lambda}\right)$

这是引理 B.14 的结果。考虑带有 $p = 1$的 (20)，取极限 $\nu \to 0$，我们知道

$$\sup_{t\in[0,T]}\|X(t) - X_\lambda(t)\| \leq \sqrt{\frac{\lambda T}{2}}\left\|m(\mathbb{A}(X_0))\right\| = \mathcal{O}\left(\sqrt{\lambda}\right).$$

(ii) $\|X_\lambda(2hk) - X_{\lambda,\delta}(2hk)\| = \mathcal{O}(\delta)$

这是命题 B.7 的结果。考虑将 (14) 中的 $\gamma = p = 1$ 用于引理 B.6，并取极限 $\nu \to 0$。然后应用引理 B.10 (iii)，可得

$$\sup_{t\in[0,T]}\|X_\lambda(t) - X_{\lambda,\delta}(t)\| \leq 2\sqrt{2}\delta\left\|\mathbb{A}_\lambda(X_0)\right\| \leq 2\sqrt{2}\delta\left\|m(\mathbb{A}(X_0))\right\| = \mathcal{O}\left(\delta\right).$$

(iii) $\left\|x_{h,\ \lambda}^{k}-x_{h}^{k}\right\|=O(\lambda)$ W

e fi第一个 s how show a 关于 Yosida逼近和预解式的一般事实。由 [15, 命题 23.7 (iv)] ，我们有 e

$$\begin{aligned}\|\mathbb{J}_{h\mathbb{A}}(x)-\mathbb{J}_{h\mathbb{A}_\lambda}(x)\| &= \left\|\mathbb{J}_{h\mathbb{A}}(x)-\left(\mathbb{I}+\frac{1}{1+\lambda}\left(\mathbb{J}_{(1+\lambda)h\mathbb{A}}-\mathbb{I}\right)\right)(x)\right\| \\ &= \left\|\mathbb{J}_{h\mathbb{A}}(x)-\left(\frac{\lambda}{1+\lambda}x+\frac{1}{1+\lambda}\mathbb{J}_{(1+\lambda)h\mathbb{A}}(x)\right)\right\| \\ &\leq \frac{1}{1+\lambda}\left\|\mathbb{J}_{h\mathbb{A}}(x)-\mathbb{J}_{(1+\lambda)h\mathbb{A}}(x)\right\|+\frac{\lambda}{1+\lambda}\left\|x-\mathbb{J}_{h\mathbb{A}}(x)\right\|\end{aligned}$$

由 [15, 命题 23.31 (iii)],，我们有

$$\left\|\mathbb{J}_{h\mathbb{A}}(x)-\mathbb{J}_{(1+\lambda)h\mathbb{A}}(x)\right\|\leq\lambda\left\|\mathbb{J}_{h\mathbb{A}}(x)-x\right\|.$$

结合这两个事实，我们得到

$$\|\mathbb{J}_{h\mathbb{A}}(x)-\mathbb{J}_{h\mathbb{A}_\lambda}(x)\|\leq\frac{2\lambda}{1+\lambda}\left\|x-\mathbb{J}_{h\mathbb{A}}(x)\right\|.$$

由引理 D.2，我们知道 (5) 中 y^k 序列的迭代等价于以下序列

$$y^{k+1}=\frac{1}{k+1}X_0+\frac{k}{k+1}\left(2\mathbb{J}_{h\mathbb{A}}-\mathbb{I}\right)(y^k).$$

利用这种交替形式，我们有

$$\begin{aligned}\left\|y_h^{k+1}-y_{h,\lambda}^{k+1}\right\| &= \frac{k}{k+1}\left\|(2\mathbb{J}_{h\mathbb{A}}-\mathbb{I})(y_h^k)-(2\mathbb{J}_{h\mathbb{A}_\lambda}-\mathbb{I})(y_{h,\lambda}^k)\right\| \\ &\leq \frac{k}{k+1}\left(\left\|(2\mathbb{J}_{h\mathbb{A}}-\mathbb{I})(y_h^k)-(2\mathbb{J}_{h\mathbb{A}_\lambda}-\mathbb{I})(y_h^k)\right\|+\left\|\mathbb{R}_{h\mathbb{A}_\lambda}(y_h^k)-\mathbb{R}_{h\mathbb{A}_\lambda}(y_{h,\lambda}^k)\right\|\right) \\ &\leq \frac{k}{k+1}\left(2\left\|\mathbb{J}_{h\mathbb{A}}(y_h^k)-\mathbb{J}_{h\mathbb{A}_\lambda}(y_h^k)\right\|+\left\|y_h^k-y_{h,\lambda}^k\right\|\right) \\ &\leq \frac{k}{k+1}\left(\frac{4\lambda}{1+\lambda}\left\|y_h^k-\mathbb{J}_{h\mathbb{A}}(y_h^k)\right\|+\left\|y_h^k-y_{h,\lambda}^k\right\|\right) \\ &\leq \frac{k}{k+1}\left(\frac{4\lambda}{1+\lambda}\frac{\|X_0-X_\star\|}{k}+\left\|y_h^k-y_{h,\lambda}^k\right\|\right).\end{aligned}$$

第一个不等式来自三角不等式。第二个不等式来自反射预解算子 $\mathrm{R}_{\mathbb{A}}=2\mathbb{J}_{\mathbb{A}}-\mathbb{I}$, [15, 推论 23.11]的非扩张性。第三个不等式来自前面给出的不等式。最后一个不等式来自 APPM[35, 定理 4.1], $\left\|y_h^k-\mathbb{J}_{h\mathbb{A}}(y_h^k)\right\|\leq\frac{\|X_0-X_\star\|}{k}$ 的收敛速率。现在将两边乘以 $k+1$，并从 0 到 k 求和，可得

$$(k+1)\left\|y_h^{k+1}-y_{h,\lambda}^{k+1}\right\|\leq\sum_{i=0}^{k}\frac{4\lambda}{1+\lambda}\left\|X_0-X_\star\right\|=(k+1)\frac{4\lambda}{1+\lambda}\left\|X_0-X_\star\right\|.$$

最后，根据 APPM 中 x^k 与 y^k 之间的关系，我们得到

$$\begin{aligned}\left\|x_h^{k+1}-x_{h,\lambda}^{k+1}\right\| &= \left\|\mathbb{J}_{h\mathbb{A}}(y_h^{k+1})-\mathbb{J}_{h\mathbb{A}_\lambda}(y_{h,\lambda}^{k+1})\right\| \\ &\leq \left\|\mathbb{J}_{h\mathbb{A}_\lambda}(y_h^{k+1})-\mathbb{J}_{h\mathbb{A}_\lambda}(y_{h,\lambda}^{k+1})\right\|+\left\|\mathbb{J}_{h\mathbb{A}}(y_h^{k+1})-\mathbb{J}_{h\mathbb{A}_\lambda}(y_h^{k+1})\right\| \\ &\leq \left\|y_h^{k+1}-y_{h,\lambda}^{k+1}\right\|+\frac{2\lambda}{1+\lambda}\left\|y_h^{k+1}-\mathbb{J}_{h\mathbb{A}}(y_h^{k+1})\right\| \\ &\leq \frac{4\lambda}{1+\lambda}\left\|X_0-X_\star\right\|+\frac{2\lambda}{1+\lambda}\frac{\|X_0-X_\star\|}{k+1} \\ &= \left(2+\frac{1}{k+1}\right)\frac{2\lambda}{1+\lambda}\left\|X_0-X_\star\right\|=O(\lambda).\end{aligned}$$

(iv) $\left\|X_{\lambda,\delta}(2hk) - X^k_{\lambda,\delta}\right\| = \mathcal{O}\left(he^{2L_{\lambda,\delta}T}\right)$

由式（23），我们可以将 $X_{\lambda,\delta}$ 视为如下定义的函数 $F\colon \mathbb{R}^n \times [0,\infty) \to \mathbb{R}^n$

$$F(X,t) = \begin{cases} -\mathbb{A}_\lambda(X) - \frac{1}{\delta}(X - X_0) & 0 \le t < \delta \\ -\mathbb{A}_\lambda(X) - \frac{1}{t}(X - X_0) & t > \delta. \end{cases} \tag{25}$$

注意，F关于 X 是 $2\max\left\{\frac{1}{\lambda},\frac{1}{\delta}\right\}$-Lipschitz 连续的。

为方便起见，记 $\alpha = 2h$。定义 $\epsilon_k := X_{\lambda,\delta}(\alpha k) - X^k_{\lambda,\delta}$。根据欧拉离散化的定义和微积分基本定理，我们有如下结果

$$\begin{aligned} X^{k+1}_{\lambda,\delta} &= X^k_{\lambda,\delta} + \alpha F(X^k_{\lambda,\delta}, \alpha k) \\ X_{\lambda,\delta}(\alpha(k+1)) &= X_{\lambda,\delta}(\alpha k) + \int_{\alpha k}^{\alpha(k+1)} \dot{X}_{\lambda,\delta}(t)dt \\ &= X_{\lambda,\delta}(\alpha k) + \int_{\alpha k}^{\alpha(k+1)} \left(\dot{X}_{\lambda,\delta}(\alpha k)) + \int_{\alpha k}^{t} \ddot{X}_{\lambda,\delta}(s)ds\right) dt \\ &= X_{\lambda,\delta}(\alpha k) + \alpha F(X_{\lambda,\delta}(\alpha k), \alpha k) + \int_{\alpha k}^{\alpha(k+1)} \int_{\alpha k}^{t} \ddot{X}_{\lambda,\delta}(s)\, ds\, dt. \end{aligned}$$

由引理 B.4 可知，$F(X_{\lambda,\delta}, t) = X_{\lambda,\delta}(t)$ 关于 t 利普希茨连续，因此 $X_{\lambda,\delta}$ 几乎处处有定义，且微积分基本定理成立。因此我们有

$$\begin{aligned} \epsilon_{k+1} &= X_{\lambda,\delta}(\alpha(k+1)) - X^{k+1}_{\lambda,\delta} \\ &= X_{\lambda,\delta}(\alpha k) - X^k_{\lambda,\delta} + \alpha\left(F(X_{\lambda,\delta}(\alpha k), \alpha k) - F(X^k_{\lambda,\delta}, \alpha k)\right) + \int_{\alpha k}^{\alpha(k+1)} \int_{\alpha k}^{t} \ddot{X}_{\lambda,\delta}(s)\, ds\, dt \end{aligned}$$

由于 F关于第一个变量是 $2\max\left\{\frac{1}{\lambda},\frac{1}{\delta}\right\}$-Lipschitz 连续的，我们有

$$\begin{aligned} \|\epsilon_{k+1}\| &\le \left\|X_{\lambda,\delta}(\alpha k) - X^k_{\lambda,\delta}\right\| + \alpha\left\|F(X_{\lambda,\delta}(\alpha k), \alpha k) - F(X^k_{\lambda,\delta}, \alpha k)\right\| + \int_{\alpha k}^{\alpha(k+1)} \int_{\alpha k}^{t} \left\|\ddot{X}_{\lambda,\delta}(s)\right\| ds\, dt \\ &\le \left(1 + 2\alpha \max\left\{\frac{1}{\lambda}, \frac{1}{\delta}\right\}\right) \|\epsilon_k\| + \int_{\alpha k}^{\alpha(k+1)} \int_{\alpha k}^{\alpha(k+1)} \left\|\ddot{X}_{\lambda,\delta}(s)\right\| ds\, dt. \end{aligned}$$

现在我们注意到 $\left\|\ddot{X}_{\lambda,\delta}\right\|$ 有界。对 $\dot{X}_{\lambda,\delta}$ 求导可得，

$$\ddot{X}_{\lambda,\delta} = -\frac{d}{dt}\mathbb{A}_\lambda(X_{\lambda,\delta}) + \frac{1}{t^2}\left(\dot{X}_{\lambda,\delta} - X_0\right) - \frac{1}{t}\dot{X}_{\lambda,\delta} = -\frac{d}{dt}\mathbb{A}_\lambda(X_{\lambda,\delta}) + \frac{1}{t}\left(-\mathbb{A}_\lambda(X_{\lambda,\delta}) - \dot{X}_{\lambda,\delta}\right) - \frac{1}{t}\dot{X}_{\lambda,\delta}$$

对于 $t > \delta$，几乎处处的 t 都有

$$\ddot{X}_{\lambda,\delta} = \begin{cases} -\frac{d}{dt}\mathbb{A}_\lambda(X_{\lambda,\delta}) + \frac{1}{\delta}\dot{X}_{\lambda,\delta} & 0 \le t < \delta \\ -\frac{d}{dt}\mathbb{A}_\lambda(X_{\lambda,\delta}) - \frac{1}{t}\mathbb{A}_\lambda(X_{\lambda,\delta}) - \frac{2}{t}\dot{X}_{\lambda,\delta} & t \ge \delta \end{cases}.$$

将 $\gamma = 1, p = 1$ 应用于引理 B.6，并由引理 B.10 (iv) 可得

$$\left\|\dot{X}_{\lambda,\delta}(t)\right\| \le \sqrt{2}\left\|\mathbb{A}_\lambda(X_0)\right\| \le \sqrt{2}\left\|m(\mathbb{A}(X_0))\right\|,$$

因此

$$\begin{aligned} \|\mathbb{A}_\lambda(X_{\lambda,\delta}(t))\| &= \left\|\dot{X}_{\lambda,\delta}(t) + \frac{1}{\max\{\delta, t\}}(X_{\lambda,\delta}(t) - X_0)\right\| \\ &\le \left\|\dot{X}_{\lambda,\delta}(t)\right\| + \frac{1}{\max\{\delta, t\}}\|X_{\lambda,\delta}(t) - X_0\| \\ &\le \left\|\dot{X}_{\lambda,\delta}(t)\right\| + \frac{1}{\max\{\delta, t\}}\int_0^t \left\|\dot{X}_{\lambda,\delta}(s)\right\| ds \le 2\sqrt{2}\left\|m(\mathbb{A}(X_0))\right\|. \end{aligned}$$

并且由于 $\mathbb{A}_\lambda$ 是 $\frac{1}{\lambda}$-Lipschitz 函数，我们知道 $\mathbb{A}_\lambda \circ X_{\lambda,\delta}$ 是 $\frac{1}{\lambda}\sqrt{2}\left\|m(\mathbb{A}(X_0))\right\|$-Lispchitz 函数，因此对于几乎所有 t,

$$\left\|\frac{d}{dt}\mathbb{A}_\lambda(X_{\lambda,\delta}(t))\right\| \leq \frac{\sqrt{2}\left\|m(\mathbb{A}(X_0))\right\|}{\lambda}.$$

利用这些事实，我们有

$$\begin{aligned}\left\|\ddot{X}_{\lambda,\delta}(t)\right\| &\leq \left\|\frac{d}{dt}\mathbb{A}_\lambda(X_{\lambda,\delta}(t))\right\| + \frac{1}{\delta}\left\|\mathbb{A}_\lambda(X_{\lambda,\delta}(t))\right\| + \frac{2}{\delta}\left\|\dot{X}_{\lambda,\delta}(t)\right\| \\ &\leq \frac{\sqrt{2}\left\|m(\mathbb{A}(X_0))\right\|}{\lambda} + \frac{4\sqrt{2}}{\delta}\left\|m(\mathbb{A}(X_0))\right\| \leq 2\max\left\{\frac{\sqrt{2}\left\|m(\mathbb{A}(X_0))\right\|}{\lambda}, \frac{4\sqrt{2}}{\delta}\left\|m(\mathbb{A}(X_0))\right\|\right\}.\end{aligned}$$

因此

$$\left\|\epsilon_{k+1}\right\| \leq \left(1 + 2\alpha\max\left\{\frac{1}{\lambda}, \frac{1}{\delta}\right\}\right)\left\|\epsilon_k\right\| + 2\max\left\{\frac{\sqrt{2}\left\|m(\mathbb{A}(X_0))\right\|}{\lambda}, \frac{4\sqrt{2}}{\delta}\left\|m(\mathbb{A}(X_0))\right\|\right\}\alpha^2 \tag{26}$$

现在，对于 $L_{\delta,\lambda} = \max\left\{\frac{1}{\lambda}, \frac{\sqrt{2}\|m(\mathbb{A}\ X_0\ \|())}{\lambda}, \frac{1}{\delta}, \frac{4\sqrt{2}}{\delta}\left\|m(\mathbb{A}(X_0))\right\|\right\}$，我们证明

$$\left\|\epsilon_k\right\| \leq he^{2L_{\lambda,\delta}T}.$$

将 (26) 乘以 $(1+2\alpha L_{\lambda,\delta})^{-(k+1)}$，得到

$$(1+2\alpha L_{\lambda,\delta})^{-(k+1)}\left\|\epsilon_{k+1}\right\| \leq (1+2\alpha L_{\lambda,\delta})^{-k}\left\|\epsilon_k\right\| + (1+2\alpha L_{\lambda,\delta})^{-(k+1)} L_{\lambda,\delta}\alpha^2$$

由于 $\left\|\epsilon_0\right\| = \left\|X_0 - X(0)\right\| = 0$，从 0 到 $k-1$ 求和，得到

$$\begin{aligned}(1+2\alpha L_{\lambda,\delta})^{-k}\left\|\epsilon_k\right\| &\leq \sum_{i=1}^{k}(1+2\alpha L_{\lambda,\delta})^{-i} L_{\lambda,\delta}\alpha^2 \\ &= \frac{(1+2\alpha L_{\lambda,\delta})^{-1}\left(1-(1+2\alpha L_{\lambda,\delta})^{-k}\right)}{1-(1+2\alpha L_{\lambda,\delta})^{-1}} L_{\lambda,\delta}\alpha^2 \\ &= \frac{1}{2}\left(1-(1+2\alpha L_{\lambda,\delta})^{-k}\right)\alpha.\end{aligned}$$

将两边乘以 $(1+2\alpha L_{\lambda,\delta})^k$，并应用 $\alpha = 2h$，得到

$$\left\|\epsilon_k\right\| \leq \frac{\alpha}{2}\left((1+2\alpha L_{\lambda,\delta})^k - 1\right) = h\left((1+4hL_{\lambda,\delta})^k - 1\right). \tag{27}$$

现在由

$$(1+4hL_{\lambda,\delta})^k \leq \left((1+4hL_{\lambda,\delta})^{\frac{1}{4hL_{\lambda,\delta}}}\right)^{4hL_{\lambda,\delta}k} \leq e^{4hL_{\lambda,\delta}k},$$

应用 $k \leq \frac{T}{2h}$ —

$$\left\|\epsilon_k\right\| \leq h\left(e^{4hL_{\lambda,\delta}k} - 1\right) \leq he^{2L_{\lambda,\delta}T}.$$

因此

$$\left\|X_{\lambda,\delta}(2hk) - X_{\lambda,\delta}^k\right\| \leq he^{2L_{\lambda,\delta}T} = \mathcal{O}\left(he^{2L_{\lambda,\delta}T}\right). \tag{28}$$

D.3 引理 D.3 中 case(v) 的证明

由于 APPM 的系数为 $\frac{1}{k+1}$，考虑到计算简便性，我们考虑 $X_{\lambda,\delta}^{k+1}$，而不是 $X_{\lambda,\delta}^k$。根据三角不等式，我们有

$$\left\|X_{\lambda,\delta}^k - x_\lambda^k\right\| \leq \left\|X_{\lambda,\delta}^k - X_{\lambda,\delta}^{k+1}\right\| + \left\|X_{\lambda,\delta}^{k+1} - x_\lambda^k\right\|.$$

我们将证明

$$\left\|X^k_{\lambda,\delta}-X^{k+1}_{\lambda,\delta}\right\|=\mathcal{O}(h)+\mathcal{O}\left(h^2L_{\lambda,\delta}e^{2L_{\lambda,\delta}T}\right)$$

$$\left\|X^{k+1}_{\lambda,\delta}-x^k_\lambda\right\|=3^{\frac{T}{\lambda}}\left(\mathcal{O}\left(\frac{h}{\lambda}\right)+\mathcal{O}\left(\frac{h}{\lambda}e^{2L_{\lambda,\delta}T}\right)+\mathcal{O}(h)+\mathcal{O}\left(\frac{h^2}{\lambda}e^{2L_{\lambda,\delta}T}\right)+\mathcal{O}\left(\frac{e^{2\delta L_{\lambda,\delta}}}{L_{\lambda,\delta}}\right)+\mathcal{O}(\delta)\right).$$

首先一个比较简单。由于 $X^{k+1}_{\lambda,\delta}=X^k_{\lambda,\delta}+2hF(X_{\lambda,\delta},2hk)$ 和 F 关于第一变量是 $2\max\left\{\frac{1}{\lambda},\frac{1}{\delta}\right\}$-Lipschitz 的，因此有

$$\begin{aligned}\left\|X^k_{\lambda,\delta}-X^{k+1}_{\lambda,\delta}\right\|&=2h\left\|F(X^k_{\lambda,\delta},2hk)\right\|\\&\le 2h\left\|F(X_{\lambda,\delta}(2hk),2hk)\right\|+2h\left\|F(X_{\lambda,\delta}(2hk),2hk)-F(X^k_{\lambda,\delta},2hk)\right\|\\&\le 2h\left\|\dot X_{\lambda,\delta}(2hk)\right\|+4h\max\left\{\frac{1}{\lambda},\frac{1}{\delta}\right\}\left\|X_{\lambda,\delta}(2hk)-X^k_{\lambda,\delta}\right\|\\&\le 2\sqrt{2}h\left\|m(\mathbb{A}_\lambda(X_0))\right\|+4h^2L_{\lambda,\delta}e^{2L_{\lambda,\delta}T}\\&=\mathcal{O}(h)+\mathcal{O}\left(h^2L_{\lambda,\delta}e^{2L_{\lambda,\delta}T}\right).\end{aligned}$$

第二个比较复杂，我们将证明分成几个小节来逐步阐述。

D.3.1 关于 $\epsilon_k=\left\|X^{k+1}_{\lambda,\delta}-x^k_{h,\lambda}\right\|$ 的递归不等式

定义 $\epsilon_k=\left\|X^{k+1}_{\lambda,\delta}-x^k_{h,\lambda}\right\|$。回想一下，$X_{\lambda,\delta}$ 是近似常微分方程的解

$$\dot X_{\lambda,\delta}(t)=F(X_{\lambda,\delta},t)=\begin{cases}-\mathbb{A}_\lambda(X_{\lambda,\delta})(t)-\frac{1}{\delta}(X(t)-X_0) & 0\le t<\delta\\ -\mathbb{A}_\lambda(X_{\lambda,\delta})(t)-\frac{1}{t}(X(t)-X_0) & t\ge\delta\end{cases}.$$

现在我们希望用 ϵ_k 来表示 ϵ_{k+1}。由于 ϵ_{k+1} 涉及 $X^{k+2}_{\lambda,\delta}$，我们先将其明确写出。

$$\begin{aligned}X^{k+2}_{\lambda,\delta}&=X^{k+1}_{\lambda,\delta}+2hF\left(X^{k+1}_{\lambda,\delta},2h(k+1)\right)\\&=\begin{cases}X^{k+1}_{\lambda,\delta}-\left(2h\mathbb{A}_\lambda(X^{k+1}_{\lambda,\delta})+\frac{2h}{\delta}(X^{k+1}_{\lambda,\delta}-X_0)\right) & 0\le k+1<\frac{\delta}{2h}\\ X^{k+1}_{\lambda,\delta}-\left(2h\mathbb{A}_\lambda(X^{k+1}_{\lambda,\delta})+\frac{1}{k+1}(X^{k+1}_{\lambda,\delta}-X_0)\right) & k+1\ge\frac{\delta}{2h}.\end{cases}\end{aligned}$$

现在我们分两种情况来推导递归不等式。

(i) $k+1\ge\frac{\delta}{2h}$回想一下，APPM

(5) 定义为

$$\begin{aligned}x^k_{h,\lambda}&=\mathbb{J}_{h\mathbb{A}_\lambda}y^{k-1}_{h,\lambda}\\y^k_{h,\lambda}&=\frac{k}{k+1}(2x^k_{h,\lambda}-y^{k-1}_{h,\lambda})+\frac{1}{k+1}X_0,\end{aligned}$$

代入 $y^{k-1}_{h,\lambda}=x^k_{h,\lambda}+h\mathbb{A}_\lambda(x^k_{h,\lambda})$ 后，我们得到一个单行表达式

$$x^{k+1}_{h,\lambda}=\frac{k}{k+1}x^k_{h,\lambda}-h\left(\mathbb{A}_\lambda(x^{k+1}_{h,\lambda})+\frac{k}{k+1}\mathbb{A}_\lambda(x^k_{h,\lambda})\right)+\frac{1}{k+1}X_0.$$

改写$X^{k+2}_{\lambda,\delta}$，以便于与上文比较，

$$\begin{aligned}X^{k+2}_{\lambda,\delta}=&\frac{k}{k+1}X^{k+1}_{\lambda,\delta}-h\left(\mathbb{A}_\lambda(X^{k+2}_{\lambda,\delta})+\frac{k}{k+1}\mathbb{A}_\lambda(X^{k+1}_{\lambda,\delta})\right)+\frac{1}{k+1}X_0\\&+h\left(\mathbb{A}_\lambda(X^{k+2}_{\lambda,\delta})-\mathbb{A}_\lambda(X^{k+1}_{\lambda,\delta})\right)-\frac{h}{k+1}\mathbb{A}_\lambda(X^{k+1}_{\lambda,\delta}).\end{aligned}$$

由于 $\mathbb{A}_\lambda$ 是$\frac{1}{\lambda}$-Lipschitz

$$\epsilon_{k+1} \le \frac{k}{k+1}\epsilon_k + \frac{h}{\lambda}\left(\epsilon_{k+1} + \frac{k}{k+1}\epsilon_k\right) + \frac{h}{\lambda}\left\|X_{\lambda,\delta}^{k+2} - X_{\lambda,\delta}^{k+1}\right\| + \frac{h}{k+1}\left\|\mathbb{A}_\lambda(X_{\lambda,\delta}^{k+1})\right\|.$$

整理后，

$$\epsilon_{k+1} \le \frac{1}{1-\frac{h}{\lambda}}\left(\frac{k}{k+1}\left(1+\frac{h}{\lambda}\right)\epsilon_k + \underbrace{h\left(\frac{1}{\lambda}\left\|X_{\lambda,\delta}^{k+2} - X_{\lambda,\delta}^{k+1}\right\| + \frac{1}{k+1}\left\|\mathbb{A}_\lambda(X_{\lambda,\delta}^{k+1})\right\|\right)}_{:=e_k}\right).$$

(ii) $k+1 < \frac{\delta}{2h}$

改写 $X_{\lambda,\delta}^{k+2}$，可得

$$\begin{aligned} X_{\lambda,\delta}^{k+2} = {} & \frac{k}{k+1}X_{\lambda,\delta}^{k+1} - h\left(\mathbb{A}_\lambda(X_{\lambda,\delta}^{k+2}) + \frac{k}{k+1}\mathbb{A}_\lambda(X_{\lambda,\delta}^{k+1})\right) + \frac{1}{k+1}X_0 \\ & + h\left(\mathbb{A}_\lambda(X_{\lambda,\delta}^{k+2}) - \mathbb{A}_\lambda(X_{\lambda,\delta}^{k+1})\right) + \frac{h}{k+1}\mathbb{A}_\lambda(X_{\lambda,\delta}^{k+1}) + \left(\frac{1}{k+1} - \frac{2h}{\delta}\right)(X_{\lambda,\delta}^{k+1} - X_0). \end{aligned}$$

唯一的区别　　与情形（i）相比，在于最后一项。因此，采用相同的计算方法，我们得到以下不等式 .

$$\epsilon_{k+1} \le \frac{1}{1-\frac{h}{\lambda}}\left(\frac{k}{k+1}\left(1+\frac{h}{\lambda}\right)\epsilon_k + \underbrace{h\left(\frac{1}{\lambda}\left\|X_{\lambda,\delta}^{k+2} - X_{\lambda,\delta}^{k+1}\right\| + \frac{1}{k+1}\left\|\mathbb{A}_\lambda(X_{\lambda,\delta}^{k+1})\right\|\right) + \left|\frac{1}{k+1} - \frac{2h}{\delta}\right|\left\|X_{\lambda,\delta}^{k+1} - X_0\right\|}_{:=e_k}\right).$$

根据情形（i）和（ii），定义

$$e_k = \begin{cases} h\left(\frac{1}{\lambda}\left\|X_{\lambda,\delta}^{k+2} - X_{\lambda,\delta}^{k+1}\right\| + \frac{1}{k+1}\left\|\mathbb{A}_\lambda(X_{\lambda,\delta}^{k+1})\right\|\right) + \left|\frac{1}{k+1} - \frac{2h}{\delta}\right|\left\|X_{\lambda,\delta}^{k+1} - X_0\right\| & 0 \le k+1 < \frac{\delta}{2h} \\ h\left(\frac{1}{\lambda}\left\|X_{\lambda,\delta}^{k+2} - X_{\lambda,\delta}^{k+1}\right\| + \frac{1}{k+1}\left\|\mathbb{A}_\lambda(X_{\lambda,\delta}^{k+1})\right\|\right) & k+1 \ge \frac{\delta}{2h}, \end{cases}$$

我们可以为所有 k 写出如下递归不等式

$$\epsilon_{k+1} \le \frac{1}{1-\frac{h}{\lambda}}\left(\frac{k}{k+1}\left(1+\frac{h}{\lambda}\right)\epsilon_k + e_k\right).$$

D.3.2 上标$0 \le i \le \frac{T}{2h}$ $\epsilon_i = 3^{\frac{T}{\lambda}}\left(\mathcal{O}\left(\frac{h}{\lambda}\right) + \mathcal{O}\left(\frac{h}{\lambda}e^{2L_{\lambda,\delta}T}\right) + \mathcal{O}(h) + \mathcal{O}\left(\frac{h^2}{\lambda}e^{2L_{\lambda,\delta}T}\right) + \mathcal{O}\left(\frac{e^{2\delta L_{\lambda,\delta}}}{L_{\lambda,\delta}}\right) + \mathcal{O}(\delta)\right)$

现在对上述不等式求和。将两边乘以$(k+1)\left(\frac{1-\frac{h}{\lambda}}{1+\frac{h}{\lambda}}\right)^{k+1}$，得到

$$\left(\frac{1-\frac{h}{\lambda}}{1+\frac{h}{\lambda}}\right)^{k+1}(k+1)\epsilon_{k+1} \le \left(\frac{1-\frac{h}{\lambda}}{1+\frac{h}{\lambda}}\right)^{k} k\epsilon_k + \frac{\left(1-\frac{h}{\lambda}\right)^{k+1}}{\left(1+\frac{h}{\lambda}\right)^{k}}(k+1)e_k.$$

将上述不等式从0到 $k-1$求和，得到

$$\left(\frac{1-\frac{h}{\lambda}}{1+\frac{h}{\lambda}}\right)^{k} k\epsilon_k \le \sum_{i=0}^{k-1}\frac{\left(1-\frac{h}{\lambda}\right)^{i+1}}{\left(1+\frac{h}{\lambda}\right)^{i}}(i+1)e_i.$$

注意， ϵ_0 因乘以 0 而消失。将各项按 ϵ_k 重新整理，可得

$$\begin{aligned} \epsilon_k &\le \left(\frac{1+\frac{h}{\lambda}}{1-\frac{h}{\lambda}}\right)^{k}\frac{1}{k}\sum_{i=0}^{k-1}\frac{\left(1-\frac{h}{\lambda}\right)^{i+1}}{\left(1+\frac{h}{\lambda}\right)^{i}}(i+1)e_i \\ &\le \left(\frac{1+\frac{h}{\lambda}}{1-\frac{h}{\lambda}}\right)^{\frac{T}{2h}}\frac{1}{k}\sum_{i=0}^{k-1}(i+1)e_i = \left(\left(\frac{1+\frac{h}{\lambda}}{1-\frac{h}{\lambda}}\right)^{\frac{\lambda}{h}}\right)^{\frac{T}{2\lambda}}\frac{1}{k}\sum_{i=0}^{k-1}(i+1)e_i \end{aligned}$$

第二个不等式由事实 $0 < \frac{h}{\lambda} < \frac{1}{2}$ 得出，这意味着 $0 < (1-\frac{h}{\lambda})^{i+1}(1+\frac{h}{\lambda})_i \leq 1$ 且 $1 < \frac{1+\frac{h}{\lambda}}{1-\frac{h}{\lambda}}$。注意，$f(x) = \left(\frac{1+x}{1-x}\right)^{\frac{1}{x}}$ 在 $x \in (0, 1)$ 上单调不减，因为

$$f'(x) = -\frac{\left(\frac{1+x}{1-x}\right)^{\frac{1}{x}}\left(\left(x^2-1\right)\log\left(\frac{1+x}{1-x}\right)+2x\right)}{x^2\left(x^2-1\right)}.$$

因此，由 $f\left(\frac{1}{2}\right) = 9$ 可得

$$\epsilon_k \leq f\left(\frac{1}{2}\right)^{\frac{T}{2\lambda}} \frac{1}{k}\sum_{i=0}^{k-1}(i+1)e_i = 3^{\frac{T}{\lambda}}\frac{1}{k}\sum_{i=0}^{k-1}(i+1)e_i.$$

