NONCONVEX ANCIENT SOLUTIONS
TO CURVE SHORTENING FLOW

YONGZHE ZHANG, CONNOR OLSON, ILYAS KHAN, AND SIGURD ANGENENT

ABSTRACT. We construct an ancient solution to planar curve shortening. The solution is at all times compact and embedded. For $t \ll 0$ it is approximated by the rotating Yin-Yang soliton, truncated at a finite angle $\alpha(t) = -t$, and closed off by a small copy of the Grim Reaper translating soliton.

1. INTRODUCTION

In [5] Daskalopoulos, Hamilton, and Sesum showed that any compact, convex, and embedded ancient solution to Curve Shortening in the plane is either a shrinking circle or the ancient paperclip solution. Qian You et.al.[10, 3] showed that there exist many other ancient solutions that are either embedded and not compact, or otherwise compact, convex, but not embedded. Here we construct an ancient solution to Plane Curve Shortening that is embedded, compact at all times, but not convex. We note that an ancient solution satisfying these properties was also constructed independently in [4]. We present a brief comparison of the two different approaches at the end of this introduction.

Our construction begins with the Yin–Yang soliton, i.e. the rotating soliton that is invariant with respect to reflection in the origin—see Figure 1 (left). This spiral shaped curve divides the plane into two congruent parts, and under Curve Shortening evolves by rotating with unit speed in the counterclockwise direction. Each of the two branches of the Yin–Yang spiral is a graph in polar coordinates, given by

$$r = \mathscr{R}(\theta + \tfrac{\pi}{2}), \quad r = \mathscr{R}(\theta - \tfrac{\pi}{2})$$

respectively. The Yin-Yang spiral is asymptotic to a Fermat spiral [9] which in polar coordinates is given by $r = a\sqrt{\theta}$. The Yin-Yang curve satisfies

$$(1.1) \qquad \mathscr{R}(\theta) = \sqrt{2\theta} + \mathcal{O}(\theta^{-3/2}), \qquad \mathscr{R}'(\theta) = \frac{1}{\sqrt{2\theta}} + \mathcal{O}(\theta^{-5/2}) \quad (\theta \to \infty).$$

We review the properties and derive expansions for $\mathscr{R}$ in appendix A.

The Yin–Yang soliton is itself an example of an embedded ancient (and in fact eternal) solution. It is however not compact. Our goal in this paper is to construct a compact embedded ancient solution which, for $t \to -\infty$ converges to the rotating

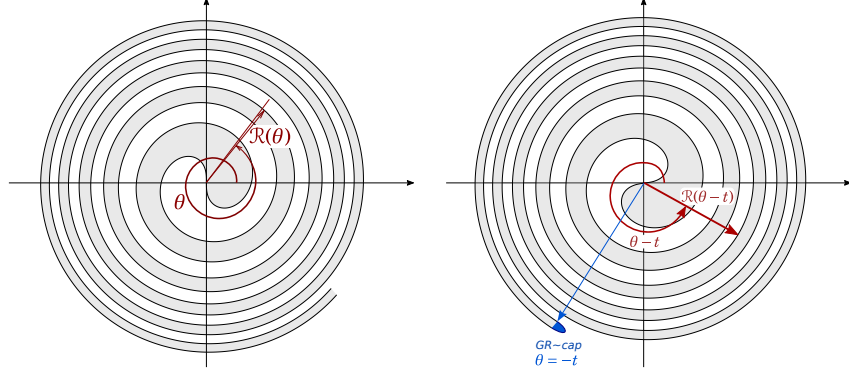Yin–Yang solution. In section 2 of this paper we construct an approximate solution to

Figure 1. Left: The Yin Yang soliton in polar coordinates. **Right:** The Truncated Yin Yang soliton at time $t \ll 0$. The bulk of the curve consists of an initial segment of the Yin–Yang curve rotated counterclockwise by an angle t, which has two ends. These ends are connected by a small cap in the shape of a translating soliton (the so-called "Grim Reaper") located at polar angle $\theta = -t$.

CSF, by truncating the Yin–Yang solution at polar angle $\alpha(t)$, and connecting the two ends with a small cap whose shape is approximately that of a rotated and rescaled copy of the Grim Reaper soliton. See Figure 1 (right).

To determine how to choose the angle $\alpha(t)$ we estimate the speed V of the cap at time t. Since the cap is approximately a Grim Reaper, its speed is related to its width w by $V = \pi/w$. The width w is approximated by

$$w \approx \mathcal{R}(\alpha(t) - t + \tfrac{\pi}{2}) - \mathcal{R}(\alpha(t) - t - \tfrac{\pi}{2}) \approx \mathcal{R}'(\alpha(t) - t)\pi,$$

so that the cap moves with speed

$$V \approx \frac{\pi}{w} \approx \frac{1}{\mathcal{R}'(\alpha(t) - t).}$$

On the other hand, far away from the center, the arms of the Yin-Yang spiral are close to circular. The cap, which moves with angular velocity $-\alpha'(t)$ along a near circle with radius $\mathcal{R}(\alpha(t) - t)$, therefore has velocity $V \approx -\alpha'(t)\mathcal{R}(\alpha(t) - t)$. It follows that $\alpha'(t) \approx -1/(\mathcal{R}\mathcal{R}')$. Since the asymptotics (1.1) imply $\mathcal{R}\mathcal{R}' \approx 1$, we end up with

$$(1.2) \qquad\qquad \alpha'(t) = -1 + o(1), \qquad \alpha(t) = -t + o(t).$$

For any $t < 0$ we let $\Omega(t) \subset \mathbb{R}^2$ be the region given in Polar Coordinates by

$$(1.3) \qquad \Omega(t) = \left\{ (r\cos\theta, r\sin\theta) \,\middle|\, t \leq \theta \leq -t, \; \mathcal{R}(\theta - t - \tfrac{\pi}{2}) \leq r \leq \mathcal{R}(\theta - t + \tfrac{\pi}{2}) \right\}$$

We will call the boundary curve $\partial\Omega(t)$ the *truncated Yin-Yang curve*. It consists of a segment of the rotating Yin-Yang curve and a straight line segment that connects the two ends at $\theta = -t$.

Main Theorem. There exists a compact ancient solution $\{C(t) \mid t < 0\}$ to Curve Shortening which for $t \to -\infty$ is uniformly close to the truncated Yin-Yang curve in

the sense that if $\Lambda(t)$ is the bounded region enclosed by $C(t)$ then[1]

$$\text{Area}(\Lambda(t) \triangle \Omega(t)) \lesssim |t|^{-1+\delta} \qquad (t \to -\infty)$$

for every $\delta > 0$.

The construction follows a pattern similar to the construction of ancient solutions in [3, 10], namely, construct a sequence of "really old solutions" $\{\mathscr{C}_n(t) \mid -n \leqslant t \leqslant t_0\}$ of Curve Shortening and extract a convergent subsequence whose limit is the desired ancient solution. In section 2 we first construct an ancient approximate solution $\{\mathscr{C}_*(t) \mid -\infty < t < 0\}$ of Curve Shortening, i.e. a family of curves for which

$$(1.4) \qquad \mathscr{E} \stackrel{\text{def}}{=} \int_{-\infty}^{t_0} \int_{\mathscr{C}_*(t)} |V - \kappa| \, ds \, dt < \infty.$$

Then we consider a sequence of solutions $\{\mathscr{C}_n(t) \mid -n \leqslant t < t_0\}$ of Curve Shortening whose initial curves $\mathscr{C}_n(-n)$ are chosen increasingly close to the approximate solution $\mathscr{C}_*(-n)$ at time $-n$, in the sense that the area between $\mathscr{C}_*(-n)$ and $\mathscr{C}_n(-n)$ tends to zero as $n \to \infty$. Arguing as in [3, 10] we observe that the area between the solutions $\mathscr{C}_n(t)$ and the approximate solution $\mathscr{C}_*(t)$ is bounded in terms of the "error" $\mathscr{E}$ and the area between the initial curves $\mathscr{C}_*(-n)$ and $\mathscr{C}_n(-n)$. Since the curves we deal with in this paper are not graphical the area estimate is a bit more complicated than in [3, 10]. In section 3 we present a more general estimate that generalizes the Altschuler-Grayson [2] area bounds for Space Curve Shortening. Finally, in section 4 we show how this bound allows us to extract a convergent subsequence of the very old solutions $\mathscr{C}_n(t)$, and provides enough control to conclude that the resulting limit satisfies the description in the Main Theorem.

In appendix A we recall the derivation of the Yin-Yang soliton, and obtain its asymptotic expansion at infinity.

Comparison with the construction in [4]. The idea to construct ancient solutions as a limit of a sequence of very old solutions is natural. In [4] the same strategy is also followed, but the core of the proof where one controls the sequence of old solutions is quite different. While the approach in our paper follows the method of area comparison established in [3], the arguments in [4] proceed by carefully controlling a sequence of old solutions via an exponential barrier for the bulk of the solution. Thereafter stability of the tip is established in [4] using a blow-up argument, combined with the uniqueness of the Grim Reaper as possible limit.

2. CONSTRUCTION OF THE CAP

2.1. **Parametrized curves and the Curve Shortening deficit.** An evolving family of curves is a map $X : (t_0, t_1) \times \mathbb{R} \to \mathbb{R}^2$ for which $X_p(t, p) \neq 0$ for all (t, p). For such a family we define

$$ds = \|X_p(t, p)\| \, dp, \qquad \frac{d}{ds} = \frac{1}{\|X_p(t, p)\|} \frac{d}{dp}.$$

The normal velocity and curvature of the family X are

$$V = \langle X_t, JX_s \rangle, \qquad \kappa = \langle X_{ss}, JX_s \rangle = \left\langle \frac{X_{pp}}{\|X_p\|^2}, \frac{JX_p}{\|X_p\|} \right\rangle$$

---

[1]For two sets A and B we denote their symmetric difference by $A \triangle B = (A \setminus B) \cup (B \setminus A)$.
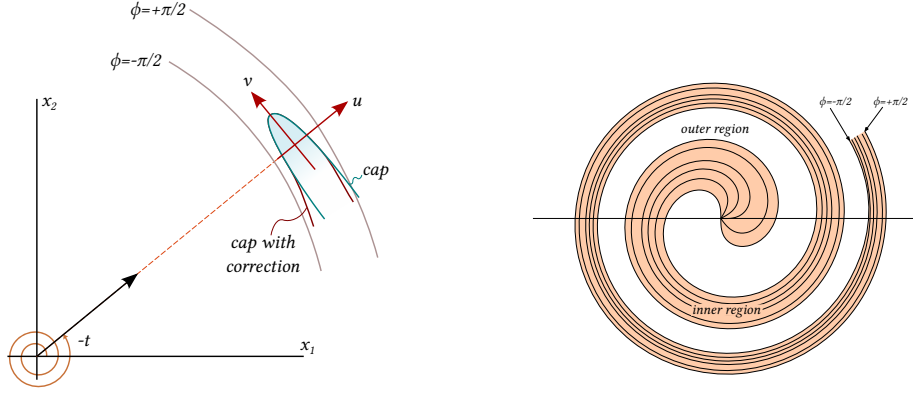
Figure 2. Left: The (u,v) coordinates. **Right:** The Yin Yang foliation at time $t = 0$.

where $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ represents counterclockwise rotation by $\frac{\pi}{2}$. By definition, the parameterized family of curves X satisfies Curve Shortening if $V = \kappa$. If it does not, then we measure the "discrepancy with curve shortening" in terms of the form

$$(2.1) \qquad |V - \kappa|\, ds = \left| \left\langle X_t - \frac{X_{pp}}{\|X_p\|^2}, J X_p \right\rangle \right| dp$$

The error $\mathcal{E}$ is obtained by integrating this form over the curve and in time.

