JOTA 稿件编号（由编辑填写）

通过热演化凸化的 Yuille-Poggio 流与多项式全局极小值点

Qiao Wang

arXiv:2301.00326v2 [math.OC] 7 May 2024

收稿日期：日期 / 接受日期：日期

摘要 本研究考察了 O. Arikan 等人在 [8] 中提出的用于多项式全局最小化的反向微分流算法的凸化版本。我们探讨了这种方法为何可能在高次多项式上失效，却能在四次多项式上奏效。我们采用热演化方法进行凸化，并结合高斯滤波；高斯滤波相当于累积形式的 Steklov 正则化。在此背景下，我们运用计算机视觉中的指纹理论。指纹理论最初由 A.L. Yuille 和 T. Poggio 于20世纪80年代为计算机视觉而提出，本文利用该理论，尤其是指纹轨迹方程，来说明极小化点的尺度（时间）演化。针对一般多项式，我们的研究提出了 Yuille-Poggio 流，并对指纹概念作出更广泛的阐释；特别是，我们确立了凸化后的反向微分流算法成功实现全局最小化的充要条件。对于四次多项式，我们的分析不仅反映了 O. Arikan 等人的结果 [8]，还提出了一种显著简化的牛顿法版本，该方法无需凸化即可始终实现四次多项式的全局最小化。

关键词 凸优化 · Yuille-Poggio 流 · 指纹理论 · 多项式优化 · 热方程 · 多尺度高斯滤波 · 计算机视觉 · 欧几里得算法

数学主题分类 (2000) 35Q90 · 46N10 · 35Q93 · 90C26

Qiao Wang 东南大学信息科学与工程学院，中国南京，210096

东南大学经济管理学院，中国南京，210096
qiaowang@seu.edu.cn

1 背景与动机

实多项式的全局优化是一个重要的非凸优化问题（参见 [1] 及其中的参考文献），并且在过去几十年中产生了许多优秀的理论。其中，N. Z. Shor [2] 于1987年首次通过二次优化将单变量多项式优化转化为凸问题，从而为这一全局优化问题提供近似解。此后，N.Z. Shor 进一步研究了它与希尔伯特第十七问题[3]之间的关系。同样在1987年，V. N. Nefnov [4] 提出了一种通过计算代数方程的根来寻找极小化点的算法。

2014 年，J. Zhu、S. Zhao 和 G. Liu [6] 提出了一种反向微分流的表述，该表述源自约束优化的 Kuhn-Tucker 方程，用于求出多项式的全局极小值点。他们考虑了充分光滑函数 $p(x)$ 的问题，

$$\begin{aligned} &\min p(x) \\ &s.t.\ \ x \in \mathcal{D} := \{x \in \mathbb{R}^n |\ \|x\| < a\} \end{aligned} \tag{1}$$

通过引入一个集合

$$G = \{\rho > 0 |\ [\nabla^2 p(x) + \rho I] > 0,\ \forall x \in \mathcal{D}\}, \tag{2}$$

以及一个初始对 $(\widehat{\rho},\ \widehat{x}) \in G \times \mathcal{D}$，使其满足

$$\nabla p(\widehat{x}) + \widehat{\rho}\widehat{x} = 0. \tag{3}$$

随后他们证明，在 $\widehat{\rho}$ 附近定义的反向微分流 $\widehat{x}(\rho)$，

$$\begin{aligned} &\frac{d\widehat{x}}{d\rho} + [\nabla^2 p(\widehat{x}) + \rho I]^{-1}\widehat{x} = 0, \\ &\widehat{x}(\widehat{\rho}) = \widehat{x} \end{aligned} \tag{4}$$

将得到式 (1) 的解。

上述工作是在该多项式的所有全局极小化点仅位于一个已知球内这一条件下进行的，因此无约束优化问题可以转化为约束优化问题。然而，O. Arikan、R.S. Burachik 和 C.Y. Kaya [7] 于 2015 年通过一个四次多项式的反例指出，[6] 中的方法可能无法收敛到全局极小值点

$$p(x) = x^4 - 8x^3 - 18x^2 + 56x. \tag{5}$$

此外，他们于 2019 年 [8] 提出了 Steklov 正则化和轨迹方法，用于优化单变量多项式。随后，2020 年，R.S. Bu- rachik 和 C.Y. Kaya [9] 将此方法扩展到多变量情形。在这些研究中，四次多项式优化是颇有趣的玩具示例。此外，正如 [8]中的各种示例和反例所展示的，六次及更高次多项式往往无法达到全局极小化点。

实际上，Steklov 正则化 [8]

$$\mu(x,t)=\frac{1}{2t}\int_{x-t}^{x+t}f(\tau)\,d\tau \tag{6}$$

从信号处理的角度来看，这是一个低通滤波器，因为我们可以写成

$$\mu(x,t)=\frac{1}{2t}\int_{x-t}^{x+t}f(\tau)\,d\tau=f(x)*1_{[-t,t]}(x) \tag{7}$$

其中

$$1_{[-t,t]}(x)=\begin{cases}\frac{1}{2t}, & x\in[-t,t]\\ 0, & x\notin[-t,t]\end{cases} \tag{8}$$

由此可以得到 $\mu_x(x,\ t)$ $\mu_{xt}(x,\ t)$ 和 $\mu_{xx}(x,\ t)$（其中下标表示偏导数）。然而，$\mu_t(x,\ t)$ 并未明确处于这一范围内，尽管我们可以表示一个微分方程

$$2\mu_t+t\mu_{tt}=t\mu_x. \tag{9}$$

显然，这给分析局部极小值点的演化带来了一些不便。因此，我们需要一种方法，既能在简单的微分方程与滤波器之间取得平衡，又能为多项式提供凸化。幸运的是，热传导方程

$$\frac{\partial p}{\partial t}=\frac{1}{2}\frac{\partial^2 p}{\partial x^2} \tag{10}$$

初始条件为

$$p(x,0)=p(x) \tag{11}$$

是实现凸化及类似优化算法的一个良好框架。此外，对所有临界点演化过程的分析也变得更易于从解析角度处理。

备注 *1.1* 需要指出的是，热方程 (10) 的初值问题等价于高斯滤波器，这将在第 2.1 小节中说明。但另一方面，Steklov 正则化的累积会导向高斯分布，因为

$$\underbrace{1_{[-t,t]}*1_{[-t,t]}*\cdots*1_{[-t,t]}}_{n}\to\frac{1}{t\sqrt{2\pi n}}\exp\left(-\frac{x^2}{2nt^2}\right) \tag{12}$$

当 n 足够大时。因此，在本文中用热演化（即高斯滤波）取代 Steklov 正则性是非常自然的。

本文旨在探讨优化偶数次首一多项式的方法

$$\min_x p(x) = x^n + \sum_{j=1}^{n} c_j x^{n-j}. \tag{13}$$

与基于 Kuhn-Tucker 方程的反向微分流不同， [6],本文介绍了一种涉及热演化的方法。通过对多项式进行热传导（高斯滤波），我们开发出一种类似于反向微分流的算法，并明确表述了演化微分方程，以帮助寻找极小化点。该算法能够有效地找到四次多项式的全局极小值点，但对于高次多项式只能取得部分成功。在 [8],通过 Steklov 正则化示例展示的研究中，已经观察到了此类模式。然而，我们将在定理 3.3 中给出一个理论判据，明确这些算法达到全局极小值点的充分必要条件。

我们的分析基于 Yuille-Poggio 指纹理论及其在计算机视觉理论中的轨迹微分方程 [12][13] ，这些理论建立于20世纪80年代。在结束本引言之前，我们将简要概述本研究贡献的动机。

多尺度高斯滤波的优雅框架等价于热传导方程模型。应用这一理论，任意偶数次首一多项式 $p(x)$ 都会在 t 足够大时[1], 通过 $p(x,\ t) = g_t(x) * p(x)$ 变为凸，其中 g_t 表示方差为 t 的高斯滤波器。此外，对于四次多项式 $p(x)$，全局极小值点 x_{min} 会沿着 $t > 0$ 连续演化，因此在每个尺度 $t \geq 0$ 上，它始终是 $p(x,\ t)$ 的全局极小值点 x_{min}^t 。因此，可以保证从 $p(x,\ t)$ 的任意全局极小值点 x_{min}^t 反向且连续地演化到 $p(x)$ 的 x_{min} 。

一个自然的问题是，次数更高的多项式的全局极小值点 x_{min} 是否也会像四次多项式情形那样，连续演化为其缩放版本 $p(x,\ t)$ 的全局极小值点? 遗憾的是，对于次数大于 4 的多项式，这一极受期待的性质通常并不成立。我们将通过一个六次多项式的反例来说明这一点。此外，我们还给出一个收敛到全局极小值点的充要条件。

多尺度高斯滤波及其等价的热传导方程，是偏微分方程、信号处理等领域的基础内容。特别是在计算机视觉领域，自20世纪50年代以来，它为我们提供了许多强有力的理论工具（参见 [10][11]）。其中，20世纪80年代提出的指纹理论（参见 [12][13]）多年来一直发挥着核心作用。

本文借鉴计算机视觉中的指纹思想，定义了缩放多项式 $p(x,\ t)$ 随尺度 t 变化的三种指纹。第一种指纹 $\mathcal{FP}_1$ 刻画了每个 t 下 $p(x,\ t)$ 的所有局部极值点，并且

[1] 我完全相信，这个非常简单的事实早就应该得到确立了。但受限于我的阅读范围，我还没有找到相关参考文献。

第二种指纹 $\mathcal{FP}_2$ 则刻画了每个 t 下 $p(x,\ t)$ 的驻点，这些驻点揭示了 $p(x,\ t)$ 在时间演化过程中的凸性区域。此外， $\mathcal{FP}_3$ 揭示了 $\mathcal{FP}_2$ 中曲线的演化。这些强大的指纹工具让我们深入理解多项式的局部和全局极值点的演化，并据此提出了通过反向轨迹算法得到全局极小值点的一个充分必要条件。

为简单起见，我们将本文中的所有符号与记法列于表 1。

符号与记法	定义	索引
$\mathcal{Z}_{t,k}$	……的实零点集 $\frac{\partial^k p(x,\ t)}{\partial x^k}$	(31)
$\mu(x,\ t)$	$p(x)$的 Steklov 正则性	(7)
$p(x,\ t)$	$p(x)$的热演化	(14)
$\mathcal{FP}_k(p)$	第k个指纹，（$k = 2$ 是 Yuille-Poggio 指纹）	(34)
	Yuille-Poggio 指纹轨迹方程	(38)
流$_{YP}$（p）	Yuille-Poggio 流	(40)
$\Omega(p)$ 和 $\Omega^c(p)$	约束区域和逃逸区	定义 3.3
t_u	指纹的合并时刻 四次多项式	(62)
t^*	演化过程超过临界时间后—— 化后的多项式变为凸函数的临界时间	(68)

表 1：符号与记法列表

总之，本研究的主要贡献是确定了反向微分流算法能够实现多项式全局极小值点的充要条件（定理 3.3）。未来的工作将探讨引入跳跳板偏微分方程来确保找到全局极小值点，前提是该点位于所定义的 可达区域内。

2 多项式的热演化与凸化

2.1 $p(x)$ 的热演化

考虑具有初始条件 (11) 的热传导方程 (10)，方程 (10) 的一般解为

$$p(x,t) = p(x) * g_t(x), \tag{14}$$

其中 $g_t(x)$表示高斯滤波器

$$g_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}},\quad t \geq 0. \tag{15}$$

在信号处理和计算机视觉中，这个时间变量 t 也称为尺度 （高斯滤波的尺度）或人工时间。注意，任何微分算子 $\mathcal{D}$ 都与卷积算子 $*$可交换，即:

$$\mathcal{D}(f * g) = f * \mathcal{D}g = \mathcal{D}f * g. \tag{16}$$

对于多项式 $p(x,\ t)$，热方程 (10) 可进一步写为

$$\frac{\partial^k p}{\partial t^k} = \frac{1}{2^k}\frac{\partial^{2k} p}{\partial x^{2k}}, \quad (k = 1, 2, \cdots) \tag{17}$$

对 (10) 两边关于 t 求导即可得到，因为光滑性有保证。然后对 $p(x,\ t)$ 关于 t 作泰勒展开，可得

$$p(x,t) = p(x,0) + t \cdot \frac{\partial p}{\partial t}\bigg|_{t=0} + \frac{t^2}{2}\frac{\partial^2 p}{\partial t^2}\bigg|_{t=0} + \cdots. \tag{18}$$

若使用推导得到的热方程 (17)，我们可以将 (18) 改写为

$$p(x,t) = p(x,0) + \frac{t}{2} \cdot \frac{\partial^2 p}{\partial x^2}\bigg|_{t=0} + \frac{t^2}{8}\frac{\partial^4 p}{\partial x^4}\bigg|_{t=0} + \cdots. \tag{19}$$