现在我们证明

$$\frac{1}{k}\sum_{i=0}^{k-1}(i+1)e_i = \mathcal{O}\left(\frac{h}{\lambda}\right) + \mathcal{O}\left(\frac{h}{\lambda}e^{2L_{\lambda,\delta}T}\right) + \mathcal{O}(h) + \mathcal{O}\left(\frac{h^2}{\lambda}e^{2L_{\lambda,\delta}T}\right) + \mathcal{O}(\delta) + \mathcal{O}\left(\frac{e^{2\delta L_{\lambda,\delta}}}{L_{\lambda,\delta}}\right)$$

对于 $0 < T < \infty$、$0 \leq k < \frac{T}{2h}$。令 $N_k = \min\left\{k, \left\lfloor \frac{\delta}{2h} \right\rfloor\right\}$。根据 e_i 的定义，我们有

$$\begin{aligned}
&\frac{1}{k}\sum_{i=0}^{k-1}(i+1)e_i \\
&= \frac{1}{k}\sum_{i=0}^{k-1}h(i+1)\left(\frac{1}{\lambda}\left\|X_{\lambda,\delta}^{i+2} - X_{\lambda,\delta}^{i+1}\right\| + \frac{1}{i+1}\left\|\mathbb{A}_\lambda(X_{\lambda,\delta}^{i+1})\right\|\right) + \frac{1}{k}\sum_{i=0}^{N_k-1}(i+1)\left|\frac{1}{i+1} - \frac{2h}{\delta}\right|\left\|X_{\lambda,\delta}^{i+1} - X_0\right\| \\
&\leq \sum_{i=0}^{k-1}h\left(\frac{1}{\lambda}\left\|X_{\lambda,\delta}^{i+2} - X_{\lambda,\delta}^{i+1}\right\| + \frac{1}{k}\left\|\mathbb{A}_\lambda(X_{\lambda,\delta}^{i+1})\right\|\right) + \frac{1}{k}\sum_{i=0}^{N_k-1}\left\|X_{\lambda,\delta}^{i+1} - X_0\right\|,
\end{aligned}$$

其中，不等式是由以下事实推出的：对于 $0 \leq i \leq k-1$，$\frac{i+1}{k} \leq 1$；对于 $0 \leq i \leq N_k - 1$，$\frac{1}{i+1} \geq \frac{2h}{\delta}$。现在我们来考察各项。由引理 B.12 和引理 D.1 可知

$$\begin{aligned}
\left\|\dot{X}_{\lambda,\delta}(t)\right\| &\leq \sqrt{2}\left\|\mathbb{A}_\lambda(X_0)\right\| \leq \sqrt{2}\left\|m(\mathbb{A}(X_0))\right\| \\
\left\|\mathbb{A}_\lambda(X(t))\right\| &\leq \left\|\mathbb{A}_\lambda(X_0)\right\| \leq \left\|m(\mathbb{A}(X_0))\right\| \\
\left\|X_{\lambda,\delta}(t) - X_0\right\| &\leq 2\left\|X_0 - X_\star\right\|.
\end{aligned}$$

令 $M = \max\left\{\sqrt{2}\left\|m(\mathbb{A}(X_0))\right\|, \ 2\left\|X_0 - X_\star\right\|\right\}$。

(i) $\frac{h}{\lambda}\left\|X_\lambda^{i+2}, {}_\delta - X_\lambda^{i+1}, {}_\delta\right\|$ 首先观察

$$\begin{aligned}
&h\sum_{i=0}^{k-1}\frac{1}{\lambda}\left\|X_{\lambda,\delta}(2h(i+2)) - X_{\lambda,\delta}(2h(i+1))\right\| \\
&= \frac{h}{\lambda}\sum_{i=0}^{k-1}\left\|\int_{2h(i+1)}^{2h(i+2)}\dot{X}_{\lambda,\delta}(t)dt\right\| \leq \frac{h}{\lambda}\sum_{i=0}^{k-1}\int_{2h(i+1)}^{2h(i+2)}\left\|\dot{X}_{\lambda,\delta}(t)\right\|dt \leq \frac{h}{\lambda}k(2h)M \leq \frac{h}{\lambda}TM.
\end{aligned}$$

因此,

$$
\begin{aligned}
&h\sum_{i=0}^{k-1}\frac{1}{\lambda}\left\|X_{\lambda,\delta}^{i+2}-X_{\lambda,\delta}^{i+1}\right\| \\
&\leq \frac{h}{\lambda}\sum_{i=0}^{k-1}\left(\left\|X_{\lambda,\delta}^{i+2}-X_{\lambda,\delta}(2h(i+2))\right\|+\|X_{\lambda,\delta}(2h(i+2))-X_{\lambda,\delta}(2h(i+1))\|+\left\|X_{\lambda,\delta}(2h(i+1))-X_{\lambda,\delta}^{i+1}\right\|\right) \\
&\leq \frac{h}{\lambda}\left(TM+2\sum_{i=0}^{k-1}he^{2L_{\lambda,\delta}T}\right) \\
&\leq \frac{h}{\lambda}\left(TM+Te^{2L_{\lambda,\delta}T}\right).
\end{aligned}
$$

因此

$$
\frac{h}{\lambda}\sum_{i=0}^{k-1}\left\|X_{\lambda,\delta}^{i+2}-X_{\lambda,\delta}^{i+1}\right\|=\mathcal{O}\left(\frac{h}{\lambda}\right)+\mathcal{O}\left(\frac{h}{\lambda}e^{2L_{\lambda,\delta}T}\right).
$$

(ii) $\frac{h}{k}\left\|\mathbb{A}_\lambda\left(X_{\lambda,\delta}^{i+1},\right)\right\|$

由于 $\mathbb{A}_\lambda$ 是$-\frac{1}{\lambda}$-Lipschitz 连续的，且根据 (28)

$$
\begin{aligned}
\sum_{i=0}^{k-1}\frac{h}{k}\left\|\mathbb{A}_\lambda(X_{\lambda,\delta}^{i+1})\right\| &\leq \frac{h}{k}\sum_{i=0}^{k-1}\left(\left\|\mathbb{A}_\lambda(X_{\lambda,\delta}^{i})-\mathbb{A}_\lambda(X_{\lambda,\delta}(2hi))\right\|+\|\mathbb{A}_\lambda(X_{\lambda,\delta}(2hi))\|\right) \\
&\leq \frac{h}{k}\sum_{i=0}^{k-1}\left(\frac{h}{\lambda}e^{2L_{\lambda,\delta}T}+M\right)=h\left(\frac{h}{\lambda}e^{2L_{\lambda,\delta}T}+M\right).
\end{aligned}
$$

因此

$$
\sum_{i=0}^{k-1}\frac{h}{k}\left\|\mathbb{A}_\lambda(X_{\lambda,\delta}^{i+1})\right\|=\mathcal{O}(h)+\mathcal{O}\left(\frac{h^2}{\lambda}e^{2L_{\lambda,\delta}T}\right).
$$

(iii) $\frac{1}{k}\sum_{i=0}^{N_k-1}\left\|X_\lambda^{i+1,\delta}-X_0\right\|$ 对于 $N_k=$最小值 $\left\{k,\lfloor\frac{\delta}{2h}\rfloor\right\}$。

首先，注意到

$$
\|X_0-X_{\lambda,\delta}(t)\|=\left\|\int_0^t\dot{X}_{\lambda,\delta}(s)ds\right\|\leq\int_0^t\left\|\dot{X}_{\lambda,\delta}(s)\right\|ds\leq tM.
$$

由式（27）

$$
\begin{aligned}
\left\|X_{\lambda,\delta}^{i+1}-X_0\right\| &\leq \left\|X_{\lambda,\delta}^{i+1}-X_{\lambda,\delta}(2h(i+1))\right\|+\|X_{\lambda,\delta}(2h(i+1))-X_0\| \\
&\leq h\left(1+4hL_{\lambda,\delta}\right)^{i+1}+2h(i+1)M.
\end{aligned}
$$

现在，令 $i+1\leq N_k=\min\left\{k,\lfloor\frac{\delta}{2h}\rfloor\right\}$,

$$
\begin{aligned}
\frac{1}{k}\sum_{i=0}^{N_k-1}\left\|X_{\lambda,\delta}^{i+1}-X_0\right\| &\leq \frac{1}{k}\sum_{i=0}^{N_k-1}h\left(\left(1+4hL_{\lambda,\delta}\right)^{i+1}+2(i+1)M\right) \\
&\leq \frac{h}{k}\frac{\left(1+4hL_{\lambda,\delta}\right)^{N_k}-1}{4hL_{\lambda,\delta}}+2hM\sum_{i=0}^{N_k-1}\frac{i+1}{k} \\
&\leq \frac{e^{4hL_{\lambda,\delta}N_k}}{kL_{\lambda,\delta}}+2hN_kM \\
&\leq \frac{e^{2\delta L_{\lambda,\delta}}}{L_{\lambda,\delta}}+2\delta M=\mathcal{O}\left(\frac{e^{2\delta L_{\lambda,\delta}}}{L_{\lambda,\delta}}\right)+\mathcal{O}\left(\delta\right).
\end{aligned}
$$

由 (i)、(ii)、(iii) 可得

$$\frac{1}{k}\sum_{i=0}^{k-1}(i+1)e_i = \mathcal{O}\left(\frac{h}{\lambda}\right) + \mathcal{O}\left(\frac{h}{\lambda}e^{2L_{\lambda,\delta}T}\right) + \mathcal{O}(h) + \mathcal{O}\left(\frac{h^2}{\lambda}e^{2L_{\lambda,\delta}T}\right) + \mathcal{O}\left(\frac{e^{2\delta L_{\lambda,\delta}}}{L_{\lambda,\delta}}\right) + \mathcal{O}\left(\delta\right).$$

因此,

$$\epsilon_k = 3^{\frac{T}{\lambda}}\left(\mathcal{O}\left(\frac{h}{\lambda}\right) + \mathcal{O}\left(\frac{h}{\lambda}e^{2L_{\lambda,\delta}T}\right) + \mathcal{O}(h) + \mathcal{O}\left(\frac{h^2}{\lambda}e^{2L_{\lambda,\delta}T}\right) + \mathcal{O}\left(\frac{e^{2\delta L_{\lambda,\delta}}}{L_{\lambda,\delta}}\right) + \mathcal{O}\left(\delta\right)\right).$$

D.4 EAG 和 FEG 导出常微分方程(4)

D.4.1 从 EAG 推导

F对于 L-Lipschitz 连续单调算子 $\mathbb{A}$ 和步长 $h>0$，EAG-C [75] 定义为

$$\begin{aligned} z^{k+\frac{1}{2}} &= z^k - \frac{1}{k+2}\left(z^k - z^0\right) - h\mathbb{A}(z^k) \\ z^{k+1} &= z^k - \frac{1}{k+2}\left(z^k - z^0\right) - h\mathbb{A}(z^{k+\frac{1}{2}}). \end{aligned}$$

将第二行除以 h 并重新整理，得到

$$\begin{aligned} \frac{z^{k+1}-z^k}{h} &= -\mathbb{A}(z^{k+\frac{1}{2}}) - \frac{1}{h(k+2)}\left(z^k - z^0\right) \\ &= -\mathbb{A}(z^k) - \frac{1}{h(k+2)}\left(z^k - z^0\right) - \left(\mathbb{A}(z^{k+\frac{1}{2}}) - \mathbb{A}(z^k)\right). \end{aligned}$$

令 $hk = t$, $z^k = X(t)$, $z^0 = X_0$。由于 z^k 是收敛序列 [76], , 因此它有界。于是我们得到 e

$$\begin{aligned} \left\|\mathbb{A}(z^{k+\frac{1}{2}}) - \mathbb{A}(z^k)\right\| &\le L\left\|z^{k+\frac{1}{2}} - z^k\right\| \\ &= L\left\|\frac{h}{h(k+2)}(z^k - z^0) + h\mathbb{A}(z^k)\right\| \\ &= Lh\left\|\frac{1}{t+2h}(X(t) - X_0) + \mathbb{A}(X(t))\right\| = \mathcal{O}\left(h\right). \end{aligned}$$

因此，取极限 $h \to 0+$ ，我们有

$$\dot{X}(t) = -\mathbb{A}(X(t)) - \frac{1}{t}(X(t) - X_0).$$

D.4.2 从 FEG 推导

对于 L-Lipschitz 连续单调算子 $\mathbb{A}$ 和步长 $h>0$，FEG [42] 定义为

$$\begin{aligned} z^{k+\frac{1}{2}} &= z^k - \frac{1}{k+1}\left(z^k - z^0\right) - \frac{k}{k+1}h\mathbb{A}(z^k) \\ z^{k+1} &= z^k - \frac{1}{k+1}\left(z^k - z^0\right) - h\mathbb{A}(z^{k+\frac{1}{2}}). \end{aligned}$$

将第二行除以 h 并重新整理，得到

$$\begin{aligned} \frac{z^{k+1}-z^k}{h} &= -\mathbb{A}(z^{k+\frac{1}{2}}) - \frac{1}{h(k+1)}\left(z^k - z^0\right) \\ &= -\mathbb{A}(z^k) - \frac{1}{h(k+1)}\left(z^k - z^0\right) - \left(\mathbb{A}(z^{k+\frac{1}{2}}) - \mathbb{A}(z^k)\right). \end{aligned}$$

令 $hk = t$、 $z^k = X(t)$、 $z^0 = X_0$。由于 z^k 是一个收敛序列 [76], ，因此它有界。由此可知

$$\begin{aligned}\left\|\mathbb{A}(z^{k+\frac{1}{2}}) - \mathbb{A}(z^k)\right\| &\leq L\left\|z^{k+\frac{1}{2}} - z^k\right\| \\ &= L\left\|\frac{h}{h(k+1)}(z^k - z^0) + h\frac{hk}{h(k+1)}\mathbb{A}(z^k)\right\| \\ &= Lh\left\|\frac{1}{t+h}(X(t) - X_0) + \frac{t}{t+h}\mathbb{A}(X(t))\right\| = \mathcal{O}(h).\end{aligned}$$

因此取极限 $h \to 0+$ ，我们有

$$\dot{X}(t) = -\mathbb{A}(X(t)) - \frac{1}{t}(X(t) - X_0).$$

E 单调算子的收敛分析证明 $\mathbb{A}$

˜E.1 将 $\mathbb{A}$ 扩展到 $[0,\infty)$

引理 E.1. 假设 $\mathbb{A}$ 是一个极大单调算子。令 X 为 (3)的解。定义 S 为

$$S = \left\{t \in [0,\infty) \mid \dot{X}(t) \in -\mathbb{A}(X(t)) - \beta(t)(X(t) - X_0) \textit{ is true}\right\}.$$

那么， 由于 X 是 (3)的解， 我们知道 $[0,\infty)\backslash S$ 是零测度。

˜定义 $\mathbb{A}(X)\colon S \to \mathbb{R}^n$ 为

$$\tilde{\mathbb{A}}(X)(t) = -\dot{X}(t) - \beta(t)(X(t) - X_0) \tag{29}$$

˜ ˜并记 $\mathbb{A}(X(t)) = \mathbb{A}(X$<style id='15'>$)(t)$。假设对每个 $T > 0$， 都存在 $M > 0$， 使得对所有 $t \in S \cap [0,T]$

$$\left\|\tilde{\mathbb{A}}(X(t))\right\| \leq M.$$

˜那么， $\mathbb{A}(X)$ 可以扩展为 $t \in [0,\infty)$ ，并满足以下性质。

˜*(i)* 对所有 $t \in [0,\infty)$，均有 $\mathbb{A}(X(t)) \in \mathbb{A}(X(t))$ 。˜*(ii)* 对于 $t \in [0,\infty)\backslash S$，存在一个序列 $\{t_k\}_{k\in\mathbb{N}}$ ，使得 $t_k \in S$，$\lim_{k\to\infty} t_k = t$ ，且 $\mathbb{A}(X(t_k))$ 收敛到˜$\mathbb{A}(X(t))$。

证明。取 $t \in [0,\infty)$，并取一个序列 $\{t_n\}_{n\in\mathbb{N}}$ ，使得 $t_n \in S$ 且 $\lim_{n\to\infty} t_n = t$。由于收敛序列$t_n$有界，存在 $T > 0$，使得 $t_n \in [0,T]$。对于该 T，根据假设，有 $\left\|\tilde{\mathbb{A}}(X(t_n))\right\| \leq M$。由于

˜$n \mapsto \mathbb{A}(X(t_n))$ 是有界序列，存在一个子序列 $\{t_{n_k}\}_{k\in\mathbb{N}}$，使得 $k \mapsto \mathbb{A}(X(t_{n_k}))$ 收敛。将其极限˜记为 $u = \lim_{k\to\infty} \mathbb{A}(X(t_{n_k}))$。另一方面，由于 X 是一条连续曲线，有 $\lim_{k\to\infty} X(t_{n_k}) = X(t)$。于是，由于 $\mathbb{A}$˜ 是极大单调的，根据 [15, 命题20.38] ，可得 $(X(t)\ u) \in \operatorname{Gra}\mathbb{A}$。定义 $\mathbb{A}(X(t))$ 为 u，即可得到所需结果。

□

˜**推论 E.2.** 设 X 为(6)的解。那么， $\mathbb{A}(X)$ 由(29)定义，其延拓满足引理 *E.1* 中所述的性质。

证明。 首先考虑 $\beta(t) = \frac{\gamma}{t^p}, p > 0, \gamma > 0$ 的情形。取 $T > 0$。只需证明存在 $M > 0$，使得对于 $t \in S \cap [0,T]$

$$\left\|\tilde{\mathbb{A}}(X(t))\right\| \leq M.$$

回顾命题 B.11 中关于 Yosida逼近 $\mathbb{A}_\lambda$的内容，我们将 X_λ 记为以下方程的解

$$\dot{X}_\lambda = -\mathbb{A}_\lambda(X_\lambda) - \frac{\gamma}{t^p}(X_\lambda - X_0),$$

并且我们已经证明，存在一个序列 $\{\lambda_n\}_{n\in\mathbb{N}}$，使得 X_{λ_n} 在 $[0,T]$上一致收敛于 X。

根据引理 B.12，我们看到对于 $h \neq 0$

$$\left\| \frac{X_\lambda(t+h) - X_\lambda(t)}{h} \right\| \leq \frac{\int_t^{t+h} \left\| \dot{X}_\lambda(s) \right\| ds}{h} \leq \frac{\int_t^{t+h} M_{dot}(T) ds}{h} = M_{dot}(T),$$

thus

$$\left\| \frac{X(t+h) - X(t)}{h} \right\| = \lim_{\lambda \to 0+} \left\| \frac{X_\lambda(t+h) - X_\lambda(t)}{h} \right\| \leq M_{dot}(T).$$

因此，对于 $t \in S \cap [0,T]$, $\dot{X}(t) = \lim_{h \to 0} \frac{X(t+h) - X(t)}{h}$ 成立的情形，我们得出

$$\left\| \dot{X}(t) \right\| = \lim_{h \to 0} \left\| \frac{X(t+h) - X(t)}{h} \right\| \leq M_{dot}(T).$$

此外，根据引理 B.12，对于 $t \in [0,T]$，我们有

$$\left\| \frac{\gamma}{t^p}(X_\lambda(t) - X_0) \right\| = \left\| \dot{X}_\lambda(t) + \mathbb{A}_\lambda(t) \right\| \leq \left\| \dot{X}_\lambda(t) \right\| + \|\mathbb{A}_\lambda(t)\| \leq M_{dot}(T) + M_{\mathbb{A}}(T),$$

因此

$$\left\| \frac{\gamma}{t^p}(X(t) - X_0) \right\| = \lim_{\lambda \to 0+} \left\| \frac{\gamma}{t^p}(X_\lambda(t) - X_0) \right\| \leq M_{dot}(T) + M_{\mathbb{A}}(T).$$

综上，对于 $t \in S \cap [0,T]$，我们有

$$\left\| \tilde{\mathbb{A}}(X(t)) \right\| = \left\| \dot{X}(t) + \frac{\gamma}{t^p}(X(t) - X_0) \right\| \leq \left\| \dot{X}(t) \right\| + \left\| \frac{\gamma}{t^p}(X(t) - X_0) \right\| \leq 2M_{dot}(T) + M_{\mathbb{A}}(T).$$

□

E.2 命题 3.2 的证明

我们通过推导膨胀坐标 $W(t) = C(t)(X(t) - X_0)$ 下的能量和来自 [67] 的守恒律来证明该定理。考虑膨胀坐标 $W(t) = C(t)(X(t) - X_0)$。由于我们考虑的是 $\mathbb{A}$ 为利普希茨连续的情形，微分包含式 (3) 变为常微分方程

$$\dot{X}(t) = -\tilde{\mathbf{A}}(X(t)) - \beta(t)(X(t) - X_0).$$

用 W 重写常微分方程，得到

$$\frac{1}{C(t)} \left(\dot{W}(t) - \beta(t) W(t) \right) = -\tilde{\mathbf{A}}(X(W(t), t)) - \frac{\beta(t)}{C(t)} W(t)$$

其中 $X(W(t), t) = X(t) = \frac{W(t)}{C(t)} + X_0$. 整理后，得到

$$0 = \dot{W}(t) + C(t)\tilde{\mathbf{A}}(X(W(t), t)).$$

根据引理 B.4，我们知道 $\mathbb{A}(X(W, t))$ 几乎处处可微。对其求导，得到一个几乎处处成立的二阶常微分方程

$$\begin{aligned} 0 &= \ddot{W} + C(t)\beta(t)\tilde{\mathbf{A}}(X(W(t), t)) + C(t)\frac{d}{dt}\tilde{\mathbf{A}}(X(W(t), t)) \\ &= \ddot{W} - \beta(t)\dot{W} + C(t)\frac{d}{dt}\tilde{\mathbf{A}}(X(W(t), t)). \end{aligned} \tag{30}$$

现在通过取内积 $\dot{\mathbf{r}}$ 与 W 相乘并积分后，我们得到一个等式，该等式在 [6中被称为守恒律 7]

$$E_1 \equiv \frac{1}{2} \left\| \dot{W}(t) \right\|^2 - \int_{t_0}^{t} \beta(s) \left\| \dot{W}(s) \right\|^2 ds + \int_{t_0}^{t} C(s) \left\langle \frac{d}{ds}\tilde{\mathbf{A}}(X(s)), \dot{W}(s) \right\rangle ds.$$

由于 $\dot{W}(t) = C(t)\left(\dot{X}(t) + \beta(t)(X(t) - X_0)\right)$，我们可以将最后一项中的被积函数改写为

$$\int_{t_0}^{t} C(s)\left\langle \frac{d}{ds}\tilde{\mathbf{A}}(X(s)), \dot{W}(s)\right\rangle ds = \int_{t_0}^{t} C(s)^2\left\langle \frac{d}{ds}\tilde{\mathbf{A}}(X(s)), \dot{X}(s)\right\rangle ds + \int_{t_0}^{t}\left\langle \frac{d}{ds}\tilde{\mathbf{A}}(X(s)), C(s)\beta(s)W(s)\right\rangle ds.$$

˜请注意，我们的目的是得到第一项，它因 $\mathbb{A}$的单调性而非负。现在对第二项进行分部积分，得到

$$\begin{aligned}
&\int_{t_0}^{t}\left\langle \frac{d}{ds}\tilde{\mathbf{A}}(X(s)), C(s)\beta(s)W(s)\right\rangle ds \\
&= \left[\left\langle \tilde{\mathbf{A}}(X(W,s)), C(s)\beta(s)W(s)\right\rangle\right]_{t_0}^{t} - \int_{t_0}^{t}\left\langle \tilde{\mathbf{A}}(X(W,s)), \left(C(s)\beta(s)^2 + C(s)\dot{\beta}(s)\right)W(s) + C(s)\beta(s)\dot{W}(s)\right\rangle ds \\
&= \left[\left\langle \tilde{\mathbf{A}}(X(W,s)), C(s)\beta(s)W(s)\right\rangle\right]_{t_0}^{t} + \int_{t_0}^{t}\beta(s)\left\|\dot{W}(s)\right\|^2 ds + \int_{t_0}^{t}\left(\beta(s)^2 + \dot{\beta}(s)\right)\left\langle \dot{W}(s), W(s)\right\rangle ds \\
&= \left[\left\langle \tilde{\mathbf{A}}(X(W,s)), C(s)\beta(s)W(s)\right\rangle\right]_{t_0}^{t} + \int_{t_0}^{t}\beta(s)\left\|\dot{W}(s)\right\|^2 ds \\
&\quad + \left[\left(\beta(s)^2 + \dot{\beta}(s)\right)\frac{1}{2}\|W(s)\|^2\right]_{t_0}^{t} - \frac{1}{2}\int_{t_0}^{t}\left(2\beta(s)\dot{\beta}(s) + \ddot{\beta}(s)\right)\|W(s)\|^2 ds
\end{aligned}$$

˜在第二个等号处，我们利用了这样一个事实：$W(t) = -C(t)\mathbf{A}(X(W(t)\ t))$。请注意，微积分基本定理适用于$\left\langle \tilde{\mathbf{A}}(X(W,\ s))\ C(s)\beta(s)W(s)\right\rangle$，因为 $\tilde{\mathbf{A}}(X(W,\ s))$是Lipschitz连续的，且 $C(s)\beta(s)W(s)$在 $[t_0, t]$上连续可微，因此它们的内积在 $[t_0, t]$上绝对连续。

注意，最后一项中的被积函数可以改写为

$$\left(2\beta(s)\dot{\beta}(s) + \ddot{\beta}(s)\right)\|W(s)\|^2 = C(s)^2\left(2\beta(s)\dot{\beta}(s) + \ddot{\beta}(s)\right)\|X(s) - X_0\|^2 = \frac{d}{ds}\left(C(s)^2\dot{\beta}(s)\right)\|X(s) - X_0\|^2.$$

汇总上述结果，可得

$$\begin{aligned}
E_1 &\equiv \frac{1}{2}\left\|\dot{W}(t)\right\|^2 - \int_{t_0}^{t}\beta(s)\left\|\dot{W}(s)\right\|^2 ds + \int_{t_0}^{t} C(s)\left\langle \frac{d}{ds}\tilde{\mathbf{A}}(X(s)), \dot{W}(s)\right\rangle ds \\
&= \frac{C(t)^2}{2}\left\|\tilde{\mathbf{A}}(X(t))\right\|^2 + \left[\left\langle \tilde{\mathbf{A}}(X(W(s),s)), C(s)\beta(s)W(s)\right\rangle\right]_{t_0}^{t} + \left[\left(\beta(s)^2 + \dot{\beta}(s)\right)\frac{1}{2}\|W(s)\|^2\right]_{t_0}^{t} \\
&\quad + \int_{t_0}^{t} C(s)^2\left\langle \frac{d}{ds}\tilde{\mathbf{A}}(X(s)), \dot{X}(s)\right\rangle ds - \frac{1}{2}\int_{t_0}^{t}\frac{d}{ds}\left(C(s)^2\dot{\beta}(s)\right)\|X(s) - X_0\|^2 ds \\
&= \frac{C(t)^2}{2}\left(\left\|\tilde{\mathbf{A}}(X(t))\right\|^2 + 2\beta(t)\left\langle \tilde{\mathbf{A}}(X(t)), X(t) - X_0\right\rangle + \left(\beta(t)^2 + \dot{\beta}(t)\right)\|X(t) - X_0\|^2\right) \\
&\quad + \int_{t_0}^{t} C(s)^2\left\langle \frac{d}{ds}\tilde{\mathbf{A}}(X(s)), \dot{X}(s)\right\rangle ds - \frac{1}{2}\int_{t_0}^{t}\frac{d}{ds}\left(C(s)^2\dot{\beta}(s)\right)\|X(s) - X_0\|^2 ds \\
&\quad \underbrace{- \frac{C(t_0)^2}{2}\left(2\beta(t_0)\left\langle \tilde{\mathbf{A}}(X(t_0)), X(t_0) - X_0\right\rangle + \left(\beta(t_0)^2 + \dot{\beta}(t_0)\right)\|X(t_0) - X_0\|^2\right)}_{=\text{constant}}
\end{aligned}$$

将常数项移至左边，并将 $E = E_1$ –命名为常数，即得所需结果。

E.3 推论3.3的证明

˜E.3.1 V在 $\mathbb{A}$ 利普希茨连续且单调时非递增

我们先来验证这一点 ˜对于 $\mathbb{A}$ 为利普希茨连续的情况成立。因此，根据命题 3.2，我们可以将 $V(t)$ 写作

$$V(t) = E - 2\int_{t_0}^{t} C(s)^2\left\langle \frac{d}{ds}\tilde{\mathbf{A}}(X(s)), \dot{X}(s)\right\rangle ds$$

其中 E 见命题 3.2。由于 $\tilde{\mathbb{A}}$ 是单调的，根据 (1) 可知，$C(s)^2\left\langle \frac{d}{ds}\tilde{\mathbb{A}}(X(s))\ \dot{X}(s)\right\rangle \geq 0$ 几乎处处成立。因此对于 $h>0$

$$V(t+h)-V(t)=-2\int_t^{t+h} C(s)^2\left\langle \frac{d}{ds}\tilde{\mathbb{A}}(X(s)),\dot{X}(s)\right\rangle ds\leq 0,$$

我们看到 V是一个非增函数。因此，对于所有的 $t>0, V(t) \leq$ 极限$_{\epsilon\to 0+} V(\epsilon)$。接下来只需证明极限 极限$_{\epsilon\to 0+} V(\epsilon)$ 存在。

˜E.3.2 单调利普希茨连续 $\mathbb{A}$ 时 $V(0) =$ 极限$_{t\to 0+} V(t)$ 的计算

˜本节计算当 $\mathbb{A}$ 利普希茨连续时，极限$_{t\to 0+} V(t)$。回顾一下，V 定义为

$$\begin{aligned}V(t)=&\frac{C(t)^2}{2}\left(\left\|\tilde{\mathbb{A}}(X(t))\right\|^2+2\beta(t)\left\langle\tilde{\mathbb{A}}(X(t)),X(t)-X_0\right\rangle+\left(\beta(t)^2+\dot{\beta}(t)\right)\|X(t)-X_0\|^2\right)\\&-\int_{t_0}^t\frac{d}{ds}\left(\frac{C(s)^2\dot{\beta}(s)}{2}\right)\|X(s)-X_0\|^2\,ds.\end{aligned}$$

从现在起，我们用 V 表示 $t_0=0$ 为 V^0 的情形。

我们先检查

$$C(t)=\begin{cases}t^\gamma & p=1\\ e^{\frac{\gamma}{1-p}t^{1-p}} & p>0,p\neq 1.\end{cases}\tag{31}$$

观察到

$$\begin{aligned}\int_1^t\frac{\gamma}{s}ds&=\gamma\log t\\ \int_0^t\frac{\gamma}{s^p}ds&=\frac{\gamma}{1-p}t^{1-p}\qquad\text{for } 0<p<1\\ \int_\infty^t\frac{\gamma}{s^p}ds&=\frac{\gamma}{1-p}t^{1-p}\qquad\text{for } p>1,\end{aligned}$$

我们看到，上面定义的 $C(t)$ 在每种情况下都与 $C(t)$ 的定义一致。注意

$$\lim_{t\to 0+}C(t)=\begin{cases}0 & p\geq 1\\ 1 & 0<p<1.\end{cases}$$

现在我们将证明

$$\lim_{t\to 0+}V^0(t)=\lim_{t\to 0+}\frac{C(t)^2}{2}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2=\begin{cases}0 & \text{if } p\geq 1\\ \frac{\|\tilde{\mathbb{A}}(X_0)\|^2}{2} & \text{if } 0<p<1.\end{cases}\tag{32}$$

为此，我们首先证明对于 $t>0$

$$\lim_{\epsilon\to 0+}\int_\epsilon^t\frac{d}{ds}\left(\frac{C(s)^2\dot{\beta}(s)}{2}\right)\|X(s)-X_0\|^2\,ds<\infty.$$

我们先给出一个初等事实作为引理。

引理 E.3. 设 $f:(0,\infty)\to\mathbb{R}$ 是一个连续函数。假设存在 $q<1, 0<l<\infty$ 使得

$$\limsup_{t\to 0+}|f(t)t^q|<l.$$

那么对于 $t>0,\ f\in L^1([0,t],\ \mathbb{R})$。

证明。 由于 $\limsup_{s\to 0+} f(s)s^q = l$，存在 $\epsilon \in (0,\ t)$ 使得

$$0 < s < \epsilon \quad \Longrightarrow \quad |f(s)s^q| < 2|l|.$$

然后，由于 $q < 1$

$$\int_0^\epsilon |f(s)|\, ds = \int_0^\epsilon |f(s)s^q|\, \frac{1}{s^q} ds \le \int_0^\epsilon \frac{2|l|}{s^q} ds = \left[\frac{2|l|}{1-q} s^{1-q}\right]_0^\epsilon = \frac{2|l|}{1-q}\epsilon^{1-q} < \infty.$$

顺便说一下，由于 $f(t)$ 在 $[\epsilon, t]$ 上连续，$M = \max_{s\in[\epsilon,t]} |f(s)|$ 存在。因此，

$$\int_0^t |f(s)| ds = \int_0^\epsilon |f(s)| ds + \int_\epsilon^t |f(s)| ds \le \frac{2|l|}{1-q}\epsilon^{1-q} + M(t-\epsilon) < \infty.$$