We write $\boldsymbol{e}_1 = (1,0)$, $\boldsymbol{e}_2 = (0,1)$ for the standard basis for $\mathbb{R}^2$. In this basis counterclockwise rotation by θ is given by

$$e^{\theta J} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

We will use the rotated frame $\{\boldsymbol{E}_1(\theta), \boldsymbol{E}_2(\theta)\} = \{e^{\theta J}\boldsymbol{e}_1, e^{\theta J}\boldsymbol{e}_2\}$, which satisfy

$$\boldsymbol{E}_1{'}(\theta) = \boldsymbol{E}_2(\theta), \qquad \boldsymbol{E}_2{'}(\theta) = -\boldsymbol{E}_1(\theta).$$

2.2. Z **or** (u,v)**-coordinates.** We expect the tip of the ancient solution to be located near the point $R(t)\boldsymbol{E}_1(-t)$, and to have width $\sim R^{-1}$, where

$$R = R(t) \overset{\text{def}}{=} \mathcal{R}(-2t).$$

Thus we introduce new, time dependent, coordinates $Z = u\boldsymbol{e}_1 + v\boldsymbol{e}_2$ related to the cartesian coordinates $X = (x_1, x_2)$ via

$$(2.2) \qquad X = e^{-tJ}\{R\boldsymbol{e}_1 + R^{-1}Z\}, \text{ i.e. } Z = R\{e^{tJ}X - R\boldsymbol{e}_1\}.$$

2.3. **The inner and outer Yin-Yang arms.** The region on one side of the Yin-Yang curve is foliated by rotated copies of the curve. At time $t \in \mathbb{R}$ the leaves of this foliation are parametrized by

$$(2.3) \qquad Y(\theta, t, y) = \mathcal{R}(\theta - t + y)\boldsymbol{E}_1(\theta)$$

where $y \in [-\pi/2, \pi/2]$ determines the leaf, and $\theta \in (t - y, \infty)$ is the polar angle on the leaf. The inner and outer arms of the region that contains our ancient solution correspond to $y = \pm\pi/2$. See Figure 2.

2.4. **Lemma.** For any $M > 0$ there is a $t_M < 0$ such that if $t \leqslant t_M$, then the segments of the Yin-Yang leaves $Y(\theta, t, y)$ with $|\theta + t| \leqslant M\tau^{-1} \ln \tau$ are graphs in (u, v) coordinates of the form $u = U_y(t, v)$, at least if $t \leqslant t_M$. Moreover, the functions $U_y(t, v)$ satisfy

$$(2.4) \qquad U_y(t, v) = y - \frac{y^2 + v^2 - 2v}{2\tau} + \mathcal{O}(\tau^{-2+\delta})$$

for all $\delta > 0$, where $\tau = -4t$.

Proof. We begin with the defining equations

$$(2.5) \qquad \mathscr{R}(\theta - t + y)\boldsymbol{E}_1(\theta) = e^{-tJ}\{R\boldsymbol{e}_1 + R^{-1}Z\} = e^{-tJ}\{(R + u/R)\boldsymbol{e}_1 + v/R\,\boldsymbol{e}_2\}.$$

Multiply with e^{tJ} on both sides:

$$\mathscr{R}(\theta - t + y)\boldsymbol{E}_1(\theta + t) = R\boldsymbol{e}_1 + R^{-1}Z = (R + u/R)\boldsymbol{e}_1 + v/R\,\boldsymbol{e}_2.$$

Here the left hand side is the polar form of the right hand side. Under our assumptions $R^2 \sim \tau$ and $|u| \ll \tau$, so $1 + u/R^2 > 0$, and hence we have

$$(2.6) \qquad \mathscr{R}(\theta - t + y)^2 = R^2 + 2u + \frac{u^2 + v^2}{R^2}, \qquad \theta + t = \arctan\frac{v/R^2}{1 + u/R^2}.$$

Since $\alpha \mapsto \mathscr{R}(\alpha)^2$ is a monotone function, the first equation in (2.6) can be solved for y. Using the asymptotic expansion in appendix A.3 we get

$$(2.7) \qquad 2(\theta - t + y) = \mathscr{R}^2 + 2\mathscr{R}^{-2} + \mathcal{O}(\mathscr{R}^{-4}),$$

where $\mathscr{R} = \mathscr{R}(\theta - t + y)$. Replace $\mathscr{R}^2$ by the expression in (2.6), to get

$$2(\theta - t + y) = R^2 + 2u + \frac{u^2 + v^2 + 2}{R^2} + \mathcal{O}(v^2 R^{-4}).$$

To eliminate θ we expand the second equation in (2.6),

$$\theta + t = \frac{v}{R^2} + \mathcal{O}(v^2 R^{-4}).$$

Hence

$$(2.8) \qquad 2y = 2(\theta - t + y) - 2(\theta + t) + 4t \qquad\qquad\qquad (-4t = \tau)$$

$$= R^2 - \tau + 2u + \frac{u^2 + v^2 - 2v + 2}{R^2} + \mathcal{O}(v^2 R^{-4}).$$

Expand $R = R(t) = \mathscr{R}(-2t)$ in powers of $\tau = -4t$ using (A.7):

$$R^2 = \mathscr{R}(-2t)^2 = \tau\left(1 - 2\tau^{-2} + \mathcal{O}(\tau^{-3})\right)$$

and substitute this in (2.8) to get

$$2y = 2u + \frac{u^2 + v^2 - 2v}{\tau} + \mathcal{O}(v^2 R^{-4}).$$

By assumption we have $|v| \lesssim \ln \tau$ and $R \sim \sqrt{\tau}$, so $v^2 R^{-4} = \mathcal{O}(\tau^{-2+\delta})$ for any $\delta > 0$.

If τ is sufficiently large then the above equation has a unique solution $u = U_y(t, v)$ with

$$u = y - \frac{u^2 + v^2 - 2v}{2\tau} + \mathcal{O}(\tau^{-2+\delta}) = y - \frac{y^2 + v^2 - 2v}{2\tau} + \mathcal{O}(\tau^{-2+\delta}). \qquad \square$$

2.5. **General *ansatz* for the cap.** We now construct the cap by assuming that it is given by

(2.9) $$X(t,p) = e^{-tJ}\{R\boldsymbol{e}_1 + \varepsilon Z(t,p)\}, \text{ with } R(t) \overset{\text{def}}{=} \mathscr{R}(-2t), \ \varepsilon(t) = \frac{1}{R(t)},$$

and by computing its deviation from Curve Shortening (2.1)

$$(V - \kappa)ds = \langle X_t - X_{ss}, JX_s\rangle ds = \left\langle X_t - \frac{X_{pp}}{\|X_p\|^2}, JX_p\right\rangle dp \overset{\text{def}}{=} W\,dp$$

In the following computations it will be convenient to abbreviate

$$R_\theta(t) = \mathscr{R}'(-2t),$$

so that $R'(t) = -2R_\theta(t)$.

The space derivatives of X are

$$X_p = e^{-tJ}\varepsilon Z_p \qquad X_{pp} = e^{-tJ}\varepsilon Z_{pp}.$$

The time derivative has a few more terms:

$$X_t = e^{-tJ}\left\{-2R_\theta\boldsymbol{e}_1 - R\boldsymbol{e}_2 + \varepsilon' Z - \varepsilon JZ + \varepsilon Z_t\right\}$$

Express $W = \langle X_t - X_{pp}/\|X_p\|^2, JX_p\rangle$ in terms of Z, keeping in mind that $\varepsilon = R^{-1}$,

(2.10) $$W = \left\langle \varepsilon^2 Z_t - \frac{Z_{pp}}{\|Z_p\|^2}, JZ_p\right\rangle - \langle\boldsymbol{e}_1, Z_p\rangle$$

$$+ 2\varepsilon R_\theta\langle\boldsymbol{e}_2, Z_p\rangle + \varepsilon\varepsilon'\langle Z, JZ_p\rangle - \varepsilon^2\langle Z, Z_p\rangle.$$

We look for a cap in the form of a normal perturbation of the Grim Reaper curve, i.e. we assume

(2.11) $$Z(t,p) = G(p) + f(t,p)JG_p(p)$$

where

$$G(p) = -\arcsin(\tanh p)\boldsymbol{e}_1 - \ln(\cosh p)\boldsymbol{e}_2$$

is the arclength parametrization of the Grim Reaper.

Since G is an arclength parametrization G_p, JG_p are unit tangent and normal to the Grim Reaper. Specifically,

$$G_p = -\frac{1}{\cosh p}\boldsymbol{e}_1 - \tanh p\,\boldsymbol{e}_2.$$

The parametrization $G(p)$ traces the Grim Reaper out from right to left. Furthermore, the curvature vector of the Grim Reaper is

$$G_{pp}(p) = \kappa(p)JG_p, \text{ where } \kappa(p) = \frac{1}{\cosh p}.$$

2.6. **Detailed computation of W on the cap.** We have

$$Z_t = f_t JG_p$$
$$Z_p = G_p + f_p JG_p + f JG_{pp} = (1 - \kappa f)G_p + f_p JG_p$$
$$JZ_p = -f_p G_p + (1 - \kappa f)JG_p$$
$$\|Z_p\|^2 = (1 - \kappa f)^2 + f_p^2 = 1 - 2\kappa f + \kappa^2 f^2 + f_p^2$$
$$Z_{pp} = -(\kappa_p f + 2\kappa f_p)G_p + (f_{pp} + \kappa - \kappa^2 f)JG_p$$

Substituting in (2.10), and using $\kappa = 1/\cosh p$, we get

$$\left\langle \varepsilon^2 Z_t, J Z_p \right\rangle = \varepsilon^2 (1 - \kappa f) f_t$$

$$\left\langle \frac{Z_{pp}}{\|Z_p\|^2}, J Z_p \right\rangle = \frac{(1 - \kappa f)(f_{pp} + \kappa(1 - \kappa f)) + \kappa_p f f_p + 2\kappa f_p^2}{1 - 2\kappa f + \kappa^2 f^2 + f_p^2}$$

$$\langle e_1, Z_p \rangle = -\kappa(1 - \kappa f) + f_p \tanh p.$$

Thus after substituting and expanding we find

$$(2.12) \quad W = \varepsilon^2 (1 - \kappa f) f_t - \frac{\kappa + f_{pp} - 2\kappa^2 f - \kappa f f_{pp} + \kappa_p f f_p + \kappa^3 f^2 + 2\kappa f_p^2}{1 - 2\kappa f + \kappa^2 f^2 + f_p^2} + \kappa(1 - \kappa f)$$

$$- f_p \tanh p - \varepsilon^2 \langle Z, Z_p \rangle + 2\varepsilon R_\theta \langle e_2, Z_p \rangle + \varepsilon \varepsilon' \langle Z, J Z_p \rangle.$$

We will choose $f(t, p)$ so as to make W integrable in space and time. To find f we linearize the expression for W and solve the resulting first order equation for f. It turns out that one solution is of the form $f(t, p) = \tau^{-1} F(p)$ for a function F that is of polynomial growth for $|p| \to \infty$. We restrict our attention to the region

$$(2.13) \qquad\qquad |p| \leqslant 2K \ln \tau, \qquad \tau \gg 1,$$

where $K = 100$ is a fixed, largish, constant. We will assume that f and its derivatives are bounded by

$$(2.14) \qquad\qquad |f| + |f_p| + |f_{pp}| + \tau |f_t| \lesssim \tau^{-1+\delta} \text{ for } |p| \leqslant 2K \ln \tau.$$

Here, and in what follows, when we write estimates for remainder terms of the form $\mathscr{O}(\tau^{-m+\delta})$, the estimate is implicitly meant to hold "for all $\delta > 0$."

The bound (2.14) will certainly hold if $f(t, p) = \tau^{-1} F(p)$ for some function $F(p)$ for which $F(p)$, $F'(p)$, and $F''(p)$ grow polynomially as $p \to \pm\infty$.

We now consider the many terms in (2.12) that add up to W. To begin, we have for $|p| \leqslant 2K \ln \tau$,

$$\varepsilon^2 (1 - \kappa f) f_t = \mathscr{O}(\tau^{-2+\delta}),$$

and also

$$\frac{-\kappa f f_{pp} + \kappa_p f f_p + \kappa^3 f^2 + 2\kappa f_p^2}{1 - 2\kappa f + \kappa^2 f^2 + f_p^2} = \mathscr{O}(\tau^{-2+\delta}) \text{ for any } \delta > 0.$$

It follows from $Z = G + f J G_p$ that

$$Z = G + \mathscr{O}(\tau^{-1+\delta}), \qquad Z_p = G_p + \mathscr{O}(\tau^{-1+\delta}).$$

Hence, for $|p| \leqslant 2K \ln \tau$,

$$\|G(p)\| \lesssim \ln \tau \qquad\qquad \varepsilon^2 \langle Z, Z_p \rangle = \tau^{-1} \langle G, G_p \rangle + \mathscr{O}(\tau^{-2+\delta})$$

$$\varepsilon R_\theta \langle e_2, Z_p \rangle = \tau^{-1} \langle e_2, G_p \rangle + \mathscr{O}(\tau^{-2+\delta}) \qquad \varepsilon \varepsilon' \langle Z, J Z_p \rangle = \mathscr{O}(\tau^{-2+\delta}).$$

In this computation we have used the expansions $\varepsilon = R(t)^{-1} = \mathscr{R}(-2t)^{-1} = \tau^{-1/2} + \mathscr{O}(\tau^{-3/2})$ and $R_\theta = \mathscr{R}'(-2t) = \tau^{-1/2} + \mathscr{O}(\tau^{-3/2})$ that follow from the expansions of $\mathscr{R}(\theta)$ and $\mathscr{R}'(\theta)$ in appendix A.