热演化对偶数次多项式的凸化由以下定理[2]刻画。

定理 2.1 对于每个偶数次首一多项式 $p(x)$，都存在一个特定的 $T^* = T^*(p)$，使得热卷积 $p(x,t)$ 关于 x 在任意 $t > T^*$处均为凸。

证明 下文中， k 在 $P_k(x)$ 和 $Q_k(x)$ 中表示多项式的次数。考虑偶数次首一多项式

$$P_{2m}(x) = x^{2m} + P_{2m-1}(x). \tag{20}$$

观察展开式

$$P_{2m}(x) * g_t(x) = x^{2m} * g_t(x) + P_{2m-1}(x) * g_t(x), \tag{21}$$

显然，我们可以写成

$$P_{2m}(x) * g_t(x) = P_{2m}(x) + \beta(x,t), \tag{22}$$

其中

$$\beta(x,t) = (2m)!!\ t^m + \sum_{k=1}^{m-1} t^{m-k} Q_{2k}(x). \tag{23}$$

利用热演化，我们有

$$\frac{1}{2}\frac{\partial^2 p(x,t)}{\partial x^2} = \frac{\partial p(x,t)}{\partial t} = \frac{\partial \beta}{\partial t}. \tag{24}$$

[2] 我再次相信，这一凸性结果在某些文献中一定已有所论述。

在我们的情形下,

$$\frac{\partial \beta}{\partial t} = m(2m)!!\ t^{m-1} + \sum_{k=1}^{m-1}(m-k)t^{m-k-1}Q_{2k}(x). \tag{25}$$

显然， $Q_{2n}(x)$ 的所有主要项都源自 $x^{2m}(x) * g_t(x) - x^{2m}$ ，且必为正。更具体地说,

$$\text{the leading coefficient of } Q_{2n}(x) = \binom{2m}{2n}(2m-2n)!! > 0. \tag{26}$$

这表明存在有限常数 K，使得

$$Q_{2n}(x) > K > -\infty, \quad (n = 2, 3, \ldots, 2m-2). \tag{27}$$

因此,

$$\frac{\partial \beta}{\partial t} > m(2m)!!\ t^{m-1} + K(t^{m-2} + t^{m-3} + \ldots + 1). \tag{28}$$

因此，存在一个值 $T^* \in (0, \ +\infty)$，使得对所有 $t > T^*$, $\frac{\partial \beta}{\partial t} > 0$ 均成立，这表明在 $t > T^*$ 之后，凸性在热演化过程中得以保持。

2.2 比较原理

热演化中最重要的机制是比较原理。由此可知，通常一个局部极小值点会在演化过程中与一个局部极大值点合并，就像极小值点与极大值点这一对之间的“湮灭”作用。

定理 2.2（比较原理）假设 x^* 是 $p\ (x, t^*)$ 的一个临界点，则对于 $t > t^*$, 该临界点的热演化满足

$$p(x^*(t), t) \geq p(x^*, t^*), \text{ if } x^* \text{ is local minimum;} \tag{29}$$

$$p(x^*(t), t) \leq p(x^*, t^*), \text{ if } x^* \text{ is local maximum.} \tag{30}$$

证明 不失一般性，设 $t^* = 0$。由于热算子U^t: $f(x) \mapsto g_t(x) * f(x)$ 构成一个半群（李群）。令 $x = x(t)$ 为 $p(x, t)$ 关于 x 的临界点积分曲线之一，则由

$$\begin{aligned}\frac{dp(x(t), t)}{dt} =& \frac{\partial p(x,t)}{\partial x}\dot{x}(t) + \frac{\partial p(x,t)}{\partial t} \\ =& 0 + \frac{\partial p(x,t)}{\partial t} \\ =& \frac{1}{2}\frac{\partial^2 p(x,t)}{\partial x^2},\end{aligned}$$

因此，我们可以得到所需的结果。注意，最后一个等式来自热传导方程。

这一比较分析表明，局部极小值点和局部极大值点可能会在整个过程中成对出现并合并。理想情况下，一个次数为 n 的多项式（其中 n 为偶数）应有 $n-1$ 个临界点。因此，我们预期全局极小值点在热演化过程中始终与任何局部极大值点保持不同。不过，在某些情况下，这一点可能并不成立，我们将深入考察这些情形。

3 全局极小值点与尺度空间指纹

3.1 尺度空间的指纹

尺度空间指纹由 A.L. Yuille 和 T.A. Poggio 于20世纪80年代提出（参见 [12] [13] 等），在计算机视觉中发挥着重要作用。在这一框架中，多项式 $p(x,t) = p(x) * g_t(x)$ 的多尺度版本 $p(x)$ 来自热传导方程，其中高斯滤波器的方差 $t \geq 0$ 也称为人工时间。

考虑到我们讨论的情形中所有多项式都是实系数的，为简单起见，我们需要将 Yuille-Poggio 对多尺度零交叉指纹的定义推广到更一般的情形，如下所示。

定义 3.1 将多项式 $p(x,t)$ 的 k 阶导数的实零点集合记为

$$\mathcal{Z}_{t,k}(p) := \left\{ x_i(t) \in \mathbb{R};\ \frac{\partial^k p(x_i(t),t)}{\partial x^k} = 0,\ i = 1, 2, \cdots . \right\}, \tag{31}$$

并将这些集合记为

$$\begin{aligned} \mathcal{FP}_k^+(p) &:= \left\{ (x,t);\ \frac{\partial^k p(x,t)}{\partial x^k} > 0,\ t \geq 0, \right\}, \\ \mathcal{FP}_k^-(p) &:= \left\{ (x,t);\ \frac{\partial^k p(x,t)}{\partial x^k} < 0,\ t \geq 0, \right\}. \end{aligned} \tag{32}$$

则多项式 $p(x)$ 的第 k 阶指纹定义为

$$\mathcal{FP}_k(p) := \overline{\mathcal{FP}_k^+(p)} \bigcap \overline{\mathcal{FP}_k^-(p)}. \tag{33}$$

在上述符号与记法中，S 表示集合 S 的拓扑闭包。在我们的情形中，这个拓扑闭包非常简单，因此我们可以用代数方程刻画 $\mathcal{FP}_k$

$$\mathcal{FP}_k(p) = \bigcup_{t:\ t \geq 0} \mathcal{Z}_{t,k}(p) = \left\{ (x,t);\ \frac{\partial^k p(x,t)}{\partial x^k} = 0 \right\}, \tag{34}$$

这归因于所有多项式都具有充分光滑性。

注 *3.1* 当 $k=2$时，所谓的零交叉 $\mathcal{FP}_2$ 指纹以及零交叉轮廓的方程由 *A.L. Yuille* 和 *T.A. Poggio* *[13]* 引入。本文将他们的指纹从 $\mathcal{FP}_2$ 推广到更一般的 $\mathcal{FP}_k$ *(*$k\geq 2$*)*。换言之，若我们考虑 $P(x)$，其导数为 $P'(x)=p(x)$，那么 $\mathcal{FP}_1(p)=\mathcal{FP}_2(P)$。也就是说，我们的 $\mathcal{FP}_k$ 框架本质上是计算机视觉理论中 *Yuille-Poggio* 指纹的一种推广。

根据此记号， $\mathcal{FP}_1$ 是极值（临界点）的指纹，而 $\mathcal{FP}_2$ 是多项式 p*(*x*)*的零交叉（凸性）[3], 。本质上，与计算机视觉理论中的做法一样，我们可以从 $\mathcal{FP}_2^+$和 $\mathcal{FP}_2^-$中获得更多信息。在本文中，我们将经典概念 $\mathcal{FP}_1$ 和 $\mathcal{FP}_2$ 推广到一般的 $\mathcal{FP}_k$，特别是纳入了 $\mathcal{FP}_3$ ，从而可以用这三种指纹来表示我们的主要结果。

我们进一步考察 $\mathcal{FP}_1$中各元素的动力学，即其轨迹。我们主要关注满足下列方程的曲线 $x=x$*(t)*:

$$\frac{\partial p(x(t),t)}{\partial x}=0, \tag{35}$$

以及初始条件

$$x(0)=x_i\in\mathcal{Z}_{0,1}(p),\quad (i=1,2,\cdots) \tag{36}$$

其中， x_i*(*$i=1,2,\cdots$*)*是 p*(x)*的临界点。为了求解这些曲线，A.L. Yuille和T.A. Poggio在 [13],中引入了如下通过改变 t 的常微分方程:

$$0=\frac{d}{dt}\left(\frac{\partial p(x(t),t)}{\partial x}\right)=\frac{\partial^2 p(x,t)}{\partial x^2}\frac{dx(t)}{dt}+\frac{\partial^2 p(x,t)}{\partial x\partial t}. \tag{37}$$

因此，我们可以将包含不同 $t>0$ 处所有极大值的指纹表征为将（37）重写为

$$\frac{dx(t)}{dt}=-\frac{\frac{\partial^2 p(x,t)}{\partial x\partial t}}{\frac{\partial^2 p(x,t)}{\partial x^2}}=-\frac{\frac{\partial^3 p(x,t)}{\partial x^3}}{2\cdot\frac{\partial^2 p(x,t)}{\partial x^2}}, \tag{38}$$

以及适当的初始条件[4]

$$x(0)\in\mathcal{Z}_{0,1}(p). \tag{39}$$

本文将这个常微分方程（38）称为*Yuille-Poggio* 方程，因为它最初由 A.L. Yuille 和 T.A. Poggio 在其开创性著作 [13] 的（3.3）式中提出。

另一方面，我们也将这个方程（38）称为轨迹方程，因为只要给定初值，反向演化算法就会沿这条曲线向后演化。给定任意初始位置，都可以通过该方程得到一条轨迹。特别是，当初始条件位于 p（x）的临界点处时，这些轨迹就构成指纹 $\mathcal{FP}_1$（p）。

我们还可以进一步将 $\mathcal{FP}_1$*(p)* 推广为 Yuille-Poggio 流。

[3] 尽管严格意义与此处的定义之间存在一定差距，本文对此不作赘述。[4] 如果 p（x）是 n 次多项式，则至多存在 $n-1$ 个不同的初始条件。

定义 3.2 对于任意 $h \in \mathbb{R}$，由 Yuille-Poggio 方程 (38) 在初值 $x(0) = h$ 下生成的积分曲线称为 Yuille-Poggio 曲线。所有这些 Yuille-Poggio 曲线构成集合

$$\mathbf{Flow}_{YP}(p) := \left\{ (x(t), t);\ \frac{dx(t)}{dt} = -\frac{\frac{\partial^3 p(x,t)}{\partial x^3}}{2\frac{\partial p^2(x,t)}{\partial x^2}},\ \ x(0) = h,\ \ \forall h \in \mathbb{R} \right\}, \quad (40)$$

并将其称为由多项式 $p(x)$ 生成的 Yuille-Poggio 流。

显然，指纹 $\mathcal{FP}_1(p)$ 中的指纹曲线是一条特殊的 Yuille-Poggio 曲线，其初值 $x(0)$ 被限制为 $\mathcal{Z}_{0,\ 1}(p)$，即满足 $p'(x(0)) = 0$。因此我们有

定理 3.1 指纹 $\mathcal{FP}_1$ 可表示为

$$\mathcal{FP}_1(p) = \left\{ (x, t);\ \frac{dx(t)}{dt} = -\frac{\frac{\partial^3 p(x,t)}{\partial x^3}}{2\frac{\partial p^2(x,t)}{\partial x^2}},\ \ x(0) \in \mathcal{Z}_{0,1} \right\}. \quad (41)$$

And

$$\mathcal{FP}_1(p) \subset \mathbf{Flow}_{YP}(p). \quad (42)$$

注意，奇点出现在分母为零处，因此我们有

引理 3.1 对于任意多项式 $p(x)$ 及其热演化 $p(x,t)$，若

$$(x', t') \in \mathcal{FP}_i \bigcap \mathcal{FP}_{i+1}, \quad (43)$$

那么， x' 必须是多项式方程 $\frac{\partial^i p(x,\ t)}{\partial x^i} = 0$的实重根，同时也是多项式方程 $\frac{\partial^{i+1} p(x,\ t)}{\partial x^{i+1}} = 0$的实根。

引理 3.2 对于 n 阶（n 为偶数）多项式 $p(x)$，集合 $\mathcal{FP}_2 \bigcap \mathcal{FP}_3$至多包含 $\frac{n}{2} - 1$ 个点 (x_i, t_i)，其中 $i = 1$、2、$\cdots$、$\frac{n}{2} - 1$。

3.2 约束区域与逃逸区

在接下来的分析中，我们将给出 $\mathcal{FP}_2 \bigcap \mathcal{FP}_3$ 的基本框架，通过改变轨迹常微分方程的初始条件，将 $\mathbb{R}$ 划分为约束区域 和逃逸区。



and $\mathcal{FP}_3$.