□

应用引理 E.3，我们将证明对于 $t > 0$

$$\left| s^2 \frac{d}{ds}\left(C(s)^2 \dot\beta(s)\right)\right| \in L^1([0,t], \mathbb{R}). \tag{33}$$

注意

$$\frac{d}{ds}\left(C(s)^2\dot\beta(s)\right) = 2C(s)^2\beta(s)\dot\beta(s) + C(s)^2\ddot\beta(s) = C(s)^2\left(-\frac{2p\gamma^2}{s^{2p+1}} + \frac{p(1+p)\gamma}{s^{p+2}}\right). \tag{34}$$

(i) $p > 1$

我们首先证明对所有 $n > 0$，$\lim_{s\to 0+} C(s)^2 \frac{1}{s^n} = 0$。取 $n > 0$。于是存在某个 $k \in \mathbb{N}$，使得 $k(p-1) > n$。

通过变量代换 $u = \frac{1}{s}$ 并应用洛必达法则，我们得到

$$\begin{aligned}\lim_{s\to 0+} C(s)^2 \frac{1}{s^n} &= \lim_{u\to\infty} \frac{u^n}{e^{\frac{2\gamma}{p-1}u^{p-1}}} \\ &= \lim_{u\to\infty} \frac{nu^{n-1}}{2\gamma u^{p-2} e^{\frac{2\gamma}{p-1}u^{p-1}}} = \frac{n}{2\gamma}\lim_{u\to\infty}\frac{u^{n+1-p}}{e^{\frac{2\gamma}{p-1}u^{p-1}}} = \cdots = \frac{\prod_{m=0}^{k-1}(n-m(p-1))}{(2\gamma)^k}\lim_{u\to\infty}\frac{u^{n-k(p-1)}}{e^{\frac{2\gamma}{p-1}u^{p-1}}} = 0.\end{aligned} \tag{35}$$

因此

$$\lim_{s\to 0+}\left| s^2\frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\right| = \lim_{s\to 0+}\left|C(s)^2\left(-\frac{2p\gamma^2}{s^{2p-1}} + \frac{p(1+p)\gamma}{s^p}\right)\right| = 0.$$

根据引理 E.3，我们得出 $\int_0^t \left|s^2 \frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\right| ds < \infty$。

(ii) $p = 1$由于 $C(s) = s^\gamma$，

我们看到

$$\lim_{s\to 0+}\left| s^2\frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\cdot s^{1-\gamma}\right| = \lim_{s\to 0+}\left|s^2\cdot\gamma(\gamma-1)s^{2\gamma-3}\cdot s^{1-\gamma}\right| = \gamma|\gamma-1|\lim_{s\to 0+} s^\gamma = 0.$$

由于 $1-\gamma < 1$，根据引理 E.3，我们得出 $\int_0^t \left|s^2 \frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\right| ds < \infty$。

(iii) $0 < p < 1$

由于 $\lim_{s\to 0+} C(s) = \lim_{s\to 0+} e^{\frac{\gamma}{1-p}s^{1-p}} = 1$，我们看到

$$\limsup_{s\to 0+}\left| s^2\frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\cdot s^p\right| = \limsup_{s\to 0+} C(s)^2\left|-2p\gamma^2 s^{1-p} + p(1+p)\gamma\right| = p(1+p)\gamma.$$

由于 $p < 1$，根据引理 E.3，我们得出 $\int_0^t \left|s^2 \frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\right| ds < \infty$。

将推论 B.8 中的界记为 $M(T)$，我们知道对于 $0 < s < T$

$$\left\| \frac{X(s) - X_0}{s} \right\| \leq \frac{\int_0^s \left\| \dot{X}(u) \right\| du}{s} \leq \frac{\int_0^s M(T) du}{s} = M(T).$$

因此，应用 (33)，对于 $0 < t < T$，我们有

$$\int_0^t \left| \frac{d}{ds} \left(\frac{C(s)^2 \dot{\beta}(s)}{2} \right) \|X(s) - X_0\|^2 \right| ds \leq M(T)^2 \int_0^t \left| s^2 \frac{d}{ds} \left(\frac{C(s)^2 \dot{\beta}(s)}{2} \right) \right| ds < \infty.$$

因此我们知道 $V^0(t)$ 定义良好，并且由于 $\lim_{t\to 0+} \int_0^t \frac{d}{ds} \left(\frac{C(s)^2 \dot{\beta}(s)}{2} \right) \|X(s) - X_0\|^2 ds = 0$，我们有

$$\lim_{t\to 0+} V^0(t) = \lim_{t\to 0+} \frac{C(t)^2}{2} \left(\left\| \tilde{\mathbb{A}}(X(t)) \right\|^2 + 2\beta(t) \left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0 \right\rangle + \left(\beta(t)^2 + \dot{\beta}(t) \right) \|X(t) - X_0\|^2 \right).$$

现在我们可以证明所需结果了。

(i) $p > 1$ 如我们所知，根据引理 D.1 和推论 B.9，$X(t) - X_0$ 和 $\mathbb{A}(X(t))$ 有界。因此，由 (35) 可得

$$0 = \lim_{t\to 0+} C(t)^2 = \lim_{t\to 0+} C(t)^2 \beta(t) = \lim_{t\to 0+} C(t)^2 \left(\beta(t)^2 + \dot{\beta}(t) \right),$$

因此 $\lim_{t\to 0+} V^0(t) = 0$。

(ii) $0 < p \leq 1$

As $\left\| \frac{X(t) - X_0}{t} \right\| \leq M(T) < \infty$ for $0 < t < T$, we see

$$\limsup_{t\to 0+} C(t)^2 \beta(t) \left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0 \right\rangle \leq \gamma \limsup_{t\to 0+} C(t)^2 t^{1-p} \left\| \tilde{\mathbb{A}}(X(t)) \right\| M(T) = 0$$

$$\limsup_{t\to 0+} C(t)^2 \left(\beta(t)^2 + \dot{\beta}(t) \right) \|X(t) - X_0\|^2 \leq \gamma \limsup_{t\to 0+} C(t)^2 \left(\gamma t^{2-2p} - t^{1-p} \right) M(T)^2 = 0.$$

因此

$$\lim_{t\to 0+} V^0(t) = \lim_{t\to 0+} \frac{C(t)^2}{2} \left\| \tilde{\mathbb{A}}(X(t)) \right\|^2 = \begin{cases} 0 & \text{if } p = 1 \\ \frac{\|\tilde{\mathbb{A}}(X_0)\|^2}{2} & \text{if } 0 < p < 1. \end{cases}$$

由 (i) 和 (ii) 可得所需结论

$$V^0(0) = \lim_{t\to 0+} V^0(t) = \begin{cases} 0 & \text{if } p \geq 1 \\ \frac{\|\tilde{\mathbb{A}}(X_0)\|^2}{2} & \text{if } 0 < p < 1, \end{cases}$$

因此，对于一般的 $t_0 \geq 0$,

$$V(0) = \lim_{t\to 0+} V(t) = \lim_{t\to 0+} V^0(t) - \int_0^{t_0} \frac{d}{ds} \left(\frac{C(s)^2 \dot{\beta}(s)}{2} \right) \|X(s) - X_0\|^2 ds$$

是良定义的。

E.3.3 $V(t) \leq \lim_{n\to\infty} V_{\lambda_n}(0)$ 对 $t \in S$ 成立，且适用于一般的极大单调 $\mathbb{A}$

按引理 E.1 中的定义定义 S。取 $t \in S$，令 $T > t_0$。设 $\{\lambda_n\}_{n\in\mathbb{N}}$ 为正数列 λ_n，使得 $\lim_{n\to\infty} \lambda_n = 0$，$X_{\lambda_n}$ 在 $[0, T]$ 上一致地收敛到 X，且 $\dot{X}_{\lambda_n}$ 在 L^2 中弱地收敛到 $\dot{X}$（$[0, T]$, $\mathbb{R}^n$）。回想一下，命题 B.11 保证了此类序列的存在性。

回顾一下，我们将 X_λ 记作常微分方程 (19) 的解。将 V_λ 记作 V，对应于 $\mathbb{A} = \mathbb{A}_\lambda$ 的情况，即。

$$V_\lambda(t) = \frac{C(t)^2}{2}\left(\|\mathbb{A}_\lambda(X_\lambda(t))\|^2 + 2\beta(t)\left\langle \mathbb{A}_\lambda\left(X_\lambda(t)\right), X_\lambda(t) - X_0\right\rangle + \left(\beta(t)^2 + \dot\beta(t)\right)\|X_\lambda(t) - X_0\|^2\right)$$
$$- \int_{t_0}^t \frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\|X_\lambda(t) - X_0\|^2\, ds.$$

注意，等式 $\dot X(t) = -\mathbb{A}(X(t)) - \beta(t)(X(t) - X_0)$ 成立，因为 $t \in S$，因此我们得到

$$V(t) = \frac{C(t)^2}{2}\left(\left\|\dot X(t)\right\|^2 + \dot\beta(t)\|X(t) - X_0\|^2\right) - \int_{t_0}^t \frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\|X(s) - X_0\|^2\, ds.$$

本节的目标是证明

$$V(t) \le \limsup_{n\to\infty} V_{\lambda_n}(t) \le \limsup_{n\to\infty} V_{\lambda_n}(0) = \lim_{n\to\infty} V_{\lambda_n}(0).$$

(1) $V(t) \le \limsup_{n\to\infty} V_{\lambda_n}(t)$
首先注意，由引理 B.12 可知

$$\frac{\|X_\lambda(s) - X_0\|}{s} \le \frac{\int_0^s \left\|\dot X_\lambda(s)\right\| ds}{s} \le \frac{\int_0^s M_{dot}(T) ds}{s} = M_{dot}(T)$$

对 $s \le T$ 成立。因此，由 (33) 可得

$$\left|\frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\|X_\lambda(s) - X_0\|^2\right| \le \left|s^2\frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\right| M_{dot}(T)^2 \in L^1([0,T],\mathbb{R}).$$

因此，应用控制收敛定理，对于 $t_0 \in [0,T]$，我们有

$$\lim_{n\to\infty}\int_{t_0}^t \frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\|X_{\lambda_n}(s) - X_0\|^2\, ds = \int_{t_0}^t \frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\|X(s) - X_0\|^2\, ds. \tag{36}$$

根据初等分析，我们可以很容易地验证，当 $\lim_{n\to\infty} b_n$ 存在时，$\limsup_{n\to\infty}(a_n + b_n) = \limsup_{n\to\infty} a_n + \lim_{n\to\infty} b_n$ 成立。因此我们有

$$\limsup_{n\to\infty} V_{\lambda_n}(t) - V(t) = \frac{C(t)^2}{2}\left(\limsup_{n\to\infty}\left\|\dot X_{\lambda_n}(t)\right\|^2 - \left\|\dot X(t)\right\|^2\right).$$

因此，只需证明以下引理。

引理 E.4. 假设对于 $T > 0$，序列 $\{\lambda_n\}_{n\in\mathbb{N}}$，$X_{\lambda_n}$ 在 $[0,T]$ 上一致收敛到 X，且 X_{λ_n} 在 $L^2([0,T], \mathbb{R}^n)$ 中弱收敛到 $\dot X$。设 $t \in S \cap [0,T]$。则以下不等式成立

$$\left\|\dot X(t)\right\| \le \limsup_{n\to\infty}\left\|\dot X_{\lambda_n}(t)\right\|.$$

证明。(i) 对几乎处处的 $s \in [0,T]$，$\left\|\dot X(s)\right\| \le \limsup_{n\to\infty}\left\|\dot X_{\lambda_n}(s)\right\|$ 成立。

设 D 为可测子集 $D \subset [0,T]$。由于 $\dot X_{\lambda_n} \rightharpoonup \dot X$ 属于 $L^2([0,T], \mathbb{R}^n)$，且 $\chi_D\dot X \in L^2([0,T], \mathbb{R}^n)$，我们有

$$\int_D \left\|\dot X(s)\right\|^2 ds = \int_0^T \left\langle \dot X(s), \chi_D(s)\dot X(s)\right\rangle ds = \lim_{n\to\infty}\int_0^T \left\langle \dot X_{\lambda_n}(s), \chi_D(s)\dot X(s)\right\rangle ds$$
$$= \limsup_{n\to\infty}\int_D \left\langle \dot X_{\lambda_n}(s), \dot X(s)\right\rangle ds \le \limsup_{n\to\infty}\int_D \left\|\dot X_{\lambda_n}(s)\right\|\left\|\dot X(s)\right\| ds.$$

该不等式等式来自柯西–施瓦茨不等式。现在根据反向 Fatou 引理，我们有 e

$$\limsup_{n\to\infty}\int_D \left\|\dot X_{\lambda_n}(s)\right\|\left\|\dot X(s)\right\| ds \le \int_D \limsup_{n\to\infty}\left\|\dot X_{\lambda_n}(s)\right\|\left\|\dot X(s)\right\| ds.$$

因此，结合这两个不等式，我们有

$$\int_D \left(\limsup_{n\to\infty} \left\| \dot{X}_{\lambda_n}(s) \right\| - \left\| \dot{X}(s) \right\| \right) \left\| \dot{X}(s) \right\| ds \geq 0.$$

由于 D 是 $[0,T]$ 的任意可测子集，我们得出结论：对于几乎所有 $s \in [0,T]$

$$\left(\limsup_{n\to\infty} \left\| \dot{X}_{\lambda_n}(s) \right\| - \left\| \dot{X}(s) \right\| \right) \left\| \dot{X}(s) \right\| \geq 0 \quad \Longrightarrow \quad \left\| \dot{X}(s) \right\| \leq \limsup_{n\to\infty} \left\| \dot{X}_{\lambda_n}(s) \right\|.$$

(ii) $\left\| \dot{X}(t) \right\| \leq$ 上极限$_{n\to\infty} \left\| \dot{X}_{\lambda_n}(t) \right\|$ 对于 $t \in S \cap [0,T]$。

Let $t \in S \cap [0,T]$。那么，对于满足 $t+h<T$ 的 $h>0$，由于 S 是零测度集，由 (i) 可得 e

$$\int_t^{t+h} \left\| \dot{X}(s) \right\| ds \leq \int_t^{t+h} \limsup_{n\to\infty} \left\| \dot{X}_{\lambda_n}(s) \right\| ds.$$

现在，对于某个 $a>0$，考虑

$$U_{\lambda_n}(s) = \left\| \dot{X}_{\lambda_n}(s) \right\|^2 + \underbrace{\frac{\gamma p}{s^{p+1}} \left\| X_{\lambda_n}(s) - X_0 \right\|^2 + \int_a^s \frac{\gamma p(1-p)}{u^{p+2}} \left\| X_{\lambda_n}(u) - X_0 \right\|^2 du}_{=f_n(s)}.$$

那么，根据引理 B.5 的证明，我们知道

$$\dot{U}_{\lambda_n}(s) = -2 \left\langle \dot{X}_{\lambda_n}(s), \frac{d}{ds} \mathbb{A}_{\lambda_n}(X(s)) \right\rangle - \frac{2\gamma}{s^p} \left\| \dot{X}_{\lambda_n}(s) - \frac{p}{s}(X_{\lambda_n}(s) - X_0) \right\|^2 \leq 0$$

对于几乎所有 s 都成立，因此 U_{λ_n} 是非递增的。

顺便，由于 X_{λ_n} 在 $[0,T]$ 上一致收敛到 X，根据控制收敛定理，我们有

$$\lim_{n\to\infty} f_n(s) = \frac{\gamma p}{s^{p+1}} \left\| X(s) - X_0 \right\|^2 + \int_a^s \frac{\gamma p(1-p)}{u^{p+2}} \left\| X(u) - X_0 \right\|^2 du.$$

记 $f(s)$ 为 $= \lim_{n\to\infty} f_n(s)$。

利用上述事实，对于 $s \in [t, t+h]$，我们有

$$\begin{aligned} \limsup_{n\to\infty} \left\| \dot{X}_{\lambda_n}(s) \right\| &= \limsup_{n\to\infty} \sqrt{U_{\lambda_n}(s) - f_n(s)} \\ &\leq \limsup_{n\to\infty} \sqrt{U_{\lambda_n}(t) - f_n(s)} = \sqrt{\limsup_{n\to\infty} U_{\lambda_n}(t) - f(s)}. \end{aligned} \tag{37}$$

因此，

$$\|X(t+h) - X(t)\| \leq \int_t^{t+h} \left\| \dot{X}(s) \right\| ds \leq \int_t^{t+h} \limsup_{n\to\infty} \left\| \dot{X}_{\lambda_n}(s) \right\| ds \leq \int_t^{t+h} \sqrt{\limsup_{n\to\infty} U_{\lambda_n}(t) - f(s)} ds.$$

注意，$\sqrt{\limsup_{n\to\infty} U_{\lambda_n}(t) - f(s)}$ 关于 s 是连续函数。因此，我们有

$$\begin{aligned} \lim_{h\to 0} \frac{1}{h} \int_t^{t+h} \sqrt{\limsup_{n\to\infty} U_{\lambda_n}(t) - f(s)} ds &= \sqrt{\limsup_{n\to\infty} U_{\lambda_n}(t) - f(t)} \\ &= \sqrt{\limsup_{n\to\infty} \left\| \dot{X}_{\lambda_n}(t) \right\|^2 + \lim_{n\to\infty} f_n(t) - f(t)} = \limsup_{n\to\infty} \left\| \dot{X}_{\lambda_n}(t) \right\|. \end{aligned}$$

最后，由于 $t \in S$，$\dot{X}(t) = \lim_{h\to 0} \frac{X(t+h) - X(t)}{h}$ 存在。因此

$$\begin{aligned} \left\| \dot{X}(t) \right\| &= \lim_{h\to 0+} \frac{\|X(t+h) - X(t)\|}{h} \\ &\leq \lim_{h\to 0+} \frac{1}{h} \int_t^{t+h} \sqrt{\limsup_{n\to\infty} U_{\lambda_n}(t) - f(s)} ds = \limsup_{n\to\infty} \left\| \dot{X}_{\lambda_n}(t) \right\|, \end{aligned}$$

由此得出所需结果。

□

(2) $\limsup_{n\to\infty} V_{\lambda_n}(t) \le \limsup_{n\to\infty} V_{\lambda_n}(0)$ 由于 $\mathbb{A}_{\lambda_n}$ 是利普希茨连续的，根据附录 E.3.1，我们知道 V_{λ_n} 是非递增的，因此

$$V_{\lambda_n}(t) \le \lim_{\epsilon\to 0+} V_{\lambda_n}(\epsilon) = V_{\lambda_n}(0).$$

对两边取上极限，即得所需结果。

(3) $\limsup_{n\to\infty} V_{\lambda_n}(0) = \lim_{n\to\infty} V_{\lambda_n}(0)$ 根据

(32) 和引理 B.10 的 (ii)，我们知道

$$\lim_{n\to\infty} V^0_{\lambda_n}(0) = \begin{cases} 0 & \text{if } p \ge 1 \\ \frac{\|m(\mathbb{A}(X_0))\|^2}{2} & \text{if } 0 < p < 1. \end{cases}$$

应用 (36) 可得

$$\lim_{n\to\infty} V_{\lambda_n}(0) = \lim_{n\to\infty} V^0_{\lambda_n}(0) - \int_0^{t_0} \frac{d}{ds}\left(\frac{C(s)^2\dot\beta(s)}{2}\right)\|X(s) - X_0\|^2\, ds.$$

由于极限 $\lim_{n\to\infty} V_{\lambda_n}(0)$ 存在，上极限与极限相同。

E.3.4 一般极大单调 $\mathbb{A}$的证明

首先，我们证明 $V(t) \le \lim_{n\to\infty} V_{\lambda_n}(0)$ 对 $t \in [0,\infty)$ 成立。接着，我们证明 $\lim_{n\to\infty} V_{\lambda_n}(0) = \lim_{t\to 0+} V(t)$。于是可得

$$V(t) \le \lim_{n\to\infty} V_{\lambda_n}(0) = \lim_{t\to 0+} V(t) = V(0),$$

这就是我们所需的结果。

(i) $V(t) \le \lim_{n\to\infty} V_{\lambda_n}(0)$ 对 $t \in [0,\infty)$ 成立

取 $t \in [0,\infty)$。我们知道当 $t \in S$ 时不等式成立，因此假设 $t \notin S$。然后根据引理 E.1，我们知道 $\tilde{\mathbb{A}}(X(t))$ 存在一个序列 $\{t_k\}_{k\in\mathbb{N}}$，使得 $t_k \in S$，$\lim_{k\to\infty} t_k = t$，且 $\mathbb{A}(X(t_k))$ 收敛到 $\tilde{\mathbb{A}}(X(t))$。由于 $t_k \in S$，$V(t_k) \le \lim_{n\to\infty} V_{\lambda_n}(0)$ 成立。由于 $\lim_{k\to\infty} V(t_k) = V(t)$，对不等式取极限 $k\to\infty$，即可得到所需结果。

(ii) $\lim_{n\to\infty} V_{\lambda_n}(0) = \lim_{t\to 0+} V(t)$

取一个序列 $\{t_k\}_{k\in\mathbb{N}}$，使得 $t_k > 0$ 且 $\lim_{k\to\infty} t_k = 0$。我们希望证明 $\lim_{k\to\infty} V(t_k) = \lim_{n\to\infty} V_{\lambda_n}(0)$。注意，附录 E.3.2 中的论证在以下条件下仍然成立：对所有 $k \in \mathbb{N}$，~~$\|\tilde{\mathbb{A}}(X(t_k))\|$~~ 和 $\|X_{(t_k)_{t_k}} - X_0\|$ 有界。$\|\tilde{\mathbb{A}}(X(t_k))\|$ 的有界性来自推论 E.2。根据引理 B.12，我们有

$$\frac{\|X(t_k) - X_0\|}{t_k} = \lim_{\lambda\to 0+} \frac{\|X_\lambda(t_k) - X_0\|}{t_k} \le \lim_{\lambda\to 0+} \frac{\int_0^{t_k} \left\|\dot X_\lambda(s)\right\| ds}{t_k} \le M_{dot}(t_k) \le \sup_{k\in\mathbb{N}} M_{dot}(t_k).$$

因此，应用附录 E.3.2 中的论证，我们有

$$\lim_{k\to\infty} V^0(t_k) = \lim_{k\to\infty} \frac{C(t_k)^2}{2}\left\|\tilde{\mathbb{A}}(X(t_k))\right\|^2.$$

因此，接下来只需证明右侧的极限存在，且等于 $\lim_{n\to\infty} V^0_{\lambda_n}(0)$。

对于 $p \ge 1$，我们知道 $\lim_{k\to\infty} C(t_k)^2 = 0$，因此 $\lim_{k\to\infty} \frac{C(t_k)^2}{2}\left\|\tilde{\mathbb{A}}(X(t_k))\right\|^2 = 0$，因为 $\|\tilde{\mathbb{A}}(X(t_k))\|$ 有界。由 (32) 可知 $\lim_{n\to\infty} V^0_{\lambda_n}(0) = 0$，证毕。

现在考虑 $0 < p < 1$ 的情形。设 $\tilde{\mathbb{A}}(X(t_{k_l}))$ 是 $\mathbb{A}(X(t_k))$ 的一个收敛子列。首先，由 (i) 可知

$$V^0(t_{k_l}) \le \lim_{n\to\infty} V^0_{\lambda_n}(0) = \frac{\|m(\mathbb{A}(X_0))\|^2}{2}.$$

由上述不等式，回想 $\lim_{l\to\infty} C(t_{k_l})^2 = 1$，我们有

$$\frac{\left\|\lim_{l\to\infty} \tilde{\mathbb{A}}(X(t_{k_l}))\right\|^2}{2} = \lim_{l\to\infty} V^0(t_{k_l}) \le \frac{\|m(\mathbb{A}(X_0))\|^2}{2}.$$

˜顺便一提，由于 $\lim_{l\to\infty} X(t_{k_l}) = X_0$ ，且 $\mathbb{A}(X(t_{k_l})) \in \mathbb{A}(X(t_{k_l}))$，根据闭图像定理，我们有˜lim $_{l\to\infty} \mathbb{A}(X(t_{k_l})) \in \mathbb{A}(X_0)$。由于 $m(\mathbb{A}(X_0))$ 是 $\mathbb{A}(X_0)$ 中范数最小的元素，我们有˜$\lim_{l\to\infty} \mathbb{A}(X(t_{k_l})) = m(\mathbb{A}($ $X_0))$。由于所有收敛子列都收敛到同一个极限 $m(\mathbb{A}(X_0))$，我们˜得出 $\lim_{k\to\infty} \mathbb{A}(X(t_k)) = m(\mathbb{A}(X_0))$。

因此

$$\lim_{k\to\infty} V^0(t_k) = \lim_{k\to\infty} \frac{C(t_k)^2}{2} \left\|\tilde{\mathbb{A}}(X(t_k))\right\|^2 = \frac{\|m(\mathbb{A}(X_0))\|^2}{2} = \lim_{n\to\infty} V^0_{\lambda_n}(0).$$

由于 t_k 是任意的正数列且收敛于 0，我们得出 $\lim_{n\to\infty} V^0_{\lambda_n}(0) = \lim_{t\to 0+} V^0(t) = V^0(0)$。因此

$$\lim_{n\to\infty} V_{\lambda_n}(0) = V^0(0) - \int_0^{t_0} \frac{d}{ds}\left(\frac{C(s)^2\dot{\beta}(s)}{2}\right) \|X(s) - X_0\|^2 \, ds = \lim_{t\to 0+} V(t).$$

E.4 引理 3.4 的证明

证明的大部分已在正文中完成。回忆一下， $\Phi(t)$ 定义为 $\Phi(t) = \left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\beta(t)\left\langle\tilde{\mathbb{A}}(X(t)),\, X(t) - X_0\right\rangle$，并且由 (8) 可得

$$\Phi(t) \geq \frac{1}{2}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 - 2\beta(t)^2 \|X_0 - X_\star\|^2.$$

另一方面，

$$\frac{2V(t)}{C(t)^2} - \Phi(t) = \left(\beta(t)^2 + \dot{\beta}(t)\right)\|X(t) - X_0\|^2 - \frac{2}{C(t)^2}\int_{t_0}^t \frac{d}{ds}\left(\frac{C(s)^2\dot{\beta}(s)}{2}\right)\|X(s) - X_0\|^2 \, ds.$$

注意，$C(t) \neq 0$ 若 $t > 0$，我们可以除以 $C(t)^2$。现在由推论 3.3 可得，对于 $t > 0$， $V(t) \geq V(0)$，因此$\frac{V(t)}{C(t)^2} - \frac{V(0)}{C(t)^2} \geq 0$。因此对于 $t > \epsilon$

$$\begin{aligned}
&\frac{2V(0)}{C(t)^2} - \left(\beta(t)^2 + \dot{\beta}(t)\right)\|X(t) - X_0\|^2 + \frac{2}{C(t)^2}\int_{t_0}^t \frac{d}{ds}\left(\frac{C(s)^2\dot{\beta}(s)}{2}\right)\|X(s) - X_0\|^2 \, ds \\
&= \frac{2V(0)}{C(t)^2} - \left(\frac{2V(t)}{C(t)^2} - \Phi(t)\right) \\
&= \left(\frac{2V(0)}{C(t)^2} - \frac{2V(t)}{C(t)^2}\right) + \Phi(t) \geq \Phi(t) \geq \frac{1}{2}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 - 2\beta(t)^2\|X_0 - X_\star\|^2.
\end{aligned}$$

将 $2\beta(t)^2\|X_0 - X_\star\|^2$ 移至等式左边并乘以 2，即得所需结果

$$\begin{aligned}
\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 \leq 4\beta(t)^2\|X_0 - X_\star\|^2 + \frac{4V(0)}{C(t)^2} - 2\left(\beta(t)^2 + \dot{\beta}(t)\right)\|X(t) - X_0\|^2 \\
+ \frac{2}{C(t)^2}\int_{t_0}^t \frac{d}{ds}\left(C(s)^2\dot{\beta}(s)\right)\|X(s) - X_0\|^2 \, ds.
\end{aligned}$$

E.5 定理 3.1 的证明

重述引理 3.4，我们知道对于 $t > 0$,

$$\begin{aligned}
\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 \leq 4\beta(t)^2\|X_0 - X_\star\|^2 + \frac{2V(0)}{C(t)^2} - 2\left(\beta(t)^2 + \dot{\beta}(t)\right)\|X(t) - X_0\|^2 \\
+ \frac{2}{C(t)^2}\int_{t_0}^t \frac{d}{ds}\left(C(s)^2\dot{\beta}(s)\right)\|X(s) - X_0\|^2 \, ds.
\end{aligned} \tag{7}$$

如引理 D.1 所述，我们知道 $\|X(t) - X_0\| \le 2\|X_0 - X_\star\|$，显然

$$\begin{aligned} 4\beta(t)^2\|X_0 - X_\star\|^2 &= \mathcal{O}\left(\beta(t)^2\right) \\ \frac{2V(0)}{C(t)^2} &= \mathcal{O}\left(\frac{1}{C(t)^2}\right) \\ 2\left(\beta(t)^2 + \dot\beta(t)\right)\|X(t) - X_0\|^2 &= \mathcal{O}\left(\beta(t)^2\right) + \mathcal{O}\left(\dot\beta(t)\right). \end{aligned} \tag{38}$$

因此，接下来只需证明

$$\frac{2}{C(t)^2}\int_{t_0}^{t}\frac{d}{ds}\left(C(s)^2\dot\beta(s)\right)\|X(s) - X_0\|^2 ds = \mathcal{O}\left(\beta(t)^2\right) + \mathcal{O}\left(\frac{1}{C(t)^2}\right) + \mathcal{O}\left(\dot\beta(t)\right).$$

我们可以验证，存在 $T > 0$，使得对于 $s > T$，$\frac{d}{ds}\left(C(s)^2\dot\beta(s)\right)$ 的符号在 $\beta(t) = \frac{\gamma}{t^p}$ 且满足 $\gamma > 0, p > 0$ 时不变。我们首先在假设满足这一条件的前提下进行证明。这个条件的作用是保证对于$t > T$，以下等式成立。

$$\left|\int_T^t \frac{d}{ds}\left(C(s)^2\dot\beta(s)\right)ds\right| = \int_T^t \left|\frac{d}{ds}\left(C(s)^2\dot\beta(s)\right)\right|ds.$$

应用上述等式并利用 $\|X(t) - X_0\| \le 2\|X_0 - X_\star\|$，可得

$$\begin{aligned} &\left|\frac{2}{C(t)^2}\int_{t_0}^{t}\frac{d}{ds}\left(C(s)^2\dot\beta(s)\right)\|X(s) - X_0\|^2 ds\right| \\ &\le \frac{2}{C(t)^2}\underbrace{\left|\int_{t_0}^{T}\frac{d}{ds}\left(C(s)^2\dot\beta(s)\right)\|X(s) - X_0\|^2 ds\right|}_{=M} + \frac{2}{C(t)^2}\int_T^t\left|\frac{d}{ds}\left(C(s)^2\dot\beta(s)\right)\|X(s) - X_0\|^2\right|ds \\ &\le \frac{2M}{C(t)^2} + \frac{2}{C(t)^2}\int_T^t\left|\frac{d}{ds}\left(C(s)^2\dot\beta(s)\right)4\|X_0 - X_\star\|^2\right|ds \\ &= \frac{2M}{C(t)^2} + \frac{8\|X_0 - X_\star\|^2}{C(t)^2}\left|C(t)^2\dot\beta(t) - C(T)^2\dot\beta(T)\right| \\ &\le \frac{2}{C(t)^2}\left(M + 4\|X_0 - X_\star\|^2 C(T)^2\left|\dot\beta(T)\right|\right) + 8\|X_0 - X_\star\|^2\left|\dot\beta(t)\right| = \mathcal{O}\left(\frac{1}{C(t)^2}\right) + \mathcal{O}\left(\dot\beta(t)\right). \end{aligned}$$

因此，由 (7) 和 (38) 可得

$$\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 = \mathcal{O}\left(\beta(t)^2\right) + \mathcal{O}\left(\frac{1}{C(t)^2}\right) + \mathcal{O}\left(\dot\beta(t)\right).$$

接下来只需证明，存在 $T > 0$，使得对于 $s > T$，$\frac{d}{ds}\left(C(s)^2\dot\beta(s)\right)$ 的符号不变。通过简单但稍显繁琐的计算即可证明这一点。由于 $C(t) = C(t)\beta(t)$，可得

$$\frac{d}{ds}\left(C(s)^2\dot\beta(s)\right) = 2C(s)^2\beta(s)\dot\beta(s) + C(s)^2\ddot\beta(s) = C(s)^2\left(2\beta(s)\dot\beta(s) + \ddot\beta(s)\right).$$

因此，只需检查 $2\beta(s)\dot\beta(s) + \ddot\beta(s)$ 的符号。注意

$$2\beta(s)\dot\beta(s) + \ddot\beta(s) = -\frac{2p\gamma^2}{s^{2p+1}} + \frac{p(1+p)\gamma}{s^{p+2}} = \begin{cases} \frac{2}{s^3}\gamma(1-\gamma) & \text{if } p = 1 \\ \frac{p(p+1)\gamma}{s^{2p+1}}\left(s^{p-1} - \frac{2\gamma}{p+1}\right) & \text{if } p \ne 1. \end{cases}$$

因此，当 $p = 1$ 时，对于所有 $s > 0$，若 $\gamma \ge 1$，则它非正；若 $0 < \gamma < 1$，则它非负。

当 $p \ne 1$时，我们看到

$$\lim_{s\to 0+} s^{p-1} = \begin{cases} \infty & \text{if } 0 < p < 1 \\ 0 & \text{if } p > 1 \end{cases}, \qquad \lim_{s\to\infty} s^{p-1} = \begin{cases} 0 & \text{if } 0 < p < 1 \\ \infty & \text{if } p > 1 \end{cases}$$

因此，根据介值定理，存在 $T > 0$ 使得 $T^{p-1} - \frac{2\gamma}{p+1} = 0$；对于该 T，当 $0 < p < 1$ 时，$s^{p-1} - \frac{2\gamma}{p+1}$ 对 $s > T$ 非正；当 $p > 1$时，$s^{p-1} - \frac{2\gamma}{p+1}$ 对 $s > T$ 非负。证明至此完成。

E.5.1 表 1 中收敛速率的证明

回忆一下，由 (31) 可知

$$C(t)=\begin{cases}t^{\gamma} & p=1\\ e^{\frac{\gamma}{1-p}t^{1-p}} & p>0,\,p\neq 1.\end{cases}$$

现在我们观察到，$\beta(t)$ 不会影响收敛速率。换言之，我们证明，与 $\beta(t)^2$ 和 $C\left(\frac{1}{t}\right)^2$ 相比，$\beta(t)$ 并非在所有情况下趋于零的最慢项。

(i) $0<p\leq 1$比较 $\dot{\beta}(t)=-\frac{p\gamma}{t^{p+1}}$ 与 $\beta(t)^2=\frac{\gamma^2}{t^{2p}}$，可见当 $0<p\leq 1$ 时，$\dot{\beta}(t)=\mathcal{O}\left(\beta(t)^2\right)$。(ii) $p>1$ 对于 $p>1$，我们有 $\lim_{t\to\infty} C\left(\frac{1}{t}\right)^2=1\neq 0$。由于 $\lim_{t\to\infty}\dot{\beta}(t)=-\lim_{t\to\infty}\frac{p\gamma}{t^{p+1}}=0$，因此有 $\dot{\beta}(t)=\mathcal{O}\left(C\left(\frac{1}{t}\right)^2\right)$。

由 (i) 和 (ii)，我们得出

$$\left\|\tilde{\mathbb{A}}(X(t))\right\|^2=\mathcal{O}\left(\beta(t)^2\right)+\mathcal{O}\left(\frac{1}{C(t)^2}\right).$$

现在，表 1 中的结果显而易见。

- $p=1$
  当 $p=1$ 时，我们有 $C(t)=t^{\gamma}$。比较 $\frac{1}{C(t)^2}=\frac{1}{t^{2\gamma}}$ 以及 $\beta(t)^2=\frac{\gamma^2}{t^2}$，
  (1) $\left\|\tilde{\mathbb{A}}(X(t))\right\|^2=\mathcal{O}\left(\beta(t)^2\right)=\mathcal{O}\left(\frac{1}{t^2}\right)$ if $\gamma\geq 1$.
  (2) $\left\|\tilde{\mathbb{A}}(X(t))\right\|^2=\mathcal{O}\left(\frac{1}{C(t)^2}\right)=\mathcal{O}\left(\frac{1}{t^{2\gamma}}\right)$ if $0<\gamma<1$.
- $0<p<1$
  当 $0<p<1$时，我们有
  $$\lim_{t\to\infty}\frac{1}{\beta(t)^2}\cdot\frac{1}{C(t)^2}=\frac{1}{\gamma^2}\lim_{t\to\infty}\frac{t^2}{e^{\frac{2\gamma}{1-p}t^{1-p}}}=0.$$
  因此，$\left\|\tilde{\mathbb{A}}(X(t))\right\|^2=\mathcal{O}\left(\beta(t)^2\right)=\mathcal{O}\left(\frac{1}{t^{2p}}\right)$。
- $p>1$
  我们此前观察到此时 $\lim_{t\to\infty} C\left(\frac{1}{t}\right)^2=1\neq 0$ 当 $p>1$。因此 $\left\|\tilde{\mathbb{A}}(X(t))\right\|^2=\mathcal{O}\left(C\left(\frac{1}{t}\right)^2\right)=\mathcal{O}(1)$.