So far we have

$$W = \mathscr{O}(\tau^{-2+\delta}) - \frac{\kappa + f_{pp} - 2\kappa^2 f}{1 - 2\kappa f + \kappa^2 f^2 + f_p^2} + \kappa(1 - \kappa f) - f_p \tanh p$$

$$+ \tau^{-1} \langle 2e_2 - G, G_p \rangle$$

We can simplify the fraction (for $|p| \leqslant 2K \ln \tau$) by using

$$\frac{1}{1 - 2\kappa f + \kappa^2 f^2 + f_p^2} = 1 + 2\kappa f - \kappa^2 f^2 - f_p^2 + \frac{(-2\kappa f + \kappa^2 f^2 + f_p^2)^2}{1 - 2\kappa f + \kappa^2 f^2 + f_p^2} = 1 + 2\kappa f + \mathcal{O}\left(\tau^{-2+\delta}\right)$$

which implies

$$\frac{\kappa + f_{pp} - 2\kappa^2 f}{1 - 2\kappa f + \kappa^2 f^2 + f_p^2} = \kappa + 2\kappa^2 f + f_{pp} - 2\kappa^2 f + \mathcal{O}\left(\tau^{-2+\delta}\right) = \kappa + f_{pp} + \mathcal{O}\left(\tau^{-2+\delta}\right),$$

and hence

$$W = -f_{pp} - \tanh(p)f_p - \kappa^2 f + \frac{1}{\tau}\langle 2\boldsymbol{e}_2 - G, G_p \rangle + \mathcal{O}(\tau^{-2+\delta}).$$

2.7. **Computation of the correction term.** We look to perturb the Grim Reaper with a term of the form

$$f(t,p) = \frac{F(p)}{\tau},$$

where F is a solution of

$$\mathscr{L}F \overset{\text{def}}{=} F_{pp} + \tanh(p)F_p + \kappa^2 F = \langle 2\boldsymbol{e}_2 - G, G_p \rangle.$$

The linear operator can be factored

$$\mathscr{L} = \frac{d^2}{dp^2} + \tanh(p)\frac{d}{dp} + \frac{1}{\cosh^2(p)} = \frac{d}{dp} \circ \frac{1}{\cosh p} \circ \frac{d}{dp} \circ \cosh(p)$$

while we also have

$$\langle 2\boldsymbol{e}_2 - G, G_p \rangle = \frac{d}{dp}\langle 2\boldsymbol{e}_2 - \tfrac{1}{2}G, G \rangle,$$

all of which allows us to solve the equation for F:

(2.15) $$F(p) = \frac{A}{\cosh p} + B\tanh p + \int_0^p \frac{\cosh r}{\cosh p}\langle 2\boldsymbol{e}_2 - \tfrac{1}{2}G(r), G(r) \rangle \, dr.$$

It appears that the first term is of no use, so we set $A = 0$. The resulting function $F(p)$ is an odd function of p. We use the asymptotic behavior of $\langle 2\boldsymbol{e}_2 - \tfrac{1}{2}G, G \rangle$ for large p to find an expansion for the integral as $p \to +\infty$.

Consider

$$I \overset{\text{def}}{=} \int_0^p \frac{\cosh r}{\cosh p}\langle 2\boldsymbol{e}_2 - \tfrac{1}{2}G(r), G(r) \rangle \, dr.$$

The explicit expression for G implies

$$\langle \boldsymbol{e}_2, G(r) \rangle = -\ln \cosh r$$

$$\|G(r)\|^2 = \left(\ln \cosh r\right)^2 + \left(\arcsin \tanh r\right)^2 = \left(\ln \cosh r\right)^2 + \frac{\pi^2}{4} + \mathcal{O}(e^{-r}) \qquad (r \to \infty).$$

To compute I we substitute $\lambda = \ln \cosh p$, $\mu = \ln \cosh r$, which leads to

$$I = -\int_0^\lambda e^{\mu - \lambda}\left\{2\mu + \tfrac{1}{2}\mu^2 + \tfrac{\pi^2}{8} + \mathcal{O}(e^{-\mu})\right\}\frac{d\mu}{\sqrt{1 - e^{-2\mu}}}$$

The integrand is singular but integrable at $\mu = 0$. To deal with this singularity split

$$(1 - e^{-2\mu})^{-1/2} = 1 + \left[(1 - e^{-2\mu})^{-1/2} - 1\right], \text{ with } 0 \leqslant (1 - e^{-2\mu})^{-1/2} - 1 \lesssim \frac{1}{\sqrt{\mu}}e^{-2\mu}.$$

Replacing $(1 - e^{-2\mu})^{-1/2}$ by 1 therefore introduces an extra term that is bounded by $\mathcal{O}(e^{-\lambda})$. Hence we have

$$
\begin{aligned}
I &= \mathcal{O}(e^{-\lambda}) - \int_0^\lambda e^{\mu-\lambda}\{2\mu + \tfrac{1}{2}\mu^2 + \tfrac{\pi^2}{8} + \mathcal{O}(e^{-\mu})\}\,d\mu \\
&= \mathcal{O}(\lambda^3 e^{-\lambda}) - \int_0^\lambda e^{\mu-\lambda}\{2\mu + \tfrac{1}{2}\mu^2 + \tfrac{\pi^2}{8}\}\,d\mu \\
&= \mathcal{O}(\lambda^3 e^{-\lambda}) - \{2\lambda + \tfrac{1}{2}\lambda^2 + \tfrac{\pi^2}{8}\} + \{2 + \lambda\} - \{1\} \\
&= -\tfrac{1}{2}\lambda^2 - \lambda - \tfrac{\pi^2}{8} + 1 + \mathcal{O}(\lambda^3 e^{-\lambda})
\end{aligned}
$$

Thus, for large p we get

$$
F(p) = B - \tfrac{1}{2}(\ln\cosh p)^2 - \ln\cosh p - \tfrac{\pi^2}{8} + 1 + \mathcal{O}\big(|p|^3 e^{-|p|}\big) \qquad (p \to +\infty).
$$

Since $F(p)$ is an odd function we also have

$$
F(p) = -\left\{B - \tfrac{1}{2}(\ln\cosh p)^2 - \ln\cosh p - \tfrac{\pi^2}{8} + 1\right\} + \mathcal{O}\big(|p|^3 e^{-|p|}\big) \qquad (p \to -\infty).
$$

Applying this to $Z = G + \tau^{-1}F(p)JG_p$ we get for the two components u and v of Z as $p \to \pm\infty$:

$$
u(t,p) = -\arcsin\tanh p + \frac{1}{\tau}F(p)\tanh p = \mp\frac{\pi}{2} + \frac{1}{\tau}F(p)\tanh p + \mathcal{O}(e^{-|p|})
$$

$$
v(t,p) = -\ln\cosh p - \frac{1}{\tau}\frac{F(p)}{\cosh p} = -\ln\cosh p + \mathcal{O}(|p|^3 e^{-|p|})
$$

We can again eliminate p when p is large by using

$$
\ln\cosh p = -v + \mathcal{O}(p^2 e^{-|p|}), \qquad (p \to \pm\infty)
$$

which leads to

$$
(2.16) \quad u(t,p) = \mp\frac{\pi}{2} - \frac{v^2 - 2v + \pi^2/4 + 2(1+B)}{2\tau} + \mathcal{O}(|p|^3 e^{-|p|}) \qquad (p \to \pm\infty)
$$

We now determine B by matching (2.16) with the representation of the Yin-Yang arms in (u,v) coordinates that we found in (2.4). Setting $y = \mp\pi/2$ in (2.4) we find for the outer and inner Yin-Yang arms

$$
(2.17) \qquad u = \mp\frac{\pi}{2} - \frac{\pi^2/4 + v^2 - 2v}{2\tau} + \mathcal{O}(\tau^{-2+\delta})
$$

If $|p| \geq \frac{K}{2}\ln\tau$ then $|p|^3 e^{-|p|} = \mathcal{O}(\tau^{-K/2+\delta}) = o(\tau^{-2+\delta})$, and therefore the two expansions (2.16),(2.17) match if

$$
(2.18) \qquad\qquad\qquad B = -1.
$$

To summarize, we choose the cap to be given by

$$
(2.19) \qquad\qquad X(t,p) = e^{-tJ}\left\{R(t)\boldsymbol{e}_1 + \frac{1}{R(t)}Z(t,p)\right\}
$$

with

$$
(2.20) \qquad\qquad Z(t,p) = G(p) + \frac{F(p)}{\tau}JG_p(p)
$$

and, from (2.15),

$$
(2.21) \qquad\qquad F(p) = -\tanh p + \int_0^p \frac{\cosh r}{\cosh p}\langle 2\boldsymbol{e}_2 - \tfrac{1}{2}G(r), G(r)\rangle\,dr
$$

2.8. **Definition of the smooth interpolation of cap and arms.** In the (u,v) coordinates, according to (2.4), the Yin-Yang arms are given by

$$u = U_\pm(t,v) \text{ for } |v| \leqslant 3K\ln\tau$$

where we abbreviate $U_\pm(t,v) = U_{\pm\pi/2}(t,v)$. In the same (u,v) coordinates the ends of the cap are given by (2.16) with, $B = -1$, i.e.

$$u = h_\pm(t,v) \text{ for } \tfrac{1}{2}K\ln\tau \leqslant -v \leqslant 2K\ln\tau.$$

We constructed the cap so that both $U_\pm$ and $h_\pm$ have the same asymptotic behavior, namely,

$$(2.22) \qquad U_\pm(t,v), h_\pm(t,v) = \pm\frac{\pi}{2} - \frac{1}{\tau}\left\{\tfrac{1}{2}v^2 - v + \pi^2/8\right\} + \mathcal{O}(\tau^{-2+\delta}).$$

Choose a smooth nondecreasing function $\eta : \mathbb{R} \to \mathbb{R}$ with $\eta(u) = 0$ for $u \leqslant \frac{1}{2}$ and $\eta(u) = 1$ for $u \geqslant 2$, and define

$$k_\pm(t,v) = \eta\left(\frac{-v}{\ln\tau}\right)U_\pm(t,v) + \left\{1 - \eta\left(\frac{-v}{\ln\tau}\right)\right\}h_\pm(t,v)$$

$$(2.23) \qquad\qquad = U_\pm(t,v) + \eta\left(\frac{-v}{\ln\tau}\right)\left(U_\pm(t,v) - h_\pm(t,v)\right).$$

The graphs of these two functions are Z-coordinate representations of curve segments that smoothly interpolate between the two ends of the cap and the two Yin-Yang arms. The two segments are parametrized by

$$X_\pm(t,v) = e^{-tJ}\left\{R\boldsymbol{e}_1 + R^{-1}Z_\pm(t,v)\right\} \text{ with } Z_\pm(t,v) = k_\pm(t,v)\boldsymbol{e}_1 + v\boldsymbol{e}_2.$$

It follows from (2.10) that the Curve Shortening Deficit for such curves is given by $(\kappa - V)ds = W_v\,dv$ with

$$(2.24) \qquad W_v[k] = -R^{-2}k_t + \frac{k_{vv}}{1 + k_v^2} - k_v + 2\frac{R_\theta}{R} + 2\frac{R_\theta}{R^3}(vk_v - k) - R^{-2}(v + kk_v).$$

2.9. **Derivative bounds for $U_\pm$, $h_\pm$, $k_\pm$.** Careful scrutiny of the construction of $U_\pm(t,v)$ and $h_\pm(t,v)$ shows that the remainder terms $\mathcal{O}(\tau^{-2+\delta})$ in (2.22) may be differentiated. This implies that the functions $U_\pm$ and $h_\pm$ satisfy

$$(2.25) \qquad |k_{vv}| + |k_v| + |k_t| \lesssim \frac{\ln\tau}{\tau} \lesssim \tau^{-1+\delta}$$

for $\frac{1}{2}\ln\tau \leqslant -v \leqslant 2K\ln\tau$, and large enough τ.