定义 3.3 设 $c \in \mathbb{R}$，如果从 $(c,\ 0)$ 出发的 Yuille-Poggio 曲线不会与任何从 $(c',\ 0) \neq (c,\ 0)$ 出发的 Yuille-Poggio 曲线相交，我们称 c 位于逃逸区。否则，我们称其位于约束区域，记作 Ω。相应地，逃逸区记作 Ω^c。

定理 3.2（约束区域与逃逸区的刻画） 约束区域 Ω 是

$$\Omega := \bigcup_{i=1}^{\frac{n}{2}-1} [X_i^{LL}, X_i^{RR}]. \tag{44}$$

其中

$$X_i^{LL} = \lim_{L(\widehat{x}_i^L,\widehat{t})\to(x_i,t_i)} x_i^{LL}, \tag{45}$$

$$X_i^{RR} = \lim_{L(\widehat{x}_i^L,\widehat{t})\to(x_i,t_i)} x_i^{RR}, \tag{46}$$

其中，极限表示点 $(\widehat{x}_i^L,\widehat{t})$（或$(\widehat{x}_i^R,\widehat{t})$）沿着局部 $\mathcal{FP}_2$ 指纹曲线 $fp_2^i(L)$（或 $fp_2^i(R)$）移动至目标点 (x_i,t_i)。此处，$(\widehat{x}_i^L,\widehat{t})$（或 $(\widehat{x}_i^R,\widehat{t})$）是与 $(x_i^{LL},0)$（或$(x_i^{RR},0)$）相连的 *Yuille-Poggio* 曲线的终点。

证明 我们首先证明

$$\Omega := \bigcup_{i=1}^{\frac{n}{2}-1} \left([X_i^{LL}, X_i^{LR}] \bigcup [X_i^{RL}, X_i^{RR}] \right), \tag{47}$$

其中我们引入两个符号与记法，

$$X_i^{LR} = \lim_{L(\widehat{x}_i^L,\widehat{t})\to(x_i,t_i)} x_i^{LR} = K, \tag{48}$$

$$X_i^{RL} = \lim_{L(\widehat{x}_i^L,\widehat{t})\to(x_i,t_i)} x_i^{RL} = K. \tag{49}$$

这里，K 表示 $\mathcal{FP}_3(p)$中的曲线与直线 $t=0$的交点$(K,\ 0)$。

下面证明式(47)右端定义良好。如图1所示，对于每个$(x_i,\ t_i) \in \mathcal{FP}_2 \bigcap \mathcal{FP}_3$，都存在 $\mathcal{FP}_2$中的一对曲线，分别对应于$(x_i+0,\ t_i-0)$和$(x_i-0,\ t_i-0)$，记为$fp_2^i(R)$和 $fp_2^i(L)$。

对于 $\in fp_2^i(R)$中的任意一点$(\widehat{x}_i^R,\ \widehat{t})$，当$(\widehat{x}_i^R,\ \widehat{t}) \neq (x_i,\ t_i)$时，存在一对满足式(38)且经过点$(\widehat{x}_i^R,\ \widehat{t})$的轨迹。我们将它们在 $t=0$ 处的端点分别记为$(x_i^{RL},\ 0)$和$(x_i^{RR},\ 0)$。这里我们假设$x_i^{RL} \leq x_i^{RR}$。

类似地，对于任意一点 $(\widehat{x}_i^L,\widehat{t}) \in fp_2^i(L)$，当 $(\widehat{x}_i^L,\widehat{t}) \neq (x_i,t_i)$ 时，存在一对满足 (38) 且经过点 $(\widehat{x}_i^L,\widehat{t})$ 的轨迹。我们分别将它们在 $t=0$ 处的端点记为 $(x_i^{LL},0)$ 和 $(x_i^{LR},0)$。这里我们假设$x_i^{LL} \leq x_i^{LR}$。

现在我们可以写成

$$x_i^{LL} \le x_i^{LR} < K < x_i^{RL} \le x_i^{RR} \tag{50}$$

由于根据 Yuille-Poggio 方程 (38) 右端的分母，Yuille-Poggio 曲线不能与 $\mathcal{FP}_3(p)$ 相交，否则会产生奇点。

此外，(45) 等式中的限制过程仍保持单调性。也就是说，从 $(\widehat{x}_i^L,\ \widehat{t})$ 移至 $(\widehat{x}_i^{L+},\ \widehat{t}^{+})$，最后到 $(x_i,\ t_i)$，我们可以观察到

$$-\infty < \widehat{x}_i^{LL+} < \widehat{x}_i^{LL} < \widehat{x}_i^{LR} < \widehat{x}_i^{LR+} < K. \tag{51}$$

这意味着所有限制 (45) 等都是良定义的。最后，我们注意到 $\forall h \in (K-\epsilon,\ K)$，必定存在一条从 $(h,\ 0)$ 出发的 Yuille-Poggio 曲线；对于 $\forall \epsilon > 0$，任意满足 $t' < t,\ x' > x_i$ 且 $\|(x',\ t') - (x_i,\ t_i)\|_2 < \epsilon$ 的点 $(x',\ t') \in fp_2^i(L)$，必定存在一条经过点 $(x',\ t')$ 的 Yuille-Poggio 曲线。也就是说，连接 $(K,\ 0)$ 和 $(x_i,\ t_i)$ 的 $\mathcal{FP}_3$ 曲线附近的 Yuille-Poggio 曲线是稠密的。为简便起见，这里省略拓扑和微分动力学的描述。

现在，（44）中的集合 Ω 定义良好。我们观察到，对于 $h \in \Omega$，任何从 $(h,\ 0)$ 出发的 Yuille-Poggio 曲线存在，当且仅当存在另一条 Yuille-Poggio 曲线，它从某个 $h' \in \Omega$ 对应的 $(h',\ 0)$ 出发。特别地，这两条曲线在 $fp_2^i(L)$ 或 $fp_2^i(R)$ 处相交。因此，当前的（ 47 ）中的 Ω 与定义 3.3 中的相一致。

现在我们陈述本文的主要贡献：

定理 3.3 设 $p(x)$ 为偶数阶且首项系数为正的多项式。假设 x^{t_0}极小值 是凸多项式（经过适当缩放的版本） $p(x,t) = p(x) * g_t(x)$ 在 $t = t_0$ 处的全局极小值点，且根据 *Yuille-Poggio* 轨迹方程 (38)，轨迹在 $t = 0$ 处的终点为 x^*。那么， $p(x)$ 的全局极小值点 x^* 可以从其凸化版本 $p(x,t_0)$的全局极小值点逆向得到，当且仅当 x^* 位于逃逸区 Ω^c 中。

证明 若最优极小化点为 $x^* \in \Omega$，则所有 Yuille-Poggio 曲线对应的 t 的最大值有界，因此这些 Yuille-Poggio 曲线均无法连接到 $\mathbb{R}_x \times \mathbb{R}_t$ 中 $t > 0$ 较大的点。由此可知，全局极小值点 $x^* \in \Omega^c$。

备注 *3.2* 虽然预计 $X_i^{LL},\ X_i^{LR},\ X_i^{RL},\ X_i^{RR}$可以显式定义，但由于对于次数为 6 或更高的多项式， $\mathcal{FP}_1$ 中相应的曲线需要处理至少为 5 次的代数方程，因此无法得到它们的代数表达式。因此，我们建议采用数值技术来确定这些值。

备注 *3.3* 通过热传导方程，特别是高斯滤波器进行凸化的解析方法，同样适用于 Steklov 正则化情形。

4 四次多项式案例研究

下面我们将给出四次多项式指纹的显式表示，并解释其几何性质。尽管部分结果与 [8] 四次多项式的结果相似，但我们发现了一种不同的算法，无需进行任何凸化处理，也无需 Steklov 正则化或热演化，即可求得四次多项式的全局极小值点。

4.1 指纹的结构

对于四次多项式

$$p(x) = x^4 + ax^3 + bx^2 + cx + d, \tag{52}$$

我们可以看到其热演化

$$\begin{aligned} p(x,t) &= p(x) + (6x^2 + 3ax + b) \cdot t + 3t^2 \\ &= x^4 + ax^3 + (b + 6t)x^2 + (c + 3at)x + (d + bt + 3t^2). \end{aligned} \tag{53}$$

继续对（53）的两边关于 x 求导，$\partial p(x,t)/\partial x$ 随时间 t 的信息可表示为——

$$\frac{\partial p(x,t)}{\partial x} = 4x^3 + 3ax^2 + (2b + 12t)x + (c + 3at) = 0. \tag{54}$$

同样，我们有

$$\frac{\partial^2 p(x,t)}{\partial x^2} = \frac{\partial^2 p(x)}{\partial x^2} + 12t = (12x^2 + 6ax + 2b) + 12t = 0, \tag{55}$$

and

$$\frac{\partial^3 p(x,t)}{\partial x^3} = 24x + 6a = 0. \tag{56}$$

这三个方程分别构成指纹 $\mathcal{FP}_1$、 $\mathcal{FP}_2$ 和$\mathcal{FP}_3$的特征描述。

4.1.1 指纹 $\mathcal{FP}_1$的结构

根据（54），指纹 $\mathcal{FP}_1$ 由以下时变三次方程刻画

$$x^3 + \frac{3a}{4}x^2 + \frac{b + 6t}{2}x + \frac{c + 3at}{4} = 0, \tag{57}$$

现在，若 x_i 是 (57) 在 $t = 0$处的一个实根，则它对应于由微分方程 (38) 描述的轨迹。更多细节如下：

引理 4.1 对于四次多项式 $p(x)$， $p(x,t)$ 关于 x 在尺度 t 下的局部极值点 $x_i^t (i = 1,2,3)$ 满足轨迹微分方程

$$\frac{dx(t)}{dt} = -\frac{12x + 3a}{12x^2 + 6ax + 2b + 12t}, \tag{58}$$

具有以下（至多三个）初始条件，

$$x_i(0) = x_i, \quad (i = 1, 2, 3). \tag{59}$$

这里， x_i 是 p（x）的局部极值。

证明 将 (55) 和 (56) 代入 (38)，即可得到所需结果。

方程 (57) 在 $t = 0$ 处至多有三个实根，对应至多三条轨迹，它们构成指纹 $\mathcal{FP}_1$。然而，从微分代数的角度来看（例如参见 [15]），微分方程 (58) 的解实际上是一条实代数曲线，即关于$x(t)$和 t 的多项式 $F(x, t)$，满足 $F(x, t) = 0$。在我们的情形中，多项式方程 (57) 描述了这条代数曲线，因此我们可以立即应用 $\mathcal{FP}_1$的代数表示：

$$\mathcal{FP}_1 = \left\{(x,t);\ t = -\frac{4x^3 + 3ax^2 + 2bx + c}{12x + 3a},\ \ x \neq -\frac{a}{4},\ \text{and } t \geq 0\right\}. \tag{60}$$

根据 A.2 小节，为了获取方程 (57) 的根的信息，我们需要其判别式，

$$\Delta(t) = \left(\frac{a^3 - 4ab + 8c}{64}\right)^2 + \left(\frac{-3a^2 + 8b}{48} + t\right)^3, \tag{61}$$

其详细说明见 (64)。

引理 4.2 方程 (57) 的判别式 $\Delta(t)$ 单调递增至无穷大。其唯一零点为

$$t_u = \frac{a^2}{16} - \frac{b}{6} - \frac{1}{16}(a^3 - 4ab + 8c)^{\frac{2}{3}}. \tag{62}$$

证明 利用 (119)，可写为

$$\begin{aligned} f(t) &= \frac{b}{2} - \frac{3a^2}{16} + 3t, \\ g(t) &= \frac{a^3}{32} - \frac{ab}{8} + \frac{c}{4}. \end{aligned} \tag{63}$$

现在，随时间变化的判别式

$$\begin{aligned} \Delta(t) =& \frac{[g(t)]^2}{4} + \frac{[f(t)]^3}{27}, \\ =& \left(\frac{a^3 - 4ab + 8c}{64}\right)^2 + \left(\frac{-3a^2 + 8b}{48} + t\right)^3, \end{aligned} \tag{64}$$

这意味着 $\Delta(t)$ 随 t 单调递增。由此立即得到 (62)。 (64)

定理 4.1（“$1+2$”结构 $\mathcal{FP}_1$） 设 t_u 是 (62) 中定义的判别式 $\Delta(t)$ 的零点。若 $t_u<0$，则 $\mathcal{FP}_1$ 仅包含一条轨迹 $x(t)$，由方程 (58) 描述，其演化为 $t\to+\infty$。若 $t_u\geq 0$，则在 $t\in[0,t_u]$ 期间，指纹 $\mathcal{FP}_1$ 包含三条由 (58) 描述的不同轨迹，其中一条在 $t>t_u$ 期间继续演化为 $+\infty$，另外两条轨迹从 $t=0$ 出发，但在 $t=t_u$ 时于点 $(x(t_u)\ t_u)$ 合并（停止）。这里，

$$x(t_u)=\left(\frac{a^3-4ab+8c}{64}\right)^{1/3}-\frac{a}{4}. \tag{65}$$

证明 根据引理 4.2，我们知道，如果 $t_u<0$，那么对于所有$t\geq 0$，$\Delta(t)>0$，这意味着方程 (57) 在每个 $t\geq 0$ 处只有一个根。当$t_u\geq 0$时，则 $\Delta(t)<0$ ($=0,\ >0,\ respectively$)，而 $t\in[0,t_u)$（分别为$t=t_u$、$t>t_u$），方程 (57) 有三个互异实根（一个实根和一对实二重根，或分别为一个实根）。特别地，我们考虑临界情形 $t=t_u$，此时 $\Delta(t)=0$。若属于这种情形，方程 (57) 在 $t=t_u$ 处有一个实根和一个实重根。由 (122) 可知，这个实重根为

$$x(t_u)=\left(\frac{g(t)}{2}\right)^{\frac{1}{3}}-\frac{1}{3}\cdot\frac{3a}{4}. \tag{66}$$