E.6 定理3.5的证明

Name

$$\bar{X}_{\star}=\Pi_{\operatorname{Zer}\mathbb{A}}(X_0)=\operatorname*{argmin}_{z\in\operatorname{Zer}\mathbb{A}}\|z-X_0\|.$$

我们首先证明

$$\limsup_{t\to\infty}\left\langle X(t)-\bar{X}_{\star},X_0-\bar{X}_{\star}\right\rangle\leq 0.$$

反证法。假设结论不成立。
那么存在 $\epsilon>0$，使得对于每个 $k\in\mathbb{N}$，都存在 $n_k\in[0,\infty)$，使得

$$\left\langle X(n_k)-\bar{X}_{\star},X_0-\bar{X}_{\star}\right\rangle>\epsilon$$

顺便说一下，由引理 D.1 可知 $X(n_k)\in B_{\|X_0-\bar{X}_{\star}\|}(X_{\star})$。因此，根据博尔扎诺-魏尔斯特拉斯定理，存在一个收敛子序列 $\{X\left(n_{k(l)}\right)\}_{l\in(N)}$。将其极限记为 X_{∞}。由于 $\lim_{t\to\infty}\left\|\tilde{\mathbb{A}}(X(t))\right\|=0$，$\{\tilde{\mathbb{A}}(X(n_{k(l)})\}_{l\in(N)}$ 收敛于 0。于是根据 [15, 命题 20.38], $(X_{\infty},\,0)\in\operatorname{Gra}\mathbb{A}$，即 $X_{\infty}\in\operatorname{Zer}\mathbb{A}$。

另一方面，由于 $\mathbb{A}$ 是极大单调的，根据 [15, 命题 23.39] Zer $\mathbb{A}$ 是闭凸集。由于 $X_\star = \Pi_{\mathrm{Zer}\,\mathbb{A}}(X_0)$，根据 [15, 定理 3.16] 可得 $\left\langle X_\infty - \bar{X}_\star,\ X_0 - \bar{X}_\star \right\rangle \le 0$。因此

$$0 < \epsilon \le \lim_{l\to\infty} \left\langle X\left(n_{k(l)}\right) - \bar{X}_\star, X_0 - \bar{X}_\star \right\rangle = \left\langle X_\infty - \bar{X}_\star, X_0 - \bar{X}_\star \right\rangle \le 0,$$

这是一个矛盾，因此我们得到了所需结果。

现在我们证明 $\lim_{t\to\infty} \left\| X(t) - \bar{X}_\star \right\| = 0$。回顾一下，对于 $t > 0$

$$\dot{X} \in -\mathbb{A}(X(t)) - \beta(t)(X - X_0).$$

由 $C(t) = C(t)\beta(t)$，以及 $\mathbb{A}$ 的单调性，我们观察到

$$\begin{aligned}
\frac{d}{dt}\left(C(t)^2 \left\| X(t) - \bar{X}_\star \right\|^2\right) &= 2C(t)^2 \left\langle \dot{X}(t), X(t) - \bar{X}_\star \right\rangle + 2C(t)^2\beta(t) \left\| X(t) - \bar{X}_\star \right\|^2 \\
&= 2C(t)^2 \left\langle -\tilde{\mathbb{A}}(X(t)) - \beta(t)(X(t) - X_0), X(t) - \bar{X}_\star \right\rangle + 2C(t)^2\beta(t) \left\| X(t) - \bar{X}_\star \right\|^2 \\
&\le -2C(t)^2\beta(t) \left\langle X(t) - X_0, X(t) - \bar{X}_\star \right\rangle + 2C(t)^2\beta(t) \left\| X(t) - \bar{X}_\star \right\|^2 \\
&= 2C(t)^2\beta(t) \left\langle X_0 - \bar{X}_\star, X(t) - \bar{X}_\star \right\rangle.
\end{aligned}$$

现在取 ϵ。由 $\limsup_{t\to\infty} \left\langle X(t) - \bar{X}_\star, X_0 - \bar{X}_\star \right\rangle \le 0$，存在 $M > 0$ 使得

$$t > M \quad \Longrightarrow \quad \left\langle X_0 - \bar{X}_\star, X(t) - \bar{X}_\star \right\rangle < \epsilon.$$

对于该 M 和 $t > M$，从 M 积分到 t，我们得到

$$\begin{aligned}
\left[C(s)^2 \left\| X(s) - \bar{X}_\star \right\|^2\right]_M^t &= C(t)^2 \left\| X(t) - \bar{X}_\star \right\|^2 - C(M)^2 \left\| X(M) - \bar{X}_\star \right\|^2 \\
&\le \int_M^t 2C(s)^2\beta(s) \left\langle X_0 - \bar{X}_\star, X(s) - \bar{X}_\star \right\rangle ds \\
&\le \int_M^t \left| 2C(s)^2\beta(s) \left\langle X_0 - \bar{X}_\star, X(s) - \bar{X}_\star \right\rangle \right| ds \\
&\le \int_M^t 2C(s)^2\beta(s)\epsilon\, ds = \epsilon \left[C(s)^2\right]_M^t = \epsilon \left(C(t)^2 - C(M)^2\right).
\end{aligned}$$

将 $C(t)^2$ 两边同时除以并整理可得，

$$\left\| X(t) - \bar{X}_\star \right\|^2 \le \epsilon \left(1 - \left(\frac{C(M)}{C(t)}\right)^2\right) + \left(\frac{C(M)}{C(t)}\right)^2 \left\| X(M) - \bar{X}_\star \right\|^2.$$

顺带一提，根据假设，我们有 $\lim_{t\to\infty} \frac{1}{C(t)} = 0$。因此我们得出

$$\lim_{t\to\infty} \left\| X(t) - \bar{X}_\star \right\|^2 \le \epsilon.$$

由于 $\epsilon > 0$ 是任意的，我们得到所需结果。

F 最坏情况示例的证明

最坏情况示例中关键地用到了线性 A 的显式解。因此，我们先给出该显式解的证明。

F.1 线性 A显式解的证明

F.1.1 引理 4.1 的证明

观察

$$X(t) = \sum_{n=0}^{\infty} \frac{(-tA)^n}{\Gamma(n+\gamma+1)} \Gamma(\gamma+1) X_0 = X_0 + \sum_{n=1}^{\infty} \frac{(-tA)^n}{\Gamma(n+\gamma+1)} \Gamma(\gamma+1) X_0,$$

我们可以验证 $X(0) = X_0$。现在利用伽马函数的性质 $\Gamma(x+1) = x\Gamma(x)$，并注意求和指标的下界，我们得到

$$\begin{aligned}\dot{X}(t) &= \sum_{n=1}^{\infty} \frac{(-nA)(-tA)^{n-1}}{\Gamma(n+\gamma+1)}\Gamma(\gamma+1)X_0 = A\sum_{n=1}^{\infty} \frac{(-n-\gamma+\gamma)(-tA)^{n-1}}{\Gamma(n+\gamma+1)}\Gamma(\gamma+1)X_0 \\ &= -A\sum_{n=1}^{\infty} \frac{(-tA)^{n-1}}{\Gamma(n+\gamma)}\Gamma(\gamma+1)X_0 + \gamma\sum_{n=1}^{\infty} \frac{(-t)^{n-1}A^n}{\Gamma(n+\gamma+1)}\Gamma(\gamma+1)X_0 \\ &= -A\sum_{n=0}^{\infty} \frac{(-tA)^{n}}{\Gamma(n+\gamma+1)}\Gamma(\gamma+1)X_0 + \frac{\gamma}{(-t)}\sum_{n=1}^{\infty} \frac{(-t)^{n}A^n}{\Gamma(n+\gamma+1)}\Gamma(\gamma+1)X_0 = -AX(t) - \frac{\gamma}{t}\left(X(t) - X_0\right).\end{aligned}$$

该解还可以写成另一种形式

$$X(t) = \gamma e^{-tA} t^{-\gamma} \int_0^t u^{\gamma-1} e^{uA}\, du\, X_0. \tag{39}$$

由于这是式 (9) 的一个特例，这里我们仅简要验证它满足常微分方程，并在引理 4.2 的证明中检查其他细节。根据微分乘积法则，

$$\begin{aligned}\dot{X} &= \gamma\left(\frac{d}{dt}e^{-tA}\right) t^{-\gamma}\int_0^t u^{\gamma-1}e^{uA}\,du\,X_0 + \gamma e^{-tA}\left(\frac{d}{dt}t^{-\gamma}\right)\int_0^t u^{\gamma-1}e^{uA}\,du\,X_0 + \gamma e^{-tA}t^{-\gamma}\frac{d}{dt}\left(\int_0^t u^{\gamma-1}e^{uA}\,du\,X_0\right) \\ &= -A(X(t)) - \frac{\gamma}{t}X(t) + \gamma e^{-tA}t^{-\gamma}\left(t^{\gamma-1}e^{tA}\right)X_0 \\ &= -A(X(t)) - \frac{\gamma}{t}\left(X(t) - X_0\right).\end{aligned}$$

F.1.2 引理 4.2 的证明

定义

$$X_*(t) = \left(I - \frac{Ae^{-tA}}{C(t)}\left(\int_0^t e^{sA}C(s)ds\right)\right)X_0.$$

我们证明 X_* 是式 (3) 在线性 $\mathbb{A}$ 下定义良好的解，然后证明 X_* 等于式 (9) 中定义的 $X(t)$。

我们先检查定义良好性。根据定义，$C(t) = e^{\int_{t_0}^t \beta(s)ds}$ 是非递减的。此外，$\left\|e^{tA}\right\|$ 也是非递减的，因为根据 A 的单调性，对所有 $x \in \mathbb{R}^n$，

$$\frac{d}{dt}\left\|e^{tA}x\right\|^2 = 2\left\langle A(e^{tA}x), e^{tA}x\right\rangle \geq 0.$$

现在我们看到 $X(t)$是定义良好的，因为

$$\begin{aligned}\left\|e^{-tA}\frac{1}{C(t)}\left(\int_0^t e^{sA}C(s)ds\right)\right\| &= \left\|\int_0^t e^{(s-t)A}\frac{C(s)}{C(t)}ds\right\| \\ &\leq \int_0^t \left\|e^{(s-t)A}\right\|\left|\frac{C(s)}{C(t)}\right|ds \leq \int_0^t (1\cdot 1)ds = t < \infty.\end{aligned}$$

此外，上述不等式意味着当 $t \to 0+$ 时，第二项趋于零，因此有 $\lim_{t\to 0+} X_*(t) = X_0$。因此，如有必要，定义$X_*(0) = X_0$，我们看到 X_*满足初始条件，且 $X_* \in \mathcal{C}([0,\infty), \mathbb{R}^n)$。

接下来验证 $X_*(t)$ 成为该解。这可直接由求导的乘积法则得到。由于 $\frac{d}{dt}\left(\frac{1}{C(t)}\right) = -\frac{1}{C(t)^2}C(t)\beta(t) = -\frac{1}{C(t)}\beta(t)$，我们有

$$\begin{aligned}\dot{X}_*(t) =& -\left(\frac{d}{dt}e^{-tA}\right)\frac{1}{C(t)}\left(\int_0^t e^{sA}C(s)ds\right)X_0 \\ &- Ae^{-tA}\left(\frac{d}{dt}\frac{1}{C(t)}\right)\left(\int_0^t e^{sA}C(s)ds\right)X_0 \\ &- Ae^{-tA}\frac{1}{C(t)}\frac{d}{dt}\left(\int_0^t e^{sA}C(s)ds\right)X_0 \\ =& -A(X_*(t) - X_0) - \beta(t)(X_*(t) - X_0) - A(X_0) = -A(X(t)) - \beta(t)(X_*(t) - X_0).\end{aligned}$$

最后，我们现在证明 $X(t) = X_*(t)$，其中 $X(t)$定义为式（9）。由分部积分可得

$$\begin{aligned} X(t) &= e^{-tA}\frac{1}{C(t)}\left(\int_0^t e^{sA}C(s)\beta(s)ds + C(0)I\right)X_0 \\ &= e^{-tA}\frac{1}{C(t)}\left(\left[e^{sA}C(s)\right]_0^t - \int_0^t Ae^{sA}C(s)ds + C(0)I\right)X_0 \\ &= \left(I - Ae^{-tA}\frac{1}{C(t)}\left(\int_0^t e^{sA}C(s)ds\right)\right)X_0 = X_*(t). \end{aligned}$$

F.2 定理 4.3 的证明
如果 $\beta(t) \equiv 0$，对于 $A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$，我们有 $X(t)$

$= e^{-tA}$，因此

$$\lim_{t\to\infty} \|A(X(t))\| = 1 \neq 0.$$

因此，考虑 β 不是 $\beta(t) \equiv 0$ 且 $\beta(t) \geq 0$的情况。对于

$$A_\xi = 2\pi\xi \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

n ˙设常微分方程 $X_\xi = -A_\xi(X_\xi) - \beta(t)(X_\xi - X_0)$ 的解为 X_ξ。根据引理 4.2，我们有

$$X_\xi(t) = \frac{e^{-tA_\xi}}{C(t)}\left(\int_0^t e^{sA_\xi}C(s)\beta(s)ds + C(0)I\right)X_0.$$

我们要证明,

$$\exists\xi \in \mathbb{R}, \qquad \lim_{t\to\infty} \|A_\xi(X_\xi(t))\| \neq 0$$

由 $\lim^{t\to\infty} c^1{}_{t()} \neq 0$，可知 $\lim_{t\to\infty} \left\| \frac{e^{-tA_\xi}}{C\,t()} \right\| = 1 \cdot \underline{c}^1{}_{t()} \neq 0$。又因为 A_ξ 除 $\xi = 0$ 情况外均可逆，所以我们有

$$\begin{aligned} \lim_{t\to\infty} \|A_\xi(X_\xi(t))\| = 0 &\iff \lim_{t\to\infty} \|X_\xi(t)\| = 0 \\ &\iff \left\| \lim_{t\to\infty} \int_0^t e^{sA_\xi}C(s)\beta(s)ds + C(0)I \right\| = 0. \end{aligned}$$

现在设 $\forall\xi \in \mathrm{R}$，$\lim_{t\to\infty} \|A_\xi(X_\xi(t))\| = 0$，从而导出矛盾。定义 $f : \mathrm{R} \to \mathrm{R}$ 如下

$$f(s) = \begin{cases} C(s)\beta(s) & s > 0 \\ 0 & s \leq 0. \end{cases}$$

于是 $\int_{-\infty}^{\infty} f(s)ds = [C(s)]_0^\infty = \lim_{t\to\infty} C(t) - C(0) < \infty$ 我们有 $f \in L^1(\mathrm{R}, \mathrm{R})$。注意，$\lim_{t\to\infty} C(t)$ 存在，因为 C 单调不减，且根据假设有界。

现在令 $X_0 = (1, 0)^T$，并由 $e^{sA_\xi} = \begin{pmatrix} \cos & & & \\ - & 2\pi\xi s & 2\pi\xi s \sin & \sin\ 2\pi\xi s \cos \end{pmatrix}$ $2\pi\xi s$，可得

$$\begin{aligned} \int_0^\infty e^{sA_\xi}C(s)\beta(s)ds\ X_0 &= \left(\int_0^\infty \cos(2\pi s\xi)f(s)ds\ ,\ \int_0^\infty \sin(-2\pi s\xi)f(s)ds\right)^T \\ &= \left(Re(\hat{f}(\xi))\ ,\ Im(\hat{f}(\xi))\right)^T \end{aligned}$$

ˆ其中 $f(\xi)$ 是 f 的傅里叶变换。
由假设可得 $\forall\xi \in \mathrm{R}, \left\|\int_0^\infty e^{sA_\xi}C(s)\beta(s)ds X_0 + C(0)X_0\right\| = 0$，因此

$$\left(Re(\hat{f}(\xi))\ ,\ Im(\hat{f}(\xi))\right)^T \equiv -C(0)X_0$$

ˆ 顺便一提，根据傅里叶理论，我们知道 f 在无穷远处趋于零，因此

$$\|C(0)X_0\| = \lim_{\xi\to\infty} \left|\hat{f}(\xi)\right| = 0.$$

ˆ 因此，对于所有 $\xi \in \mathrm{R}$，我们有 $f(\xi) = 0$。

ˆ现在 $f \equiv 0$ 显然 在 $L^1(\mathrm{R},\mathrm{R})$ 上，我们可以应用 傅里叶反演公式，并得出 $f \equiv 0$ 几乎处处 re.

然而，由于 $\beta(s)$ 并非常数零，$f(s) = C(s)\beta(s)$ 不可能几乎处处为零。因此产生矛盾，所需结论得证。

F.3 定理 4.4 的证明

F.3.1 当 $p = 1$、$\gamma \geq 1$ 时的证明

由（39）可知

$$X(t) = \gamma e^{-tA} t^{-\gamma} \int_0^t e^{sA} s^{\gamma-1} ds \; X_0.$$

不失一般性，我们考虑 $X_0 = (1,\ 0)^T$ 的情况。

(i) $\gamma = 1$

代入 $\gamma = 1$ 可得

$$X(t) = \frac{e^{-tA}}{t} A^{-1} \left[e^{sA}\right]_0^t \; X_0 = \frac{1}{t} A^{-1} \left(I - e^{-tA}\right) (1,0)^T = \frac{1}{t} A^{-1} \left(1 - \cos t, \sin t\right)^T$$

因此

$$\lim_{t\to\infty} \|tA(X(t))\| = \lim_{t\to\infty} \left\|(1 - \cos t, \sin t)^T\right\| \neq 0.$$

(ii) $\gamma > 1$

我们将展示 极限$_{t\to\infty} \|tA(X(t))\| = \gamma$。通过变量代换 $s = tv$ 和分部积分，我们有

$$\begin{aligned}
\left(\frac{1}{\gamma}\right) tA(X(t)) &= e^{-tA} A t^{-(\gamma-1)} \int_0^t e^{sA} s^{\gamma-1} ds \; X_0 \\
&= e^{-tA} tA \int_0^1 e^{tvA} v^{\gamma-1} dv \; X_0 \\
&= e^{-tA} \left[e^{tvA} v^{\gamma-1}\right]_0^1 \; X_0 - (\gamma-1) e^{-tA} \int_0^1 e^{tvA} v^{\gamma-2} dv \; X_0 \\
&= X_0 - (\gamma-1) e^{-tA} \left(\int_0^1 \cos(tv) v^{\gamma-2} dv \;,\; \int_0^1 \sin(tv) v^{\gamma-2} dv\right)^T
\end{aligned}$$

顺便地，由 $\gamma > 1$ 我们有 $v^{\gamma-2} \in L^1[0,1]$，根据黎曼-勒贝格引理，我们有

$$\lim_{t\to\infty} \int_0^1 \cos(tv) v^{\gamma-2} dv = \lim_{t\to\infty} \int_0^1 \sin(tv) v^{\gamma-2} dv = 0.$$

注意到 e^{-tA} 是一个旋转，我们知道

$$\left\| e^{-tA} \left(\int_0^1 \cos(tv) v^{\gamma-2} dv \;,\; \int_0^1 \sin(tv) v^{\gamma-2} dv\right) \right\| = \left\| \left(\int_0^1 \cos(tv) v^{\gamma-2} dv \;,\; \int_0^1 \sin(tv) v^{\gamma-2} dv\right) \right\|.$$

取极限 $t \to \infty$，我们知道右侧收敛到零点，因此得到

$$\lim_{t\to\infty} e^{-tA} \left(\int_0^1 \cos(tv) v^{\gamma-2} dv \;,\; \int_0^1 \sin(tv) v^{\gamma-2} dv\right) = 0.$$

因此，

$$\begin{aligned}
\lim_{t\to\infty} \|tA(X(t))\| &= \left\| \gamma X_0 - \gamma(\gamma-1) \lim_{t\to\infty} e^{-tA} \left(\int_0^1 \cos(tv) v^{\gamma-2} dv \;,\; \int_0^1 \sin(tv) v^{\gamma-2} dv\right)^T \right\| \\
&= \gamma \|X_0\| = \gamma
\end{aligned}$$

F.3.2 对情形 $p = 1, \gamma < 1$的证明

由于 A和 e^{tA} 均可逆，我们可以看出

$$\lim_{t\to\infty} t^{2\gamma} \|A(X(t))\|^2 \neq 0 \iff \lim_{t\to\infty} t^{\gamma} \|A(X(t))\| \neq 0 \iff \lim_{t\to\infty} t^{\gamma} \left\| \frac{1}{\gamma} e^{tA} X(t) \right\| \neq 0.$$

因此，只需观察 $\frac{1}{\gamma} e^{tA} t^{\gamma} X(t)$。对于 $X_0 = (1, 0)^T$，同样如此。

$$\begin{aligned} \frac{1}{\gamma} e^{tA} t^{\gamma} X(t) &= \int_0^t e^{sA} s^{\gamma-1} ds \; X_0 \\ &= \left(\int_0^t \cos(s) s^{\gamma-1} ds \; , \; \int_0^t \sin(s) s^{\gamma-1} ds \right)^T. \end{aligned}$$

由于 $s^{\gamma-1}$ 单调递减并趋于零，我们可以用类似的论证，并应用交错级数判别法。D

定义 $a_n = \int_{(n-1)\pi}^{n\pi} \sin(s) s^{\gamma-1} ds$，并记为 $S_n = \sum_{k=1}^{n} a_n$。那么对于 $m \in \mathbb{N}$，$t > 2m\pi$ 我们 s ee

$$S_{2m} \leq \int_0^t \sin(s) s^{\gamma-1} ds \leq S_{2m-1}.$$

顺便一提，根据交错级数判别法，$S = \lim_{n\to\infty} S_n$ 存在，因此我们有

$$\lim_{t\to\infty} \int_0^t \sin(s) s^{\gamma-1} ds = S$$

根据夹逼定理。注意 $S \geq S_2 > 0$。用类似的论证，我们可以得出 $\lim_{t\to\infty} \int_0^t \cos(s) s^{\gamma-1} ds$ 也存在。因此我们得出 $\lim_{t\to\infty} \|t^{\gamma} X(t)\| \neq 0$，因为

$$\lim_{t\to\infty} \|t^{\gamma} X(t)\| = \lim_{t\to\infty} \gamma \left\| \left(\int_0^t \cos(s) s^{\gamma-1} ds \; , \; \int_0^t \sin(s) s^{\gamma-1} ds \right)^T \right\| \geq \gamma S > 0.$$

F.3.3 $0 < p < 1$**情形的证明**

回顾式 (9)，我们有

$$X(t) = \frac{e^{-tA}}{C(t)} \left(\int_0^t e^{sA} C(s) \beta(s) ds + C(0) I \right) X_0.$$

我们的目标是证明

$$\lim_{t\to\infty} t^{2p} \|A(X(t))\|^2 \neq 0.$$

我们首先注意到，只需证明

$$\lim_{t\to\infty} \frac{1}{\beta(t) C(t)} \int_0^t C(s) \beta(s) \sin s \, ds \neq 0.$$

这可由以下事实推出。

(1) 由于 $\beta(t) = \frac{\gamma}{t^p}$和 A 是线性的且可逆，

$$\lim_{t\to\infty} t^{2p} \|A(X(t))\|^2 = \lim_{t\to\infty} \frac{\gamma^2}{\beta(t)^2} \|A(X(t))\|^2 \neq 0 \quad \iff \quad \lim_{t\to\infty} \left\| \frac{1}{\beta(t)} X(t) \right\| \neq 0.$$

(2) 回顾 (31)，我们有 $C(t) = e^{\frac{\gamma}{1-p} t^{1-p}}$当 $\beta(t) = \frac{\gamma}{t^p}$，所以 $C(0) = \lim_{t\to 0+} C(t) = 1$ 且 $\lim_{t\to\infty} \frac{1}{\beta(t) C(t)} = 0$。因此$\lim_{t\to\infty} \frac{1}{\beta(t) C(t)} \|C(0) I\| = 0$，第二项可忽略。因此问题归结为

$$\lim_{t\to\infty} \left\| \frac{e^{-tA}}{\beta(t) C(t)} \left(\int_0^t C(s) \beta(s) e^{-sA} ds \right) X_0 \right\| \neq 0.$$

(3) 由于 e^{tA} 是线性且可逆的，问题简化为

$$\lim_{t\to\infty}\left\|\frac{1}{\beta(t)C(t)}\left(\int_0^t C(s)\beta(s)e^{-sA}ds\right)X_0\right\|\neq 0.$$

(4) 同样不失一般性，令 $X_0=(1,\ 0)^T$。回忆 $e^{-sA}=\begin{pmatrix}\cos s & \sin s\\ -\sin s & \cos s\end{pmatrix}$，聚焦于一个分量

我们可以看到，只需证明

$$\lim_{t\to\infty}\frac{1}{\beta(t)C(t)}\int_0^t C(s)\beta(s)\sin s\,ds\neq 0.$$

现在我们证明该命题。为方便起见，记 $D(t)=\beta(t)C(t)$。我们要证明

$$\lim_{t\to\infty}\underbrace{\int_0^t\frac{D(s)}{D(t)}\sin s\,ds}_{=S(t)}\neq 0.$$

我们首先观察到

$$\lim_{t\to\infty}\sup_{\delta\in[0,\delta]}\left|\frac{D(t+\delta)}{D(t+\pi)}-1\right|=0.$$

为此，我们先证明 $\lim_{t\to\infty}\frac{D(t+\delta)}{D(t+\pi)}=1$ 对于 $\delta\in[0,\pi]$。取 $\delta\in[0,\pi]$。观察到

$$\frac{D(t+\delta)}{D(t+\pi)}=\frac{\beta(t+\delta)C(t+\delta)}{\beta(t+\pi)C(t+\pi)}=\left(\frac{t+\pi}{t+\delta}\right)^p\frac{e^{\frac{\gamma}{1-p}(t+\delta)^{1-p}}}{e^{\frac{\gamma}{1-p}(t+\pi)^{1-p}}}=\left(\frac{t+\pi}{t+\delta}\right)^p e^{\frac{\gamma}{1-p}(t+\delta)^{1-p}\left(1-\left(\frac{t+\pi}{t+\delta}\right)^{1-p}\right)}.$$

考虑指数中的洛必达法则，我们得到

$$\lim_{t\to\infty}(t+\delta)^{1-p}\left(1-\left(\frac{t+\pi}{t+\delta}\right)^{1-p}\right)=\lim_{t\to\infty}\frac{1-\left(\frac{t+\pi}{t+\delta}\right)^{1-p}}{(t+\delta)^{p-1}}=\lim_{t\to\infty}\frac{\frac{-\pi(p-1)\left(\frac{t+\pi}{t+\delta}\right)^{-p}}{(t+\delta)^2}}{(p-1)(t+\delta)^{p-2}}=\lim_{t\to\infty}\frac{-\pi}{(t+\pi)^p}=0.$$

当指数趋于零时，我们得出 $\lim_{t\to\infty}\frac{D(t+\delta)}{D(t+\pi)}=1$。

现在由

$$\dot{D}(t)=C(t)\left(\beta(t)^2+\dot{\beta}(t)\right)=e^{\frac{\gamma}{1-p}t^{1-p}}\gamma t^{-2p-1}\left(\gamma t-pt^p\right),$$

我们看到 $D(t)$ 关于 $t\geq\left(\frac{\gamma}{p}\right)^{\frac{1}{p-1}}$ 单调不减，因为 $0<p<1$。因此，对于 $t\geq\left(\frac{\gamma}{p}\right)^{\frac{1}{p-1}}$，我们有 $\sup_{\delta\in[0,\delta]}\left|\frac{D(t+\delta)}{D(t+\pi)}-1\right|=\max\left\{\left|\frac{D(t)}{D(t+\pi)}-1\right|,\ \left|\frac{D(t+\delta)}{D(t+\pi)}-1\right|\right\}$，因此当 $t\to\infty$ 时它趋于零点。

现在我们证明所需的结论。反证法。假设 $\lim_{t\to\infty}S(t)=0$。那么对于 $0<\epsilon<\frac{1}{2}$，存在 $T_1>0$ 使得 $t>T_1$ 蕴含 $|S(t)|<\epsilon$。现在对于 $t>T_1$，注意到

$$\begin{aligned}2\epsilon>|S(t+\pi)-S(t)|&=\left|\int_0^{t+\pi}\frac{D(s)}{D(t+\pi)}\sin s\,ds-\int_0^t\frac{D(s)}{D(t)}\sin s\,ds\right|\\&=\left|\frac{1}{D(t+\pi)}\left(\int_0^{t+\pi}D(s)\sin s\,ds-\int_0^t D(s)\sin s\,ds\right)+\left(\frac{1}{D(t+\pi)}-\frac{1}{D(t)}\right)\int_0^t D(s)\sin s\,ds\right|\\&\geq\left|\int_t^{t+\pi}\frac{D(s)}{D(t+\pi)}\sin s\,ds\right|-\left|\left(\frac{D(t)}{D(t+\pi)}-1\right)\int_0^t\frac{D(s)}{D(t)}\sin s\,ds\right|.\end{aligned}$$

另一方面，存在 T_2 使得 $t>T_2$ 蕴含 $\sup_{\delta\in[0,\pi]}\left|\frac{D(t+\delta)}{D(t+\pi)}-1\right|\leq\frac{\epsilon}{2}$。现在对于 $t=2n\pi>\max\{T_1,T_2\}$,

$$2\epsilon>\left|\int_{2n\pi}^{(2n+1)\pi}\frac{D(s)}{D(t+\pi)}\sin s\,ds\right|-\left|\frac{D(t)}{D(t+\pi)}-1\right|\times|S(t)|>\int_{2n\pi}^{(2n+1)\pi}(1-\epsilon)\left|\sin s\right|ds-\epsilon=2-2\epsilon.$$

这与事实 $\epsilon<\frac{1}{2}$ 矛盾，因此我们证明假设 $\lim_{t\to\infty}\int_0^t S(t)=0$ 不成立。因此 $\lim_{t\to\infty}\int_0^t S(t)\neq 0$。

G 离散对应方法的收敛分析证明

G.1 定理 5.1 中的离散方法与连续模型（6）之间的对应关系

为检验该方法与式 (6) 的对应关系，我们给出一个非正式推导。假设算子 $\mathbb{A}$ 连续。利用 $y^{k-1}=x^k+\mathbb{A}x^k$，代入 y^k 和 y^{k-1} 后，该方法可等价地写为

$$x^{k+1}+\mathbb{A}x^{k+1}=\frac{k^p}{k^p+\gamma}(x^k-\mathbb{A}x^k)+\frac{\gamma}{k^p+\gamma}x^0. \tag{40}$$

这可以视为下述方法的一个特例，其中 $h=1$。

$$x^{k+1}+h\mathbb{A}x^{k+1}=\frac{k^p}{k^p+h^{1-p}\gamma}(x^k-h\mathbb{A}x^k)+\frac{h^{1-p}\gamma}{k^p+h^{1-p}\gamma}x^0.$$

两边除以 h并重新整理，得到

$$\frac{x^{k+1}-x^k}{h}=-\mathbb{A}x^{k+1}-\frac{h^pk^p}{h^pk^p+h\gamma}\mathbb{A}x^k-\frac{\gamma}{h^pk^p+h\gamma}\left(x^k-x^0\right).$$

令 $hk=t$、 $x^k=X$（t），并取 $h\to 0$，得到

$$\dot{X}(t)=-2\mathbb{A}(X(t))-\frac{\gamma}{t^p}\left(X-X_0\right).$$

由于标量倍数仍保持单调性，将 $2\mathbb{A}$ 重新缩放为 $\mathbb{A}$ 不会改变该常微分方程所涵盖的算子类别。为简化记号，将 $2\mathbb{A}$ 替换为 $\mathbb{A}$，得到

$$\dot{X}(t)=-\mathbb{A}(X(t))-\frac{\gamma}{t^p}\left(X-X_0\right).$$

G.2 $\left\|x^k\right\|$有界性的证明

在证明 定理 5.1 时，我们需要为 $\left\|x^k\right\|$找到一个上界。因此，我们首先证明以下引理 a.