The derivatives of the gluing function $\eta(v/\ln\tau)$ are

$$\eta_v = -\frac{\eta'(-v/\ln\tau)}{\ln\tau}, \quad \eta_{vv} = \frac{\eta''(-v/\ln\tau)}{(\ln\tau)^2}, \quad \eta_t = -\frac{4v\eta'(-v/\ln\tau)}{(\ln\tau)^2},$$

so they are bounded by

$$|\eta_v| + |\eta_{vv}| + |\eta_t| \lesssim \left(\ln\tau\right)^{-1}$$

for $\frac{1}{2}\ln\tau \leqslant -v \leqslant 2K\ln\tau$, and large enough τ.

It follows that the interpolating functions $k_\pm = \eta U_\pm + (1-\eta)h_\pm$ also satisfy (2.25).

2.10. **Estimating** $W_v[k_\pm]$**.** We show that

$$(2.26) \qquad \left|W_v[k_\pm]\right| \lesssim \tau^{-2+\delta}$$

holds in the region $\frac{1}{2}K\ln\tau \le -v \le 2K\ln\tau$, for sufficiently large τ.

If k is any of the functions $U_\pm, h_\pm, k_\pm$ then we have

$$\frac{|k_v^2 k_{vv}|}{1+k_v^2} \le |k_v^2 k_{vv}| \lesssim \tau^{-3+\delta}, \quad R^{-2}|kk_v| \lesssim \tau^{-3+\delta}, \text{ and } -R^{-2}|k_t| \lesssim \tau^{-2+\delta}.$$

Furthermore $|k| \lesssim 1$, $|vk_v| \lesssim \tau^{-1+\delta}$, and $R_\theta \sim R^{-1} \lesssim \tau^{-1/2}$ lead to

$$\frac{R_\theta}{R^3}|vk_v - k| \lesssim \tau^{-2}.$$

Hence

$$W_v[k] = k_{vv} - k_v + 2\frac{R_\theta}{R} - vR^{-2} + \mathcal{O}(\tau^{-2+\delta})$$

holds for all six functions $k \in \{U_\pm, h_\pm, k_\pm\}$.

To simplify our notation we drop the $\pm$ subscript for now and expand the derivatives of $k = \eta U + (1-\eta)h$ (with $U = U_\pm$ and $h = h_\pm$),

$$k_v = \eta U_v + (1-\eta)h_v + \eta_v(U - h),$$
$$k_{vv} = \eta U_{vv} + (1-\eta)h_{vv} + 2\eta_v(U - h)_v + \eta_{vv}(U - h).$$

Since we have matched the two cap ends with the Yin-Yang arms, it follows from (2.22) that the difference $U - h$ and its derivatives are bounded by

$$|U - h| + |(U - h)_v| + |(U - h)_{vv}| \lesssim \tau^{-2+\delta}.$$

Together these inequalities give us the desired estimate for $W_v[k_\pm]$, namely

$$W_v[k] = W_v[\eta U + (1-\eta)h]$$

$$= \left(\eta U + (1-\eta)h\right)_{vv} - \left(\eta U + (1-\eta)h\right)_v + 2\frac{R_\theta}{R} - vR^{-2} + \mathcal{O}(\tau^{-2+\delta})$$

$$= \eta W_v[U] + (1-\eta)W_v[h] + \mathcal{O}(\tau^{-2+\delta})$$

$$(2.27) \qquad = \mathcal{O}(\tau^{-2+\delta}).$$

Definition 2.1 (The Approximate Solution)**.** Let $\mathscr{C}_*(t) : (-\infty, -T) \times \mathbb{R} \to \mathbb{R}^2$ for some sufficiently large $T > 0$ be the family of smooth curves formed by the concatenation of the Yin-Yang leaves $Y(\theta, t, -\frac{\pi}{2})$, $Y(\theta, t, +\frac{\pi}{2})$ cut off in a neighborhood of $R(t)\boldsymbol{E}_1(-t)$ and glued to the cap $X(t,p)$ defined by the ansatzes (2.9) and (2.11), with $f(t,p) = \tau^{-1}F(p)$ and $F(p)$ given by (2.15) with $A = 0$ and $B = -1$. The gluing between the arms of the cap and the two Yin-Yang segments is given by the interpolation $k_\pm(t,v)$ in (2.23), which is done in a neighborhood of $R(t)\boldsymbol{E}_1(-t)$.

Lemma 2.1. The error

$$\mathscr{E}(T) = \int_{-\infty}^{-T} \int_{\mathscr{C}_*(t)} |V - \kappa|\, ds\, dt < \infty$$

is finite on the approximate solution $\mathscr{C}_*(t) : (-\infty, -T) \times \mathbb{R} \to \mathbb{R}^2$ given by Definition 2.1.

Proof. It suffices to show that the Curve Shortening Deficit $|V - \kappa|ds$ is L^1-integrable in space and time on three regions: the cap, the transition region discussed in §2.8, and the unmodified Yin-Yang curve. Since the Yin-Yang curve is a solution to Curve

Shortening Flow, the deficit $|V - \kappa|ds = 0$, which leaves only cap and the transition region as contributing to the error.

The cap, given by the expressions (2.19), (2.20), and (2.21), is defined on the region $|p| \leq 2K \ln \tau$, where p is the arc-length coordinate p for the grim reaper $G(p)$. In §2.6-§2.7, the Curve Shortening Deficit $|V - \kappa|ds$ is written in terms of p as $W dp$, and it is shown to be $W = \mathcal{O}(\tau^{-2+\delta})$ for $\delta > 0$. Integrating over the cap, we have

$$\int_{\text{cap}} |V - \kappa|ds \leq \int_{-2K \ln \tau}^{2K \ln \tau} W dp \leq 4K \ln \tau \, \mathcal{O}(\tau^{-2+\delta}) = \mathcal{O}(\tau^{-2+\delta}).$$

This quantity is integrable in time, and thus the contribution to the error $\mathcal{E}$ on the cap is bounded.

In §2.8, the Curve Shortening Deficit on the transition region is written in terms of the parameter v as $W_v[k]dv$ on the interval $\frac{1}{2} \leq -v \leq 2K \ln \tau$. Furthermore, in (2.27) it is shown that $W_v[k] = \mathcal{O}(\tau^{-2+\delta})$. Integrating over both curves in the transition region, we have

$$\int_{\text{trns. reg.}} |V - \kappa|ds = 2 \int_{\frac{1}{2}K \ln \tau}^{2K \ln \tau} W_v[k]dv \leq 3K \ln \tau \, \mathcal{O}(\tau^{-2+\delta}) = \mathcal{O}(\tau^{-2+\delta}).$$

As before, this is integrable in time and thus the contribution to the error is bounded. Therefore, the sum of the integrals of the deficit over each region is $\mathcal{O}(\tau^{-2+\delta})$. This completes the proof of the lemma. □

3. Area decreasing property of Space Curve Shortening

In 1991 Altschuler and Grayson [2] observed that for two solutions of space curve shortening the area of the minimal surface spanning them is non increasing. Here we elaborate on this and prove a similar result without using the existence of the minimal surface.

3.1. **Moving space curves.** For an immersed curve $X : \mathbb{R} \to \mathbb{R}^n$ one defines the arc length one-form ds and the arc length derivative ∂_s of any quantity $f : \mathbb{R} \to \mathbb{R}$ by

$$ds = \|X_p\| dp, \text{ and } \frac{\partial f}{\partial s} = \frac{1}{\|X_p\|} \frac{\partial f}{\partial p}.$$

The unit tangent and curvature of the curve are $X_s = \frac{X_p}{\|X_p\|}$ and X_{ss}.

A moving family of space curves is a map $X : (t_0, t_1) \times \mathbb{R} \to \mathbb{R}^n$. The family evolves by Curve Shortening if it satisfies $X_t^{\perp} = X_{ss}$, i.e., if for some smooth function $\lambda(t, p)$ one has

$$(3.1) \qquad X_t = X_{ss} + \lambda X_s = \frac{1}{\|X_p\|} \frac{\partial}{\partial p} \left(\frac{X_p}{\|X_p\|} \right) + \lambda \frac{X_p}{\|X_p\|}$$

Since $X_s \perp X_{ss}$, one can always find λ from $\lambda = \langle X_t, X_s \rangle$.

3.2. **Evolution of arc length and the commutator $[\partial_t, \partial_s]$.** The following are commonly used relations. We record them here for completeness, and also because we allow the velocity X_t of the parametrizations to have a nonvanishing tangential component. Assuming that $X_t = X_{ss} + \lambda X_s$ one has

$$(3.2) \qquad \partial_t \|X_p\| = (\lambda_s - \kappa^2)\|X_p\|, \qquad \frac{\partial}{\partial t} ds = (\lambda_s - \kappa^2)ds = d\lambda - \kappa^2 ds,$$

and

$$(3.3) \qquad [\partial_t, \partial_s] = (-\lambda_s + \kappa^2)\partial_s.$$

Proof. We have

$$\frac{\partial}{\partial t}\|X_p\| = \left\langle \frac{X_p}{\|X_p\|}, X_{tp} \right\rangle = \langle X_s, X_{tp}\rangle = \langle X_s, X_{ts}\rangle\|X_p\|,$$

and hence

$$\frac{\partial}{\partial t}ds = \frac{\partial}{\partial t}\big(\|X_p\|dp\big) = \langle X_s, X_{tp}\rangle dp = \langle X_s, X_{ts}\rangle ds.$$

The evolution equation $X_t = X_{ss} + \lambda X_s$ then implies

$$\langle X_s, X_{ts}\rangle = \langle X_s, X_t\rangle_s - \langle X_{ss}, X_t\rangle = \lambda_s - \langle X_{ss}, X_{ss} + \lambda X_s\rangle = \lambda_s - \|X_{ss}\|^2 = \lambda_s - \kappa^2,$$

which directly implies (3.2). Using (3.2) we get

$$[\partial_t, \partial_s] = \big[\partial_t, \|X_p\|^{-1}\partial_p\big] = -\frac{\partial_t\|X_p\|}{\|X_p\|^2}\partial_p = -(\lambda_s - \kappa^2)\partial_s. \qquad \square$$

3.3. **Dependence on a parameter.** Let $[\epsilon_0, \epsilon_1] \subset \mathbb{R}$ be some parameter interval, and let $X : [\epsilon_0, \epsilon_1] \times (t_0, t_1) \times \mathbb{R} \to \mathbb{R}^n$ be a family of moving curves that depends on a parameter $\epsilon \in [\epsilon_0, \epsilon_1]$. We compute the evolution of the first variation

$$X_\epsilon = \partial_\epsilon X(\epsilon, t, p).$$

Throughout the computation we will assume that the parametrization X is such that

(3.4) $$X_\epsilon \perp X_p(\epsilon, t, p) \text{ for all } (\epsilon, t, p)$$

For any given parametrization $\widetilde{X}$ one can find a reparametrization $\varphi(\epsilon, t, p)$ so that $X(\epsilon, t, p) = \widetilde{X}(\epsilon, t, \varphi(\epsilon, t, p))$ satisfies (3.4).

If $X : [\epsilon_0, \epsilon_1] \times [0, 1] \to \mathbb{R}^n$ is injective with $X_\epsilon \perp X_p$, then the double integral

$$\int_{\epsilon_0}^{\epsilon_1} \int_{p=0}^{1} \|X_\epsilon\|\,ds\,d\epsilon \qquad \text{(where } ds = \|X_p\|dp)$$

is the area (2-dim Hausdorff measure) of the image $X([\epsilon_0, \epsilon_1] \times [0, 1])$. If X merely satisfies $X_\epsilon \perp X_p$, without necessarily being injective, then the area formula implies that the double integral is bounded from below, by

(3.5) $$\ell(X) \stackrel{\text{def}}{=} \int_{\epsilon_0}^{\epsilon_1} \int_{p=0}^{1} \|X_\epsilon\|\,ds\,d\epsilon \geqslant \mathcal{H}^2(X([\epsilon_0, \epsilon_1] \times [0, 1]))$$

We will call the integral $\ell(X)$ the *length of the homotopy* X, and we will show that Curve Shortening decreases the length of homotopies.