将 (63) 代入此公式可得 (65)。

4.1.2 $\mathcal{FP}_2$ 和 $\mathcal{FP}_3$的结构

对于四次多项式，$\mathcal{FP}_3$ 的结构非常简单，因为由此可写为 (56)

$$\mathcal{FP}_3=\left\{(x,t);\ x=-\frac{a}{4},\ t\geq 0\right\}. \tag{67}$$

为了分析 $\mathcal{FP}_2$的结构，我们有如下引理。

引理 4.3 记

$$t^*=\frac{a^2}{16}-\frac{b}{6}, \tag{68}$$

则在每个 t >max(t^*,0)处，式(53) 中定义的多项式 $p(x,t)$ 关于 x 是凸的。此外，这个 t^* 无法进一步改进。

证明 考虑 $t=0$ 处 (55) 的下界，

$$\begin{aligned}\frac{\partial^2 p(x)}{\partial x^2}&=12x^2+6ax+2b\\&=12\left(x+\frac{a}{4}\right)^2-\frac{3a^2}{4}+2b\\&\geq-\frac{3a^2}{4}+2b=-12t^*.\end{aligned} \tag{69}$$

将其与 (55) 结合，我们得到

$$\frac{\partial^2 p(x,t)}{\partial x^2} \geq -\frac{3a^2}{4} + 2b + 12t = 12(t - t^*), \tag{70}$$

这意味着所需的结果。

另一方面，在任意固定的 $t' < t^*$ 下，注意到在 $x = -\frac{a}{4}$ 处，我们有

$$\begin{aligned} \frac{\partial^2 p(x,t')}{\partial x^2} &= 12x^2 + 6ax + 2b + 12t' \\ &= 12\left(x + \frac{a}{4}\right)^2 - 12(t^* - t') \\ &= -12(t^* - t') < 0, \end{aligned} \tag{71}$$

在该 $x = -\frac{a}{4}$ 下并不凸，因此 t^* 是最优的，无法进一步改进。

定理 4.2（$\mathcal{FP}_2$）关于 $\mathcal{FP}_2$ 四次多项式 $p(x)$ 的结构，

1. 当 $t^* < 0$ 时，指纹 $\mathcal{FP}_2$ 为空。*2.* 当 $t^* = 0$ 时，

$$\mathcal{FP}_2 = \left\{(x,t) = (-\frac{a}{4}, 0)\right\}$$

只有一个元素；

3. 当 $t^* > 0$ 时，指纹 $\mathcal{FP}_2$ 由两条曲线组成：左侧的是

$$x_L(t) = -\frac{a}{4} - \sqrt{t^* - t}, \quad (t^* \geq t \geq 0), \tag{72}$$

右侧的是

$$x_R(t) = -\frac{a}{4} + \sqrt{t^* - t}, \quad (t^* \geq t \geq 0). \tag{73}$$

具体而言，这两条曲线必须在 $t = t^*$ 处相交，即在点

$$(x_L(t^*), t^*) = (x_R(t^*), t^*) = \left(-\frac{a}{4}, t^*\right). \tag{74}$$

证明 (1) 源于这样一个事实：对于每个 $t \geq 0$，所有的 $\frac{\partial^2 p}{\partial x^2} > 0$。也就是说，$\mathcal{FP}_2^+ = \{(x,\ t);\ x \in \mathbb{R},\quad t \in [0, +\infty)\}$，但 $\mathcal{FP}_2^- = \emptyset$。(2) 是显然的结果，而 (3) 则由二次方程 (70) 得出。

4.1.3 指纹之间的交集

根据引理 3.1，我们可以总结指纹的交集。

定理4.3 对于首一四次多项式 $p(x) = x^4 + ax^3 + bx^2 + cx$，两个交集

$$\mathcal{FP}_2 \bigcap \mathcal{FP}_3 = \left\{ (-\frac{a}{4}, t^*) \right\}, \tag{75}$$

and

$$\mathcal{FP}_1 \bigcap \mathcal{FP}_2 = \{(x(t_u), t_u)\}, \tag{76}$$

其中， t^* 定义于(68)， t_u 和 $x(t_u)$ 分别定义于(62) 和。 (65)

备注*4.1*（时间演化的三个阶段）一般而言，多项式 $p(x)$的演化可根据 $0 \leq t_u \leq t^*$分为三个阶段。在第一阶段， t 从0演化至 t_u，而 $\mathcal{FP}_1$ 包含三条不同轨迹。其中两条将在 $t = t_u$处合并。

然后在第二阶段， $t_u < t < t^*$， $\mathcal{FP}_1$ 仅包含一条轨迹，但 $p(x, t)$不是凸的。

最后，在第三阶段 $t > t^*$， $\mathcal{FP}_1$ 只包含一条轨迹，且 $p(x,\ t)$是凸的。

4.2 约束区域

现在我们来计算约束区域。

定理4.4 四次多项式 $p(x)$ 的约束区域为区间

$$\Omega = \left[-\frac{a}{4} - \sqrt{3t^*}, -\frac{a}{4} + \sqrt{3t^*} \right], \tag{77}$$

其中 t^* 定义于(68)。

证明 我们将(62)改写为

$$t_u(c) = t^* - \frac{1}{16}(a^3 - 4ab + 8c)^{\frac{2}{3}}. \tag{78}$$

显然，我们应当对 c 进行变分，使得 $t_u(c) = t^*$，即，

$$a^3 - 4ab + 8c = 0 \implies c = \frac{ab}{2} - \frac{a^3}{8}. \tag{79}$$

将此 c 代入轨迹代数曲线方程 (60)，并令$t = 0$，得到方程

$$4x^3 + 3ax^2 + 2bx + \frac{ab}{2} - \frac{a^3}{8} = 0. \tag{80}$$

该方程的三个根为

$$x_1 = -\frac{a}{4},\ x_{2,3} = -\frac{a}{4} \pm \sqrt{3t^*}, \tag{81}$$

这将产生两对从 t^* 出发、反向演化至 $t = 0$ 的轨迹，其四个终点构成约束区域

$$\left[-\frac{a}{4} - \sqrt{3t^*}, -\frac{a}{4}\right] \bigcup \left[-\frac{a}{4}, -\frac{a}{4} + \sqrt{3t^*}\right] = \left[-\frac{a}{4} - \sqrt{3t^*}, -\frac{a}{4} + \sqrt{3t^*}\right] \tag{82}$$

4.3 跨尺度临界点的微分方程

现在我们用指纹方程 $\frac{\partial p\ (x,\ t)}{\partial x} = 0$ 的根来表示 t^* 和 t_u。

引理 4.4 设 x_1、x_2、x_3 是 (13) 的临界点，t^* 在 (68) 中定义，t_u 在 (62) 中定义，则有

$$t^* = \left(\frac{x_1 + x_2 + x_3}{3}\right)^2 - \frac{x_1x_2 + x_2x_3 + x_3x_1}{3}, \tag{83}$$

and

$$t_u = t^* - \left[\frac{32}{27}(2x_1 - x_2 - x_3)(2x_2 - x_3 - x_1)(2x_3 - x_1 - x_2)\right]^{\frac{2}{3}}. \tag{84}$$

证明 x_1, x_2, x_3 满足 $p'(x) = 0$, 即

$$x^3 + \frac{3a}{4}x^2 + \frac{b}{2}x + \frac{c}{4} = 0. \tag{85}$$

thus

$$\begin{aligned} a =& -\frac{4}{3}(x_1 + x_2 + x_3), \\ b =& 2(x_1x_2 + x_2x_3 + x_3x_1), \\ c =& -4x_1x_2x_3. \end{aligned} \tag{86}$$

将其代入 (62) 和 (68) 即可得到所有这些结果。

定理 4.5 方程 (58) 的奇点仅出现在

$$x^{t_u} = x(t_u) = \left(\frac{a^3 - 4ab + 8c}{64}\right)^{1/3} - \frac{a}{4}, \quad t = t_u. \tag{87}$$

证明 奇点出现在微分方程(58)处，该方程描述$\mathcal{FP}_1$，因此必须满足 (57)。同时，(58) 右侧的分母实际上是 $\mathcal{FP}_2$的指纹，它应满足 (70)。因此，我们可以结合这两个代数方程来求解 (x,t)。

将 (70) 两边乘以 $x+\frac{a}{4}$，再从 (57) 中减去所得结果，可得

$$t=\frac{(3a^2-8b)x-(6c-ab)}{48x+12a},\quad x\neq-\frac{a}{4}. \tag{88}$$

将其代入 (70)，得到关于 x 的三次方程，

$$48x^3+36ax^2+9a^2x+(3ab-6c)=0. \tag{89}$$

该三次方程只有一个实数解 (87)。将这个 x 代入 (88)，可得 $t=t_u$。

当 $-\frac{a}{4}$ 不是临界点时，点 $(x^{t_u},\ t_u)$ 仅位于从 (58) 得到的两条 $\mathcal{FP}_1$ 积分曲线上，它们分别始于一个局部极小值和一个局部极大值点。其中一条曲线对应于 $\dot{x}(t_u)=+\infty$，另一条则对应于 $\dot{x}(t_u)=-\infty$。值得注意的是，从全局最小值出发的积分曲线不会经过该点 $(x^{t_u},\ t_u)$；这一点将在第 4.4 节进一步阐述。

注 *4.2* 如果存在一个临界点 x'，使其位于 $t=0$ 且满足 $x'=-\frac{a}{4}$，那么 (58) 表明其指纹曲线为 $x(t)\equiv-\frac{a}{4}$。当 x' 是局部极大值点，且另外两个临界点 x_1 和 x_3 满足 $x_1+x_3=2x'$ 时，就会出现这种情况。在这种情况下，当 $t=t_u$ 时，三条指纹曲线都会在 $x'=-\frac{a}{4}$ 相交；这一点将在后续章节中详细说明。

例 *4.1* 考虑多项式 $p(x)=x^4+0.2x^3-0.5x^2+0.01x$，图示见图 2。

4.4 热 四次多项式临界点的演化 式

现在我们研究临界点的演化，尤其是四次多项式的情形。本质上，我们关注这样的情形：$t_u>0$ 存在三个互不相同的临界点 $x_1<x_2<x_3$，它们相应地演化为临界点 $x_1^t,\ x_2^t,\ x_3^t$，当 $0<t<t_u$ 时。一般来说，x_1 和 x_3 都是局部极小值点，而 x_2 是局部极大值点。我们主要关注由方程（58）刻画的热演化相关行为。

接下来我们关注热演化过程中两个局部极小值之间的比较原理。令人惊讶的是，对于四次多项式的热演化，我们得到了完全符合预期的结果：

定理 4.6 对于首一四次多项式 $p(x)$，假设 $t_u>0$，并将其三个临界点记为 $x_1<x_2<x_3$ （或 $x_1>x_2>x_3$）。若 $p(x_1)<p(x_3)$，则 $p(x_1^t,t)<p(x_2^t,t)$。

在证明这一结果之前，我们需要几个引理。

引理 4.5 设 x_1、x_2、x_3 是四次多项式 $p(x)$ 的临界点，则有

$$p(x_3) - p(x_1) = -(x_3 - x_1)^3 \cdot (x_1 + x_3 - 2x_2)/3. \tag{90}$$

证明 由 (86)，用 x_1, x_2, x_3 表示 a, b, c。因此可得 $p(x_1) - p(x_3)$。将其因式分解即可得到所需结果。

由引理 4.5 可知

引理 4.6 对于任意 $t \in [0, t_u)$、$p(x_3^t, t) = p(x_1^t, t)$，当且仅当 $x_1^t + x_3^t = 2x_2^t$ 时成立。

因此，我们看到

引理 4.7 对于任意 $t \in [0, t_u)$，$x_1^t + x_3^t = 2x_2^t$ 当且仅当 $x_2^t = -\frac{a}{4}$。 –