引理 G.1. 设 $\mathbb{A}$ 是一个极大单调算子。考虑一种方法

$$\begin{aligned}x^k&=\mathbb{J}_{\mathbb{A}}y^{k-1}\\ y^k&=(1-\beta_k)\left(2x^k-y^{k-1}\right)+\beta_kx^0\end{aligned}$$

对于 $k=1$，2，……，其序列为 $\{\beta_k\}_{k\in\mathbb{N}}$、 $0\le\beta_k\le 1$，初始条件为 $y^0=x^0\in\mathbb{R}^n$。则有以下结论

$$\left\|x^k-x^\star\right\|\le\left\|x^0-x^\star\right\|.$$

对于 $x^\star\in\mathrm{Zer}\mathbb{A}$。因此， $\left\|x^k-x^0\right\|\le 2\left\|x^0-x^\star\right\|$。

证明。 回顾引理 D.2，对于 $\mathbb{T}=\mathrm{R}_{\mathbb{A}}=2\mathbb{J}_{\mathbb{A}}-\mathbb{I}$，给定的方法等价于以下方法

$$y^k=\beta_ky^0+(1-\beta_k)\,\mathbb{T}y^{k-1}.$$

注意， $x^\star\in\mathrm{Zer}\ \mathbb{A}\Leftrightarrow 0\in\mathbb{A}x^\star\Leftrightarrow x^\star=\mathbb{J}_{\mathbb{A}}x^\star\Leftrightarrow x^\star=\mathbb{T}x^\star\Leftrightarrow x^\star\in\mathrm{Fix}\,\mathbb{T}$。我们先通过归纳法证明 $\left\|y^k-x^\star\right\|\le\left\|y^0-x^\star\right\|$。当 $k=0$ 时，该命题显然成立。现在假设该命题对 $k\in\mathrm{N}$ 成立。那么

$$\begin{aligned}\left\|y^{k+1}-x^\star\right\|&=\left\|\beta_{k+1}\left(y^0-x^\star\right)+(1-\beta_{k+1})\left(\mathbb{T}y^k-x^\star\right)\right\|\\ &\le\beta_{k+1}\left\|y^0-x^\star\right\|+(1-\beta_{k+1})\left\|\mathbb{T}y^k-x^\star\right\|\\ &\le\beta_{k+1}\left\|y^0-x^\star\right\|+(1-\beta_{k+1})\left\|y^k-x^\star\right\|\\ &\le\beta_{k+1}\left\|y^0-x^\star\right\|+(1-\beta_{k+1})\left\|y^0-x^\star\right\|=\left\|y^0-x^\star\right\|.\end{aligned}$$

第一个不等式仅由三角不等式得出，第二个不等式来自这样一个事实： $x^\star\in\mathrm{Fix}\ \mathbb{T}$和 $\mathbb{T}$是非扩张的，最后一个不等式则来自归纳假设。由归纳法可得，对于 $\left\|y^k-x^\star\right\|\le\left\|y^0-x^\star\right\|$，1，……。

定义 $\tilde{\mathbb{A}}x^k = y^{k-1} - x^k$，那么由于 $\mathbb{J}_{\mathbb{A}} y^{k-1} = x^k$，有 $\tilde{\mathbb{A}}x^k \in \mathbb{A}x^k$。注意，对于 $k = 1, 2, \ldots\ldots$，根据单调不等式，我们有

$$\left\|y^{k-1} - x^\star\right\|^2 = \left\|\tilde{\mathbb{A}}x^k + (x^k - x^\star)\right\|^2 = \left\|\tilde{\mathbb{A}}x^k\right\|^2 + 2\left\langle \tilde{\mathbb{A}}x^k, x^k - x^\star\right\rangle + \left\|x^k - x^\star\right\|^2 \geq \left\|x^k - x^\star\right\|^2.$$

因此，我们得出结论

$$\left\|x^k - x^\star\right\| \leq \left\|y^{k-1} - x^\star\right\| \leq \left\|y^0 - x^\star\right\| = \left\|x^0 - x^\star\right\|.$$

后一个陈述直接由三角不等式得出，

$$\left\|x^k - x^0\right\| \leq \left\|x^k - x^\star\right\| + \left\|x^\star - x^0\right\| \leq 2\left\|x^0 - x^\star\right\|.$$

□

G.3 定理 5.1 的证明

证明思路源自连续情形证明。为简化计算，我们不直接推导命题 3.2 的离散对应形式，而是针对每种情形考虑重标度守恒律。通过对 $p > 1$（$X - X_0$）对应的 $W_1 = X - X_0$、 $W_2 = t^p$，对 $0 < p < 1$、 $W_3 = t$（$X - X_0$）对应的 $p = 1$、 $\gamma \geq 1$，以及对 $W_4 = t^\gamma$（$X - X_0$）对应的 $p = 1$、 $0 < \gamma < 1$ 考虑膨胀坐标，我们得到以下守恒律。

$$\begin{aligned}
E_1 \equiv{}& \left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + \frac{2\gamma}{t^p}\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle + \left(\frac{\gamma^2}{t^{2p}} + \frac{\gamma p}{t^{p+1}}\right)\left\|X(t) - X_0\right\|^2 \\
&+ \int_0^t \left\langle \frac{d}{ds}\tilde{\mathbb{A}}(X(s)), \dot{X}(s)\right\rangle ds + \int_0^t \frac{\gamma}{s^p}\left\|\dot{X}(s) + \frac{p}{s}(X(s) - X_0)\right\|^2 ds + \frac{1}{2}\int_0^t \frac{\gamma p(p-1)}{s^{p+2}}\left\|X(s) - X_0\right\|^2 ds \\
E_2 \equiv{}& t^{2p}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\gamma t^p\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle + \left(\gamma^2 - \gamma p t^{p-1}\right)\left\|X(t) - X_0\right\|^2 \\
&+ \int_0^t 2s^{2p}\left\langle \frac{d}{ds}\tilde{\mathbb{A}}(X(s)), \dot{X}(s)\right\rangle ds + \int_0^t 2s^p\left(\gamma - ps^{p-1}\right)\left\|\dot{X}(s)\right\|^2 ds + \int_0^t \gamma p(p-1)s^{p-2}\left\|X(s) - X_0\right\|^2 ds \\
E_3 \equiv{}& t^2\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\gamma t\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle + \gamma(\gamma - 1)\left\|X(t) - X_0\right\|^2 \\
&+ \int_{t_0}^t 2s^2\left\langle \frac{d}{ds}\tilde{\mathbb{A}}(X(s)), \dot{X}(s)\right\rangle ds + \int_{t_0}^t 2s(\gamma - 1)\left\|\dot{X}(s)\right\|^2 ds \\
E_4 \equiv{}& t^{2\gamma}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\gamma t^{2\gamma-1}\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle + \gamma(\gamma-1)t^{2\gamma-2}\left\|X(t) - X_0\right\|^2 \\
&+ \int_0^t 2s^{2\gamma}\left\langle \frac{d}{ds}\tilde{\mathbb{A}}(X(s)), \dot{X}(s)\right\rangle ds + \int_0^t 2\gamma(\gamma-1)s^{2\gamma-3}\left\|X(s) - X_0\right\|^2 ds.
\end{aligned}$$

通过考虑以下函数，可以得到李雅普诺夫型证明

$$\begin{aligned}
U_1(t) &= \left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + \frac{2\gamma}{t^p}\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle + \left(\frac{\gamma^2}{t^{2p}} + \frac{\gamma p}{t^{p+1}}\right)\left\|X(t) - X_0\right\|^2 \\
U_2(t) &= t^{2p}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\gamma t^p\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle + \left(\gamma^2 - \gamma p t^{p-1}\right)\left\|X(t) - X_0\right\|^2 \\
U_3(t) &= t^2\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\gamma t\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle + \gamma(\gamma - 1)\left\|X(t) - X_0\right\|^2 \\
U_4(t) &= t^{2\gamma}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + \gamma t^{2\gamma-1}\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle + \gamma(\gamma-1)t^{2\gamma-2}\left\|X(t) - X_0\right\|^2.
\end{aligned}$$

计算的主要部分与连续情形相对应，但由于离散化， $\tilde{\mathbb{A}}x^k$ 和 $x^k - x^0$方面会出现一些“误差”。证明通过展示这些“误差”不会影响结论来完成。

(0) 准备工作。对于所有情形，我们都将考虑以下形式的函数

$$U^k = a_k\left\|\tilde{\mathbb{A}}x^k\right\|^2 + b_k\left\langle \tilde{\mathbb{A}}x^k, x^k - x^0\right\rangle + c_k\left\|x^k - x^0\right\|^2, \tag{41}$$

并考虑 $U^{k+1}-U^k$。由于类似的计算会反复出现，我们先整理这些重复的计算。也就是说，我们证明以下等式成立。

$$
\begin{aligned}
&U^{k+1}-U^k+\lambda_k\left\langle\tilde{\mathbb{A}}x^{k+1}-\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle-\tau_k\left\langle\tilde{\mathbb{A}}x^{k+1}-\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle \qquad (42)\\
&=\left(a_{k+1}-\lambda_k\frac{k^p+\gamma}{k^p}\right)\|\tilde{\mathbb{A}}x^{k+1}\|^2+\left(\lambda_k\frac{k^p}{k^p+\gamma}-a_k\right)\|\tilde{\mathbb{A}}x^k\|^2\\
&\quad+\left(b_{k+1}-\lambda_k\frac{\gamma}{k^p}-\tau_k-c_{k+1}\frac{k^p+\gamma}{k^p}\right)\langle\tilde{\mathbb{A}}x^{k+1},x^{k+1}-x^0\rangle+(\tau_k-c_{k+1})\left\langle\tilde{\mathbb{A}}x^{k+1},x^k-x^0\right\rangle\\
&\quad+(\tau_k-c_{k+1})\left\langle\tilde{\mathbb{A}}x^k,x^{k+1}-x^0\right\rangle+\left(\lambda_k\frac{\gamma}{k^p+\gamma}-b_k-\tau_k-c_{k+1}\frac{k^p}{k^p+\gamma}\right)\langle\tilde{\mathbb{A}}x^k,x^k-x^0\rangle\\
&\quad-c_{k+1}\frac{\gamma}{k^p}\left\|x^{k+1}-x^0\right\|^2-c_{k+1}\frac{\gamma}{k^p+\gamma}\left\|x^k-x^0\right\|^2+(c_{k+1}-c_k)\left\|x^k-x^0\right\|^2.
\end{aligned}
$$

请注意，通过检查以下等式是否成立即可证明这一点。

$$
\begin{aligned}
&\left\langle\tilde{\mathbb{A}}x^{k+1}-\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle \qquad (43)\\
&\quad=-\frac{k^p+\gamma}{k^p}\|\tilde{\mathbb{A}}x^{k+1}\|^2-\frac{\gamma}{k^p}\langle\tilde{\mathbb{A}}x^{k+1},x^{k+1}-x^0\rangle+\frac{k^p}{k^p+\gamma}\|\tilde{\mathbb{A}}x^k\|^2+\frac{\gamma}{k^p+\gamma}\langle\tilde{\mathbb{A}}x^k,x^k-x^0\rangle\\
&\left\langle\tilde{\mathbb{A}}x^{k+1}-\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle \qquad (44)\\
&\quad=\left\langle\tilde{\mathbb{A}}x^{k+1},x^{k+1}-x^0\right\rangle+\left\langle\tilde{\mathbb{A}}x^k,x^k-x^0\right\rangle-\left(\left\langle\tilde{\mathbb{A}}x^{k+1},x^k-x^0\right\rangle+\left\langle\tilde{\mathbb{A}}x^k,x^{k+1}-x^0\right\rangle\right)\\
&c_{k+1}\left\|x^{k+1}-x^0\right\|^2-c_k\left\|x^k-x^0\right\|^2 \qquad (45)\\
&\quad=-c_{k+1}\frac{k^p+\gamma}{k^p}\left\langle\tilde{\mathbb{A}}x^{k+1},x^{k+1}-x^0\right\rangle-c_{k+1}\left\langle\tilde{\mathbb{A}}x^{k+1},x^k-x^0\right\rangle\\
&\qquad-c_{k+1}\left\langle\tilde{\mathbb{A}}x^k,x^{k+1}-x^0\right\rangle-c_{k+1}\frac{k^p}{k^p+\gamma}\left\langle\tilde{\mathbb{A}}x^k,x^k-x^0\right\rangle\\
&\qquad-c_{k+1}\frac{\gamma}{k^p}\left\|x^{k+1}-x^0\right\|^2-c_{k+1}\frac{\gamma}{k^p+\gamma}\left\|x^k-x^0\right\|^2+(c_{k+1}-c_k)\left\|x^k-x^0\right\|^2.
\end{aligned}
$$

– (43)的证明

回顾一下，该方法定义如下:

$$
\begin{aligned}
x^k&=\mathbb{J}_{\mathbb{A}}y^{k-1}\\
y^k&=\frac{k^p}{k^p+\gamma}(2x^k-y^{k-1})+\frac{\gamma}{k^p+\gamma}x^0.
\end{aligned}
$$

利用 $y^{k-1}=x^k+\tilde{\mathbb{A}}x^k$，代入 y^k 和 y^{k-1}，我们可以将该方法改写为

$$
x^{k+1}+\tilde{\mathbb{A}}x^{k+1}=\left(1-\frac{\gamma}{k^p+\gamma}\right)(x^k-\tilde{\mathbb{A}}x^k)+\frac{\gamma}{k^p+\gamma}x^0. \qquad (46)
$$

将 $k^p+\gamma$ 乘到等式两边并重新整理，可得:

$$
(k^p+\gamma)\{(x^{k+1}-x^0)+\tilde{\mathbb{A}}x^{k+1}\}=k^p\{(x^k-x^0)-\tilde{\mathbb{A}}x^k\}.
$$

分别从等式两边减去 $k^p(x^{k+1}-x^0)$ 和 $(k^p+\gamma)(x^k-x^0)$ 后，上述等式可改写为

$$
\begin{aligned}
-k^p(\tilde{\mathbb{A}}x^k+(x^{k+1}-x^k))&=(k^p+\gamma)\tilde{\mathbb{A}}x^{k+1}+\gamma(x^{k+1}-x^0)\\
-(k^p+\gamma)(\tilde{\mathbb{A}}x^{k+1}+(x^{k+1}-x^k))&=k^p\tilde{\mathbb{A}}x^k+\gamma(x^k-x^0).
\end{aligned}
$$

由上可得

$$
\begin{aligned}
&(k^p+\gamma)k^p\left\langle\tilde{\mathbb{A}}x^{k+1}-\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle\\
&=(k^p+\gamma)k^p\left\langle\tilde{\mathbb{A}}x^{k+1},x^{k+1}-x^k\right\rangle+(k^p+\gamma)k^p\left\langle\tilde{\mathbb{A}}x^{k+1},\tilde{\mathbb{A}}x^k\right\rangle\\
&\quad-(k^p+\gamma)k^p\left\langle\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle-(k^p+\gamma)k^p\left\langle\tilde{\mathbb{A}}x^{k+1},\tilde{\mathbb{A}}x^k\right\rangle\\
&=-(k^p+\gamma)\langle\tilde{\mathbb{A}}x^{k+1},-k^p(\tilde{\mathbb{A}}x^k+(x^{k+1}-x^k))\rangle+k^p\langle\tilde{\mathbb{A}}x^k,-(k^p+\gamma)(\tilde{\mathbb{A}}x^{k+1}+(x^{k+1}-x^k))\rangle\\
&=-(k^p+\gamma)\langle\tilde{\mathbb{A}}x^{k+1},(k^p+\gamma)\tilde{\mathbb{A}}x^{k+1}+\gamma(x^{k+1}-x^0)\rangle+k^p\langle\tilde{\mathbb{A}}x^k,k^p\tilde{\mathbb{A}}x^k+\gamma(x^k-x^0)\rangle\\
&=-(k^p+\gamma)\left((k^p+\gamma)\|\tilde{\mathbb{A}}x^{k+1}\|^2+\gamma\langle\tilde{\mathbb{A}}x^{k+1},x^{k+1}-x^0\rangle\right)+k^p\left(\|\tilde{\mathbb{A}}x^k\|^2+\gamma\langle\tilde{\mathbb{A}}x^k,x^k-x^0\rangle\right).
\end{aligned}
$$

现在将等式两边同时除以 $k^p(k^p+\gamma)$，即可得到所需结果。– (44) 的证明：只需展开左侧的内积即可验证。– (45) 的证明：首先注意到

$$\begin{aligned}
&c_{k+1}\left\|x^{k+1}-x^0\right\|^2-c_k\left\|x^k-x^0\right\|^2\\
&=c_{k+1}\left(\left\|x^{k+1}-x^0\right\|^2-\left\|x^k-x^0\right\|^2\right)+(c_{k+1}-c_k)\left\|x^k-x^0\right\|^2\\
&=c_{k+1}\left\langle x^{k+1}-x^k,\left(x^{k+1}-x^0\right)+\left(x^k-x^0\right)\right\rangle+(c_{k+1}-c_k)\left\|x^k-x^0\right\|^2\\
&=c_{k+1}\left(\left\langle x^{k+1}-x^k,x^{k+1}-x^0\right\rangle+\left\langle x^{k+1}-x^k,x^k-x^0\right\rangle\right)+(c_{k+1}-c_k)\left\|x^k-x^0\right\|^2.
\end{aligned}$$

重新整理（40），我们得到 $x^{k+1}-x^k$ 的两个不同表达式。

$$\begin{aligned}
x^{k+1}-x^k&=-\left(\tilde{\mathbb{A}}x^{k+1}+\frac{k^p}{k^p+\gamma}\tilde{\mathbb{A}}x^k\right)-\frac{\gamma}{k^p+\gamma}\left(x^k-x^0\right)\\
x^{k+1}-x^k&=-\left(\frac{k^p+\gamma}{k^p}\tilde{\mathbb{A}}x^{k+1}+\tilde{\mathbb{A}}x^k\right)-\frac{\gamma}{k^p}\left(x^{k+1}-x^0\right).
\end{aligned}$$

现在将这些代入前一个等式，即可得到所需结果。

$$\begin{aligned}
&c_{k+1}\left\|x^{k+1}-x^0\right\|^2-c_k\left\|x^k-x^0\right\|^2\\
&=-c_{k+1}\left(\left\langle\frac{k^p+\gamma}{k^p}\tilde{\mathbb{A}}x^{k+1}+\tilde{\mathbb{A}}x^k+\frac{\gamma}{k^p}\left(x^{k+1}-x^0\right),x^{k+1}-x^0\right\rangle\right.\\
&\qquad\left.+\left\langle\tilde{\mathbb{A}}x^{k+1}+\frac{k^p}{k^p+\gamma}\tilde{\mathbb{A}}x^k+\frac{\gamma}{k^p+\gamma}\left(x^k-x^0\right),x^k-x^0\right\rangle\right)+(c_{k+1}-c_k)\left\|x^k-x^0\right\|^2\\
&=-c_{k+1}\frac{k^p+\gamma}{k^p}\left\langle\tilde{\mathbb{A}}x^{k+1},x^{k+1}-x^0\right\rangle-c_{k+1}\left\langle\tilde{\mathbb{A}}x^{k+1},x^k-x^0\right\rangle\\
&\quad-c_{k+1}\left\langle\tilde{\mathbb{A}}x^k,x^{k+1}-x^0\right\rangle-c_{k+1}\frac{k^p}{k^p+\gamma}\left\langle\tilde{\mathbb{A}}x^k,x^k-x^0\right\rangle\\
&\quad-c_{k+1}\frac{\gamma}{k^p}\left\|x^{k+1}-x^0\right\|^2-c_{k+1}\frac{\gamma}{k^p+\gamma}\left\|x^k-x^0\right\|^2+(c_{k+1}-c_k)\left\|x^k-x^0\right\|^2
\end{aligned}$$

(i) 对于 $p>0$、$\gamma>0$，$\left\|\tilde{\mathbb{A}}(x^k)\right\|^2=\mathcal{O}(1)$。代入

$$\begin{aligned}
a_k&=1+\frac{\gamma}{2k^p}, & b_k&=\frac{\gamma}{k^p}, & c_{k+1}&=\frac{\gamma k^p}{4\left(k^p+\frac{\gamma}{2}\right)}\left(\frac{\gamma}{k^{2p}}-\left(\frac{1}{k^p}-\frac{1}{(k+1)^p}\right)\right)\\
\lambda_k&=1+\frac{\gamma}{2k^p}, & \tau_k&=c_{k+1}
\end{aligned}$$

由式 (41) 和 (42)，我们得到

$$\begin{aligned}
&U^{k+1}-U^k+\left(1+\frac{\gamma}{2k^p}-\frac{\gamma k^p}{4\left(k^p+\frac{\gamma}{2}\right)}\left(\frac{\gamma}{k^{2p}}-\left(\frac{1}{k^p}-\frac{1}{(k+1)^p}\right)\right)\right)\left\langle\tilde{\mathbb{A}}x^{k+1}-\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle\\
&=-\left(\frac{\gamma\left(2k^p+\gamma\right)}{2k^{2p}}+\frac{\gamma}{2}\left(\frac{1}{k^p}-\frac{1}{(k+1)^p}\right)\right)\|\tilde{\mathbb{A}}x^{k+1}\|^2-\frac{\gamma\left(2k^p+\gamma\right)}{2k^p\left(k^p+\gamma\right)}\|\tilde{\mathbb{A}}x^k\|^2\\
&\quad-\left(\frac{\gamma^2}{k^{2p}}+\frac{\gamma}{2}\left(\frac{1}{k^p}-\frac{1}{(k+1)^p}\right)\right)\left\langle\tilde{\mathbb{A}}x^{k+1},x^{k+1}-x^0\right\rangle\\
&\quad-\left(\frac{\gamma^2}{k^p\left(k^p+\gamma\right)}+\frac{\gamma k^p}{2(k^p+\gamma)}\left(\frac{1}{k^p}-\frac{1}{(k+1)^p}\right)\right)\left\langle\tilde{\mathbb{A}}x^k,x^k-x^0\right\rangle\\
&\quad-\frac{\gamma^2}{2\left(2k^p+\gamma\right)}\left(\frac{\gamma}{k^{2p}}-\left(\frac{1}{k^p}-\frac{1}{(k+1)^p}\right)\right)\left\|x^{k+1}-x^0\right\|^2
\end{aligned}$$

$$
\begin{aligned}
&- \frac{\gamma^2 k^p}{2\left(k^p+\gamma\right)\left(2k^p+\gamma\right)}\left(\frac{\gamma}{k^{2p}} - \left(\frac{1}{k^p} - \frac{1}{(k+1)^p}\right)\right)\left\|x^k - x^0\right\|^2 + \left(c_{k+1} - c_k\right)\left\|x^k - x^0\right\|^2 \\
&= -\frac{\gamma\left(2k^p+\gamma\right)}{2k^{2p}}\left\|\tilde{\mathbf{A}}x^{k+1} + \frac{1}{2k^p+\gamma}\left(\gamma + \frac{k^{2p}}{2}\left(\frac{1}{k^p} - \frac{1}{(k+1)^p}\right)\right)\left(x^{k+1} - x^0\right)\right\|^2 \\
&\quad - \frac{\gamma\left(2k^p+\gamma\right)}{2k^p\left(k^p+\gamma\right)}\left\|\tilde{\mathbf{A}}x^{k} + \frac{1}{2k^p+\gamma}\left(\gamma + \frac{k^{2p}}{2}\left(\frac{1}{k^p} - \frac{1}{(k+1)^p}\right)\right)\left(x^{k} - x^0\right)\right\|^2 \\
&\quad - \frac{1}{2}\left(\frac{\gamma}{k^p} - \frac{\gamma}{(k+1)^p}\right)\left\|\tilde{\mathbf{A}}x^{k+1}\right\|^2 \\
&\quad \underbrace{- \frac{\gamma}{2\left(2k^p+\gamma\right)}\left(2\gamma\left(\frac{1}{(k+1)^p} - \frac{1}{k^p}\right) + \frac{k^{2p}}{4}\left(\frac{1}{k^p} - \frac{1}{(k+1)^p}\right)^2\right)}_{s_{k,1}}\left\|x^{k+1} - x^0\right\|^2 \\
&\quad \underbrace{- \frac{\gamma k^p}{2\left(k^p+\gamma\right)\left(2k^p+\gamma\right)}\left(2\gamma\left(\frac{1}{(k+1)^p} - \frac{1}{k^p}\right) + \frac{k^{2p}}{4}\left(\frac{1}{k^p}\right)^2 - \frac{1}{(k+1)^p}\right)}_{s_{k,0}}\left\|x^{k} - x^0\right\|^2 \\
&\quad + \left(c_{k+1} - c_k\right)\left\|x^k - x^0\right\|^2.
\end{aligned}
$$

上述等式的连续对应形式为

$$
\begin{aligned}
&\dot{U}_1(t) + \left\langle \frac{d}{dt}\tilde{\mathbf{A}}(X(t)), \dot{X}(t)\right\rangle \\
&\quad = -\frac{\gamma}{t^p}\left\|\tilde{\mathbf{A}}(X(t)) + \left(\frac{\gamma}{t^p} + \frac{p}{t}\right)\left(X(t) - X_0\right)\right\|^2 - \frac{1}{2}\frac{\gamma p(p-1)}{t^{p+2}}\left\|X(t) - X_0\right\|^2,
\end{aligned}
$$

这可以通过对 E_1 的守恒律求导并重新整理得到。注意，由于 $\frac{k^{2p}}{2}\left(\frac{1}{k^p} - \frac{1}{(k+1)^p}\right) = \mathcal{O}\left(k^{2p}\right)\mathcal{O}\left(\frac{1}{k^{p+1}}\right) = \mathcal{O}\left(\frac{1}{k^{1-p}}\right)$，项 $\frac{1}{2k^p+\gamma}\left(\gamma + \frac{k^{2p}}{2}\left(\frac{1}{k^p} - \frac{1}{(k+1)^p}\right)\right)$ 的阶对应于 $\left(\frac{\gamma}{t^p} + \frac{p}{t}\right)$。因此可以看出，离散情形等式右侧前两项之和对应于连续情形等式右侧的第一项。

我们可以观察到 $s_{k,1}$，$s_{k,0} = \mathcal{O}\left(\frac{1}{k^{2p+1}}\right) + \mathcal{O}\left(\frac{1}{k^{p+2}}\right)$。并且由于 $c_k = \mathcal{O}\left(\frac{1}{k^{2p}}\right) + \mathcal{O}\left(\frac{1}{k^{p+1}}\right)$，我们也有 $c_{k+1} - c_k = \mathcal{O}\left(\frac{1}{k^{2p+1}}\right) + \mathcal{O}\left(\frac{1}{k^{p+2}}\right)$。因此，$\left\|x^{k+1} - x^0\right\|^2$ 和 $\left\|x^k - x^0\right\|^2$ 的系数对于 $p > 0$ 是可求和的。根据引理 G.1，我们知道 $\left\|x^{k+1} - x^0\right\|^2$ 和 $\left\|x^k - x^0\right\|^2$ 有界，且上界为 $4\left\|x^0 - x^\star\right\|^2$。

由 $\tilde{\mathbf{A}}$ 的单调性可知，左侧的项 $\left\langle \tilde{\mathbf{A}}x^{k+1} - \tilde{\mathbf{A}}x^k,\ x^{k+1} - x^k\right\rangle$ 非负，且其系数也非负。右侧中 $\left\|\tilde{\mathbf{A}}x^{k+1}\right\|^2$ 的系数 $-\frac{1}{2}\left(\frac{\gamma}{k^p} - \frac{\gamma}{(k+1)^p}\right)$ 非正。因此，我们得到以下不等式。

$$
\begin{aligned}
U^{k+1} &\le U^k + \frac{1}{2}\left(\frac{\gamma}{k^p} - \frac{\gamma}{(k+1)^p}\right)\left\|x^{k+1} - x^0\right\|^2 + \left(c_{k+1} - c_k\right)\left\|x^k - x^0\right\|^2 \\
&\le U^1 + 2\left\|x^0 - x^\star\right\|^2 \sum_{m=1}^{\infty}\left(\left(\frac{\gamma}{k^p} - \frac{\gamma}{(k+1)^p}\right) + 2\left(c_{k+1} - c_k\right)\right) = M.
\end{aligned}
$$

与式（8）中的做法相同，由于 A的单调性和杨氏不等式

$$
\begin{aligned}
M \ge U^k &= a_k\left\|\tilde{\mathbf{A}}x^k\right\|^2 + b_k\left\langle \tilde{\mathbf{A}}x^k, x^k - x^0\right\rangle + c_k\left\|x^k - x^0\right\|^2 \\
&\ge a_k\left\|\tilde{\mathbf{A}}x^k\right\|^2 + b_k\left\langle \tilde{\mathbf{A}}x^k, x^\star - x^0\right\rangle \\
&\ge a_k\left\|\tilde{\mathbf{A}}x^k\right\|^2 - \frac{1}{2}\left(\left\|\tilde{\mathbf{A}}x^k\right\|^2 + b_k^2\left\|x^\star - x^0\right\|^2\right) = \frac{1}{2}\left(1 + \frac{\gamma}{k^p}\right)\left\|\tilde{\mathbf{A}}x^k\right\|^2 + \frac{\gamma^2}{2k^{2p}}\left\|x^\star - x^0\right\|^2.
\end{aligned}
$$

重新整理后，即可得到所需结果

$$\begin{aligned}\left\|\tilde{\mathbb{A}}x^k\right\|^2 &\leq 2\left(1-\frac{\gamma}{k^p}\right)^{-1}\left(M+\frac{\gamma^2}{2k^{2p}}\left\|x^\star-x^0\right\|^2\right)\\ &= 2\left(1+\frac{\gamma}{k^p-\gamma}\right)\left(M+\frac{\gamma^2}{2k^{2p}}\left\|x^\star-x^0\right\|^2\right)=\mathcal{O}\left(1\right).\end{aligned}$$