The following improvement of the inequality $\|X_\epsilon\|_s \leqslant \|X_{\epsilon s}\|$ (which follows from the Cauchy-Schwarz inequality) will be useful.

3.4. **Lemma.** *Assuming* (3.4) *we have*

$$\big(\|X_\epsilon\|_s\big)^2 \leqslant \|X_{\epsilon s}\|^2 - \langle X_{ss}, X_\epsilon\rangle^2.$$

Proof. Split $X_{\epsilon s}$ into tangential and orthogonal components:

$$X_{\epsilon s} = P + \langle X_{\epsilon s}, X_s\rangle X_s$$

Since $X_\epsilon \perp X_s$ we have $\langle X_{\epsilon s}, X_s\rangle = \partial_s\langle X_\epsilon, X_s\rangle - \langle X_\epsilon, X_{ss}\rangle = -\langle X_\epsilon, X_{ss}\rangle$. Therefore

$$X_{\epsilon s} = P - \langle X_\epsilon, X_{ss}\rangle X_s,$$

and thus

$$\|X_{\epsilon s}\|^2 = \|P\|^2 + \langle X_\epsilon, X_{ss}\rangle^2.$$

On the other hand

$$\|X_\epsilon\|_s = \left\langle \frac{X_\epsilon}{\|X_\epsilon\|}, X_{\epsilon s} \right\rangle = \left\langle \frac{X_\epsilon}{\|X_\epsilon\|}, P \right\rangle \leq \|P\|,$$

where we have again used that $X_s \perp X_\epsilon$. Combining these observations we arrive at

$$\left(\|X_\epsilon\|_s\right)^2 - \|X_{\epsilon s}\|^2 \leq \|P\|^2 - \|X_{\epsilon s}\|^2 = -\langle X_\epsilon, X_{ss}\rangle^2,$$

as claimed. $\qquad\qquad\square$

3.5. **The commutator** $[\partial_\epsilon, \partial_s]$. *Assuming* (3.4) *one has*

$$[\partial_\epsilon, \partial_s] = \langle X_{ss}, X_\epsilon\rangle \partial_s \text{ and } [\partial_\epsilon, \partial_s^2] = 2\langle X_{ss}, X_\epsilon\rangle \partial_s^2 + \langle X_{ss}, X_\epsilon\rangle_s \partial_s$$

Proof. The computation follows the same pattern as the derivation of (3.3). Here we have no equation for X_ϵ, but we do know that $X_\epsilon \perp X_s$. Thus

$$\partial_\epsilon \|X_p\| = \left\langle \frac{X_p}{\|X_p\|}, X_{p\epsilon} \right\rangle = \langle X_s, X_{\epsilon p}\rangle = \langle X_s, X_{\epsilon s}\rangle \|X_p\| = -\langle X_{ss}, X_\epsilon\rangle \|X_p\|.$$

Apply this to $\partial_s = \|X_p\|^{-1}\partial_p$ to get the commutator $[\partial_\epsilon, \partial_s]$. The other commutator follows from expanding $[\partial_\epsilon, \partial_s^2] = [\partial_\epsilon, \partial_s]\partial_s + \partial_s[\partial_\epsilon, \partial_s]$. $\qquad\square$

3.6. **Lemma.** *The length of the first variation X_ϵ satisfies the differential inequality*

$$\partial_t \|X_\epsilon\| - \lambda \partial_s \|X_\epsilon\| \leq \partial_s^2 \|X_\epsilon\| + \kappa^2 \|X_\epsilon\|.$$

Proof. Differentiating the evolution equation (3.1) for X we get

$$\begin{aligned}
\partial_t X_\epsilon = \partial_\epsilon X_t &= \partial_\epsilon(X_{ss} + \lambda X_s) \\
&= X_{\epsilon ss} + 2\langle X_{ss}, X_\epsilon\rangle X_{ss} + \langle X_{ss}, X_\epsilon\rangle_s X_s + \lambda_\epsilon X_s + \lambda \partial_\epsilon X_s.
\end{aligned}$$

Hence

$$\begin{aligned}
\partial_t X_\epsilon - \lambda \partial_s X_\epsilon &= X_{\epsilon ss} + 2\langle X_{ss}, X_\epsilon\rangle X_{ss} + \langle X_{ss}, X_\epsilon\rangle_s X_s + \lambda_\epsilon X_s + \lambda[\partial_\epsilon, \partial_s]X \\
&= X_{\epsilon ss} + 2\langle X_{ss}, X_\epsilon\rangle X_{ss} + \{\langle X_{ss}, X_\epsilon\rangle_s + \lambda_\epsilon + \lambda\langle X_{ss}, X_\epsilon\rangle\} X_s
\end{aligned}$$

We next compute the evolution of $\|X_\epsilon\|^2$, keeping in mind that $X_\epsilon \perp X_s$:

$$\begin{aligned}
(\partial_t - \lambda\partial_s)\|X_\epsilon\|^2 &= 2\langle X_\epsilon, \partial_t X_\epsilon - \lambda\partial_s X_\epsilon\rangle \\
&= 2\langle X_\epsilon, X_{\epsilon ss}\rangle + 4\langle X_{ss}, X_\epsilon\rangle^2 \\
&= \left(\|X_\epsilon\|^2\right)_{ss} - 2\|X_{\epsilon s}\|^2 + 4\langle X_{ss}, X_\epsilon\rangle^2 \\
&= 2\|X_\epsilon\| \|X_\epsilon\|_{ss} + 2\left(\|X_\epsilon\|_s\right)^2 - 2\|X_{\epsilon s}\|^2 + 4\langle X_{ss}, X_\epsilon\rangle^2.
\end{aligned}$$

At this point we use Lemma 3.4, to get

$$(\partial_t - \lambda\partial_s)\|X_\epsilon\|^2 \leq 2\|X_\epsilon\| \|X_\epsilon\|_{ss} + 2\langle X_{ss}, X_\epsilon\rangle^2.$$

Since $(\partial_t - \lambda\partial_s)\|X_\epsilon\|^2 = 2\|X_\epsilon\|(\partial_t - \lambda\partial_s)\|X_\epsilon\|$, we have

$$\partial_t \|X_\epsilon\| - \lambda\partial_s \|X_\epsilon\| \leq \|X_\epsilon\|_{ss} + \left\langle X_{ss}, \frac{X_\epsilon}{\|X_\epsilon\|}\right\rangle^2 \|X_\epsilon\| \leq \|X_\epsilon\|_{ss} + \kappa^2 \|X_\epsilon\|. \qquad\square$$

3.7. **Contractive property of Curve Shortening.** If $X^0, X^1 : [t_0, t_1] \times [0,1] \to \mathbb{R}^n$ are two solutions of Curve Shortening (3.1), then a homotopy $\{X^\epsilon : 0 \leq \epsilon \leq 1\}$ of solutions to Curve Shortening connecting them is, by definition, a map $X : [0,1] \times [t_0, t_1] \times [0,1] \to \mathbb{R}^n$ such that $(t,p) \mapsto X(\epsilon, t, p)$ is a solution of Curve Shortening with $X^\epsilon(t,p) = X(\epsilon, t, p)$ for $\epsilon \in \{0, 1\}$.

Given any homotopy X^ϵ between solutions X^0, X^1 of Curve Shortening, one can always find a reparametrization $\widetilde{X}^\epsilon(t,p) = X^\epsilon(t, \varphi(\epsilon, t, p))$ for which $\partial_\epsilon \widetilde{X}^\epsilon \perp \partial_s \widetilde{X}^\epsilon$ holds pointwise. We will call such a homotopy a *normal homotopy*.

Our main observation in this section is the following: *if X^ϵ is a normal homotopy between solutions X^0, X^1 of Curve Shortening, then one has for each $\epsilon \in [0,1]$ and $t \in (t_0, t_1)$*

$$(3.6) \qquad \frac{d}{dt} \int_{p=0}^1 \|\partial_\epsilon X(\epsilon, t, p)\| \, ds \leq 0$$

and

$$(3.7) \qquad \int_{\epsilon=0}^1 \int_{p=0}^1 \|\partial_\epsilon X(\epsilon, t_1, p)\| \, ds \, d\epsilon \leq \int_{\epsilon=0}^1 \int_{p=0}^1 \|\partial_\epsilon X(\epsilon, t_0, p)\| \, ds \, d\epsilon$$

Proof. We use (3.2) and Lemma 3.6 to differentiate under the integral:

$$\frac{d}{dt} \int_{p=0}^1 \|X_\epsilon\| \, ds \leq \int \left\{ \|X_\epsilon\|_{ss} + \kappa^2 \|X_\epsilon\| + \lambda \|X_\epsilon\|_s + (\lambda_s - \kappa^2)\|X_\epsilon\| \right\} ds$$

$$= \int \left\{ \|X_\epsilon\|_s + \lambda \|X_\epsilon\| \right\}_s ds = 0.$$

This implies (3.6). Integration in ϵ and in time then leads to (3.7). $\qquad \square$

3.8. **Deviation from an approximate solution.** We now consider the case of two moving curves $X^0, X^1 : [t_0, t_1] \times [0,1] \to \mathbb{R}^n$ with the same initial value, i.e. with $X^0(t_0, p) = X^1(t_0, p)$ for all $p \in [0,1]$. We assume that X^1 is a solution of Curve Shortening but allow X^0 to be a general moving curve. We measure its deviation from Curve Shortening in terms of

$$(3.8) \qquad \Delta \overset{\text{def}}{=} \int_{t_0}^{t_1} \int_{p=0}^1 \left\| (\partial_t X^0)^\perp - \partial_s^2 X^0 \right\| ds \, dt.$$

In the case of plane curves $X : [t_0, t_1] \times [0,1] \to \mathbb{R}^2$, we have $\partial_t(X^0)^\perp = VJX_s^0$, where V is the normal velocity of the curve X^0. Therefore the integrand in (3.8) is

$$\left\| (\partial_t X^0)^\perp - \partial_s^2 X^0 \right\| ds = |V - \kappa| \, ds.$$

The quantity Δ therefore coincides with the "error" $\mathscr{E}$ defined in (1.4).

Assume that for each $\epsilon \in [t_0, t_1]$ there is a smooth solution $(t,p) \mapsto X(\epsilon, t, p)$ of Curve Shortening that is defined for $t \in [\epsilon, t_1]$, and that has initial value $X(\epsilon, \epsilon, p) = X^0(\epsilon, p)$. After reparametrizing we may assume that $X_\epsilon \perp X_p$ holds point-wise. Then the final values of these solutions, i.e. the curves $p \mapsto X(\epsilon, t_1, p)$ form a normal homotopy from $X^0(t_1, \cdot)$ to $X^1(t_1, \cdot)$. We will now show that

$$(3.9) \qquad \int_{\epsilon=t_0}^{t_1} \int_{p=0}^1 \|\partial_\epsilon X(\epsilon, t_1, p)\| \, ds \, d\epsilon \leq \Delta.$$

Proof. Our argument is a nonlinear version of the Variation of Constants Formula, or of Duhamel's principle.

For $\epsilon \in [t_0, t_1]$ we consider

$$E(\epsilon) \overset{\text{def}}{=} \int_{\bar{\epsilon}=t_0}^{\epsilon} \int_0^1 \|\partial_\epsilon X(\bar{\epsilon}, t_1, p)\| \, ds \, d\bar{\epsilon}.$$

Then

$$E'(\epsilon) = \int_0^1 \|\partial_\epsilon X(\epsilon, t_1, p)\| \, ds$$

The contraction property (3.6) implies

$$\int_0^1 \|\partial_\epsilon X(\epsilon, t_1, p)\| \, ds \leqslant \int_0^1 \|\partial_\epsilon X(\epsilon, \epsilon, p)\| \, ds.$$

We compute $X_\epsilon(\epsilon, \epsilon, p)$ by differentiating the relation $X(\epsilon, \epsilon, p) = X^0(\epsilon, p)$ with respect to ϵ:

$$\left(\partial_t X^0\right)(\epsilon, p) = \frac{\partial X^0(\epsilon, p)}{\partial \epsilon} = \frac{\partial X(\epsilon, \epsilon, p)}{\partial \epsilon} = X_\epsilon(\epsilon, \epsilon, p) + X_t(\epsilon, \epsilon, p).$$

Since $(t, p) \mapsto X(\epsilon, t, p)$ evolves by Curve Shortening, we have $X_t(\epsilon, \epsilon, p) = X_{ss}(\epsilon, \epsilon, p) + \lambda X_s$. By definition of X we have $X(\epsilon, \epsilon, p) = X^0(\epsilon, p)$, so $X_{ss}(\epsilon, \epsilon, p) = X_{ss}^0(\epsilon, p)$.