证明 注意，x^3 在 $p(x, t)$ 中的系数不随 t变化，x^2 在 $\frac{\partial p}{\partial x}$ 中的系数也不随 t 变化。根据附录 A, 一我们有

$$x_1^t + x_2^t + x_3^t = -\frac{3a}{4}. \tag{91}$$

应用引理 4.6 即可得出结论。

引理 4.8 设 $x_1 < x_2 < x_3$ 是首一四次多项式 $p(x)$的三个临界点，若 $p(x_1) = p(x_3)$，则对于所有 $t \in (-\infty, t_u)$，有

$$p(x_1^t, t) = p(x_3^t, t). \tag{92}$$

证明 回顾式 (57)，并应用式 (86)，我们实际上可以将 $\frac{\partial p_{(x,t)}}{\partial x} = 0$ 表示为关于 x_1, x_2, x_3 的形式:

$$\begin{aligned} x^3 - (x_1 + x_2 + x_3)x^2 + (x_1x_2 + x_2x_3 + x_3x_1 + 3t)x \\ -[x_1x_2x_3 + (x_1 + x_2 + x_3)t] = 0. \end{aligned} \tag{93}$$

现在，若 $p(x_1) = p(x_3)$，引理 4.6 告诉我们 $x_3 = 2x_2 - x_1$，因此我们可以将上面的方程简化为

$$x^3 - 3x_2x^2 + (2x_1x_2 + 2x_2^2 - x_1^2 + 3t)x - (2x_1x_2^2 - x_1^2x_2 + 3x_2t) = 0, \tag{94}$$

其解为

$$\begin{aligned} &x_2^t = x_2, \\ &x_1^t = x_2 - \sqrt{(x_1 - x_2)^2 - 3t}, \\ &x_3^t = x_2 + \sqrt{(x_1 - x_2)^2 - 3t}, \qquad -\infty < t < \min\left\{\frac{(x_1 - x_2)^2}{3}, t_u\right\} \end{aligned} \tag{95}$$

这表明 $x_1^t + x_3^t = 2x_2^t$。根据引理 4.6，可得$p(x_1^t, t) = p(x_3^t, t)$。

现阶段，我们将所有引理总结如下，

定理 4.7 在与定理 *4.6* 相同的假设下，并记 $x_1^t < x_2^t < x_3^t$ 为 $p(x,t)$的临界点。则以下陈述等价：

1. $p(x_1) = p(x_3)$，*2.* $p(x_1^t,t) = p(x_3^t,t)$ $\forall t \in [0,t_u)$；*3.* $x_1 + x_3 = 2x_2$；*4.* $x_1^t + x_3^t = 2x_2^t$，$\forall t \in [0,t_u)$；*5.* $x_2 = -\frac{a}{4}$；*6.* $x_2^t = -\frac{a}{4}$，$\forall t \in [0,t_u)$。*7.* $t^* = t_u$。

证明 我们将证明 (1) $\Longrightarrow$(3) $\Longrightarrow$(5) $\Longrightarrow$(6) $\Longrightarrow$(4) $\Longrightarrow$(2) $\Longrightarrow$(1)。此外，(4) $\Longleftrightarrow$ (7)。事实上，这一过程与前面的证明部分重复。

(1) $\Longrightarrow$ (3)（也就是 (4) $\Longrightarrow$ (2)）来自引理 4.6，(3) $\Longrightarrow$ (5)（也就是 (6) $\Longrightarrow$ (4) ）来自引理 4.7，(5) $\Longrightarrow$ (6) 来自微分方程 (58)。最后，(4) $\Longleftrightarrow$ (7) 来自引理 4.4 中的 (84)，以及条件$x_1 < x_2 < x_3$。

证明（定理 *4.6*）动力学方程 (58) 表明，当 $t \in [0,t_u$ 时，三个临界点的演化是连续的。因此，如果 $p(x_1) < p(x_3)$，那么对于 $t \in [0,t_u)$，必有 $p(x_1^t,t) < p(x_3^t,t)$；否则，必定存在

$$p(x_1^{t'},t') = p(x_3^{t'},t')$$

某个 $t' \in (0,t_u)$。然而，若情况如此，定理 4.7（引理 4.8）表明 $p(x_1) = p(x_3)$，因为 t 是可逆的，这与初始假设矛盾。

为直观地解释这一结果，我们建议对每个 t 使用图 12 中的三角形表示，其中三角形的顶点在 $t < t_u$ 时为 $(x_i^t, p(x_i^t,t))$，$(i = 1,2,3)$。请注意，当 $0 \le t \le t_u$ 时，三角形序列以及延续曲线 $\dot{x}(t)$ 实际上在每个 $t \ge 0$ 处都与 $p(x,t)$ 的全局最小值相连。

最后，我们讨论一个有趣的问题：如果 $x_1 < x_2 < x_3$ 是四次多项式 $p(x)$ 的三个临界点，我们能否不计算所有这些 $p(x_i)$，就判断其中哪一个是全局极小值点？答案是可以。

定理 4.8 设 $x_1 < x_2 < x_3$ 是首一四次多项式 $p(x)$ 的三个互异临界点，则以下陈述等价：

1. x_3 （分别地， x_1）是全局极小值点；
2. $x_1 + x_3 > 2x_2$ （分别地， $x_1 + x_3 < 2x_2$）；
3. $x_2 < -a/4$ （分别地， $x_2 > -a/4$）。

证明 应用引理 4.5。

定理 4.8 启发了以下这种无需热卷积的非常简单的四次多项式欧拉算法。

定理 4.9 对于任意首一四次多项式 $p(x)$，其中 a x^3 的系数为 x^3，以固定初始位置 $x^{(0)} = -\frac{a}{4}$ 运行欧拉算法，

$$x^{(k+1)} = x^{(k)} - \Delta x \cdot p'(x^{(k)}), \tag{96}$$

必须收敛到 $p(x)$ 的全局极小值点。

4.5 算法

实际上，欧拉算法可以从 $t > t_u$ 出发。幸运的是，我们知道在 $t \geq t_u$ 处， $p(x, t)$ 关于 x 只有唯一极小值。回顾公式 (116)，我们可以看到，指纹三次方程 (57) 的所有实根之和在 $0 \leq t \geq t_u$ 下应保持不变，因此，由于我们已从定理 4.5 获得了 $(x(t_u)\ t_u)$ 的信息，便可以求出 $t = t_u$ 处剩余的临界点 x_{init}。这意味着我们可以采用

$$\begin{aligned} x_{init} &= -\frac{3a}{4} - 2x(t_u) \\ &= -\frac{a}{4} - 2\left(\frac{a^3 - 4ab + 8c}{64}\right)^{1/3}, \end{aligned} \tag{97}$$

以 $t = t_u$ 作为初始点，然后对方程 (58) 执行欧拉算法，最终得到 $p(x)$ 的全局最小值。这意味着有以下结果。

定理 4.10 对于四次多项式 (13)，若 $t_u \leq 0$，则该多项式只有一个临界点。若 $t_u > 0$，该多项式有三个互不相同的临界点，则如果所有临界点都满足 $x \neq -\frac{a}{4}$，我们反向求解微分方程

$$\frac{dx(t)}{dt} = -\frac{12x + 3a}{12x^2 + 6ax + 2b + 12t}, \tag{98}$$

初始条件为

$$x^{t_u} = x(t_u) = -\frac{a}{4} - \frac{(a^3 - 4ab + 8c)^{1/3}}{2} \tag{99}$$

从 $t = t_u > 0$ 到 $t = 0$，必须在 $t = 0$处取得(13) 的全局极小值点。最后，如果一个临界点等于 $-\frac{a}{4}$，那么 $p(x)$ 有两个全局极小化点，它们是二次多项式的根

$$\frac{p(x)}{x + \frac{a}{4}}. \tag{100}$$

到目前为止，我们可以先验证 $-\frac{a}{4}$ 是否是三次多项式$-\frac{\partial p_{(x)}}{\partial x}$ 的根。如果 $-\frac{a}{4}$ 是根，则可立即得到全局极小值点。否则，设 $x^{(0)} = x_{init}$，且 $t^{(0)} = t_u$。然后受（58）启发，迭代过程如下

$$\begin{aligned} x^{(i+1)} =& x^{(i)} - \Delta t \cdot \frac{12x^{(i)} + 3a}{12{x^{(i)}}^2 + 6ax^{(i)} + 2b + 12t^{(i)}}, \\ t^{(i+1)} =& t^{(i)} - \Delta t. \end{aligned} \tag{101}$$

这里，给定的步长 $\Delta t > 0$ 足够小，我们可以在 $t^{(n)} \approx 0$时停止迭代。最后，该算法提供了

$$\lim_{i\to\infty} x^{(i)} = x_{min}. \tag{102}$$

除了从 x_{init}开始，我们也可以先充分演化$p(x,\ t)$。这会增加迭代步数。

4.6 数值实验

例 *4.2* 这个反例 [7] 是针对 [6], 提出的“反向微分流”方法的反例，其中 $p(x) = x^4 - 8x^3 - 18x^2 + 56x$。在我们的热传导框架中，有 $p(x,\ t) = x^4 - 8x^3 - (18 - 6t)x^2 + 32x - (18t - 3t^2)$。注意，图 12 展示了由临界点构成的三角形序列。

例 *4.3* 设 $p(x) = x^4 + 0.2114x^3 - 2.6841x^2 - 0.1110x + 1.2406$，则在图 3 中，最左侧的曲线 $x_1(t)$ 是相应的 $p(x,\ t)$ 在每个 $t \geq 0$处的全局极小值点，图 4 展示了指纹 $\mathcal{FP}_1$。理论上的极小化点是 $x_1 = -1.2307$，而我们的迭代算法给出了 $x_1 = -1.2308$。

例 *4.4* 考虑 $p(x) = x^4 - 4x^3 - 2x^2 + 12x$，那么实际上有三个临界点 $x_1 = -1$、$x_2 = 1$、$x_3 = 3$。注意，$x_1 + x_3 = 2x_2$ 因此$p(x_1,\ t) = p(x_3,\ t)$，并且 $x_2(t) = x_2 = 1$ 对所有 $x \in [0, t_u]$都成立。进一步可以验证，在这个对称情形下，必有 $t^* = t_u$。详细解释可参见定理 4.7。

4.7 四次多项式情形总结

对于四次多项式 $p(x)$ 的全局极小值点，在借助方差从 $t = 0$ 到 $+\infty$的高斯滤波器 $g_t(x)$ 生成其多尺度版本 $p(x,\ t) = p(x) * g_t(x)$ 时，我们将看到：

– 如果 $t_u < 0$, $p(x)$ 本身不必是凸的，但它有唯一的临界点。因此，在任意 $t \geq 0$ 下，每个 $p(x,t)$ 都只有一个临界点；– 如果进一步要求 $t_u \leq t^* < 0$, $p(x)$ 必须是凸的，那么在任意 $t \geq 0$ 下，每个 $p(x,t)$ 关于 x 都是凸的；

– 如果 $t_u > 0$，当 $0 \le t < t_u$ 时，多项式 $p(x)$ 有三个不同的临界点 $x_1 < x_2 < x_3$；– 当 $t = t_u$ 时，对应于全局极小值点的临界点将从 t_u 连续演化至 t^*，而局部极小化点将与局部极大值 x_2^t 相遇。此外，这两个临界点将在 $t = t_u$ 停止演化；– 当 $t_u < t < t^*$ 时，多项式 $p(x,t)$ 在每个 t 处都有唯一极小值点。– 当 $t \ge t^*$ 时，多项式 $p(x,t)$ 将关于 x 变为凸的，并具有唯一极小值点。

5 六次多项式的案例研究

5.1 演化与指纹

现在我们考虑六次首一多项式

$$p(x) = x^6 + bx^4 + cx^3 + dx^2 + ex + f, \tag{103}$$

为简化起见，这里我们将 x^5 的系数设为零以进行正则化，这是处理代数方程的一种标准技巧。因此，热演化为

$$\begin{aligned} p(x,t) =& p(x) + t \cdot \frac{\partial p}{\partial t} + \frac{t^2}{2}\frac{\partial^2 p}{\partial t^2} + \frac{t^3}{6}\frac{\partial^3 p}{\partial t^3} \\ =& p(x) + \frac{t}{2} \cdot \frac{\partial^2 p}{\partial x^2} + \frac{t^2}{8}\frac{\partial^4 p}{\partial x^4} + \frac{t^3}{48}\frac{\partial^6 p}{\partial x^6} \\ =& x^6 + b(t)x^4 + c(t)x^3 + d(t)x^2 + e(t)x + f(t), \end{aligned} \tag{104}$$

其中

$$\begin{aligned} b(t) &= b + 15t, \\ c(t) &= c, \\ d(t) &= d + 6bt + 45t^2, \\ e(t) &= e + 3ct, \\ f(t) &= f + dt + 3bt^2 + 15t^3. \end{aligned} \tag{105}$$

(103) 的临界点满足五次方程

$$0 = \frac{1}{6}\frac{\partial p(x,t)}{\partial x} = x^5 + B(t)x^3 + C(t)x^2 + D(t)x + E(t), \tag{106}$$

其中

$$\begin{aligned} B(t) &= \frac{2b}{3} + 10t, \\ C(t) &= \frac{c}{2}, \\ D(t) &= \frac{d}{3} + 2bt + 15t^2, \\ E(t) &= \frac{e}{6} + \frac{ct}{2}. \end{aligned} \tag{107}$$