(ii) $\left\|\tilde{\mathbb{A}}(x^k)\right\|^2=\mathcal{O}\left(\frac{1}{k^{2p}}\right)$ 对于 $0<p<1$, $\gamma>0$。代入

$$a_{k+1}=k^p\left(k^p+\frac{\gamma}{2}\right),\qquad b_{k+1}=\gamma k^p,\qquad c_{k+1}=\frac{k^p\gamma^2}{4\left(k^p+\frac{\gamma}{2}\right)}$$

$$\lambda_k=a_{k+1},\qquad \tau_k=c_{k+1}$$

并结合 $p=1$代入 (41) 和 (42)，我们得到

$$\begin{aligned}&U^{k+1}-U^k+\left(k^p\left(k^p+\frac{\gamma}{2}\right)-\frac{k^p\gamma^2}{4\left(k^p+\frac{\gamma}{2}\right)}\right)\left\langle\tilde{\mathbb{A}}x^{k+1}-\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle\\ &=-\gamma\left(k^p+\frac{\gamma}{2}\right)\|\tilde{\mathbb{A}}x^{k+1}\|^2-\gamma^2\langle\tilde{\mathbb{A}}x^{k+1},x^{k+1}-x^0\rangle-\frac{\gamma^3}{4\left(k^p+\frac{\gamma}{2}\right)}\left\|x^{k+1}-x^0\right\|^2\\ &\quad+\underbrace{\left(\frac{k^{2p}\left(k^p+\frac{\gamma}{2}\right)}{k^p+\gamma}-(k-1)^p\left((k-1)^p+\frac{\gamma}{2}\right)\right)}_{=q_k}\left\|\tilde{\mathbb{A}}x^k\right\|^2+\underbrace{\frac{\gamma\left(-(k-1)^pk^p+k^{2p}-\gamma(k-1)^p\right)}{k^p+\gamma}}_{=r_k}\langle\tilde{\mathbb{A}}x^k,x^k-x^0\rangle\\ &\quad-\frac{\gamma^3k^p}{4\left(k^p+\gamma\right)\left(k^p+\frac{\gamma}{2}\right)}\left\|x^k-x^0\right\|^2+\frac{\gamma^3\left(k^p-(k-1)^p\right)}{8\left((k-1)^p+\frac{\gamma}{2}\right)\left(k^p+\frac{\gamma}{2}\right)}\left\|x^k-x^0\right\|^2\\ &=-\gamma\left(k^p+\frac{\gamma}{2}\right)\left\|\tilde{\mathbb{A}}x^{k+1}+\frac{\gamma}{2k^p+\gamma}\left(x^{k+1}-x^0\right)\right\|^2+q_k\left\|\tilde{\mathbb{A}}x^k+\frac{r_k}{2q_k}\left(x^k-x^0\right)\right\|^2\\ &\quad-\underbrace{\left(\frac{r_k^2}{4q_k}+\frac{\gamma^3k^p}{4\left(k^p+\gamma\right)\left(k^p+\frac{\gamma}{2}\right)}+\frac{\gamma^3\left(k^p-(k-1)^p\right)}{8\left((k-1)^p+\frac{\gamma}{2}\right)\left(k^p+\frac{\gamma}{2}\right)}\right)}_{s_k}\left\|x^k-x^0\right\|^2.\end{aligned}$$

该等式的连续对应形式为

$$\begin{aligned}&\dot{U}_2(t)+2t^{2p}\left\langle\frac{d}{dt}\tilde{\mathbb{A}}(X(t)),\dot{X}(t)\right\rangle\\ &\quad=-2t^p\left(\gamma-pt^{p-1}\right)\left\|\tilde{\mathbb{A}}(X(t))+\frac{\gamma}{t^p}\left(X(t)-X_0\right)\right\|^2-\gamma p(p-1)t^{p-2}\left\|X(t)-X_0\right\|^2\end{aligned}$$

这可以通过对 E_2 的守恒律求导得到。右侧的第一项对应于离散等式右侧的前两项之和。因此，我们可以预期各项系数的顺序会相匹配，下面将验证这一预期确实成立。

经过一些计算，我们可以观察到

$$s_k=\frac{2\gamma^2(k-1)^pk^{2p}\left((k-1)^p-k^p\right)^2}{4\left((k-1)^p+\frac{\gamma}{2}\right)\left(k^p+\frac{\gamma}{2}\right)d_k},$$

其中

$$d_k=2k^p(k-1)^p\left((k-1)^p+\frac{\gamma}{2}\right)-2k^{3p}-\gamma k^{2p}+2\gamma(k-1)^p\left((k-1)^p+\frac{\gamma}{2}\right).$$

通过考虑牛顿展开并由 $0<p<1$可知

$$\begin{aligned}d_k&=2k^{3p}+\gamma k^{2p}+\mathcal{O}\left(k^{3p-1}\right)-2k^{3p}-\gamma k^{2p}+2\gamma k^{2p}+\mathcal{O}\left(k^p\right)\\ &=2\gamma k^{2p}+\mathcal{O}\left(k^{3p-1}\right)+\mathcal{O}\left(k^p\right)=\mathcal{O}\left(k^{2p}\right).\end{aligned}$$

因此，我们可以验证分子的主导阶为 $p+2p+(2p-2) = 5p-2$，分母的主导阶为 $p+p+2p = 4p$。由于 $5p-2-4p = p-2$，我们得到 $s_k \in \mathcal{O}\left(k^{p-2}\right)$，这与连续对应形式相符。因此 $\sum_{k=1}^{\infty} s_k < \infty$。

另一方面，我们可以看到

$$\begin{aligned} q_k &= \frac{k^{2p}\left(k^p+\frac{\gamma}{2}\right)}{k^p+\gamma} - (k-1)^p\left((k-1)^p+\frac{\gamma}{2}\right) \\ &\leq k^{2p} - \left(k^p - pk^{p-1} + \mathcal{O}\left(k^{p-2}\right)\right)\left(\left(k^p - pk^{p-1} + \mathcal{O}\left(k^{p-2}\right)\right) + \frac{\gamma}{2}\right) \\ &= -\frac{\gamma}{2}k^p + 2pk^{2p-1} + \mathcal{O}\left(k^{2p-2}\right). \end{aligned}$$

由于 $p > 2p-1$，我们有 $\lim_{k\to\infty} q_k = -\infty$。注意，这也与连续对应形式相符。因此，存在 $N > 0$，使得对于 $k > N$，$q_k < 0$。现在，对于 $k > N$，我们有

$$U_{k+1} \leq U_k + s_k \left\|x^k - x^0\right\|^2 \leq U_N + 4\left(\sum_{m=N}^{\infty} s_m\right)\left\|x^0 - x^\star\right\|^2 = M.$$

因此，对于 $k > N$，根据 $\mathbb{A}$ 的单调性和杨氏不等式

$$\begin{aligned} M \geq U^{k+1} &= k^p\left(k^p+\frac{\gamma}{2}\right)\left\|\tilde{\mathbb{A}}x^{k+1}\right\|^2 + \gamma k^p\left\langle\tilde{\mathbb{A}}x^{k+1}, x^{k+1}-x^0\right\rangle + \frac{k^p\gamma^2}{4\left(k^p+\frac{\gamma}{2}\right)}\left\|x^{k+1}-x^0\right\|^2 \\ &\geq k^p\left(k^p+\frac{\gamma}{2}\right)\left\|\tilde{\mathbb{A}}x^{k+1}\right\|^2 + \gamma k^p\left\langle\tilde{\mathbb{A}}x^{k+1}, x^\star-x^0\right\rangle \\ &\geq k^p\left(k^p+\frac{\gamma}{2}\right)\left\|\tilde{\mathbb{A}}x^{k+1}\right\|^2 - \frac{k^p}{2}\left(\left(k^p+\frac{\gamma}{2}\right)\left\|\tilde{\mathbb{A}}x^{k+1}\right\|^2 + \frac{\gamma^2}{k^p+\frac{\gamma}{2}}\left\|x^\star-x^0\right\|^2\right) \\ &= \frac{k^p\left(k^p+\frac{\gamma}{2}\right)}{2}\left\|\tilde{\mathbb{A}}x^{k+1}\right\|^2 - \frac{k^p\gamma^2}{2k^p+\gamma}\left\|x^\star-x^0\right\|^2. \end{aligned}$$

重新整理后，即得所需结果。

$$\left\|\tilde{\mathbb{A}}x^{k+1}\right\|^2 \leq \frac{2}{k^p\left(k^p+\frac{\gamma}{2}\right)}\left(M + \frac{\gamma^2}{2}\left\|x^0-x^\star\right\|^2\right) = \mathcal{O}\left(\frac{1}{k^{2p}}\right).$$

(iii) $\left\|\tilde{\mathbb{A}}(x^k)\right\|^2 = \mathcal{O}\left(\frac{1}{k^2}\right)$ 对于 $p = 1$，$\gamma \geq 1$。代入

$$\begin{aligned} a_{k+1} &= k^2, & b_{k+1} &= \gamma\left(k - \frac{1}{2}(\gamma-1)\right), & c_{k+1} &= \frac{1}{4}\gamma(\gamma-1) \\ \lambda_k &= k(k+1), & \tau_k &= c_{k+1} \end{aligned}$$

(41) 和 (42)，可得

$$
\begin{aligned}
&U^{k+1}-U^k+\left(k(k+1)-\frac{1}{4}\gamma(\gamma-1)\right)\left\langle\tilde{\mathbb{A}}x^{k+1}-\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle\\
&=-(\gamma-1)(k+1)\|\tilde{\mathbb{A}}x^{k+1}\|^2-\frac{k^2(\gamma-1)}{k+\gamma}\|\tilde{\mathbb{A}}x^k\|^2\\
&\quad-\frac{\gamma(\gamma-1)\left(k+\frac{\gamma}{4}\right)}{k}\langle\tilde{\mathbb{A}}x^{k+1},x^{k+1}-x^0\rangle-\frac{\gamma(\gamma-1)\left(k-\frac{\gamma}{4}\right)}{k+\gamma}\langle\tilde{\mathbb{A}}x^k,x^k-x^0\rangle\\
&\quad-\frac{\gamma^2(\gamma-1)}{4k}\left\|x^{k+1}-x^0\right\|^2-\frac{\gamma^2(\gamma-1)}{4(k+\gamma)}\left\|x^k-x^0\right\|^2\\
&=-(\gamma-1)(k+1)\left\|\tilde{\mathbb{A}}x^{k+1}+\frac{\gamma}{2(k+1)}\left(1+\frac{\gamma}{4k}\right)\left(x^{k+1}-x^0\right)\right\|^2\\
&\quad-\frac{k^2(\gamma-1)}{k+\gamma}\left\|\tilde{\mathbb{A}}x^k+\frac{\gamma}{2k}\left(1-\frac{\gamma}{4k}\right)\left(x^k-x^0\right)\right\|^2\\
&\quad-\underbrace{\frac{\gamma^2(\gamma-1)}{64k(k+1)}\left(8(\gamma-2)-\frac{\gamma^2}{k}\right)}_{=s_{k,1}=\mathcal{O}\left(\frac{1}{k^2}\right)}\left\|x^{k+1}-x^0\right\|^2-\underbrace{\frac{\gamma^3(\gamma-1)}{64k(k+\gamma)}\left(8-\frac{\gamma}{k}\right)}_{=s_{k,0}=\mathcal{O}\left(\frac{1}{k^2}\right)}\left\|x^k-x^0\right\|^2.
\end{aligned}
$$

上述等式的连续对应形式为

$$
\dot{U}_3(t)+2t^2\left\langle\frac{d}{dt}\tilde{\mathbb{A}}(X(t)),\dot{X}(t)\right\rangle=-2t(\gamma-1)\left\|\tilde{\mathbb{A}}(X(t))+\frac{\gamma}{t}(X(t)-X_0)\right\|^2
$$

这可以通过对 E_3的守恒律求导得到。注意，除了两个 $\left\|x^{k+1}-x^0\right\|^2$和 $\left\|x^k-x^0\right\|^2$之外，各项的系数阶数都相同，而这些项是可求和的，因为$s_{k,\ 1}$，$s_{k,\ 0}=\mathcal{O}\left(\frac{1}{k^2}\right)$。因此我们有

$$
\begin{aligned}
U^{k+1}&\leq U^k-s_{k,1}\left\|x^{k+1}-x^0\right\|^2-s_{k,0}\left\|x^k-x^0\right\|^2\\
&\leq U^1+4\left\|x^0-x^\star\right\|^2\sum_{m=1}^{\infty}(s_{m,1}+s_{m,0})=M.
\end{aligned}
$$

因此，根据 A 的单调性和杨氏不等式

$$
\begin{aligned}
M\geq U^k&=k^2\left\|\tilde{\mathbb{A}}x^{k+1}\right\|^2+\gamma\left(k-\frac{1}{2}(\gamma-1)\right)\left\langle\tilde{\mathbb{A}}x^k,x^k-x^0\right\rangle+\frac{1}{4}\gamma(\gamma-1)\left\|x^k-x^0\right\|^2\\
&=k^2\left\|\tilde{\mathbb{A}}x^{k+1}\right\|^2+\gamma\left(k-\frac{1}{2}(\gamma-1)\right)\left\langle\tilde{\mathbb{A}}x^k,x^\star-x^0\right\rangle\\
&\geq k^2\left\|\tilde{\mathbb{A}}x^k\right\|^2-\frac{1}{2}\left(\left(k-\frac{1}{2}(\gamma-1)\right)^2\left\|\tilde{\mathbb{A}}x^k\right\|^2+\gamma^2\left\|x^\star-x^0\right\|^2\right)\\
&\geq\frac{k^2}{2}\left\|\tilde{\mathbb{A}}x^k\right\|^2-\frac{\gamma^2}{2}\left\|x^\star-x^0\right\|^2.
\end{aligned}
$$

重新整理后，得到所需结果。

$$
\left\|\tilde{\mathbb{A}}x^k\right\|^2\leq\frac{2}{k^2}\left(M+\frac{\gamma^2}{2}\left\|x^0-x^\star\right\|^2\right)=\mathcal{O}\left(\frac{1}{k^2}\right).
$$

（iv） $\left\|\tilde{\mathbb{A}}\left(x^k\right)\right\|^2=\mathcal{O}\left(\frac{1}{k^{2\gamma}}\right)$ 对于 $p=1$、$0<\gamma<1$。代入

$$
\begin{aligned}
&a_k=k^{2\gamma},&&b_k=\gamma k\left(k-\frac{1}{4}\right)^{2\gamma-2},&&c_{k+1}=\frac{1}{4}\gamma(\gamma-1)k^{2\gamma-2}\\
&\lambda_k=k^{2\gamma-1}(k+\gamma),&&\tau_k=c_{k+1}&&
\end{aligned}
$$

（41）和（42），并令 $p=1$，可得

$$
\begin{aligned}
&U^{k+1}-U^k+\left(k^{2\gamma-1}(k+\gamma)+\frac{1}{4}\gamma(1-\gamma)k^{2\gamma-2}\right)\left\langle\tilde{\mathbb{A}}x^{k+1}-\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle\\
&=\underbrace{\left((k+1)^{2\gamma}-k^{2\gamma-2}(k+\gamma)^2\right)}_{=q_k}\left\|\tilde{\mathbb{A}}x^{k+1}\right\|^2\\
&\quad+\underbrace{\left(\gamma(k+1)\left(k+\frac{3}{4}\right)^{2\gamma-2}-\gamma k^{2\gamma-2}(k+\gamma)-\left(1+\frac{k+\gamma}{k}\right)\frac{1}{4}\gamma(\gamma-1)k^{2\gamma-2}\right)}_{=s_{k,1}}\left\langle\tilde{\mathbb{A}}x^{k+1},x^{k+1}-x^0\right\rangle\\
&\quad+\underbrace{\left(\gamma k^{2\gamma-1}-\gamma k\left(k-\frac{1}{4}\right)^{2\gamma-2}-\left(1+\frac{k}{k+\gamma}\right)\frac{1}{4}\gamma(\gamma-1)k^{2\gamma-2}\right)}_{=s_{k,0}}\left\langle\tilde{\mathbb{A}}x^{k},x^{k}-x^0\right\rangle\\
&\quad-\frac{1}{4}\gamma^2(\gamma-1)k^{2\gamma-3}\left\|x^{k+1}-x^0\right\|^2-\frac{1}{4}\gamma^2(\gamma-1)k^{2\gamma-3}\frac{1}{1+\frac{\gamma}{k}}\left\|x^{k}-x^0\right\|^2\\
&\quad+\frac{1}{4}\gamma(\gamma-1)\left(k^{2\gamma-2}-(k-1)^{2\gamma-2}\right)\left\|x^{k}-x^0\right\|^2
\end{aligned}
$$

这个等式的连续对应形式为

$$
\dot{U}_4(t)+2t^{2\gamma}\left\langle\frac{d}{dt}\tilde{\mathbb{A}}(X(t)),\dot{X}(t)\right\rangle ds=-2\gamma(\gamma-1)t^{2\gamma-3}\left\|X(t)-X_0\right\|^2
$$

它可以通过对关于 E_4 的守恒律求导得到。因此，我们可以预期匹配项的阶数相同，而连续版本中没有出现的项不会影响我们想要得到的结论。我们来验证这一预期是否成立。

项 $\left\|x^{k+1}-x^0\right\|^2$ 和 $\left\|x^{k}-x^0\right\|^2$ 对应于 $\|X-X_0\|^2$。$\left\|x^{k+1}-x^0\right\|^2$ 和 $\left\|x^{k}-x^0\right\|^2$ 的系数显然是 $\mathcal{O}\left(k^{2\gamma-3}\right)$，这与连续对应形式的阶数相同。由于 $\gamma<1$，我们知道这些项是可求和的。

接下来我们观察 $q_k\le 0$。观察

$$
\begin{aligned}
q_k\le 0 &\iff \frac{(k+1)^{2\gamma}}{k^{2\gamma}}\le\frac{(k+\gamma)^2}{k^2}\\
&\iff \left(1+\frac{1}{k}\right)^{2\gamma}\le\left(1+\frac{\gamma}{k}\right)^2\iff\left(1+\frac{1}{k}\right)^{\gamma}\le 1+\frac{\gamma}{k}.
\end{aligned}
$$

为了检验 l 最后一个不等式是否成立，考虑 $f(x)=x^\gamma$。由于该函数在 $0<\gamma<1$ 上是凹的，我们可知 e

$$
\left(1+\frac{1}{k}\right)^{\gamma}=f\left(1+\frac{1}{k}\right)\le f(1)+\frac{1}{k}f'(1)=1+\frac{\gamma}{k}.
$$

最后我们关注 $s_{k,\ 0}$ 和 $s_{k,\ 1}$。由于连续版本中没有交叉项，我们可以预期这些项“很小”，或者用数学术语来说，是可求和的。根据柯西-施瓦茨不等式，我们知道

$$
\begin{aligned}
\left\langle\tilde{\mathbb{A}}x^{k+1},x^{k+1}-x^0\right\rangle&\le\left\|\tilde{\mathbb{A}}x^{k+1}\right\|\left\|x^{k+1}-x^0\right\|\\
\left\langle\tilde{\mathbb{A}}x^{k},x^{k}-x^0\right\rangle&\le\left\|\tilde{\mathbb{A}}x^{k}\right\|\left\|x^{k}-x^0\right\|.
\end{aligned}
$$

由于根据情形 (i) 的证明，$\left\|\tilde{\mathbb{A}}x^{k+1}\right\|$ 和 $\left\|\tilde{\mathbb{A}}x^{k}\right\|$ 有界，因此两个内积项有界。于是，只要证明 $\sum_{k=1}^{\infty}|s_{k,0}|$，$\sum_{k=1}^{\infty}|s_{k,1}|<\infty$，就可以得出 U^k 有界。考虑牛顿展开，可得

$$
\begin{aligned}
&\gamma(k+1)\left(k+\frac{3}{4}\right)^{2\gamma-2}-\gamma k^{2\gamma-2}(k+\gamma)\\
&=\gamma k^{2\gamma-1}+\gamma\left(1+\frac{3}{2}(\gamma-1)\right)k^{2\gamma-2}+\mathcal{O}\left(k^{2\gamma-3}\right)-\gamma k^{2\gamma-1}-\gamma^2k^{2\gamma-2}\\
&=\frac{1}{2}\gamma(\gamma-1)k^{2\gamma-2}+\mathcal{O}\left(k^{2\gamma-3}\right)
\end{aligned}
$$

因此

$$s_{k,1}=\frac{1}{2}\gamma(\gamma-1)k^{2\gamma-2}+\mathcal{O}\left(k^{2\gamma-3}\right)-\left(2+\frac{\gamma}{k}\right)\frac{1}{4}\gamma(\gamma-1)k^{2\gamma-2}=\mathcal{O}\left(k^{2\gamma-3}\right).$$

用类似的论证

$$\begin{aligned}s_{k,0}&=\gamma k^{2\gamma-1}-\gamma k\left(k-\frac{1}{4}\right)^{2\gamma-2}-\left(1+\frac{k}{k+\gamma}\right)\frac{1}{4}\gamma(\gamma-1)k^{2\gamma-2}\\&=\gamma k^{2\gamma-1}-\gamma k\left(k^{2\gamma-2}-(2\gamma-2)\frac{1}{4}k^{2\gamma-3}+\mathcal{O}\left(k^{2\gamma-4}\right)\right)-\left(2-\frac{\gamma}{k+\gamma}\right)\frac{1}{4}\gamma(\gamma-1)k^{2\gamma-2}\\&=\frac{1}{2}\gamma(\gamma-1)k^{2\gamma-2}+\mathcal{O}\left(k^{2\gamma-3}\right)-\left(\frac{1}{2}\gamma(\gamma-1)k^{2\gamma-2}+\mathcal{O}\left(k^{2\gamma-3}\right)\right)=\mathcal{O}\left(k^{2\gamma-3}\right)\end{aligned}$$

因此，我们有

$$\begin{aligned}U^{k+1}&\leq U^k+s_{k,1}\left\|\tilde{\mathbb{A}}x^{k+1}\right\|\left\|x^{k+1}-x^0\right\|+s_{k,0}\left\|\tilde{\mathbb{A}}x^k\right\|\left\|x^k-x^0\right\|\\&\quad-c_{k+1}\frac{\gamma}{k}\left\|x^{k+1}-x^0\right\|^2-c_{k+1}\frac{\gamma}{k+\gamma}\left\|x^k-x^0\right\|^2+(c_{k+1}-c_k)\left\|x^k-x^0\right\|^2\\&\leq\sum_{m=1}^{\infty}\left(|s_{m,1}|\left\|\tilde{\mathbb{A}}x^{k+1}\right\|\left\|x^{k+1}-x^0\right\|+|s_{m,0}|\left\|\tilde{\mathbb{A}}x^{k+1}\right\|\left\|x^{k+1}-x^0\right\|\right)\\&\quad+\sum_{m=1}^{\infty}\left(\left|c_{m+1}\frac{\gamma}{m}\right|\left\|x^{k+1}-x^0\right\|^2+\left(\left|c_{m+1}\frac{\gamma}{m+\gamma}\right|+|c_{k+1}-c_k|\right)\left\|x^k-x^0\right\|^2\right)=M_1.\end{aligned}$$

根据杨氏不等式

$$\begin{aligned}M_1+\frac{1}{4}\gamma(1-\gamma)k^{2\gamma-2}\left\|x^k-x^0\right\|^2&\geq k^{2\gamma}\|\tilde{\mathbb{A}}x^k\|^2+\gamma k\left(k-\frac{1}{4}\right)^{2\gamma-2}\left\langle\tilde{\mathbb{A}}x^k,x^k-x^0\right\rangle\\&\geq k^{2\gamma}\|\tilde{\mathbb{A}}x^k\|^2+\gamma k\left(k-\frac{1}{4}\right)^{2\gamma-2}\left\langle\tilde{\mathbb{A}}x^k,x^\star-x^0\right\rangle\\&=k^{2\gamma}\|\tilde{\mathbb{A}}x^k\|^2+\left\langle k^\gamma\tilde{\mathbb{A}}x^k,\gamma k^{1-\gamma}\left(k-\frac{1}{4}\right)^{2\gamma-2}\left(x^\star-x^0\right)\right\rangle\\&\geq k^{2\gamma}\|\tilde{\mathbb{A}}x^k\|^2-\frac{1}{2}k^{2\gamma}\|\tilde{\mathbb{A}}x^k\|^2-\frac{\gamma^2}{2}\left(k^{1-\gamma}\left(k-\frac{1}{4}\right)^{2\gamma-2}\right)^2\left\|x^\star-x^0\right\|^2.\end{aligned}$$

由于 $k^{1-\gamma}\left(k-\frac{1}{4}\right)^{2\gamma-2}=\mathcal{O}\left(k^{\gamma-1}\right)$ 和 $\gamma<1$，当 $k\to\infty$ 时，这些项趋于零，因此存在某个 $M_2>0$，使得对所有 $k\geq 1$ 都有 $\frac{1}{2}\left(\gamma k^{1-\gamma}\left(k-\frac{1}{4}\right)^{2\gamma-2}\right)^2\left\|x^k-x^\star\right\|^2\leq M_2$。重新整理各项，得到所需结果

$$\|\tilde{\mathbb{A}}x^k\|^2\leq\frac{2}{k^{2\gamma}}\left(M_1+M_2+\gamma(1-\gamma)k^{2\gamma-2}\left\|x^\star-x^0\right\|^2\right)=\mathcal{O}\left(\frac{1}{k^{2\gamma}}\right).$$

H 强单调 $\mathbb{A}$的收敛分析证明

H.1 命题6.1的证明

我们按照命题 3.2 的证明中所做的那样，取膨胀坐标 $W(t)=C(t)(X(t)-X_0)$。回顾式 (30)，常微分方程的二阶形式写为 $0=\ddot{W}-\beta(t)\dot{W}+C(t)\frac{d}{dt}\tilde{\mathbb{A}}(X(t))$。现在，我们在常微分方程两边乘以 $R(t)^2$，得到

$$0=R(t)^2\left(\ddot{W}-\beta(t)\dot{W}+C(t)\frac{d}{dt}\tilde{\mathbb{A}}(X)\right).$$

现在与 W 取内积并积分，得到

$$E_1 \equiv \frac{R(t)^2}{2}\left\|\dot{W}(t)\right\|^2 - \underbrace{\int_{t_0}^{t} R(s)^2\left(\frac{\dot{R}(s)}{R(s)}\left\|\dot{W}(s)\right\|^2\right)ds}_{*} - \int_{t_0}^{t}\beta(s)R(s)^2\left\|\dot{W}(s)\right\|^2 ds$$

$$+\int_{t_0}^{t} C(s)R(s)^2\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)),\dot{W}(s)\right\rangle ds.$$

注意，通过分部积分得到的第二项，在 $R(t)=1$ 如命题 3.2 中那样时不会出现。这是利用条件 A 是强单调的关键项。

现在再次利用 $\dot{W}(t)=C(t)\left(\dot{X}(t)+\beta(s)(X(t)-X_0)\right)$，我们将最后一项改写为

$$\int_{t_0}^{t} C(s)R(s)^2\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)),\dot{W}(s)\right\rangle ds$$

$$=\int_{t_0}^{t} C(s)^2R(s)^2\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)),\dot{X}(s)\right\rangle ds+\int_{t_0}^{t}\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)),C(s)R(s)^2\beta(s)W(s)\right\rangle ds.$$

对第二项进行分部积分，我们得到

$$\int_{t_0}^{t}\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)),C(s)R(s)^2\beta(s)W(s)\right\rangle ds-\left[\left\langle\tilde{\mathbb{A}}(X(W(s),s)),C(s)R(s)^2\beta(s)W(s)\right\rangle\right]_{t_0}^{t}$$

$$=-\int_{t_0}^{t}\left\langle\tilde{\mathbb{A}}(X(W,t)),\left(\beta(s)^2+2\frac{\dot{R}(s)}{R(s)}\beta(s)+\dot{\beta}(s)\right)C(s)R(s)^2W(s)+C(s)R(s)^2\beta(s)\dot{W}(s)\right\rangle ds$$

$$=\int_{t_0}^{t}\beta(s)R(s)^2\left\|\dot{W}(s)\right\|^2 ds+\int_{t_0}^{t}\left(\beta(s)^2+\underbrace{2\frac{\dot{R}(s)}{R(s)}\beta(s)}_{*}+\dot{\beta}(s)\right)R(s)^2\left\langle\dot{W}(s),W(s)\right\rangle ds.$$

事实 $C(t)\tilde{\mathbb{A}}(X(W,t))=-W(t)$ 被应用于第二个等式。注意，微积分基本定理是成立的，因为 $C(s)R(s)^2\beta(s)W(s)$ 可微，$\tilde{\mathbb{A}}(X(W(s),t))$ 根据引理 B.4 在 $[t_0,t]$ 上是 Lipschitz 连续的，因此它们的内积在 $[t_0,t]$ 上是绝对连续的。

现在考虑第二个被积函数中除标有 * 的项以外的部分。根据分部积分，我们有

$$\int_{t_0}^{t}\left(\beta(s)^2+\dot{\beta}(s)\right)R(s)^2\left\langle\dot{W}(s),W(s)\right\rangle ds-\left[\left(\beta(s)^2+\dot{\beta}(s)\right)R(s)^2\frac{1}{2}\|W(s)\|^2\right]_{t_0}^{t}$$

$$=-\frac{1}{2}\int_{t_0}^{t}\left(\left(2\beta(s)\dot{\beta}(s)+\ddot{\beta}(s)\right)R(s)^2+\left(\underbrace{\beta(s)^2}_{*}+\dot{\beta}(s)\right)2R(s)\dot{R}(s)\right)\|W(s)\|^2 ds.$$

除标有 * 的项以外，*被积函数可改写为

$$\frac{1}{2}\int_{t_0}^{t}\left(2\beta(s)\dot{\beta}(s)R(s)+\ddot{\beta}(s)R(s)+2\dot{\beta}(s)\dot{R}(s)\right)R(s)\|W(s)\|^2 ds$$

$$=\frac{1}{2}\int_{t_0}^{t}C(s)^2\left(2\beta(s)\dot{\beta}(s)R(s)+\ddot{\beta}(s)R(s)+2\dot{\beta}(s)\dot{R}(s)\right)R(s)\|X(s)-X_0\|^2 ds$$

$$=\frac{1}{2}\int_{t_0}^{t}\frac{d}{ds}\left(C(s)^2R(s)^2\dot{\beta}(s)\right)\|X(s)-X_0\|^2 ds.$$

现在将标有 * 的项收集起来，得到

$$-\int_{t_0}^{t}R(s)^2\left(\frac{\dot{R}(s)}{R(s)}\left\|\dot{W}(s)\right\|^2\right)ds+\int_{t_0}^{t}2\frac{\dot{R}(s)}{R(s)}\beta(s)R(s)^2\left\langle\dot{W}(s),W(s)\right\rangle ds-\int_{t_0}^{t}\beta(s)^2R(s)\dot{R}(s)\|W(s)\|^2 ds$$

$$=-\int_{t_0}^{t}R(s)^2\frac{\dot{R}(s)}{R(s)}\left\|\dot{W}(s)-\beta(s)W(s)\right\|^2 ds=-\int_{t_0}^{t}R(s)^2C(s)^2\frac{\dot{R}(s)}{R(s)}\left\|\dot{X}(s)\right\|^2 ds.$$

汇总所有结果，我们得到

$$
\begin{aligned}
E_1 \equiv & \frac{R(t)^2}{2}\left\|\dot{W}(t)\right\|^2+\left[\left\langle\tilde{\mathbb{A}}(X(W,s)),C(s)R(s)^2\beta(s)W(s)\right\rangle\right]_{t_0}^t+\left[\left(\beta(s)^2+\dot{\beta}(s)\right)R(s)^2\frac{1}{2}\|W(s)\|^2\right]_{t_0}^t \\
& +\int_{t_0}^t C(s)^2R(s)^2\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)),\dot{X}(s)\right\rangle ds-\int_{t_0}^t R(s)^2C(s)^2\frac{\dot{R}(s)}{R(s)}\left\|\dot{X}(s)\right\|^2 ds \\
& +\frac{1}{2}\int_{t_0}^t\frac{d}{ds}\left(C(s)^2R(s)^2\dot{\beta}(s)\right)\|X(s)-X_0\|^2 ds \\
= & \frac{C(t)^2R(t)^2}{2}\left(\|\mathbb{A}(X(t))\|^2+2\beta(t)\left\langle\mathbb{A}(X(t)),X(t)-X_0\right\rangle+\left(\beta(t)^2+\dot{\beta}(t)\right)\|X(t)-X_0\|^2\right) \\
& +\int_{t_0}^t C(s)^2R(s)^2\left(\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)),\dot{X}(s)\right\rangle-\frac{\dot{R}(s)}{R(s)}\left\|\dot{X}(s)\right\|^2\right)ds-\int_{t_0}^t\frac{d}{ds}\left(\frac{C(s)^2R(s)^2\dot{\beta}(s)}{2}\right)\|X(s)-X_0\|^2 ds \\
& \underbrace{-\frac{C(t_0)^2R(t_0)^2}{2}\left(2\beta(t_0)\left\langle\mathbb{A}(X(t_0)),X(t_0)-X_0\right\rangle+\left(\beta(t_0)^2+\dot{\beta}(t_0)\right)\|X(t_0)-X_0\|^2\right)}_{\text{constant}}.
\end{aligned}
$$

将常数重命名为 $E=E_1-$，即可得到所需结果。

H.2 定理 6.2 的证明

H.2.1 不等式 $V(t)\leq V(0)$ 的证明

证明的基本结构与附录 E.3 相同。这里不再重复完整证明，而是检查附录 E.3 中的各个步骤对于定理 6.2 的设定是否同样成立。

˜(i) V对于 Lipschitz 连续的 μ-强单调 $\mathbb{A}$是非递增的。回顾一下，定理 6.2 中的 V定义为

$$
V(t)=\frac{(e^{\mu t}-e^{-\mu t})^2}{2}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2+2\mu(1-e^{-2\mu t})\left\langle\tilde{\mathbb{A}}(X(t)),X(t)-X_0\right\rangle-2\mu^2\left(1-e^{-2\mu t}\right)\|X(t)-X_0\|^2. \tag{47}
$$