We have parameterized the homotopy $X(\epsilon, t, p)$ so that $X_\epsilon \perp X_p$, and therefore

$$\partial_\epsilon X(\epsilon, \epsilon, p) = \left(\partial_t X^0(\epsilon, p)\right)^{\perp} - X_{ss}^0(\epsilon, p).$$

Hence

$$E'(\epsilon) \leqslant \int_0^1 \|\partial_\epsilon X(\epsilon, \epsilon, p)\| \, ds = \int_0^1 \left\|\left(\partial_t X^0(\epsilon, p)\right)^{\perp} - X_{ss}^0(\epsilon, p)\right\| \, ds$$

Integrate over $\epsilon \in [t_0, t_1]$ to recover (3.9). $\qquad\qquad\qquad\qquad\qquad\square$

3.9. **Application to plane Curve Shortening.** Let $X^0, X^1 : [t_0, t_1] \times [0, 1] \to \mathbb{R}^2$ be two moving curves that are embedded at all time. Assume X^1 evolves by Curve Shortening, and assume that initially X^1 lies in the closed region enclosed by X^0, i.e. for all $p \in [0, 1]$ the point $X^1(t_0, p)$ lies in the region enclosed by the simple curve $p \mapsto X^0(t_0, p)$.

Assume furthermore that at each time $t_* \in [t_0, t_1]$ the area enclosed by $p \mapsto X^0(t_*, p)$ is at least $2\pi(t_1 - t_*)$. By the Gage-Hamilton-Grayson theorem this guarantees that the solution to Curve Shortening starting at $X^0(t_*, \cdot)$ exists until time t_1.

We now consider two homotopies A and B of evolving curves. The first is the homotopy defined in the proof in the previous section 3.8, i.e. for each $\epsilon \in [t_0, t_1]$ we consider the solution $(t, p) \mapsto X(\epsilon, t, p)$ to Curve Shortening defined for $t \in [\epsilon, t_1]$ and starting from $X(\epsilon, \epsilon, p) = X^0(\epsilon, p)$. Our first homotopy is then the family of final curves $A(\epsilon, p) = X(\epsilon, t_1, p)$ of these solutions. In section 3.8 we showed that the length of the homotopy A is bounded from above by

$$\ell(A) = \int_{t_0}^{t_1} \int_0^1 \|A_\epsilon(\epsilon, p)\| \, ds \, d\epsilon \leqslant \int_{t_0}^{t_1} \int_0^1 \left\|(\partial_t X^0)^{\perp} - X_{ss}^0\right\| \, ds \, dt.$$

The second homotopy is constructed by evolving a homotopy between the two initial curves $p \mapsto X^j(t_0, p)$ ($j = 0, 1$). Since X^1 initially lies inside X^0 we can choose the homotopy $(\epsilon, p) \mapsto \bar{X}(\epsilon, t_0, p)$ so that its length is exactly the area of the region between the two initial curves, and so that the curve $p \mapsto \bar{X}(\epsilon, t_0, p)$ lies inside the curve $p \mapsto \bar{X}(\epsilon', t_0, p)$ if $0 \leqslant \epsilon \leqslant \epsilon' \leqslant 1$. Given this initial homotopy let $(t, p) \mapsto \bar{X}(\epsilon, t, p)$ be the solution to Curve Shortening starting at $\bar{X}(\epsilon, t_0, p)$. Since all initial curves enclose $X^1(t_0, p)$ the corresponding solutions exist for $t \in [t_0, t_1]$, and possibly longer. Our

second homotopy is now $B(\epsilon, p) = \bar{X}(\epsilon, t_1, p)$. As explained in section 3.7, the length of the homotopy B is bounded by the length of the initial homotopy $(\epsilon, p) \mapsto \bar{X}(\epsilon, t_0, p)$. We had chosen this initial homotopy so that its length is exactly the area between the two curves $p \mapsto X^0(t_0, p)$ and $p \mapsto X^1(t_0, p)$.

By concatenating the two homotopies A and B we obtain a combined homotopy $A\#B$ between $p \mapsto X^0(t_1, p)$ and $p \mapsto X^1(t_1, p)$. The length of this homotopy is bounded by

$$\ell(A\#B) = \ell(A) + \ell(B) \leqslant \text{Area between } X^0(t_0, \cdot) \text{ and } X^1(t_0, \cdot)$$
$$+ \int_{t_0}^{t_1} \int_0^1 \left\| (\partial_t X^0)^\perp - X^0_{ss} \right\| ds\, dt$$

If $\Omega^0(t)$ and $\Omega^1(t)$ are the regions enclosed by $X^0(t, \cdot)$ and $X^1(t, \cdot)$, then the homotopy $A\#B$ between the two curves at time t_1 must pass through each point in the interior of the symmetric difference $\Omega^0(t_1) \triangle \Omega^1(t_1)$, as one sees by considering the winding numbers of the curves in the homotopy around any point in $\Omega^0(t_1) \triangle \Omega^1(t_1)$. It follows that the area of $\Omega^0(t_1) \triangle \Omega^1(t_1)$ is a lower bound for the length of the homotopy $A\#B$, and thus we conclude that

$$(3.10) \quad \text{Area of } \Omega^0(t_1) \triangle \Omega^1(t_1) \leqslant$$
$$\text{Area between } X^0(t_0, \cdot) \text{ and } X^1(t_0, \cdot) + \int_{t_0}^{t_1} \int_0^1 \left\| (\partial_t X^0)^\perp - X^0_{ss} \right\| ds\, dt.$$

4. Convergence

In this section, we obtain uniform curvature bounds on a sequence of "really old solutions" $\{\mathscr{C}_j(t)\}$ and extract a subsequence of solutions that converges locally smoothly to an ancient solution of curve shortening flow (CSF).

Theorem 4.1. There exists a T such that for any $T^* > T$, the curvatures $|\kappa_j|$ of the "really old solutions" $\{\mathscr{C}_j(t)\}_j$ are bounded independently of j on the interval $[-T^* - 1/4, -T^*]$.

The strategy to obtain these bounds is as follows: a) decompose an element $\mathscr{C}_j(t)$ of this sequence into the union of several graph representations, b) use the L^1 bound on the error to obtain L^∞ estimates for these graphs, and c) apply the standard estimates for divergence-form quasilinear parabolic equations to establish a uniform curvature bound.

Let $T^* > 0$ be a large positive number, which may be increased as necessary throughout this section. The obvious candidates for the sequence of "really old solutions" are the CSF solutions defined on $[-j, -T^*]$ starting at $\mathscr{C}_*(-j)$ at time $-j$ — call these $\Gamma_j(t)$. At any time $t \in [-j, -T^*]$, the unsigned area enclosed by the curves $\mathscr{C}_*(t)$ and $\Gamma_j(t)$ is bounded by the quantity

$$\int_{-j}^t \int_{\mathscr{C}_*(\tau)} |V - \kappa| ds\, d\tau.$$

By Lemma 2.1, this quantity is in L^1 and given $\epsilon > 0$, we can find $T^* > 0$ such that

$$\mathscr{E}(T^*) = \int_{-\infty}^{-T^*} \int_{\mathscr{C}_*(\tau)} |V - \kappa| ds\, d\tau < \epsilon.$$

By (3.10), this estimate gives a uniform bound on the unsigned area between $\Gamma_j(t)$ and $\mathscr{C}_*(t)$ for any sufficiently large j and $t < -T^*$, whenever defined.

For simplicity of calculation, we consider an alternative sequence of "really old solutions" $\{\mathscr{C}_j(t)\}$, called "square-profile approximations", which do **not** satisfy the initial condition $\mathscr{C}_j(-j) = \mathscr{C}_*(-j)$. They are constructed as follows: Let $[\vartheta_*^-(t), \vartheta_*^+(t)]$ be the interval of polar angles that contain the approximate solution at a given time t. Let $r = \mathscr{R}(\theta)$ be the central branch of the Yin-Yang foliation, so the Yin-Yang solution is given by the two branches $r = \mathscr{R}(\theta + \pi/2 - t)$ and $r = \mathscr{R}(\theta - \pi/2 - t)$. We define $\mathscr{C}_j(t)$ to be the solution of curve shortening which at time $t = -j$ is given by

- Two arms of the Yin-Yang soliton $r = \mathscr{R}(\theta - \pi/2 - t) = \mathscr{R}(\theta - \pi/2 + j)$, and $r = \mathscr{R}(\theta + \pi/2 - t) = \mathscr{R}(\theta + \pi/2 + j)$ truncated at $\theta = \vartheta_*^+(-j)$;
- a straight line segment connecting the two arms of the Yin-Yang soliton. This segment is part of the ray $\theta = \vartheta_*^+(-j)$.

Notice that at each time t along the flow, the square-profile approximations $\mathscr{C}_j(t)$ enclose the curves $\Gamma_j(t)$, the CSF solutions starting from $\mathscr{C}_*(-j)$, and that the area bounded by these two solutions stays constant along the flow, for all $t \in [-j, -T^*]$. The area A_j between $\mathscr{C}_j(-j)$ and $\mathscr{C}_*(-j)$ is small and goes to zero as $j \to \infty$. Thus, the area between the old solution $\mathscr{C}_j(t)$ with the "square initial data" and the approximate solution $\mathscr{C}_*(t)$ is bounded by

$$\text{Area}(\mathscr{C}_j(t), \mathscr{C}_*(t)) \leqslant A_j + \int_{-j}^{t} \int_{\mathscr{C}_*(\tau)} |V - \kappa| \, ds \, d\tau.$$

In order to improve these area bounds to L^∞ bounds, we will use the geometry of the $\mathscr{C}_j(t)$ and several properties of CSF. In particular, we often appeal to the maximum principle and the following Sturmian property for intersections of curve shortening flows.

Theorem 4.2. Consider two CSF solutions $\gamma_0, \gamma_1 : [T_1, T_2) \times [0, 1] \to \mathbb{R}^2$, for which

$$\partial\gamma_0(t) \cap \gamma_1(t) = \partial\gamma_1(t) \cap \gamma_0(t) = \varnothing$$

holds for any $t \in [T_1, T_2)$. Then the number of intersections of $\gamma_0(t)$ and $\gamma_1(t)$ is a finite and non-increasing function of $t \in (T_1, T_2)$. It decreases whenever γ_0 and γ_1 have a tangency.

There is a useful related theorem for inflections points.

Theorem 4.3. Let $\gamma : [T_1, T_2) \times S^1 \to \mathbb{R}^2$ be a solution of CSF. Then, for any $t \in (T_1, T_2)$, $\gamma(t)$ has at most a finite number of inflection points, and this number does not increase with time. In fact, it drops whenever the curvature κ has a multiple zero.

While the curves $\mathscr{C}_j(t)$ are not convex, we do have a one sided curvature bound.

Theorem 4.4. If κ is the curvature of a counterclockwise oriented parametrization X of the curves $\mathscr{C}_j(t)$, then

$$\kappa - \langle X, X_s \rangle > 0.$$

Proof. Assuming that the parametrization X is normal ($X_t \perp X_s$), the curvature evolves by

$$\kappa_t = \kappa_{ss} + \kappa^3.$$

A short computation using $X_t = X_{ss}$ and $\|X_s\| = 1$ shows that

$$\left(\partial_t - \partial_s^2\right)\|X\|^2 = 2\langle X_t, X\rangle - 2\langle X_{ss}, X\rangle - 2\|X_s\|^2 = -2.$$

Differentiating with respect to arclength, using the commutator $[\partial_t, \partial_s] = \kappa^2 \partial_s$, and also $\partial_s \|X\|^2 = 2\langle X, X_s\rangle$ we get

$$\partial_t \langle X, X_s\rangle = \partial_s^2 \langle X, X_s\rangle + \kappa^2 \langle X, X_s\rangle.$$

Hence κ and $\langle X, X_s\rangle$ satisfy the same linear equation. Therefore $\sigma = \kappa - \langle X, X_s\rangle$ also satisfies

$$\sigma_t = \sigma_{ss} + \kappa^2 \sigma.$$

The quantity σ vanishes on the rotating soliton (see the appendix).