（106）的解为 $t \geq 0$ $\mathcal{FP}_1$ 指纹。不幸的是，这个五次方程的根在代数上无法求解[18]。

同样，我们可以将 $\mathcal{FP}_2$ 的方程写成如下形式，

$$\frac{1}{30}\frac{\partial^2 p}{\partial x^2} = x^4 + \left(\frac{2b}{5} + 6t\right)x^2 + \frac{c}{5}x + \frac{d}{15} + \frac{2b}{5}t + 3t^2 = 0. \tag{108}$$

那么 $\mathcal{FP}_3$ 的方程为

$$\frac{1}{120}\frac{\partial^3 p}{\partial x^3} = x^3 + \left(\frac{b}{5} + 3t\right)x + \frac{c}{20} = 0. \tag{109}$$

我们关注的是交集 $\mathcal{FP}_2 \bigcap \mathcal{FP}_3$，它提供了合并信息。与四次多项式的情况不同，目前我们还没有任何显式表示。显然，四次方程 (108) 的实重根必定是 (108) 和 (109) 的公共根。一般而言，存在两个 $0 \leq t_1 < t_2$，使得对应的 x_1 和 x_2 是这两个实重根。定理 5.1 说明了数值求解的过程。

定理 5.1 对于任意六次首一多项式 $p(x)$，集合 $\mathcal{FP}_2 \bigcap \mathcal{FP}_3$ 包含一对元素 (x_i, t_i) $i = 1,2$，或一个元素 (x_1, t_1)，或为空。具体而言，

(1) 任意 t_i 都必须是判别式的零点

$$\Delta(t) = 27648t'^6 + \frac{1728c^2}{25}t'^3 - \frac{256}{625}h^2t'^2 - \frac{288}{625}c^2ht' - \frac{256h^3}{75^3} - \frac{27c^4}{625} \tag{110}$$

其中

$$t' = t + \frac{b}{15}, \quad h = b^2 - 5d. \tag{111}$$

（2）任意 x_i 都通过函数依赖于 t_i

$$x = -c \cdot \frac{1800\left(t + \frac{b}{15}\right)^2 - 4(b^2 - 5d)}{36000\left(t + \frac{b}{15}\right)^3 + 80(b^2 - 5d)\left(t + \frac{b}{15}\right) + 45c^2}. \tag{112}$$

证明 我们将在附录 B 中给出详细分析。根据该分析可知，在一般情形下，由四次方程（108）和（131）的判别式方程 $\Delta(t) = 0$，至多存在两个合并时刻 t_1 和 t_2。利用（130）和（131）可以立即验证函数（110），但此处省略详细计算。

现在，我们可以利用欧几里得算法，降低关于 x 的多项式的次数，以处理式 (108) 和 (109)。首先，将式 (109) 乘以 x，再分别从式 (108) 两边相减，得到一个二次多项式

$$\left(t + \frac{b}{15}\right)x^2 + \frac{c}{20}x + \left(t + \frac{b}{15}\right)^2 - \frac{b^2 - 5d}{225} = 0. \tag{113}$$

再次将式 (113) 两边乘以 x，然后从式 (109) 两边乘以 $t+\frac{b}{15}$ 后的结果中减去，即得

$$-\frac{c}{20}x^2+\left[2\left(t+\frac{b}{15}\right)^2+\frac{b^2-5d}{225}\right]x+\frac{c}{20}\left(t+\frac{b}{15}\right)=0. \tag{114}$$

最后，将 (114) 与 (113) 结合以消去二次项，即可得到 (112)。

如附录 B 所述，我们首先可以从判别式 $\Delta(t)=0$ 中得到合适的 t，然后将其代入上述 (112) 式，再从 (112) 中选择 x。

5.2 六次多项式的数值例子

将热演化算法推广到六次或更高偶次多项式，以求得全局极小值点，是极具意义的。然而，定理 4.6 无法推广到更高次数的多项式。在此，我们将说明正面和负面的例子。

例*5.1* 我们说明六次多项式的指纹

$$p(x)=x^6-0.3726x^4+0.0574x^3+0.0306x^2-0.0084x$$

如图 8 和图 9 所示。我们指出，全局极小值点（图 8(a) 中的“*”）不会随 t 增大而演化，这意味着针对该 $p(x)$ 的凸卷积不会收敛到其全局极小值点。

例 *5.2* $p(x)$ 的指纹 $\mathcal{FP}_1 = x^6+0.6987x^5-1.0908x^4-0.4216x^3+0.2177x^2+0.1071x$ 如图 10 所示，其中全局极小化点包含在通向凸 $p(x,\ t)$ 的积分曲线中。

例 *5.3* $p(x)$ 的指纹 $\mathcal{FP}_1 = x^6-0.8529x^5-0.4243x^4-0.2248x^3+0.0916x^2-0.0074x$ 如图 11 所示，其中全局极小化点未包含在通向凸 $p(x,\ t)$ 的积分曲线中。

6 结论

受 O.Arikan 等人的 [8] Stekelov 凸化版反向微分流算法启发，为了搜索多项式的全局极小值点，我们引入了一个自然的热演化框架，并应用 Yuille-Poggio 的指纹理论以及他们在计算机视觉中的 Yuille-Poggio 轨迹方程，以刻画 $p(x,\ t)$ 的临界点演化。我们提出了多项式 $p(x)$ 的逃逸区 和约束区域 的概念，并据此发现，多项式 $p(x)$ 的全局极小值点 x^*

可以从其凸化版本的全局极小值点$p(x,\ t)\ = p(x) * g_t(x)$ 反向演化，**当且仅当**该 x^* 位于多项式$p(x)$ 的逃逸区内。

具体而言，对于四次多项式，我们不仅得到了与 O. Arikan 等人类似的结果 [8]，还提出了一种明显更简单的牛顿法版本，无需凸化即可求得四次多项式的全局极小值点。

然而，根据伽罗瓦理论，多项式 $p(x)$ 的逃逸区和可达区通常无法通过代数方法明确表示。因此，我们迫切需要高效的数值方法，以及用于判定这些区域的各种准则。

此外，如果全局极小值点 x^* 位于逃逸区之外，我们设计了一种偏微分方程方法，从 $p(x)$ 的拟线性修改版本的全局极小值点出发，求得该点 x^*；相关内容将于近期公布。另外，关于多变量情形的研究结果也将在我们即将发表的论文中介绍。

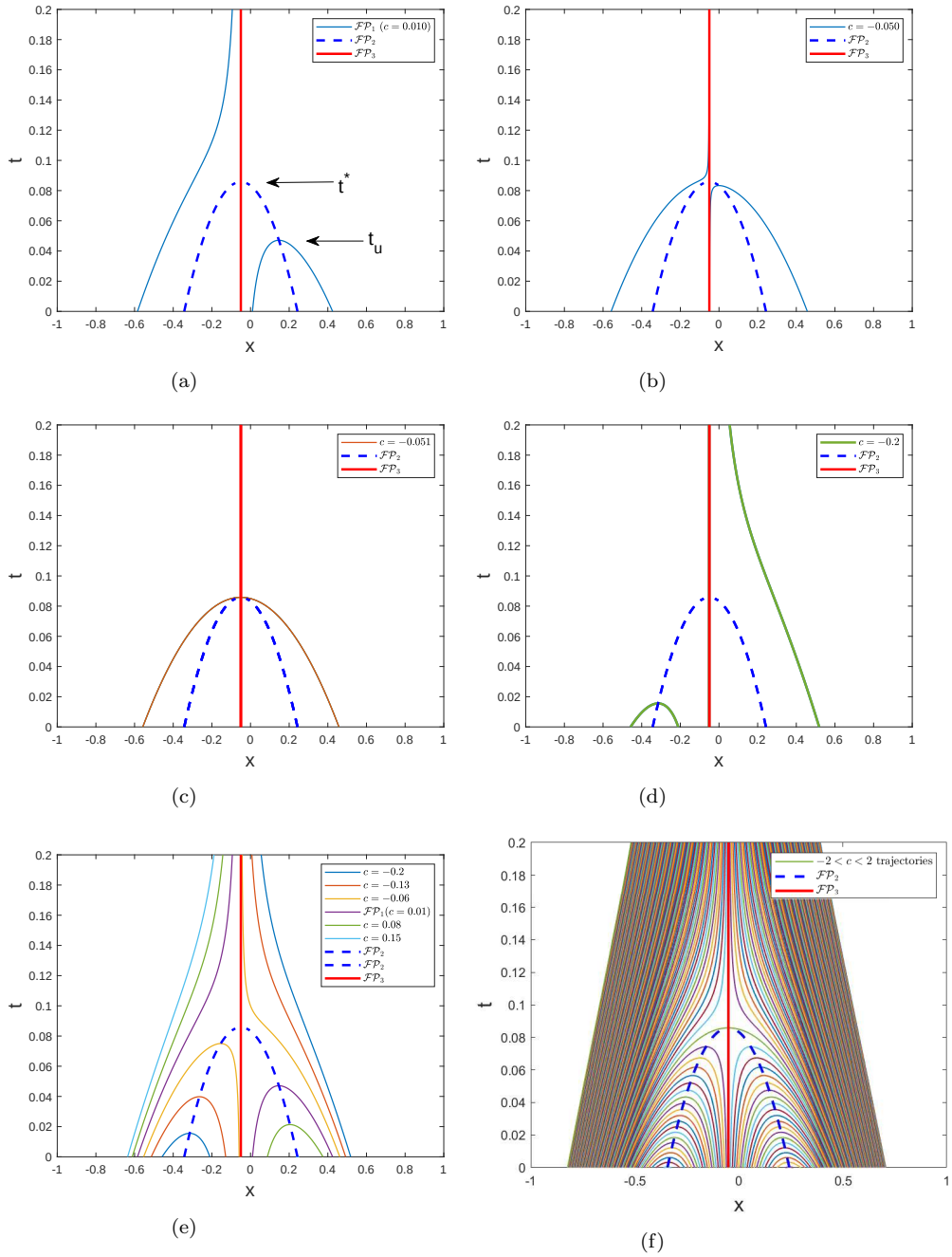


图2：例 4.1 中指纹和轨迹的示意图。为了观察多项式中系数 c 变化时轨迹的变化，我们在$c \in [-2, 2]$中改变它。(a)： $\mathcal{FP}_i$，$(i = 1, 2, 3)$， t_u 和 t^*。注意，$\mathcal{FP}_1$ 对应于$c = 0.01$。(b)： $\mathcal{FP}_2$， $\mathcal{FP}_3$ 以及 $c = -0.05$的轨迹。(c)：当 $c = -0.051$时的临界轨迹，它们关于 $x = -\frac{a}{4}$对称。(d)： $\mathcal{FP}_2$， $\mathcal{FP}_3$ 以及 $c = -0.2$的轨迹。(e)：通过改变 c得到的轨迹。(f)：通过改变 c得到的更多轨迹。

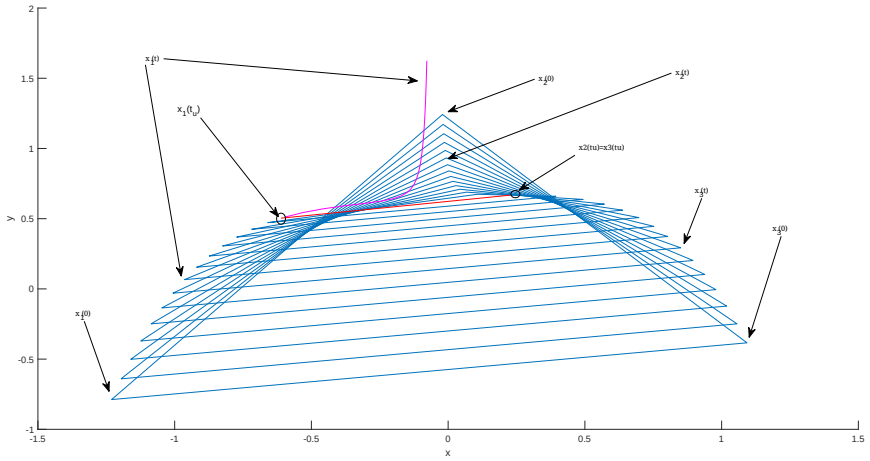


图 3：(x, y) 坐标系中的三角形序列示例，$p(x) = x^4 + 0.2114x^3 - 2.6841x^2 - 0.1110x + 1.2406$。更多详情请参见示例 4.3。

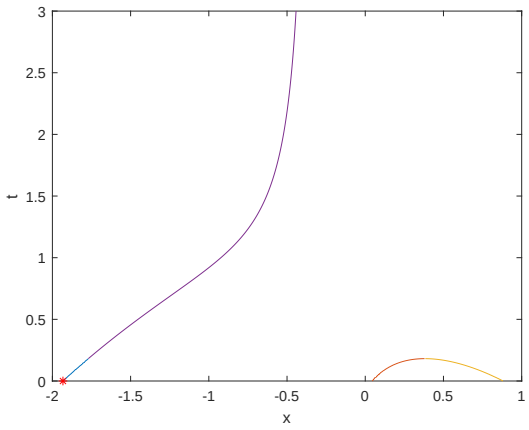


图 4：(x, t) 坐标系中的指纹 $\mathcal{FP}_1$，$p(x) = x^4 + 0.2114x^3 - 2.6841x^2 - 0.1110x + 1.2406$。更多详情请参见示例 4.3。