我们先检查下面的等式是否成立。

$$
V(t)=\frac{C(t)^2R(t)^2}{2}\left(\left\|\tilde{\mathbb{A}}(X(t))\right\|^2+2\beta(t)\left\langle\tilde{\mathbb{A}}(X(t)),X(t)-X_0\right\rangle+\left(\beta(t)^2+\dot{\beta}(t)\right)\|X(t)-X_0\|^2\right). \tag{48}
$$

回顾一下，我们正在考虑 (10)，$\beta(t)=\frac{2\mu}{e^{2\mu t}-1}$。由于 $\frac{d}{dt}\log\left(1-e^{-2\mu t}\right)=\frac{2\mu e^{-2\mu t}}{1-e^{-2\mu t}}=\frac{2\mu}{e^{2\mu t}-1}=\beta(t)$，我们有

$$
C(t)=e^{\int_\infty^t\frac{2\mu}{e^{2\mu s}-1}ds}=e^{\log\left(1-e^{-2\mu t}\right)}=1-e^{-2\mu t}.
$$

由于 $\dot{\beta}(t)=-\frac{4\mu^2e^{2\mu t}}{(e^{2\mu t}-1)^2}$ 和 $R(t)=e^{\mu t}$，

$$
\begin{aligned}
\frac{C(t)^2R(t)^2}{2}&=\frac{1}{2}\left(1-e^{-2\mu t}\right)^2e^{2\mu t}=\frac{\left(e^{\mu t}-e^{-\mu t}\right)^2}{2} \\
C(t)^2R(t)^2\beta(t)&=e^{-2\mu t}\left(e^{2\mu t}-1\right)^2\frac{2\mu}{e^{2\mu t}-1}=2\mu\left(1-e^{-2\mu t}\right) \\
\frac{C(t)^2R(t)^2}{2}\left(\beta(t)^2+\dot{\beta}(t)\right)&=\frac{\left(e^{\mu t}-e^{-\mu t}\right)^2}{2}\left(\left(\frac{2\mu}{e^{2\mu t}-1}\right)^2-\frac{4\mu^2e^{2\mu t}}{\left(e^{2\mu t}-1\right)^2}\right)=2\mu^2\left(e^{-2\mu t}-1\right).
\end{aligned}
$$

这就证明了所需的等式。现在我们证明

$$
V(t)=E-\int_{t_0}^t C(s)^2R(s)^2\left(\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)),\dot{X}(s)\right\rangle-\frac{\dot{R}(s)}{R(s)}\left\|\dot{X}(s)\right\|^2\right)ds, \tag{49}
$$

其中 E 来自命题 6.1，其定义为

$$E = \frac{C(t)^2R(t)^2}{2}\left(\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\beta(t)\left\langle\tilde{\mathbb{A}}(X(t)), X(t)-X_0\right\rangle + \left(\beta(t)^2+\dot{\beta}(t)\right)\|X(t)-X_0\|^2\right)$$
$$+\int_{t_0}^{t} C(s)^2R(s)^2\left(\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)),\dot{X}(s)\right\rangle - \frac{\dot{R}(s)}{R(s)}\left\|\dot{X}(s)\right\|^2\right)ds - \int_{t_0}^{t}\frac{d}{ds}\left(\frac{C(s)^2R(s)^2\dot{\beta}(s)}{2}\right)\|X(t)-X_0\|^2\,ds.$$

由 (48) 和 E 的定义，只需证明 $\frac{d}{ds}\left(\frac{C(s)^2R(s)^2\dot{\beta}(s)}{2}\right) = 0$。由于

$$\frac{C^2(t)R^2(t)\dot{\beta}(t)}{2} = \left(1-e^{-2\mu t}\right)^2 e^{2\mu t}\left(-\frac{4\mu^2e^{2\mu t}}{\left(e^{2\mu t}-1\right)^2}\right) = -4\mu^2,$$

我们看到 $\frac{d}{ds}\left(\frac{C(s)^2R(s)^2\dot{\beta}(s)}{2}\right) = 0$。现在，由于 $\mathbb{A}$ 是利普希茨连续的，根据命题 6.1，我们知道 E 是常数。因此，由 (49)，对于 $t>0$、$|h|<t$，有

$$V(t+h)-V(t) = \int_t^{t+h} C(s)^2R(s)^2\left(\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)),\dot{X}(s)\right\rangle - \frac{\dot{R}(s)}{R(s)}\left\|\dot{X}(s)\right\|^2\right)ds.$$

由于 $\frac{\dot{R}(s)}{R(s)} = \mu$，且 $\tilde{\mathbb{A}}$ 是 μ-强单调的，由 (2) 可知

$$\left\langle\frac{d}{ds}\tilde{\mathbb{A}}(X(s)),\dot{X}(s)\right\rangle - \frac{\dot{R}(s)}{R(s)}\left\|\dot{X}(s)\right\|^2 \geq 0.$$

因此，对于 $h>0$，有 $V(t+h)-V(t)\geq 0$，即可得到所需结果。

(ii) 计算 Lipschitz 连续单调 $\mathbb{A}$ 的 $V(0)$。将 $t=0$ 代入 (47)，即可得到 $V(0)=0$。

(iii) $V(t)\leq 0$ 对所有 $t\in[0,\infty)$ 都成立，且对一般的最大 μ-强单调 $\mathbb{A}$ 也成立

我们检查了附录 E.3.3 和附录 E.3.4 中使用的论证在此处同样有效。

按引理 E.1 中的定义定义 S。取 $t\in S$，令 $T>t$。如附录 C 中所检查的，用于证明命题 B.11 的论证对 $\beta(t) = \frac{2\mu}{e^{2\mu t}-1}$ 的情形同样有效。这一事实给出了所需的序列 $\{X_{\lambda_n}\}_{n\in\mathbb{N}}$，其中 X_{λ_n} 在 $[0,T]$ 上一致地收敛到 X，而 $\dot{X}_{\lambda_n}$ 在 $L^2([0,T],\mathbb{R}^n)$ 中弱地收敛到 $\dot{X}$。与附录 E.3.3 中一样，将 $\mathbb{A}_\lambda$ 对应的解 V 记作 V_λ。由 (i) 可知，对所有 $n\in\mathbb{N}$，V_{λ_n} 都是非递增的，因此 $\limsup_{n\to\infty} V_{\lambda_n}(t)\leq \limsup_{n\to\infty} V_{\lambda_n}(0) = 0$。

此外，我们还可以验证，引理 E.1 中定义的 $\mathbb{A}$ 的延拓同样成立。由 (21) 和 (22) 可知，对于所有 $\lambda>0$，$t\in[0,T]$，$\left\|\dot{X}_\lambda(t)\right\|\leq \frac{e^{\mu T}}{\sqrt{2}}\|m(\mathbb{A}(X_0))\|$ 且 $\|\mathbb{A}_\lambda(X_\lambda(t))\|\leq\sqrt{e^{2\mu T}+1}\,\|m(\mathbb{A}(X_0))\|$。因此，对于 $\beta(t)=\frac{2\mu}{e^{2\mu t}-1}$ 这一情形，我们可以用同样的证明来证明推论 E.2，只需将 $M_{\mathbb{A}}(T)$ 替换为 $\sqrt{e^{2\mu T}+1}\,\|m(\mathbb{A}(X_0))\|$，并将 $M_{dot}(T)$ 替换为 $\frac{e^{\mu T}}{\sqrt{2}}\|m(\mathbb{A}(X_0))\|$。因此，$\tilde{\mathbb{A}}(X(t))$ 对 $t=0$ 是良定义的；将 $t=0$ 代入 (47)，可得 $V(0)=0$。

因此我们有

$$\limsup_{n\to\infty} V_{\lambda_n}(t)\leq\limsup_{n\to\infty} V_{\lambda_n}(0) = 0 = V(0),$$

还需证明 $V(t)\leq\limsup_{n\to\infty} V_{\lambda_n}(t)$。注意，由于 $t\in S$，等式 $\dot{X}(t) = -\mathbb{A}(X(t)) - \beta(t)(X(t)-X_0)$ 成立，由 (48) 可得

$$V(t) = \frac{C(t)^2R(t)^2}{2}\left(\left\|\dot{X}(t)\right\|^2 + \dot{\beta}(t)\|X(t)-X_0\|^2\right).$$

因此，只需检查引理 E.4 在此处对某个 U_{λ_n} 是否成立。对于某个 $a>0$，定义 $U_{\lambda_n}:[0,\infty)\to\mathbb{R}$ 为

$$U_{\lambda_n}(t) = \left\|\dot{X}_{\lambda_n}(t)\right\|^2 \underbrace{-\dot{\beta}(t)\|X_{\lambda_n}(t)-X_0\|^2 + \int_a^t\left(\ddot{\beta}(s) - \frac{2\dot{\beta}(s)^2}{\beta(s)}\right)\|X_{\lambda_n}(s)-X_0\|^2\,ds}_{f_n(t)}.$$

我们采用与引理 B.5 类似的论证。对 $X_{\lambda_n}(t) = -\mathbb{A}_{\lambda_n}(X(t)) - \beta(t)(X_{\lambda_n}(t) - X_0)$ 求导，可得对于几乎所有 $t > 0$

$$\ddot{X}_{\lambda_n}(t) = -\frac{d}{dt}\mathbb{A}_{\lambda_n}(X(t)) - \dot{\beta}(t)(X_{\lambda_n}(t) - X_0) - \beta(t)\dot{X}_{\lambda_n}(t).$$

因此，对于几乎所有 $t > 0$,

$\dot{U}_{\lambda_n}(t)$

$$\begin{aligned}
&= 2\left\langle \dot{X}_{\lambda_n}(t), \ddot{X}_{\lambda_n}(t)\right\rangle - \ddot{\beta}(t)\left\|X_{\lambda_n}(t) - X_0\right\|^2 - 2\dot{\beta}(t)\left\langle \dot{X}_{\lambda_n}(t), X_{\lambda_n}(t) - X_0\right\rangle + \left(\ddot{\beta}(t) - \frac{2\dot{\beta}(t)^2}{\beta(t)}\right)\left\|X_{\lambda_n}(t) - X_0\right\|^2 \\
&= 2\left\langle \dot{X}_{\lambda_n}(t), -\frac{d}{dt}\mathbb{A}_{\lambda_n}(X(t)) - \dot{\beta}(t)(X_{\lambda_n}(t) - X_0) - \beta(t)\dot{X}(t)\right\rangle \\
&\quad - 2\dot{\beta}(t)\left\langle \dot{X}_{\lambda_n}(t), X_{\lambda_n}(t) - X_0\right\rangle - \frac{2\dot{\beta}(t)^2}{\beta(t)}\left\|X_{\lambda_n}(t) - X_0\right\|^2 \\
&= -2\left\langle \dot{X}_{\lambda_n}(t), \frac{d}{dt}\mathbb{A}_{\lambda_n}(X(t))\right\rangle - 2\beta(t)\left\|\dot{X}_{\lambda_n}(t) + \frac{\dot{\beta}(t)}{\beta(t)}\left(X_{\lambda_n}(t) - X_0\right)\right\|^2 \le 0.
\end{aligned}$$

因此，U_{λ_n} 是非递增的，我们可以用引理 E.4 中相同的论证证明 $\left\|\dot{X}(t)\right\| \le \limsup_{n\to\infty}\left\|\dot{X}_{\lambda_n}(t)\right\|^2$。因此，对于 $t \in S$，我们有 $V(t) \le V(0) = 0$。将结果推广至 $t \in [0,\infty)$，可采用附录 E.3.4 中相同的论证。

H.2.2 收敛率证明

回顾

$$V(t) = \frac{(e^{\mu t} - e^{-\mu t})^2}{2}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\mu(1 - e^{-2\mu t})\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle - 2\mu^2\left(1 - e^{-2\mu t}\right)\left\|X(t) - X_0\right\|^2.$$

注意

$$\begin{aligned}
\frac{2V(t)}{1 - e^{-2\mu t}} &= (e^{2\mu t} - 1)\|\tilde{\mathbb{A}}(X(t))\|^2 + 4\mu\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle - 4\mu^2\left\|X(t) - X_0\right\|^2 \\
&= (e^{\mu t} - 1)\underbrace{\left(\|\tilde{\mathbb{A}}(X(t))\|^2 - 4\mu\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle + 4\mu^2\left\|X(t) - X_0\right\|^2\right)}_{=\left\|\tilde{\mathbb{A}}(X(t)) - 2\mu(X(t) - X_0)\right\|^2} \\
&\quad + e^{\mu t}(e^{\mu t} - 1)\|\tilde{\mathbb{A}}(X(t))\|^2 + \underbrace{e^{\mu t}\left(4\mu\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle - 4\mu^2\left\|X(t) - X_0\right\|^2\right)}_{=p(t)}.
\end{aligned}$$

根据余弦定理，我们有 $\|X(t) - X_0\|^2 = \|X(t) - X_\star\|^2 - 2\langle X(t) - X_0,\ X_0 - X_\star\rangle - \|X_\star - X_0\|^2$。将其应用于 $p(t)$，可得

$$\begin{aligned}
p(t) &= 4\mu e^{\mu t}\big(\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_\star\rangle - \langle \tilde{\mathbb{A}}(X(t)), X_0 - X_\star\rangle - \mu\|X(t) - X_0\|^2\big) \\
&= 4\mu e^{\mu t}\left(\langle \tilde{\mathbb{A}}(X(t)) + \mu(X(t) - X_\star), X(t) - X_\star\rangle - \langle \tilde{\mathbb{A}}(X(t)) - 2\mu(X(t) - X_0), X_0 - X_\star\rangle - \mu\|X_0 - X_\star\|^2.\right)
\end{aligned}$$

Thus

$$\begin{aligned}
\frac{2V(t)}{1 - e^{-2\mu t}} &= (e^{\mu t} - 1)\left\|\tilde{\mathbb{A}}(X(t)) - 2\mu(X(t) - X_0)\right\|^2 + e^{\mu t}(e^{\mu t} - 1)\|\tilde{\mathbb{A}}(X(t))\|^2 + p(t) \\
&= (e^{\mu t} - 1)\|\tilde{\mathbb{A}}(X(t)) - 2\mu(X(t) - X_0)\|^2 - 4\mu e^{\mu t}\langle \tilde{\mathbb{A}}(X(t)) - 2\mu(X(t) - X_0), X_0 - X_\star\rangle \\
&\quad + e^{\mu t}(e^{\mu t} - 1)\|\tilde{\mathbb{A}}(X(t))\|^2 + 4\mu e^{\mu t}\langle \tilde{\mathbb{A}}(X(t)) - \mu(X(t) - X_\star), X(t) - X_\star\rangle + 4\mu^2 e^{\mu t}\|X_0 - X_\star\|^2 \\
&= (e^{\mu t} - 1)\left\|\tilde{\mathbb{A}}(X(t)) - 2\mu(X(t) - X_0) - \frac{2\mu e^{\mu t}}{e^{\mu t} - 1}(X_0 - X_\star)\right\|^2 - \frac{4\mu^2 e^{2\mu t}}{e^{\mu t} - 1}\|X_0 - X_\star\|^2 \\
&\quad + e^{\mu t}(e^{\mu t} - 1)\|\tilde{\mathbb{A}}(X(t))\|^2 + 4\mu e^{\mu t}\left(\left\langle \tilde{\mathbb{A}}(X(t)), X(t) - X_\star\right\rangle - \mu\left\|X(t) - X_\star\right\|^2\right) + 4\mu^2 e^{\mu t}\|X_0 - X_\star\|^2.
\end{aligned}$$

现在，由于 $\tilde{\mathbb{A}}$ 是 μ-强单调的，$\left\langle \tilde{\mathbb{A}}(X(t))\ X(t)-X_\star\right\rangle - \mu\left\|X(t)\ -X_\star\right\|^2 \geq 0$。根据上一节，我们知道$0 = V(0) \geq V(t)$ 对所有 $t>0$都成立。因此，对所有 $t>0$

$$\begin{aligned} 0 \geq \frac{2V(t)}{1-e^{-2\mu t}} &\geq e^{\mu t}(e^{\mu t}-1)\|\tilde{\mathbb{A}}(X(t))\|^2 + \left(4\mu^2 e^{\mu t} - \frac{4\mu^2 e^{2\mu t}}{e^{\mu t}-1}\right)\|X_0 - X_\star\|^2 \\ &= e^{\mu t}(e^{\mu t}-1)\|\tilde{\mathbb{A}}(X(t))\|^2 - \frac{4\mu^2 e^{\mu t}}{e^{\mu t}-1}\|X_0 - X_\star\|^2. \end{aligned}$$

整理后，我们得到

$$\|\tilde{\mathbb{A}}(X(t))\|^2 \leq 4\left(\frac{\mu}{e^{\mu t}-1}\right)^2 \|X_0 - X_\star\|^2.$$

H.2.3 从该方法非正式推导常微分方程

假设 $\mathbb{A}\colon \mathrm{R}^n \to \mathrm{R}^n$ 为一个连续单调算子。在 [54], 中，方法 OS-PPM 表示为

$$\begin{aligned} x^k &= \mathbb{J}_{h\mathbb{A}} y^{k-1} \\ y^k &= \left(1-\frac{1}{s_k}\right)\left\{x^k - \frac{1}{\nu}(y^{k-1}-x^k)\right\} + \frac{1}{s_k}y^0 \end{aligned}$$

其中 $y^0 = x^0$、 $\nu = 1+2h\mu$ 和 $s_k = 1+\nu^2+\cdots+\nu^{2k} = \frac{\nu^{2k+2}-1}{\nu^2-1}$。利用 $y^{k-1} = x^k + h\mathbb{A}x^k$，代入 y^k 和 y^{k-1}，该方法可写成一行，

$$x^{k+1} + h\mathbb{A}x^{k+1} = \left(1-\frac{1}{s_k}\right)\left(x^k - \frac{h}{\nu}\mathbb{A}x^k\right) + \frac{1}{s_k}x^0.$$

重新整理并将两边除以 h，我们有

$$\frac{x^{k+1}-x^k}{h} = -\mathbb{A}x^{k+1} - \left(\frac{1}{\nu} - \frac{1}{\nu s_k}\right)\mathbb{A}x^k - \frac{1}{hs_k}(x^k - x^0).$$

令 $x^0 = X_0$、 $2hk = t$、 $x^k = X(t)$,

$$\frac{X(t+2h)-X(t)}{h} = -\mathbb{A}(X(t+2h)) - \left(\frac{1}{\nu} - \frac{1}{\nu s_k}\right)\mathbb{A}(X(t)) - \frac{1}{hs_k}(X(t)-X_0).$$

现在观察

$$\begin{aligned} \lim_{h\to 0^+} hs_k &= \lim_{h\to 0^+} h\frac{\nu^{2k+2}-1}{\nu^2-1} = \lim_{h\to 0^+} h\frac{(1+2h\mu)^{2k+2}-1}{(1+2h\mu)^2-1} \\ &= \lim_{h\to 0^+} \frac{(1+2h\mu)^2(1+2h\mu)^{t/h}-1}{4\mu(1+h\mu)} = \frac{e^{2\mu t}-1}{4\mu} \\ \lim_{h\to 0^+} \frac{1}{\nu} &= \lim_{h\to 0^+} \frac{1}{1+2h\mu} = 1 \\ \lim_{h\to 0^+} \frac{1}{\nu s_k} &= \lim_{h\to 0^+} \frac{1}{\nu}\frac{1}{hs_k}h = 1\times\frac{4\mu}{e^{2\mu t}-1}\times 0 = 0. \end{aligned}$$

取极限 $h\to 0^+$ 并整理,

$$2\dot{X}(t) = -\mathbb{A}(X(t)) - (1+0)\mathbb{A}(X(t)) - \frac{4\mu}{e^{2\mu t}-1}(X(t)-X_0).$$

两边同时除以 2，即得到所需的锚定常微分方程

$$\dot{X}(t) = -\mathbb{A}(X(t)) - \frac{2\mu}{e^{2\mu t}-1}(X(t)-X_0).$$

I 第 7 节中省略的证明

I.1 定理 7.1 的证明

我们首先检验

$$\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 \leq 4\beta(t)^2 \left\|X_0 - X_\star\right\|^2 = \mathcal{O}\left(\beta(t)^2\right),$$

其中

$$\dot{X}(t) = -\tilde{\mathbb{A}}(X(t)) - \beta(t)(X(t) - X_0)$$
$$\beta(t) = -\frac{\left\|\tilde{\mathbb{A}}(X(t))\right\|^2}{2\left\langle\tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle}.$$

根据 $\beta(t)$ 的定义，在 (8) 中定义的 $\Phi(t)$ 变为零，即

$$0 \equiv \Phi(t) = \left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\beta(t)\left\langle\tilde{\mathbb{A}}(X(t)), X - X_0\right\rangle.$$

将 $\Phi(t) = 0$ 代入不等式 (8)，可得

$$0 \geq \frac{1}{2}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 - 2\beta(t)^2\left\|X_0 - X_\star\right\|^2.$$

重新整理可得　　期望的结果。其他结果需要一些工作，我们分小节逐步给出证明。　　.

I.1.1 $\beta(t) > 0$ 用于 $t > 0$

对常微分方程两边取与 $\tilde{\mathbb{A}}(X)$ 的内积，并应用 $\frac{1}{2}\Phi(t) = 0$，可得

$$\left\langle\dot{X}(t), \tilde{\mathbb{A}}(X(t))\right\rangle = -\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 - \beta(t)\left\langle\tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle = -\frac{1}{2}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2.$$

由于假设 $\mathbb{A}$ 连续，取极限 $t \to 0+$，可得

$$\lim_{t\to 0+}\left\langle\dot{X}(t), \tilde{\mathbb{A}}(X(t))\right\rangle = \left\langle\lim_{t\to 0+}\dot{X}(t), \tilde{\mathbb{A}}(X_0)\right\rangle = -\frac{1}{2}\lim_{t\to 0+}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 = -\frac{1}{2}\left\|\tilde{\mathbb{A}}(X_0)\right\|^2.$$

另一方面，根据假设，我们有 $\lim_{t\to 0+}\dot{X}(t) = \lim_{t\to 0+}{}^{X_{(t)}-X_0}_{t}$，且 $\left\|\tilde{\mathbb{A}}(X_0)\right\| \neq 0$，因此

$$\lim_{t\to 0+}\frac{1}{t}\left\langle\tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle = \left\langle\tilde{\mathbb{A}}(X_0), \lim_{t\to 0+}\dot{X}(t)\right\rangle = -\frac{1}{2}\left\|\tilde{\mathbb{A}}(X_0)\right\|^2 < 0. \tag{50}$$

因此，存在 $\epsilon > 0$ 使得 $\left\langle\tilde{\mathbb{A}}(X(t))\ X(t) - X_0\right\rangle < 0$ 对于 $t \in (0, \epsilon)$ 成立；于是，对于 $t \in (0, \epsilon)$，我们有 $\left\langle\tilde{\mathbb{A}}(X(t))\ X(t) - X_0\right\rangle \neq 0$，因此 $\beta(t)$ 定义良好且满足

$$\beta(t) = -\frac{\left\|\tilde{\mathbb{A}}(X(t))\right\|^2}{2\left\langle\tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle} > 0.$$

注意到，当 $\left\|\tilde{\mathbb{A}}(X(t))\right\| = 0$ 时，$\beta(t)$ 的分母为零。由于假设 β 有定义，因此对于所有 $t > 0$，都有 $\left\|\tilde{\mathbb{A}}(X(t))\right\| \neq 0$。由于 β 关于 $\mathbb{A}$ 连续且 X 连续，依据介值定理，对于所有 $t > 0$，都有 $\left\|\tilde{\mathbb{A}}(X(t))\right\| > 0$。

I.1.2 主要结论的证明

我们先证明以下引理。

引理 I.1. 以下等式几乎处处成立 $t > 0$。

$$\left(\dot{\beta}(t) + \beta(t)^2\right)\left\langle\tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle = \left\langle\frac{d}{dt}\tilde{\mathbb{A}}(X(t)), \dot{X}(t)\right\rangle. \tag{51}$$

˜ ˜证明。 根据假设， $\mathbb{A}$ 是 Lipschitz 连续的。由引理 B.4 可知， $\mathbb{A}(X(t))$几乎处处可微。对 $\Phi(t)$求导，得到

$$
\begin{aligned}
0 &= \dot{\Phi}(t) \\
&= 2\left\langle \frac{d}{dt}\tilde{\mathbb{A}}(X(t)), \tilde{\mathbb{A}}(X(t)) \right\rangle + 2\dot{\beta}(t)\left\langle \tilde{\mathbb{A}}(X(t)), X(t)-X_0 \right\rangle + 2\beta(t)\left\langle \frac{d}{dt}\tilde{\mathbb{A}}(X(t)), X(t)-X_0 \right\rangle + 2\beta(t)\left\langle \tilde{\mathbb{A}}(X(t)), \dot{X}(t) \right\rangle \\
&= 2\left\langle \frac{d}{dt}\tilde{\mathbb{A}}(X(t)), \tilde{\mathbb{A}}(X(t)) + \beta(t)(X(t)-X_0) \right\rangle + 2\dot{\beta}(t)\left\langle \tilde{\mathbb{A}}(X(t)), X(t)-X_0 \right\rangle \\
&\quad + 2\beta(t)\left\langle \tilde{\mathbb{A}}(X(t)), -\tilde{\mathbb{A}}(X(t)) - \beta(t)(X(t)-X_0) \right\rangle \\
&= -2\left\langle \frac{d}{dt}\tilde{\mathbb{A}}(X(t)), \dot{X}(t) \right\rangle - 2\beta(t)\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\left(\dot{\beta}(t) - \beta(t)^2\right)\left\langle \tilde{\mathbb{A}}(X(t)), X(t)-X_0 \right\rangle \\
&= -2\left\langle \frac{d}{dt}\tilde{\mathbb{A}}(X(t)), \dot{X}(t) \right\rangle + 4\beta(t)^2\left\langle \tilde{\mathbb{A}}(X(t)), X(t)-X_0 \right\rangle + 2\left(\dot{\beta}(t) - \beta(t)^2\right)\left\langle \tilde{\mathbb{A}}(X(t)), X(t)-X_0 \right\rangle \\
&= -2\left\langle \frac{d}{dt}\tilde{\mathbb{A}}(X(t)), \dot{X}(t) \right\rangle + 2\left(\dot{\beta}(t) + \beta(t)^2\right)\left\langle \tilde{\mathbb{A}}(X(t)), X(t)-X_0 \right\rangle.
\end{aligned}
$$

. ˜对于几乎处处的 $t \in (0,\infty)$。这些等式来自常微分方程 $\dot{X}(t) = -\mathbb{A}(X(t)) - \beta(t)(X(t) - X_0)$ 以及事实$\Phi(t) = 0$。重新整理后，即得所需的方程 (51)。 □

现在我们分别证明单调情形和强单调情形下 $\beta(t)$ 的上界。由 (51) 和 $\beta(t)$ 的定义可知，对于几乎所有 $t \in (0,\infty)$,

$$
-\frac{\dot{\beta}(t)+\beta(t)^2}{2\beta(t)}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 = \left\langle \frac{d}{dt}\tilde{\mathbb{A}}(X), \dot{X} \right\rangle. \tag{52}
$$

˜(i) 当 $\mathbb{A}$ 单调时。
由 (52) 和 (1) 可知，对于几乎所有 $t \in (0,\infty)$,

$$
-\frac{\dot{\beta}(t)+\beta(t)^2}{2\beta(t)}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 = \left\langle \frac{d}{dt}\tilde{\mathbb{A}}(X), \dot{X} \right\rangle \geq 0.
$$

由于 $\beta(t) > 0$，我们有

$$
\dot{\beta}(t) + \beta(t)^2 \leq 0.
$$

几乎处处。现在将两边除以 $\beta(t)^2$，得到

$$
1 \leq -\frac{\dot{\beta}(t)}{\beta(t)^2}
$$

几乎处处成立。由于 $\frac{\dot{\beta}(t)}{\beta(t)^2} = \frac{d}{dt}\left(\frac{1}{\beta(t)}\right)$，将上述不等式两边从 δ 到 t 积分，得到

$$
t - \delta \leq \frac{1}{\beta(t)} - \frac{1}{\beta(\delta)} \Longrightarrow \beta(t) \leq \frac{1}{t - \delta + \frac{1}{\beta(\delta)}}.
$$

顺便说一下，由（50）可得

$$
\lim_{t\to 0+} t\beta(t) = -\lim_{t\to 0+} \frac{\left\|\tilde{\mathbb{A}}(X(t))\right\|^2}{2\left\langle \tilde{\mathbb{A}}(X(t)), \frac{X(t)-X_0}{t} \right\rangle} = 1,
$$

因此，极限为 $\lim_{\delta\to 0+}\beta(\delta) = \infty$，所以极限为$\lim_{\delta\to 0+}\frac{1}{\beta(\delta)} = 0$。因此，取极限 $\delta \to 0+$，我们得到

$$
\beta(t) \leq \lim_{\delta\to 0+} \frac{1}{t - \delta + \frac{1}{\beta(\delta)}} = \frac{1}{t}.
$$

˜(ii) 当 $\mathbb{A}$ 是 μ-强单调时。
由 (52) 和 (2)，对于几乎所有的 $t \in (0,\infty)$，我们有

$$-\frac{\dot\beta(t)+\beta(t)^2}{2\beta(t)}\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 = \left\langle \frac{d}{dt}\tilde{\mathbb{A}}(X), \dot X\right\rangle \geq \mu\left\|\dot X(t)\right\|^2. \tag{53}$$

另一方面，注意到

$$\begin{aligned}\left\|\dot X(t)\right\|^2 &= \left\|\tilde{\mathbb{A}}(X(t))+\beta(t)(X(t)-X_0)\right\|^2\\ &= \underbrace{\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 + 2\beta(t)\left\langle \tilde{\mathbb{A}}(X(t)), X(t)-X_0\right\rangle}_{=\Phi(t)=0} + \beta(t)^2\left\|X(t)-X_0\right\|^2 = \beta(t)^2\left\|X(t)-X_0\right\|^2.\end{aligned}$$

由柯西-施瓦茨不等式可知

$$\left\|\tilde{\mathbb{A}}(X(t))\right\|^2 = -2\beta(t)\left\langle \tilde{\mathbb{A}}(X(t)), X(t)-X_0\right\rangle \leq 2\beta(t)\left\|\tilde{\mathbb{A}}(X(t))\right\|\left\|X(t)-X_0\right\|,$$

因此 $\left\|\tilde{\mathbb{A}}(X(t))\right\| \leq 2\beta(t)\left\|X(t)-X_0\right\|$。综合以上观察，对于几乎所有的 $t>0$

$$\frac{4\left\|\dot X(t)\right\|^2}{\left\|\tilde{\mathbb{A}}(X(t))\right\|^2} = \frac{4\beta(t)^2\left\|X(t)-X_0\right\|^2}{\left\|\tilde{\mathbb{A}}(X(t))\right\|^2} \geq 1. \tag{54}$$

˙由 (53) 和 (54)，我们得到关于 $\beta(t)$ 和 $\beta(t)$ 的一个不等式

$$-\frac{\dot\beta(t)+\beta(t)^2}{2\beta(t)} = \frac{\dot\beta(t)}{2\beta(t)} + \frac{\beta(t)}{2} \geq \frac{\mu}{4}.$$

将 $\beta(t)$ 2移到右侧并整理可得，对几乎所有 $t>0$

$$1 \leq -\frac{\dot\beta(t)}{\beta(t)(\frac{\mu}{2}+\beta(t))} = -\frac{\dot\beta(t)}{\frac{\mu}{2}}\left(\frac{1}{\beta(t)} - \frac{1}{\frac{\mu}{2}+\beta(t)}\right) \quad\Longrightarrow\quad \frac{\mu}{2} \leq -\frac{\dot\beta(t)}{\beta(t)} + \frac{\dot\beta(t)}{\frac{\mu}{2}+\beta(t)}.$$

对上述不等式两边从 δ 到 t 积分可得

$$\frac{\mu}{2}(t-\delta) \leq \left[-\log\beta(t) + \log\left(\frac{\mu}{2}+\beta(t)\right)\right]_\delta^t = \log\frac{\frac{\mu}{2}+\beta(t)}{\beta(t)} - \log\frac{\frac{\mu}{2}+\beta(\delta)}{\beta(\delta)}.$$

如情形 (i) 中所述，我们知道 $\lim_{\delta\to 0+}\beta(\delta)=\infty$，对两边取 $\delta\to 0+$ 极限可得 e

$$\frac{\mu t}{2} \leq \log\frac{\frac{\mu}{2}+\beta(t)}{\beta(t)} = \log\left(1+\frac{\mu/2}{\beta(t)}\right) \Longrightarrow e^{\mu t/2} \leq 1+\frac{\mu/2}{\beta(t)}.$$

整理可得所需结果。

$$\beta(t) \leq \frac{\mu/2}{e^{\mu t/2}-1}.$$

I.2 常微分方程（11）与定理 7.2 中离散方法之间的对应关系

请注意，对于下文的情况，$h=1$所用的方法与定理 7.2 中给出的方法相对应 .