The square-profile initial curves $\mathscr{C}_j(-j)$ consist of two arcs. One is the Yin-Yang soliton, so on this arc we have $\sigma = 0$. The other arc is the radial line segment on the ray $\theta = \vartheta_*^+(-j)$. On this segment we clearly have $\kappa = 0$. Since we orient $\mathscr{C}_j$ counterclockwise, X and X_s are parallel with opposite directions; i.e. $-\langle X, X_s\rangle > 0$. Hence $\sigma > 0$ on the line segment. Finally, the initial curve $\mathscr{C}_j(-j)$ is not smooth, having two corners where the line segment and Yin-Yang arms meet. If one rounds these corners off by replacing them with small circle arcs with radius $\rho \ll 1$, then the curvature of these arcs will be $\kappa = \rho^{-1} \gg 1$, so that $\sigma > 0$ on the circular arcs, provided ρ is sufficiently small. The resulting curve has $\sigma = 0$ on the Yin-Yang arms, and $\sigma > 0$ on the line segment, as well as the small circular arcs. The solution to CS starting from the modified initial curve therefore has $\sigma > 0$. Letting $\rho \searrow 0$ we conclude that $\sigma > 0$ also holds on $\mathscr{C}_j(t)$. □

With Theorem 4.2, we can decompose the solutions $\mathscr{C}_j(t)$ into exactly two graphs over the polar angle parameter.

Lemma 4.5. For any $t \in (-j, -T^*]$, there is an interval $[\vartheta_j^-(t), \vartheta_j^+(t)]$ such that the curve $\mathscr{C}_j(t)$ can be written as the union of two graphs of polar functions, $R_j^-(\theta, t)$ and $R_j^+(\theta, t)$ defined for $\theta \in [\vartheta_j^-(t), \vartheta_j^+(t)]$. The functions $t \mapsto \vartheta_j^-(t)$ and $t \mapsto \vartheta_j^+(t)$ are strictly increasing and decreasing, respectively.

Proof. By the maximum principle, the "really old solutions" $\mathscr{C}_j(t)$ will be contained inside of the Yin-Yang curve. The Sturmian property, Theorem 4.2, tells us that the number of intersections of $\mathscr{C}_j(t)$ and the rays $\theta = \theta_0 \in \mathbb{R}$ is non-increasing, and only decreases when there is a tangency. This implies that the desired graph decomposition exists. These two graphs are bounded above and below by the branches of the Yin-Yang soliton on their polar interval of definition, $[\vartheta_j^-(t), \vartheta_j^+(t)]$. □

Similarly, we can always write each $\mathscr{C}_j(t)$ as a union of two graphs taking values in θ, the polar angle. Recall that the images of $\mathscr{R}(\theta - t)$ for $t \in [-\pi, \pi)$ foliate the punctured plane $\mathbb{R}^2 \setminus \{0\}$. See Figure 2.

Lemma 4.6. For all t, $\mathscr{C}_j(t)$ can be decomposed into two graphs of two functions which take leaves of the foliation as inputs and have their range in the set of polar angles. More specifically, for $T \ll 0$ there exist $y_{j,1}, y_{j,2} : (-\infty, T] \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and functions

$$\Theta_j^\pm : \{(t, y) : t < T, y_{j,1}(t) < y < y_{j,2}(t)\} \to \mathbb{R}$$

such that the very old solution $\mathscr{C}_j(t)$ is the union of the two curves

$$Y(\Theta_j^\pm(t, y), t, y) = \mathscr{R}\left(\Theta_j^\pm(t, y) - t + y\right)\boldsymbol{E}_1(\theta),$$

where Y is given by (2.3).

Proof. The initial square-profile curve $\mathscr{C}_j(-j)$ is tangent to the graphs of $\mathscr{R}(\theta \pm \pi/2 + j)$ and intersects the graphs of $\mathscr{R}(\theta + y + j)$, $y \in (-\pi/2, \pi/2)$ twice: once at the origin and once on the line segment connecting the two branches of $\mathscr{C}_j(-j)$. Then, by the Sturmian theorem, for all subsequent $t > -j$, $\mathscr{C}_j(t)$ can be split into two graphs corresponding to the "upper" and "lower" intersection points with the leaves of the Yin-Yang foliation. At each time t, these graphs split at two unique leaves of the foliation, marked by values $y_{j,1}(t), y_{j,2}(t) \in (-\pi/2, \pi/2)$, so that $\mathscr{C}_j(t)$ is tangent to the curves $\{r = \mathscr{R}(\theta + y_{j,1}(t) - t)\}$ and $\{r = \mathscr{R}(\theta + y_{j,2}(t) - t)\}$. We know that these two points are unique since a greater number of tangencies would introduce more than two intersection points for other curves $\{r = \mathscr{R}(\theta + y - t)\}$. We call the coordinate system $(y, \theta) \in (-\pi, \pi) \times (0, \infty)$ the "Yin-Yang polar coordinate system" and denote the two functions giving the upper and lower graphs comprising $\mathscr{C}_j(t)$ by $\Theta_j^-(y, t)$ and $\Theta_j^+(y, t)$ respectively, defined on the interval $(y_{j,1}(t), y_{j,2}(t)) \subset (-\pi/2, \pi/2)$. □

Lemma 4.7. There exist $T < 0$ and $C > 0$ such that $\vartheta_j^+(t) \leqslant \vartheta_*^+(t) + C$ for all $j \in \mathbb{N}$ and all $t \in [-j, -T]$.

Proof. Assume that $\epsilon < \pi/16$.

For any $t \in [-j, T]$ at which $\vartheta_j^+(t) > \vartheta_*^+(t)$ we consider the area $\mathscr{A}_j(t)$ of the "really old solution" $\mathscr{C}_j(t)$ inside the polar interval $[\vartheta_*^+(t), \vartheta_j^+(t)]$, where $\vartheta_*^+(t)$ and $\vartheta_j^+(t)$ are the endpoints of the intervals of definition of the approximate solution $\mathscr{C}_*(t)$ and $\mathscr{C}_j(t)$ respectively. This area measures the "tail" of the $\mathscr{C}_j(t)$ that may form between the tip of $\mathscr{C}_j(t)$ and the tip of $\mathscr{C}_*(t)$. Note that the area $\mathscr{A}_j(t)$ is bounded above by the error

$$\mathscr{A}_j(t) \leqslant A_j + \mathscr{E}(T^*) = A_j + \int_{-\infty}^{-T^*} \int_{\mathscr{C}_*(\tau)} |V - \kappa| \, ds \, d\tau < \epsilon$$

To calculate this area, first consider the function $\Theta_{t,j}(y) := \max\{\Theta_j^+(y, t) - \vartheta_*^+(t), 0\}$ over the interval $(y_{j,1}(t), y_{j,2}(t))$. Then in the (θ, y) "Yin-Yang coordinates," we can integrate to find the area:

$$\mathscr{A}_j(t) = \int_{y_{j,1}(t)}^{y_{j,2}(t)} \int_{\vartheta_*^+}^{\Theta_{t,j}(y)+\vartheta_*^+} \mathscr{R}(\theta + y - t) \det(D\mathscr{T}) d\theta \, dy,$$

where $\mathscr{T} : (0, \infty) \times (-\pi/2, \pi/2) \to (0, \infty) \times (0, \infty)$ is the coordinate transformation given by $\mathscr{T}(\theta, y) = (\theta, \mathscr{R}(\theta + y - t))$. Clearly, $\det \mathscr{T} = \mathscr{R}'(\theta + y - t)$, so

$$\mathscr{A}_j(t) = \int_{y_{j,1}(t)}^{y_{j,2}(t)} \int_{\vartheta_*^+}^{\Theta_{t,j}(y)+\vartheta_*^+} \mathscr{R}(\theta + y - t) \mathscr{R}'(\theta + y - t) d\theta \, dy \approx \int_{y_{j,1}(t)}^{y_{j,2}(t)} \Theta_{t,j}(y) dy,$$

by the asymptotic expansions in (1.1).

We argue that given a small $\delta > 0$, it is possible to pick an angle θ_0 independent of j such that the measure $|\{y : \Theta_{t,j}(y) > \theta_0\}| < \delta$. Indeed, it follows from

$$\theta_0 |\{y : \Theta_{t,j}(y) > \theta_0\}| \leqslant \int_{y_{j,1}(t)}^{y_{j,2}(t)} \Theta_{t,j}(y) dy < \epsilon$$

that if $\theta_0 < \frac{\epsilon}{\delta}$, then $|\{y : \Theta_{t,j}(y) > \theta_0\}| < \delta$ holds for all t, j.

The two points intersection of $\mathscr{C}_j(t)$ with the ray $\theta = \vartheta_*^+(t) + \theta_0$ are

$$P^\pm(t) = R_j^\pm(\vartheta_*(t) + \theta_0, t).$$

Let $\gamma(t)$ be the arc on $\mathscr{C}_j(t)$ on which $\theta \geq \vartheta_*^+(t) + \theta_0$, and whose endpoints therefore are $P^\pm(t)$. Consider the area $A(t)$ of the region enclosed by $\gamma(t)$ and the line segment connecting $P^\pm(t)$. This area changes because the arc $\gamma(t)$ moves, and also because the line segment P^-P^+ moves. The rate of change is therefore the sum of $-\int_{\gamma(t)} \kappa\, ds$ and the rate at which the segment P^-P^+ sweeps out area.
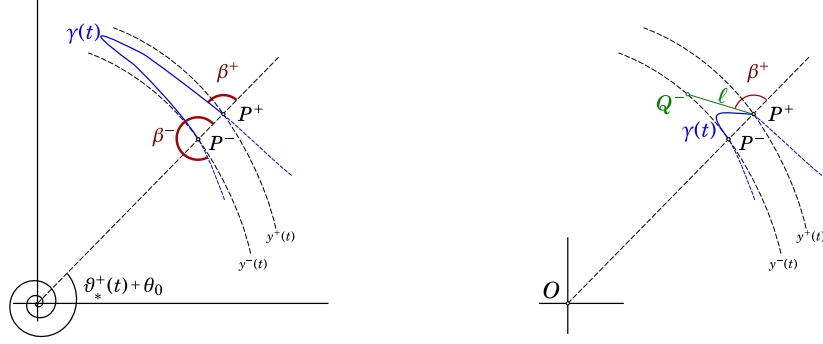
Figure 3. Left: The arc $\gamma(t)$. **Right:** the angles $\beta^\pm$

If $\phi : \gamma(t) \to \mathbb{R}$ is the tangent angle along the arc (i.e. $X_s = \mathbf{E}_1(\phi)$), then the curvature integral is

$$\int_{\gamma(t)} \kappa\, ds = \phi_{P^-(t)} - \phi_{P^+(t)}.$$

The line segment P^+P^- moves with angular velocity $\frac{d}{dt}\vartheta_*^+(t)$ and therefore adds area to the region enclosed by $\gamma(t)$ at the rate

$$\frac{1}{2}\left\{(R_j^+)^2 - (R_j^-)^2\right\}\frac{d\vartheta_*^+(t)}{dt}$$

in which $R_j^\pm$ are evaluated at $\theta = \vartheta_*^+(t) + \theta_0$. Our construction of the cap implies that $\vartheta_*^+(t) = -t + o(1)$, and that this relation may be differentiated: $\frac{d}{dt}\vartheta_*^+(t) = -1 + o(t)$.