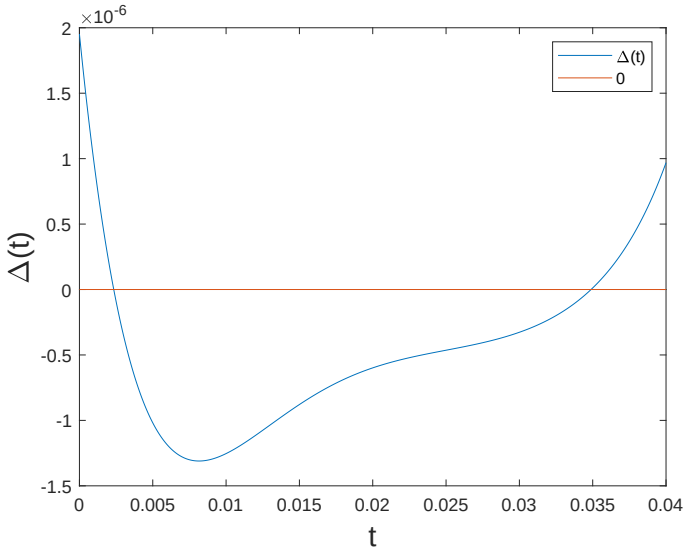


图5：四次方程的判别式 $\Delta(t)$ 由 $\partial^2 p(x,t)\,\partial x^2 \quad = 0$生成，并在 (110) 中定义。这里的数据来自例5.1。 $\Delta(t)$ 的两个实根分别为 $t_1 = 0.002341$ 和 $t_2 = 0.034887$；在 $t_1 = 0.002341$ 处，四次方程有一个实重根和两个互异实根；在 $t_2 = 0.034887$ 处，四次方程有一个实重根和一对共轭复根。

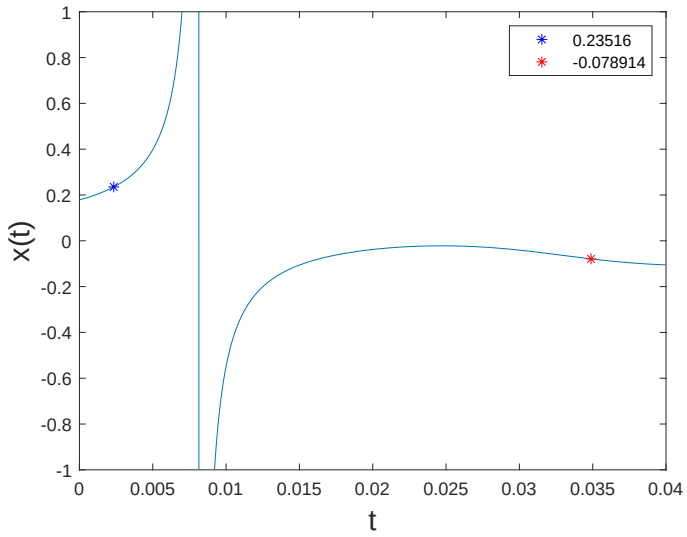


图6：函数 $x(t)$ 定义见（112），其中数据来自例5.1。这里我们得到两个解（t_1，x_1）$=(0.0023, 0.23516)$、（t_2，x_2）$=(0.03489, -0.078914)$，它们构成集合 $\mathcal{FP}_2 \bigcap \mathcal{FP}_3$。

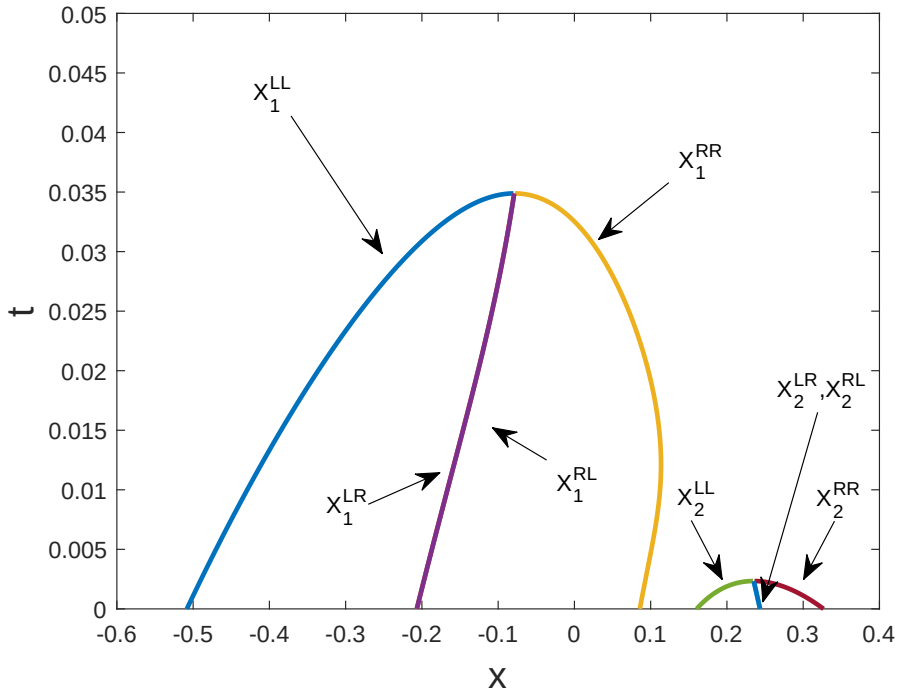


图7：数值上，通过 MATLAB 常微分方程工具包中的 @ode25 得到了例5.1 中定义的与 $p(x)$ 相关的约束区域 和逃逸区 的划分，且约束区域 为 $[-0.5082, 0.0858] \bigcup [0.1603, 0.3267]$。

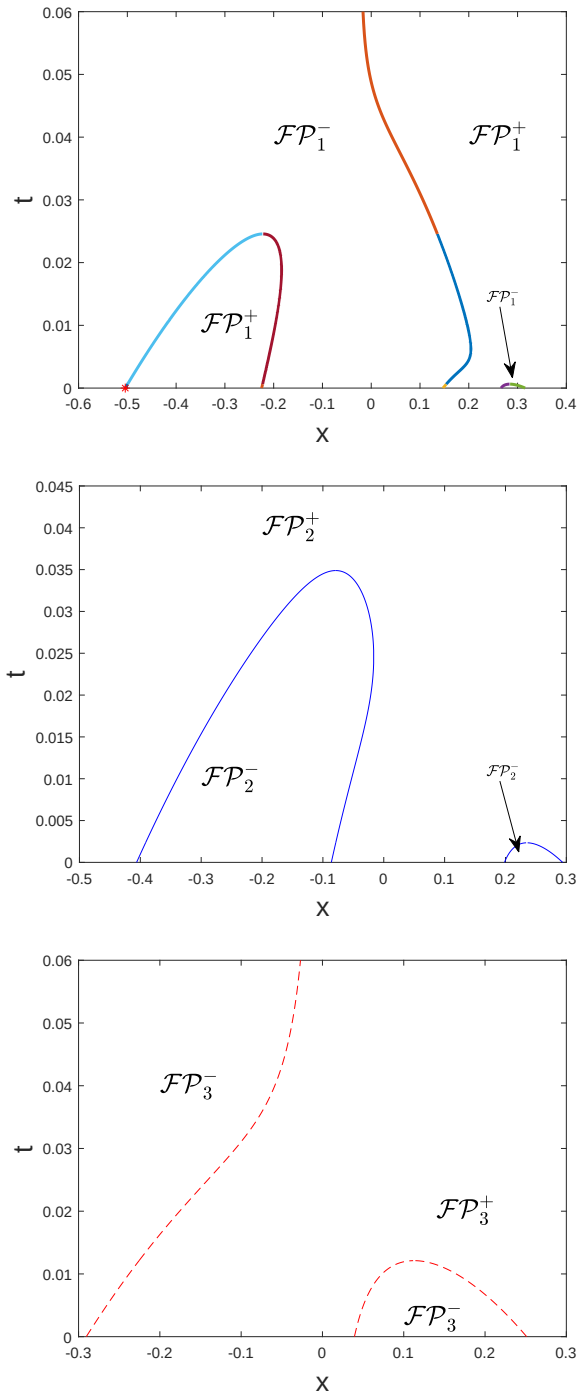


图8：例5.1中的指纹分别展示。

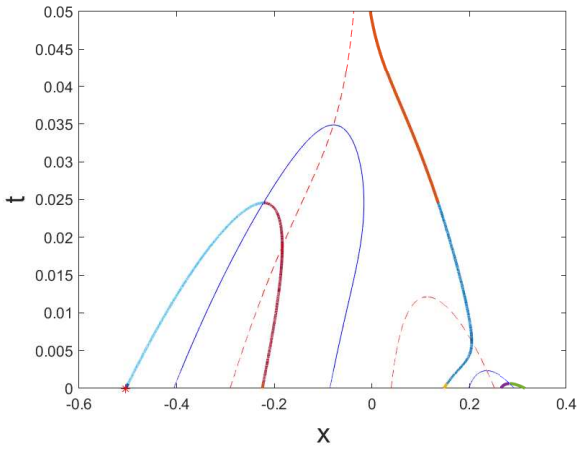


图9：例5.1中前几幅图的指纹 $\mathcal{FP}_1$、 $\mathcal{FP}_2$ 和 $\mathcal{FP}_3$ 的联合示意图。

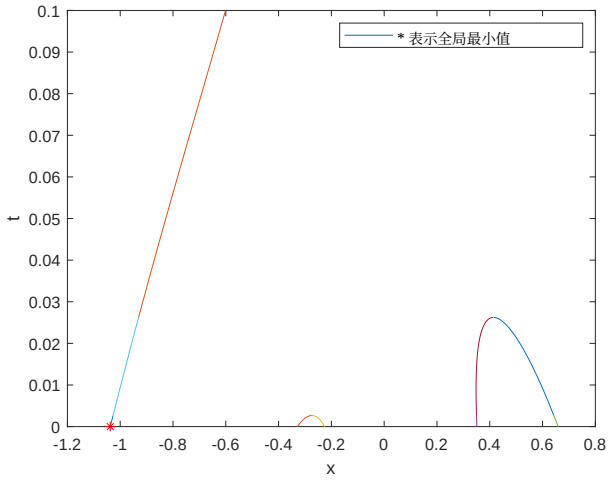


图10： $p\left(x\right) = x^6 + 0.6987x^5 - 1.0908x^4 - 0.4216x^3 + 0.2177x^2 + 0.1071x$ 的指纹示例。这里 $*$ 表示全局极小值点，沿临界点指纹从较大的 t 出发应用欧拉方法，将回溯至真正的全局极小值点。

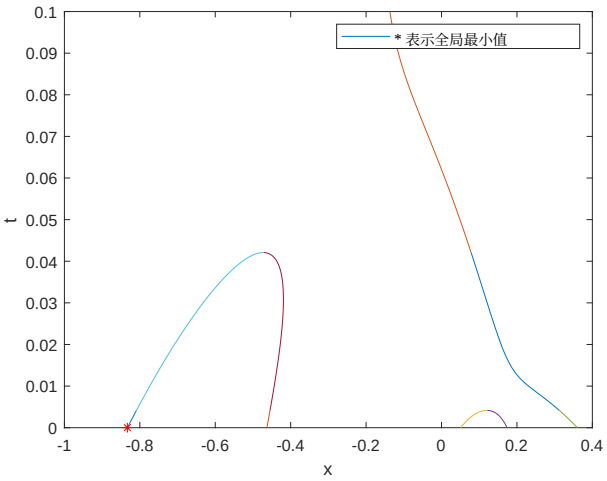
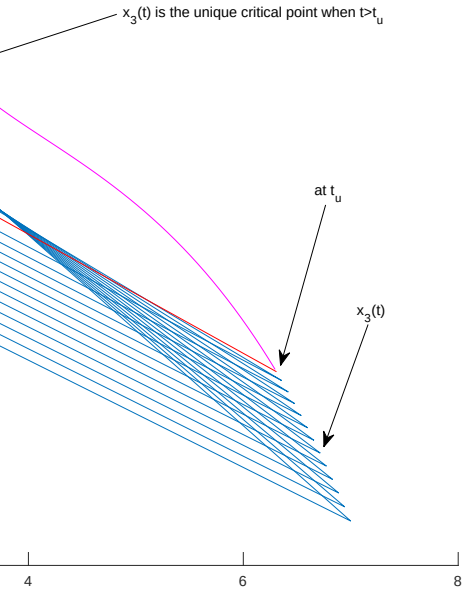