$$\begin{aligned} x^k &= \mathbb{J}_{h\mathbb{A}} y^{k-1}\\ \beta_k &= \begin{cases} \dfrac{\|h\tilde{\mathbb{A}}x^k\|^2}{-\langle h\tilde{\mathbb{A}}x^k,\, x^k-x^0\rangle + \|h\tilde{\mathbb{A}}x^k\|^2} & \text{if } \|\tilde{\mathbb{A}}x^k\|^2 \neq 0\\ 0 & \text{if } \|\tilde{\mathbb{A}}x^k\|^2 = 0 \end{cases}\\ y^k &= (1-\beta_k)(2x^k - y^{k-1}) + \beta_k x^0. \end{aligned}$$

现在说明，当 $\mathbb{A}$ 连续且对于所有 $k \geq 0$ 均满足 $\|\tilde{\mathbb{A}}x^k\|^2 \neq 0$ 时，取极限 $h \to 0+$ 后可得到常微分方程 (11)。˜请注意，当 $\mathbb{A}$ 连续时， $\mathbb{A}$ 等于 $\mathbb{A}$。

利用 $y^{k-1} = h\tilde{\mathbb{A}}x^k + x^k$，代入 y^k 和 y^{k-1}，得到一个单行表达式

$$h\tilde{\mathbb{A}}x^{k+1} + x^{k+1} = (1-\beta_k)(x^k - h\tilde{\mathbb{A}}x^k) + \beta_k x^0.$$

重新整理并将两边除以 h，得到

$$\frac{x^{k+1}-x^k}{h} = -\tilde{\mathbb{A}}x^{k+1} - (1-\beta_k)\tilde{\mathbb{A}}x^k - \frac{\beta_k}{h}(x^k - x^0).$$

令 $x^0 = X_0$、 $2hk = t$ 和 $x^k = X$（t）。于是可得

$$\frac{\beta_k}{h} = \frac{h^2\|\tilde{\mathbb{A}}x^k\|^2}{-h^2\langle\tilde{\mathbb{A}}x^k,\, x^k - x^0\rangle + h^3\|\tilde{\mathbb{A}}x^k\|^2} = \frac{\|\tilde{\mathbb{A}}(X(t))\|^2}{-\langle\tilde{\mathbb{A}}(X(t)),\, X(t) - X_0\rangle + h\|\tilde{\mathbb{A}}(X(t))\|^2}.$$

因此 $\beta_k = \mathcal{O}$（h），我们有 $\lim_{h\to 0+}\beta_k = 0$。现在取 $h \to 0+$ 的极限，得到

$$2\dot{X}(t) = -2\tilde{\mathbb{A}}(X(t)) - \frac{\left\|\tilde{\mathbb{A}}(X(t))\right\|^2}{\left\langle\tilde{\mathbb{A}}(X(t)), X(t) - X_0\right\rangle}(X(t) - X_0).$$

将两边除以 2，得到 (11)。

I.2.1 定理 7.1 与定理 7.2 中收敛速率的对应关系

根据上述对应关系，我们看到 β(t) 对应于 $\frac{\beta_k}{2h}$。因此，根据对应关系 $x^0 = X_0$、 $x^\star = X_\star$、 $2hk = t$ 和 $x^k = X$(t)，我们得到

$$\left\|h\tilde{\mathbb{A}}(x^{k+1})\right\|^2 \leq \beta_k^2\left\|x^0 - x^\star\right\|^2 = 4h^2\frac{\beta_k^2}{(2h)^2}\left\|x^0 - x^\star\right\|^2 \quad \xrightarrow{\text{divide by } h^2,\ h\to 0+} \quad \left\|\tilde{\mathbb{A}}(X(t))\right\|^2 \leq 4\beta(t)^2\left\|X_0 - X_\star\right\|^2.$$

我们还可以验证，单调情形下的界 $\beta_k \leq \frac{1}{k+1}$ 等价于 $\frac{\beta_k}{2h} \leq \frac{1}{2hk+\mathcal{O}(h)}$，并且也对应于 β(t) $\leq \frac{1}{t}$。

现在假设 $\mathbb{A}$ 是 μ-强单调且 L-利普希茨连续的。那么 $h\mathbb{A}$ 是 $h\mu$-强单调且 hL-利普希茨连续的，由定理 7.2 可得以下不等式

$$\frac{\beta_k}{2h} \leq \frac{\frac{\mu}{2(1+(hL)^2)}}{\left(1+\frac{h\mu}{1+(hL)^2}\right)^k - 1 + \frac{h\mu}{1+(hL)^2}}. \tag{55}$$

回顾对应关系 $2hk = t$，我们看到

$$\left(1+\frac{h\mu}{1+(hL)^2}\right)^k = \left(1+\frac{h\mu}{1+(hL)^2}\right)^{\frac{t}{2h}} = \left(\left(1+\frac{h\mu}{1+(hL)^2}\right)^{\frac{1+(hL)^2}{h\mu}}\right)^{\frac{\mu t}{2(1+(hL)^2)}} \quad \xrightarrow{h\to 0+} \quad e^{\mu t/2}.$$

因此，将 β(t) 与 $= \frac{\beta_k}{2h}$ 对应，并将极限 $h \to 0+$ 代入不等式 (55)，我们得到

$$\beta(t) \leq \frac{\mu/2}{e^{\mu t/2} - 1}.$$

I.3 定理 7.2 的证明

回顾一下，该方法定义为

$$\begin{aligned}
x^k &= \mathbb{J}_{\mathbb{A}}y^{k-1} \\
\beta_k &= \begin{cases} \dfrac{\|\tilde{\mathbb{A}}x^k\|^2}{-\langle\tilde{\mathbb{A}}x^k,\, x^k - x^0\rangle + \|\tilde{\mathbb{A}}x^k\|^2} & \text{if } \|\tilde{\mathbb{A}}x^k\|^2 \neq 0 \\ 0 & \text{if } \|\tilde{\mathbb{A}}x^k\|^2 = 0 \end{cases} \\
y^k &= (1-\beta_k)(2x^k - y^{k-1}) + \beta_k x^0.
\end{aligned}$$

首先，我们假设对所有 $k \geq 0$ 都有 $\tilde{\mathbf{A}}x^k \neq 0$。定义

$$\tilde{\Phi}^k = (1-\beta_k)\,\|\tilde{\mathbf{A}}x^k\|^2 + \beta_k\langle \tilde{\mathbf{A}}x^k,\, x^k - x^0\rangle$$

对于 $k = 1, 2, \ldots,$ 和

$$\Phi^k = \|\tilde{\mathbf{A}}x^k\|^2 + \beta_{k-1}\langle \tilde{\mathbf{A}}x^k,\, x^k - x^0\rangle$$

对于 $k = 2,\ 3,\ \ldots$。根据 β_k 的定义，我们有 $\tilde{\Phi}^k = 0$。我们的目标是证明

$$\Phi^{k+1} \leq 0.$$

然后，采用与 (8) 相同的论证，我们可以得出 $\|\tilde{\mathbf{A}}x^{k+1}\|^2 \leq \beta_k^2\|x^0 - x^\star\|^2$. 为此，我们首先证明以下引理。**引理 I.2.** 对于 $k \geq 1,$ 以下结论成立。

$$(1-\beta_k)\tilde{\Phi}^k - \Phi^{k+1} = (1-\beta_k)\langle \tilde{\mathbf{A}}x^{k+1} - \tilde{\mathbf{A}}x^k,\, x^{k+1} - x^k\rangle. \tag{56}$$

证明。 利用 $y^{k-1} = x^k + \tilde{\mathbf{A}}x^k$，代入 y^k 和 y^{k-1} 后，该方法等价于

$$x^{k+1} + \tilde{\mathbf{A}}x^{k+1} = (1-\beta_k)(x^k - \tilde{\mathbf{A}}x^k) + \beta_k x^0.$$

由此可得 $(1-\beta_k)(x^{k+1} - x^k)$ 的两种不同表达式。

$$\begin{aligned}(1-\beta_k)(x^{k+1} - x^k) &= -(1-\beta_k)(\tilde{\mathbf{A}}x^{k+1} + \tilde{\mathbf{A}}x^k) - \beta_k\left(\tilde{\mathbf{A}}x^{k+1} + (x^{k+1} - x^0)\right)\\ &= -(1-\beta_k)\left[(\tilde{\mathbf{A}}x^{k+1} + \tilde{\mathbf{A}}x^k) - \beta_k\left(\tilde{\mathbf{A}}x^k - (x^k - x^0)\right)\right].\end{aligned}$$

整理后，第一个等式可通过在等式两边减去 $\beta_k(x^{k+1} - x^0)$ 得到，第二个等式可通过将等式两边乘以 $(1-\beta_k)$ 得到。

由此可得

$$\begin{aligned}&(1-\beta_k)\langle \tilde{\mathbf{A}}x^{k+1} - \tilde{\mathbf{A}}x^k,\, x^{k+1} - x^k\rangle\\ &= \langle \tilde{\mathbf{A}}x^{k+1},\, (1-\beta_k)(x^{k+1} - x^k)\rangle - \langle \tilde{\mathbf{A}}x^k,\, (1-\beta_k)(x^{k+1} - x^k)\rangle\\ &= \left\langle \tilde{\mathbf{A}}x^{k+1},\, -(1-\beta_k)(\tilde{\mathbf{A}}x^{k+1} + \tilde{\mathbf{A}}x^k) - \beta_k\{\tilde{\mathbf{A}}x^{k+1} + (x^{k+1} - x^0)\}\right\rangle\\ &\quad - (1-\beta_k)\left\langle \tilde{\mathbf{A}}x^k,\, -(\tilde{\mathbf{A}}x^{k+1} + \tilde{\mathbf{A}}x^k) + \beta_k\{\tilde{\mathbf{A}}x^k - (x^k - x^0)\}\right\rangle\\ &= -(1-\beta_k)\|\tilde{\mathbf{A}}x^{k+1}\|^2 - \beta_k\langle \tilde{\mathbf{A}}x^{k+1},\, \tilde{\mathbf{A}}x^{k+1} + (x^{k+1} - x^0)\rangle\\ &\quad + (1-\beta_k)\left(\|\tilde{\mathbf{A}}x^k\|^2 - \beta_k\langle \tilde{\mathbf{A}}x^k,\, \tilde{\mathbf{A}}x^k - (x^k - x^0)\rangle\right)\\ &= (1-\beta_k)\left((1-\beta_k)\|\tilde{\mathbf{A}}x^k\|^2 + \beta_k\langle \tilde{\mathbf{A}}x^k,\, x^k - x^0\rangle\right) - \|\tilde{\mathbf{A}}x^{k+1}\|^2 - \beta_k\langle \tilde{\mathbf{A}}x^{k+1},\, x^{k+1} - x^0\rangle\\ &= (1-\beta_k)\tilde{\Phi}^k - \Phi^{k+1}.\end{aligned}$$

□

由于 $\mathbf{A}$ 是单调的，我们有 $\langle \tilde{\mathbf{A}}x^{k+1} - \tilde{\mathbf{A}}x^k,\ \ x^{k+1} - x^k\rangle \geq 0$，因此可知，当 $1 - \beta_k \geq 0$时，(56) 的右侧大于或等于 0。因此只需证明 $1 - \beta_k \geq 0$。由于指标相当令人困惑，为避免在后续证明过程中产生混淆，我们将其作为一个引理给出。

引理 I.3。 若对所有 $k \geq 1,$ 都有 $\tilde{\mathbf{A}}x^k \neq 0,$ 则

$$\left\langle \tilde{\mathbf{A}}x^k, x^k - x^0\right\rangle < 0, \quad \beta_k \in (0,1), \quad \Phi^{k+1} \leq 0$$

对所有 $k \geq 1$均成
立。

证明。 归纳证明。由 $x^1 = \mathbb{J}_{\mathbb{A}}y^0 = \mathbb{J}_{\mathbb{A}}x^0$，我们有 $x^1 + \tilde{\mathbf{A}}x^1 = x^0$，因此 $x^1 - x^0 = -\tilde{\mathbf{A}}x^1$。应用这些事实，我们得到

$$\begin{aligned}\langle \tilde{\mathbf{A}}x^1,\, x^1 - x^0\rangle &= -\|\tilde{\mathbf{A}}x^1\|^2 < 0,\\ \beta_1 &= \frac{\|\tilde{\mathbf{A}}x^1\|^2}{-\langle \tilde{\mathbf{A}}x^1,\, x^1 - x^0\rangle + \|\tilde{\mathbf{A}}x^1\|^2} = \frac{1}{2} \in (0,1).\end{aligned}$$

由于 $1-\beta_1>0$，由（56）可得

$$(1-\beta_1)\tilde{\Phi}^1-\Phi^2=(1-\beta_1)\langle\tilde{\mathbb{A}}x^2-\tilde{\mathbb{A}}x^1,\,x^2-x^1\rangle\geq 0.$$

根据定义， $\tilde{\Phi}^1=0$，我们有 $\Phi^2\leq 0$。因此，该命题对 $k=1$ 成立。

现在假设这些命题对 k成立。根据归纳假设，我们知道

$$0\geq\Phi^{k+1}=\|\tilde{\mathbb{A}}x^{k+1}\|^2+\beta_k\langle\tilde{\mathbb{A}}x^{k+1},\,x^{k+1}-x^0\rangle.$$

As β根据归纳假设有 $_k>0$，并根据假设有 $\tilde{\mathbb{A}}x^{k+1}\neq 0$，重新整理 $\Phi^k\leq 0$，我们有 e

$$\langle\tilde{\mathbb{A}}x^{k+1},\,x^{k+1}-x^0\rangle\leq-\frac{\|\tilde{\mathbb{A}}x^{k+1}\|^2}{\beta_k}<0.$$

因此

$$\beta_{k+1}=\frac{\|\tilde{\mathbb{A}}x^{k+1}\|^2}{\underbrace{-\langle\tilde{\mathbb{A}}x^{k+1},\,x^{k+1}-x^0\rangle}_{>0}+\|\tilde{\mathbb{A}}x^{k+1}\|^2}\in(0,1).$$

由于 $1-\beta_{k+1}>0$，由 (56) 可得

$$(1-\beta_{k+1})\tilde{\Phi}^{k+1}-\Phi^{k+2}=(1-\beta_{k+1})\langle\tilde{\mathbb{A}}x^{k+2}-\tilde{\mathbb{A}}x^{k+1},\,x^{k+2}-x^{k+1}\rangle\geq 0.$$

根据定义， $\tilde{\Phi}^{k+1}=0$，我们有 $\Phi^{k+2}\leq 0$。因此这些命题对 $k+1$ 成立；由归纳法可得所需结果。

□

假设对于所有 $k\geq 1$， $\|\tilde{\mathbb{A}}x^k\|^2\neq 0$。由引理可知，对于 $k\geq 1$， $\Phi^{k+1}\leq 0$，因此用与 (8) 相同的论证，我们对 $x^\star\in$ 有 ZerA

$$\begin{aligned}0\geq\Phi^{k+1}&=\|\tilde{\mathbb{A}}x^{k+1}\|^2+\beta_k\langle\tilde{\mathbb{A}}x^{k+1},\,x^{k+1}-x^\star\rangle-\beta_k\langle\tilde{\mathbb{A}}x^{k+1},\,x^0-x^\star\rangle\\&\geq\|\tilde{\mathbb{A}}x^{k+1}\|^2-\beta_k\langle\tilde{\mathbb{A}}x^{k+1},\,x^0-x^\star\rangle\\&\geq|\tilde{\mathbb{A}}x^{k+1}\|^2-\left(\frac{1}{2}\|\tilde{\mathbb{A}}x^{k+1}\|^2+\frac{\beta_k^2}{2}\|x^0-x^\star\|^2\right)=\frac{1}{2}\|\tilde{\mathbb{A}}x^{k+1}\|^2-\frac{\beta_k^2}{2}\|x^0-x^\star\|^2.\end{aligned}$$

整理可得

$$\|\tilde{\mathbb{A}}x^{k+1}\|^2\leq\beta_k^2\|x^0-x^\star\|^2.$$

˜现在证明 β_k 的上界。由 (56) 及事实 $\Phi=0$可知,

$$\begin{aligned}(1-\beta_k)\left\langle\tilde{\mathbb{A}}x^{k+1}-\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle&=(1-\beta_k)\tilde{\Phi}^k-\Phi^{k+1}\\&=\frac{\beta_k}{\beta_{k+1}}\tilde{\Phi}^{k+1}-\Phi^{k+1}=\left(\frac{\beta_k}{\beta_{k+1}}-\beta_k-1\right)\left\|\tilde{\mathbb{A}}x^{k+1}\right\|^2.\end{aligned}$$

由于根据引理 I.3， $1-\beta_k\neq 0$ 对 $k\geq 1$ 成立，两边除以 $1-\beta_k$，可得 (52) 的离散对应形式,

$$\left\langle\tilde{\mathbb{A}}x^{k+1}-\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle=\frac{\frac{\beta_k}{\beta_{k+1}}-\beta_k-1}{1-\beta_k}\left\|\tilde{\mathbb{A}}x^{k+1}\right\|^2.\tag{57}$$

(i) 当 $\mathbb{A}$ 单调时。

由 (57) 及 $\mathbb{A}$ 的单调性可得

$$0\leq\left\langle\tilde{\mathbb{A}}x^{k+1}-\tilde{\mathbb{A}}x^k,x^{k+1}-x^k\right\rangle=\frac{\frac{\beta_k}{\beta_{k+1}}-\beta_k-1}{1-\beta_k}\left\|\tilde{\mathbb{A}}x^{k+1}\right\|^2.$$

由引理 I.3 可得 $1-\beta_k>0$，因此

$$\frac{\beta_k}{\beta_{k+1}}-\beta_k-1\geq 0\quad\Longrightarrow\quad\frac{1}{\beta_k}+1\leq\frac{1}{\beta_{k+1}}.$$

综上，作为 $\beta_1=\frac{1}{2}$和 $\frac{1}{\beta_k}\succ 0$，我们有

$$\frac{1}{\beta_1}+(k-1)=k+1\leq\frac{1}{\beta_k}\quad\Longrightarrow\quad\beta_k\leq\frac{1}{k+1}.$$

(ii) 当 $\mathbb{A}$ 是 μ-强单调时。

由于 $\mathbb{A}$ 是 μ-强单调的，由 (57) 可得

$$\mu\left\|x^{k+1}-x^{k}\right\|^{2} \leq\left\langle\tilde{\mathbb{A}} x^{k+1}-\tilde{\mathbb{A}} x^{k}, x^{k+1}-x^{k}\right\rangle=\frac{\frac{\beta_{k}}{\beta_{k+1}}-\beta_{k}-1}{1-\beta_{k}}\left\|\tilde{\mathbb{A}} x^{k+1}\right\|^{2}.$$

定义 $r_k = \frac{\left\|x^{k+1}-x^{k}\right\|^{2}}{\left\|\tilde{\mathbb{A}} x^{k+1}\right\|^{2}}$，其中 $k=0$，1，...。将两边除以 $\left\|\tilde{\mathbb{A}} x^{k+1}\right\|^{2}$ 并整理，得到

$$r_{k} \mu \leq \frac{\frac{\beta_{k}}{\beta_{k+1}}-\beta_{k}-1}{1-\beta_{k}} \quad \Longrightarrow \quad r_{k} \mu(1-\beta_{k}) \leq \frac{\beta_{k}}{\beta_{k+1}}-\beta_{k}-1=\beta_{k}\left(\frac{1}{\beta_{k+1}}-1\right)-\beta_{k}-(1-\beta_{k}).$$

将两边除以 β_k 并重新整理，可得关于 $\frac{1}{\beta_k}-1$ 的递归不等式

$$\left(1+r_{k} \mu\right)\left(\frac{1}{\beta_{k}}-1\right)+1 \leq \frac{1}{\beta_{k+1}}-1. \tag{58}$$

现在我们证明由上述不等式可得 β_k 的一个上界。

引理 I.4。设 $\mathbb{A}$ 是一个 μ-强单调算子。令 β_k 为按定理 *7.2* 定义的序列，并令$r_k = \frac{\left\|x^{k+1}-x^{k}\right\|^{2}}{\left\|\tilde{\mathbb{A}} x^{k+1}\right\|^{2}}$，其中 $k=\overline{0,\ 1,\ \ldots}$。则对于 $k=1,\ 2,\ \ldots$，以下结论成立。

$$\beta_{k} \leq \frac{1}{\sum_{j=1}^{k-1} \prod_{i=j}^{k-1}\left(1+r_{i} \mu\right)+2}.$$

证明。 首先注意到，该陈述等价于

$$\frac{1}{\beta_{k}}-1 \geq \sum_{j=1}^{k-1} \prod_{i=j}^{k-1}\left(1+r_{i} \mu\right)+1.$$

利用（58）可以通过归纳法完成证明。

当 $k=1$ 时，回顾引理 I.3 的证明，我们可以验证该不等式成立。

现在假设对于 $k=m$，该不等式成立。那么由（58）可得

$$\begin{aligned}
\frac{1}{\beta_{m+1}}-1 & \geq\left(1+r_{m} \mu\right)\left(\frac{1}{\beta_{m}}-1\right)+1 \\
& \geq\left(1+r_{m} \mu\right)\left(\sum_{j=1}^{m-1} \prod_{i=j}^{m-1}\left(1+r_{i} \mu\right)+1\right)+1 \\
& =\left(\sum_{j=1}^{m-1} \prod_{i=j}^{m}\left(1+r_{i} \mu\right)+\left(1+r_{m} \mu\right)\right)+1=\sum_{j=1}^{m} \prod_{i=j}^{m}\left(1+r_{i} \mu\right)+1.
\end{aligned}$$

因此，我们得到了所需的结果。 □

在附录 I.2 中所考虑的对应关系下， r_k 的连续对应形式为

$$r_{k}=\frac{\left\|x^{k+1}-x^{k}\right\|^{2}}{\left\|h \tilde{\mathbb{A}} x^{k+1}\right\|^{2}}=4 \frac{\left\|x^{k+1}-x^{k}\right\|^{2}}{(2 h)^{2}} \frac{1}{\left\|\tilde{\mathbb{A}} x^{k+1}\right\|^{2}} \quad \xrightarrow{h \rightarrow 0+} \quad \frac{4\left\|\dot{X}(t)\right\|^{2}}{\left\|\tilde{\mathbb{A}}(X(t))\right\|^{2}},$$

根据式 (54)，其大于或等于 1。由此，我们在连续情形下得到了指数收敛速率。同理，如果 r_k 存在正下界，我们也可以在离散情形下得到指数收敛速率；我们将其作为引理 I.4 的推论给出。

推论 I.5。 考虑引理 *I.4* 中的情形。假设存在 $r \geq 0$，使得对于 $k=0,\ r_k \geq r$，1，.....。那么，对于 $k=1$，2，.....，以下结论成立。

$$\beta_k \leq \frac{r\mu}{(1+r\mu)^k - 1 + r\mu}.$$

证明。 由 $r_k \geq r$ 可得

$$\begin{aligned}\sum_{j=1}^{k-1}\prod_{i=j}^{k-1}(1+r_i\mu)+2 &\geq \sum_{j=1}^{k-1}\prod_{i=j}^{k-1}(1+r\mu)+2\\ &= \sum_{j=1}^{k-1}(1+r\mu)^{k-j}+2 = \sum_{l=0}^{k-1}(1+r\mu)^l+1 = \frac{(1+r\mu)^k-1+r\mu}{r\mu}.\end{aligned}$$

应用引理 I.4，即得所需结果。 □

现在证明，当 $\mathbb{A}$进一步为 L-利普希茨连续时， $r_k \geq \frac{1}{1+L^2}$ 成立。

(iii) 当 $\mathbb{A}$ 是 μ-强单调且 L-Lipschitz 连续时 us.

回顾引理 I.2 的证明，我们知道

$$x^{k+1}-x^k = -\mathbb{A}x^{k+1} - (1-\beta_k)\,\mathbb{A}x^k - \beta_k(x^k-x^0).$$

在等式两边与 $\tilde{\mathbb{A}}x^k$ 取内积，得到

$$\left\langle \tilde{\mathbb{A}}x^k, x^{k+1}-x^k\right\rangle = -\left\langle \mathbb{A}x^{k+1}, \tilde{\mathbb{A}}x^k\right\rangle - \tilde{\Phi}^k = -\left\langle \mathbb{A}x^{k+1}, \tilde{\mathbb{A}}x^k\right\rangle.$$

由上面的等式可知

$$\left\|x^{k+1}-x^k\right\|^2 + \left\|\mathbb{A}x^{k+1}-\tilde{\mathbb{A}}x^k\right\|^2 = \left\|\mathbb{A}x^{k+1}\right\|^2 + \left\|x^{k+1}-x^k+\tilde{\mathbb{A}}x^k\right\|^2 \geq \left\|\mathbb{A}x^{k+1}\right\|^2.$$

由于 $\mathbb{A}$ 是 L-Lipschitz 连续的，我们有

$$\left(1+L^2\right)\left\|x^{k+1}-x^k\right\|^2 \geq \left\|x^{k+1}-x^k\right\|^2 + \left\|\mathbb{A}x^{k+1}-\tilde{\mathbb{A}}x^k\right\|^2 \geq \left\|\mathbb{A}x^{k+1}\right\|^2.$$

两边同时除以 $\left(1+L^2\right)\left\|\mathbb{A}x^{k+1}\right\|^2$，得到 r_k 的一个下界

$$r_k = \frac{\left\|x^{k+1}-x^k\right\|^2}{\left\|\mathbb{A}x^{k+1}\right\|^2} \geq \frac{1}{1+L^2}.$$

应用推论 I.5 可得所需结果

$$\beta_k \leq \frac{\mu/(1+L^2)}{\left(1+\mu/(1+L^2)\right)^k - 1 + \mu/(1+L^2)}.$$

注意，若取极限 $\mu \to 0^+$，并代入 $\alpha = \frac{\mu}{1+L^2}$，则有—

$$\lim_{\mu\to 0^+}\frac{\mu/(1+L^2)}{\left(1+\mu/(1+L^2)\right)^k - 1 + \mu/(1+L^2)} = \frac{1}{\lim_{\alpha\to 0^+}\frac{1}{\alpha}\left((1+\alpha)^k-1\right)+1} = \frac{1}{k+1},$$

这就是单调情形的界。

I.3.1 若存在 $k \geq 0$ 使得 $\tilde{\mathbb{A}}x^k = 0$

假设对于某个 k，$\tilde{\mathbb{A}}x^k = 0$。令 x^N 为满足 $\tilde{\mathbb{A}}x^N = 0$的第一个迭代点。根据前面的论证，我们知道定理 7.2 对 $k < N$成立。因此，接下来只需证明这些陈述对 $k \geq N$成立。

由 $\tilde{\mathbb{A}}x^N = 0$可知 $y^{N-1} = x^N + \tilde{\mathbb{A}}x^N = x^N$。又由于 $\tilde{\mathbb{A}}x^N = 0$蕴含 $\beta_N = 0$，根据该方法的定义，我们有 $y^N = 2x^N - y^{N-1} = x^N$。因此

$$x^{N+1} = \mathbb{J}_{\mathbb{A}}y^N = \mathbb{J}_{\mathbb{A}}x^N = x^N,$$

我们得出 $x^k = x^N \in \mathrm{Zer}\mathbb{A}$对所有 $k \geq N$都成立。因此，对于所有 $k \geq N$，定理 7.2 对 $k \geq N$显然成立。

J 第7.1节实验详情

我们求解 Shi 等人提出的压缩感知问题 [63]，该问题表述为一个 ℓ_1正则化最小二乘问题

$$\underset{x\in\mathbb{R}^d}{\text{minimize}} \quad \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{1}{2}\|A_{(i)}x-b_i\|^2+\rho\|x\|_1\right\}.$$

由于问题设置中局部智能体网络如图 2 所示，且每个智能体仅与其邻居（即通过边与该智能体相连的节点）通信，因此我们以去中心化方式求解该问题。我们将 Metropolis-Hastings 矩阵用作混合矩阵 $W\in\mathbf{R}^{n\times n}$，并应用 PG-EXTRA。令 $W_{i,\ j}$表示 W 的第（$i,\ j$）个元素，$N_i\subseteq\{1,2,\ldots,n\}$表示智能体 i邻域中智能体的索引。考虑

$$\mathbf{x}_i^k=\mathrm{Prox}_{\alpha\rho\|\cdot\|_1}\left(\sum_{j\in N_i}W_{i,j}\mathbf{x}_j^{k-1}-\alpha A_{(i)}^{\intercal}\left(A_{(i)}\mathbf{x}_i^{k-1}-b_{(i)}\right)-\mathbf{w}_i^{k-1}\right)$$

$$\mathbf{w}_i^k=\mathbf{w}_i^{k-1}+\frac{1}{2}\left(x_i^{k-1}-\sum_{j\in N_i}W_{i,j}\mathbf{x}_j^{k-1}\right),\quad k=1,2,\ldots \tag{PG-EXTRA}$$

对于上述问题。在参数选择恰当时，PG-EXTRA 可视为关于度量矩阵 M 的平均算子的不动点迭代，如 $\|\cdot\|_M$ [74, 定理 2], 所述，其中该度量矩阵定义为

$$M=\begin{bmatrix}(1/\alpha)I & U^{\intercal}\\ U & \alpha I\end{bmatrix},$$

U是一个对称正定矩阵，满足 $U^2=\frac{1}{2}(I-W)$。也就是说，将$\mathbf{x}^k,\ \mathbf{w}^k\in\mathbf{R}^{d\times n}$分别视为$\mathbf{x}_i^k$和$\mathbf{w}_i^k$的纵向堆叠向量 [58, 第 11 章。PG-EXTRA 可以重写为（$\mathbf{x}^k,\ \mathbf{w}^k$）$=\mathbb{T}$（$\mathbf{x}^{k-1},\ \mathbf{w}^{k-1}$）。基于此 $\mathbb{T}$，我们采用 Halpern 方法进行实验

$$(\mathbf{x}^k,\mathbf{w}^k)=\beta_k(\mathbf{x}^0,\mathbf{w}^0)+(1-\beta_k)\,\mathbb{T}(\mathbf{x}^{k-1},\mathbf{w}^{k-1}).$$

当 $\mathbb{T}$是平均算子时，$\frac{1}{2}(\mathbb{T}+\mathbb{I})$ 它是坚定非扩张的。我们知道，对于某个极大单调算子 $\mathbb{A}$ [15, ，有 $\mathbb{T}=2\mathbf{J}_{\mathbb{A}}-\mathbb{I}=\mathbf{R}_{\mathbb{A}}$。命题2 3.8]。考虑引理 D.2 中讨论的等价关系，可以看出，上述 Halpern 方法等价于我们提出的以下形式的算法

$$\begin{aligned}x^k&=\mathbb{J}_{\mathbb{A}}y^{k-1}\\ y^k&=(1-\beta_k)\,(2x^k-y^{k-1})+\beta_k x^0,\end{aligned}$$

由相应的 $y^k=(\mathbf{x}^k,\mathbf{w}^k)$ 给出。注意，算子范数 $\left\|\tilde{\mathbf{A}}x^k\right\|_M^2$可通过考虑下式计算

$$\frac{1}{2}\left(\mathbb{T}y^{k-1}-y^{k-1}\right)=\mathbb{J}_{\mathbb{A}}y^{k-1}-y^{k-1}=x^k-\left(\tilde{\mathbf{A}}x^k-x^k\right)=-\tilde{\mathbf{A}}x^k.$$

我们采用锚定系数 $\beta_k=\frac{1}{k+1}$，$\beta_k=\frac{\gamma}{k^p+\gamma}$，其中 $p=1.5,\gamma=2.0$，并在定理 7.2 中使用 M-范数下的自适应选择 β_k。注意，在实验中，自适应系数通过考虑下式计算

$$\begin{aligned}\frac{1}{2}\frac{\left\|\mathbb{T}y^{k-1}-y^{k-1}\right\|_M^2}{\|\mathbb{T}y^{k-1}-y^{k-1}\|_M^2+\langle\mathbb{T}y^{k-1}-y^{k-1},y^{k-1}-x^0\rangle_M}&=\frac{1}{2}\frac{4\left\|\tilde{\mathbf{A}}x^k\right\|_M^2}{4\left\|\tilde{\mathbf{A}}x^k\right\|_M^2+\left\langle-2\tilde{\mathbf{A}}x^k,\tilde{\mathbf{A}}x^k+x^k-x^0\right\rangle_M}\\ &=\frac{\|\tilde{\mathbf{A}}x^k\|_M^2}{-\left\langle\tilde{\mathbf{A}}x^k,x^k-x^0\right\rangle_M+\|\tilde{\mathbf{A}}x^k\|_M^2}.\end{aligned}$$

我们选择信号维数 $d=100$、智能体数量 $n=20$、每个智能体的测量数量 $m_i=4$、ℓ_1-正则化参数 $\rho=0.01$以及算法参数 $\alpha=0.01$。

K 更广泛的影响

我们的工作聚焦于凸优化算法的理论方面。我们预计，我们的理论结果不会产生负面的社会影响。

L 局限性

我们的分析涉及凸优化。尽管这一假设在优化理论中很常见，但机器学习实践中出现的许多函数并非凸函数。