图 11： $p(x) = x^6 - 0.8529x^5 - 0.4243x^4 - 0.2248x^3 + 0.0916x^2 - 0.0074x$的指纹反例。这里，$*$ 表示全局极小值点，但最右侧的曲线从较大的 $t > 0$开始，只连接到局部极小值点。



$+ 56x$ is analyzed in Example 4.2.

and convexity of heat evolved version

A 三次方程的实数解

A.1 根的表示

回顾三次代数方程的经典理论（参见 [14]）

$$x^3+\alpha x^2+\beta x+\gamma=0, \tag{115}$$

根据牛顿法，我们有

引理 A.1 *Let $x_i(i=1,2,3)$ be the roots (real or complex) of polynomial equation*(115) *. We have the following propositions,*

$$\begin{aligned}
x_1+x_2+x_3&=-\alpha,\\
x_1x_2+x_2x_3+x_3x_1&=\beta,\\
x_1x_2x_3&=-\gamma,\\
x_1^2+x_2^2+x_3^2&=\alpha^2-2\beta,\\
x_1^3+x_2^3+x_3^3&=-\alpha^3+3\alpha\beta-3\gamma,\\
x_1^4+x_2^4+x_3^4&=\alpha^4-4\alpha^2\beta+4\alpha^2+2\beta^2.
\end{aligned} \tag{116}$$

当方程的系数为实数时，该方程的判别式为

$$\Delta=(x_1-x_2)^2(x_2-x_3)^2(x_3-x_1)^2, \tag{117}$$

这等价于

$$\Delta=\frac{g^2}{4}+\frac{f^3}{27}, \tag{118}$$

其中

$$f=\beta-\frac{\alpha^2}{3}\quad\text{and}\quad g=\frac{2\alpha^3}{27}-\frac{\alpha\beta}{3}+\gamma. \tag{119}$$

A.2 判别式所描述的实根

那么，该三次方程的解如下:

如果 $\Delta<0$，方程 (115) 有三个互异实根。

$$\begin{aligned}
x_1&=\frac{2}{\sqrt{3}}\sqrt{-f}\sin(\theta)-\frac{\alpha}{3},\\
x_2&=-\frac{2}{\sqrt{3}}\sqrt{-f}\sin\left(\theta+\frac{\pi}{3}\right)-\frac{\alpha}{3},\\
x_3&=\frac{2}{\sqrt{3}}\sqrt{-f}\cos\left(\theta+\frac{\pi}{6}\right)-\frac{\alpha}{3},
\end{aligned} \tag{120}$$

其中

$$\theta=\frac{1}{3}\arcsin\left(\frac{3\sqrt{3}g}{2(\sqrt{-f})^3}\right). \tag{121}$$

如果 $\Delta=0$，解中包含一个根和两个重根

根，

$$\begin{aligned}
x_1&=-2\left(\frac{g}{2}\right)^{\frac{1}{3}}-\frac{\alpha}{3},\\
x_2=x_3&=\left(\frac{g}{2}\right)^{\frac{1}{3}}-\frac{\alpha}{3}.
\end{aligned} \tag{122}$$

最后，当 $\Delta>0$ 时，该方程只有一个实根

$$x=\left(-\frac{g}{2}+\sqrt{\Delta}\right)^{\frac{1}{3}}+\left(-\frac{g}{2}-\sqrt{\Delta}\right)^{\frac{1}{3}}. \tag{123}$$

B 四次多项式的实重根及$\mathcal{FP}_2 \bigcap \mathcal{FP}_3$的结构

我们考虑实系数四次方程的正则化形式

$$x^4 + \beta x^2 + \gamma x + \delta = 0. \tag{124}$$

其根的结构由 [19] 描述。当且仅当其判别式 $\Delta = 0$ 时，它有重根，其中

$$\Delta = 256\gamma^3 - 128\beta^2\delta^2 + 144\beta\gamma^2\delta - 27\gamma^4 + 16\beta^4\delta - 4\beta^3\gamma^2. \tag{125}$$

此外，我们还需要另外四个多项式，

$$\begin{aligned} P =& 8\beta, \\ R =& 8\gamma, \\ \Delta_0 =& \beta^2 + 12\delta, \\ D =& 64\delta - 16\beta^2. \end{aligned} \tag{126}$$

方程 (124) 有一个实二重根和另外两个互异实根，当且仅当

$$\Delta = 0 \text{ and } \mathrm{P} < 0 \text{ and } \mathrm{D} < 0 \text{ and } \Delta_0 \neq 0. \tag{127}$$

方程 (124) 有一个实二重根和一对复根，当且仅当

$$(\Delta = 0 \text{ and } \mathrm{D} > 0) \text{ or } (\Delta = 0 \text{ and } \mathrm{P} > 0 \text{ and } (\mathrm{D} \neq 0 \text{ or } \mathrm{R} \neq 0)). \tag{128}$$

为了将上述结果应用于四次多项式演化，我们可以按照以下方式表示变量：

$$\begin{aligned} \beta =& \frac{2b}{5} + 6t, \\ \gamma =& \frac{c}{5}, \\ \delta =& \frac{d}{15} + \frac{2b}{5}t + 3t^2. \end{aligned} \tag{129}$$

最后，我们看到

$$\Delta(t) = \sum_{k=0}^{6} c_{6-k} t^k, \tag{130}$$

其中

$$\begin{aligned} c_0 =& 27648, \\ c_1 =& \frac{55296b}{5}, \\ c_2 =& \frac{9216b^2}{5}, \\ c_3 =& \frac{4096b^3 + 1728c^2}{25}, \\ c_4 =& \frac{4864b^4}{625} + \frac{512db^2 + 1728bc^2}{125} - \frac{256d^2}{25}, \\ c_5 =& \frac{32(b^2 + 5d)(48b^3 - 80db + 135c^2)}{9375}, \\ c_6 =& \frac{256b^4 d}{9375} - \frac{32b^3c^2}{3125} - \frac{512b^2d^2}{5625} + \frac{96bc^2d}{625} - \frac{27c^4}{625} + \frac{256d^3}{3375} \end{aligned} \tag{131}$$

这里我们应明确写出式 (127) 和 (128) 中的条件。实际上，根据式 (129)，我们有

$$\begin{aligned} P<0 &\iff t<-\frac{b}{15},\\ R\neq 0 &\iff c\neq 0,\\ \Delta_0\neq 0 &\iff d>\frac{b^2}{5},\\ D>0\ (=0,\ <0,\ \text{resp.}) &\iff \left(t+\frac{b}{15}\right)^2-\frac{1}{90}\left(d-\frac{b^2}{5}\right)>0\ (=0,\ <0,\ \text{resp.}). \end{aligned} \tag{132}$$

这意味着

$$(127)\iff -\frac{b}{15}-\frac{1}{\sqrt{90}}\sqrt{d-\frac{b^2}{5}}<t<-\frac{b}{15},\ d-\frac{b^2}{5}>0,\ \text{and}\ \Delta(\mathrm{t})=0.$$

如果我们记

$$\widehat{t}=\frac{1}{\sqrt{90}}\sqrt{d-\frac{b^2}{5}}-\frac{b}{15},$$

那么我们可以看到

– 若 $d-\frac{b^2}{5}<0$，则 (128) $\iff$ $\Delta(t)=0$.– 若 $d-\frac{b^2}{5}=0$，则 $D>0$ 等价于 $t\neq-\frac{b}{15}$。因此 (128) $\iff$ $t\neq-\frac{b}{15}$ 且 $\Delta(t)=0$.

– 若 $d-\frac{b^2}{5}>0$，我们可以将其分为两种情况。第一种情况，若 $d\geq\frac{3b^2}{5}$，则 $D>0$ 等价于

$$0\leq t<-\frac{b}{15}-\frac{1}{\sqrt{90}}\sqrt{d-\frac{b^2}{5}},\quad \text{or t}>-\frac{\text{b}}{15}+\frac{1}{\sqrt{90}}\sqrt{\text{d}-\frac{\text{b}^2}{5}}. \tag{133}$$

注意到

$$\begin{aligned} &D>0\ \vee\ (P>0\ \wedge\ (D\neq 0\ \vee\ R\neq 0))\\ =&(D>0\ \vee\ P>0)\ \wedge\ (D>0\ \vee\ D\neq 0\ \vee\ R\neq 0)\\ =&(D>0\ \vee\ P>0)\ \wedge\ (D\neq 0\ \vee\ R\neq 0) \end{aligned} \tag{134}$$

这表明 (128) $\iff$ $\Delta(t)=0$，且 $\mathrm{t}\in\left[0,\ -\frac{\text{b}}{15}-\frac{1}{\sqrt{90}}\sqrt{\text{d}-\frac{\text{b}^2}{5}}\right)\bigcup\left(-\frac{\text{b}}{15},\ +\infty\right)$，但排除 $D=0$ 和 $c=0$ 同时成立的情况。不过如果 $\frac{3b^2}{5}>d>\frac{b^2}{5}$，那么 $D>0$ 等价于

$$t>-\frac{b}{15}+\frac{1}{\sqrt{90}}\sqrt{d-\frac{b^2}{5}}. \tag{135}$$

那么 (128) $\iff$ $\Delta(t)=0$，且排除 $D=0$ 和 $c=0$ 同时成立的情况。

总之，我们可以得到在 $0\leq t_1\leq t_2$ 处，分别对应于方程 (124) 的一对实根 x_1 和 x_2，因此

$$\{(x_1,t_1),(x_2,t_2)\}=\mathcal{FP}_2\bigcap\mathcal{FP}_3,$$

其中 $t_2\geq t_1$ 和 $0\leq t_1<-\frac{b}{15}$。

参考文献

1. J. B. Lasserre, 《多项式全局优化与矩问题》, SIAM J. Optim. 第 11 卷第 3 期, 第 796–817 页, 2001 年。2. N.Z. Shor, 《二次优化问题》, Soviet J. Comput. Systems Sci., 25（1987）, 第 1–11 页。3. N.Z. Shor, 《不可微优化与多项式问题》, Kluwer Academic Publishers, 1998 年。4. V.N. Nefedov, 《多项式优化问题》, U.S.S.R. Comput. Maths. Math. Phys., 第 27 卷, 21, 1987 年, 5 3, 3No. pp.l - 。J. Zhu 和 X. Zhang, 《关于多项式全局优化》, Optimization Letters, （2008）2：239–249。6. J. Zhu、S. Zhao 和 G. Liu, 《通过反向微分流实现多项式全局最小化的解》, J. Optim Theory Appl（2014）161：828–836。7. O. Arikan、R.S. Burachik 和 C.Y. Kaya, 《“反向微分流”可能无法收敛到多项式的全局极小值点》, J. Optim Theory Appl（2015）：167：401–408。8. O. Arikan、R.S. Burachik 和 C.Y. Kaya, 《用于单变量全局优化的 Steklov 正则化与轨迹方法》, J. Global Optimization, 2019 年。9. Burachik, R.S., Kaya, C.Y.《用于多变量四次多项式全局优化的 Steklov 凸化与轨迹方法》。Math. Program.（2020）。https://doi.org/10 8 10 1007 10107 020 01536 . /s - - - . T. Iijima, 《模式观测基础理论》, 日本电子信息通信学会自动机与自动控制技术组论文集, 1959 年（日文）。11. T. Iijima, 《模式规范化基础理论（典型一维模式的情形）》, 电气试验所公报, 第 26 卷：368–388, 1962 年。12. A. L. Yuille 和 T. Poggio, 《零交叉点的指纹定理》, J. Opt. Soc. Am. A, 2 5 683 692, 1985 = 10 1364 2 000683 ( ): - 。doi . /JOSAA. . . 13 A. L. Yuille 和 T. A. Poggio, 《零交叉点的尺度定理》, IEEE Trans. on Pattern Analysis and Machine Intelligence, 第 8 卷第 1 期, 第 15–25 页, 1986 年 1 月, doi: 10.1109/TPAMI.1986.4767748。14. E. Zeidler, 《牛津数学用户指南》, Oxford Univ. Press, 2013 年。15. Irving Kaplansky, 《微分代数导论》, Hermann, 巴黎, 1957 年。16. Fritz John, 《偏微分方程》, Springer-Verlag, 第 4 版, 1982 年。17. W.A. Strauss, 《偏微分方程》。John Wiley and Sons Inc., 1992 年。18. I.N. Stewart, 《伽罗瓦理论》, Chapman and Hall/CRC, 第 4 版, 2015 年。19. E. L. Rees, 《四次方程根的图解讨论》, The American Mathematical Monthly, 29 11986100 20 2, 51 55, 1922 10 1080 00029890 1922 3 : - 。DOI: . / . . . Birkhoff, G. 和 Mac Lane, S.《现代代数概论》, 第 3 版。纽约：Macmillan, 1965 年。21. Nickalls, R.W.D, 《四次方程：揭示不变量与欧拉解法》, The Mathematical Gazette, 第 93 卷第 526 期, 第 66–75 页, 2009 年。