The radii $R_j^\pm(\vartheta_*^+(t) + \theta_0, t)$ are given in terms of their Yin-Yang coordinates $y^\pm(t)$ via

$$R_j^\pm(\vartheta_*^+(t) + \theta_0, t) = \mathscr{R}(\vartheta_*^+(t) + \theta_0 - t + y^\pm).$$

It follows that at $\vartheta_*^+(t) + \theta_0$

$$\frac{1}{2}\left\{(R_j^+)^2 - (R_j^-)^2\right\} = \frac{1}{2}\left\{\mathscr{R}(\vartheta_*^+(t) + \theta_0 - t + y^+)^2 - \mathscr{R}(\vartheta_*^+(t) + \theta_0 - t + y^-)^2\right\}$$

$$= \mathscr{R}\mathscr{R}'(y^+ - y^-)$$

in which $\mathscr{R}, \mathscr{R}'$ are evaluated at $\vartheta_*(t) + \theta_0 + \widetilde{y}$ for some $\widetilde{y} \in [y^-, y^+]$ that is provided by the mean value theorem. The asymptotics of $\mathscr{R}$ imply that $\mathscr{R}\mathscr{R}' = 1 + o(1) < 2$. Our choice of θ_0 was such that $0 < y^+ - y^- \leq \delta$. Hence

$$\left|\frac{1}{2}\left\{(R_j^+)^2 - (R_j^-)^2\right\}\frac{d\vartheta_*^+(t)}{dt}\right| \leq 2\delta.$$

In total, the rate at which the area $A(t)$ enclosed by the arc $\gamma(t)$ grows is bounded by

$$\frac{dA}{dt} \leq -\left(\phi_{P^-(t)} - \phi_{P^+(t)}\right) + 2\delta.$$

We estimate the change in tangent angle across the arc $\gamma(t)$. Let β^+ be the counterclockwise angle from the ray $\theta = \vartheta_*^+(t) + \theta_0$ to the tangent X_s to γ at P^+, and similarly, let β^- be the counterclockwise angle from the same ray to the tangent to γ at P^- (see Figure 3). We have $0 < \beta^+ < \pi < \beta^-$ and $\phi_{P^-} - \phi_{P^+} = \beta^- - \beta^+$.

Recall that $\kappa - \langle X, X_s \rangle > 0$ along $\mathscr{C}_j(t)$. Since $\kappa = \phi_s$ and $\langle X, X_s \rangle = \frac{1}{2} \frac{d}{ds} r^2$, where $r = \|X\|$, it follows that $\phi - \frac{1}{2} r^2$ increases as one traverses γ from P^+ to P^-. Thus, at any point with polar coordinates (θ, r) on γ one has

$$\phi > \phi_{P^+} + \frac{1}{2}(r^2 - r_{P^+}^2)$$

The lowest value r has on γ occurs at the point P_-, and we have just shown that $r_{P^+}^2 - r_{P_-}^2 \leqslant 2\delta$. Hence $\phi > \phi_{P^+} - 2\delta$ on the entire arc γ.

It follows that if $\beta^+ > \frac{3}{4}\pi$, then the angle between the tangent to γ and the ray OP^+ (O is the origin) will always be at least $\frac{3}{4}\pi - 2\delta$, i.e. more than $\frac{5}{8}\pi$, provided we choose $\delta < \frac{\pi}{16}$.

Consider the line ℓ through P^+ whose angle with OP^+ is $\frac{5}{8}\pi$. The euclidean distance between P^- and P^+ is $r_{P^+} - r_{P^-} \leqslant 2\delta/(r_{P^-} + r_{P^+}) \leqslant C\delta|t|^{-1/2}$, since $r_{P^+} > r_{P^-} \gtrsim |t|^{1/2}$.

At this scale the Yin-Yang leaves will be almost straight lines near $P^{\pm}$, so that the line ℓ then intersects the Yin-Yang leaf with $y = y_-(t)$ at a point $Q^-(t)$, also at a distance $d(P^-, Q^-) \lesssim \delta|t|^{-1/2}$.

$$\vartheta_j^+(t) - \vartheta_*^+(t) - \theta_0 \sim \frac{d(P^-, Q^-)}{r_{P^-}} \lesssim \delta|t|^{-1} \ll \theta_0.$$

Hence the largest polar angle on γ will be at most

$$\vartheta_j^+(t) \leqslant \vartheta_*^+(t) + \theta_0 + \delta|t|^{-1} \leqslant \vartheta_*^+(t) + 2\theta_0.$$

Thus we find that if $|t|$ is sufficiently large, then either $\vartheta_j(t) < \vartheta_*(t)$, or else $\beta^+ < \frac{5}{8}\pi$. In the latter case the area enclosed by $\gamma(t)$ decreases faster than

$$\frac{dA}{dt} \leqslant -\beta^- + \beta^+ + 2\delta \leqslant -\pi + \frac{5}{8}\pi + 2\delta = -\frac{3}{8}\pi + 2\delta < -\frac{\pi}{4},$$

again assuming that $\delta < \pi/16$.

We now finally prove that $\vartheta_j^+(t) - \vartheta_*^+(t)$ is uniformly bounded for all $t \in [-j, T]$ and j.

At $t = -j$ we have $\vartheta_j^+(t) < \vartheta_*^+(t) + \theta_0$, by definition of the initial curve $\mathscr{C}_j(-j)$. Hence, if at any time $t_1 < T$ one has $\vartheta_j^+(t) > \vartheta_*^+(t) + 2\theta_0$, then there is a largest interval $(t_2, t_3) \ni t_1$ on which $\vartheta_j^+(t) > \vartheta_*^+(t) + 2\theta_0$. In particular, at $t = t_2$ one has $\vartheta_j^+(t) = \vartheta_*^+(t) + 2\theta_0$.

Define the arc $\gamma(t)$ as above. Its enclosed area is at most ε, where we may assume that $\varepsilon < \pi/4$. During the time interval (t_2, t_3) the area decreases at a rate of at least $\pi/4$, and therefore the length $t_3 - t_2$ of the time interval cannot exceed $\varepsilon/(\pi/4) < 1$. At time $t = t_2$ we had $\vartheta_j^+(t) = \vartheta_*^+(t) + 2\theta_0$. Since $\vartheta_j^+(t), \vartheta_*^+(t)$ are nonincreasing functions,

we have throughout (t_2, t_3)

$$\vartheta_j^+(t) - (\vartheta_*^+(t) + 2\theta_0) \leqslant \vartheta_j^+(t_2) - (\vartheta_*^+(t_2) + 2\theta_0)$$
$$= \vartheta_*^+(t_2) + 2\theta_0 - (\vartheta_*^+(t) + 2\theta_0)$$
$$\leqslant \vartheta_*^+(t_2) + 2\theta_0 - (\vartheta_*^+(t_3) + 2\theta_0)$$
$$\leqslant \vartheta_*^+(t_2) - \vartheta_*^+(t_3).$$

Since $\frac{d}{dt}\vartheta_*^+(t) = 1 + o(1)$ we find that

$$\vartheta_j^+(t) - (\vartheta_*^+(t) + 2\theta_0) \leqslant 1 + o(1) < 2$$

for all $t \in (t_2, t_3)$. $\qquad\square$

To summarize, we can now decompose every very-old solution $\mathscr{C}_j(t)$, for $t \in [-T^* - 1/2, -T^*]$ into four graphs in two different coordinate systems, $R_j^\pm(\theta, t)$ in polar coordinates, and $\Theta_j^\pm(y, t)$ in Yin-Yang polar coordinates.

Lemma 4.8 (Curvature bounds). *There exist $T < 0$ such that for any $T' < T$ the lengths $L_j(t)$ and curvatures of $\mathscr{C}_j(t)$ are uniformly bounded for all j and all $t \in [T', T-2]$.*

Proof. The length bounds follow from the fact that in (y, θ) coordinates each $\mathscr{C}_j(t)$ is contained in a uniformly bounded rectangle $|y| \leqslant \pi/2$, $-t \leqslant \theta \leqslant \vartheta_*^+(t) + 2\theta_0 + 2$, and the fact that $\mathscr{C}_j(t)$ decomposes into four segments on each of which both y and θ are monotone.

Consider a given value $T' < T$. Assume that our Lemma fails, and that along some subsequence j the maximal curvature of $\mathscr{C}_j(t)$ with $t \in [T', T-2]$ becomes unbounded.

For $t \in [T'-2, T]$ the lengths $L_j(t)$ of $\mathscr{C}_j(t)$ are uniformly bounded by some $L > 0$. It follows that

$$\int_{T'-2}^{T'-1} \int_{\mathscr{C}_j(t)} \kappa^2 \, ds \, dt = \left[-L_j(t)\right]_{T'-2}^{T'-1} < L.$$

Therefore, there is a sequence $t_j \in [T'-2, T'-1]$ such that

$$\int_{\mathscr{C}_j(t_j)} \kappa^2 \, ds < L.$$

By a Sobolev embedding theorem this implies that the curves $\mathscr{C}_j(t_j)$ are uniformly $C^{1,1/2}$, i.e. they are continuously differentiable, and their tangent angles ϕ_j are uniformly Hölder continuous — in fact, for any two points at arclength coordinates s_1, s_2 in $\mathscr{C}_j(t_j)$ one has

$$|\phi(s_2) - \phi(s_1)| = \left|\int_{s_1}^{s_2} \kappa \, ds\right| \leqslant \sqrt{s_2 - s_1}\sqrt{\int_{s_1}^{s_2} \kappa^2 ds} \leqslant \sqrt{L}\sqrt{s_2 - s_1}.$$

It follows that all $\mathscr{C}_j(t_j)$ are uniformly locally Lipschitz curves. Now consider the solutions to curve shortening with $\mathscr{C}_j(t_j)$ as initial data, i.e. consider $\widetilde{\mathscr{C}}_j(t) = \mathscr{C}_j(t_j + t)$. These solutions all exist for $0 \leqslant t \leqslant T - t_j \geqslant T - T'$. Supposing that along some subsequence of t_j the curvatures of the $\widetilde{\mathscr{C}}_j$ are not bounded for $1 \leqslant t \leqslant T - T'$, we pass to a further subsequence for which the initial curves $\widetilde{\mathscr{C}}_j(0)$ converge in C^1 to some limit curve $\widetilde{\mathscr{C}}_*$. The enclosed areas of the $\widetilde{\mathscr{C}}_j(0)$ then also converge, and hence, by Grayson's theorem [6] the evolution by Curve Shortening $\mathscr{C}_*(t)$ starting from $\mathscr{C}_*$ exists for $0 \leqslant t \leqslant T - T'$. By continuous dependence on initial data it follows that the solutions $\widetilde{\mathscr{C}}_j(t)$ converge in C^∞ to $\mathscr{C}_*(t)$ on any time interval $[\delta, T - T']$ with $\delta > 0$.

This implies that the curvatures of the $\widetilde{\mathscr{C}}_j(t)$ are uniformly bounded for $t \in [1, T - T']$, which then implies that the curvatures of $\mathscr{C}_j(t)$ are uniformly bounded after all for $t \in [t_j + 1, t_j + T - T'] \subset [T', T - 2]$. $\square$

APPENDIX A. THE YIN-YANG SOLITON

Hungerbühler and Smoczyk [8] proved uniqueness and existence of a rotating soliton for curve shortening that contains the origin. See also Halldorson [7] and Altschuler et.al. [1]. Here we derive its more detailed asymptotic behavior, which we use in our construction of the approximate solution.

A.1. **The Yin–Yang soliton in polar coordinates.** For an evolving family of curves written in polar coordinates $X(t, \theta) = r(t, \theta)\boldsymbol{E}_1(\theta)$ the Curve Shortening Deficit is

$$(V - \kappa)\, ds = \left\langle X_t - \frac{X_{\theta\theta}}{\|X_\theta\|^2}, JX_\theta \right\rangle d\theta = \left\{ -rr_t + \frac{r(r_{\theta\theta} - r) - 2r_\theta^2}{r^2 + r_\theta^2} \right\} d\theta.$$

It follows that X is a solution of CSF if and only if $r(\theta, t)$ satisfies

$$(A.1) \qquad r\frac{\partial r}{\partial t} = \frac{rr_{\theta\theta} - r_\theta^2}{r^2 + r_\theta^2} - 1 = \frac{\partial}{\partial \theta}\left(\arctan \frac{r_\theta}{r}\right) - 1.$$

If we look for solutions of the form $r(\theta, t) = \mathscr{R}(\theta - t)$ we get an ODE for $\mathscr{R}(\theta)$

$$(A.2) \qquad -\mathscr{R}\mathscr{R}' = \frac{\mathscr{R}\mathscr{R}'' - \mathscr{R}'^2}{\mathscr{R}^2 + \mathscr{R}'^2} - 1.$$

Hungerbühler and Smoczyk [8] observed that this equation can be integrated once. By suitably rotating the curve around the origin we can ensure that the resulting integration constant vanishes, and we therefore have

$$(A.3) \qquad \frac{1}{2}\mathscr{R}(\theta)^2 - \theta + \arctan\frac{\mathscr{R}'(\theta)}{\mathscr{R}(\theta)} = 0.$$

We consider the soliton that passes through the origin. When this happens $\mathscr{R} \to 0$ and $\mathscr{R}' \to \infty$, so that (A.3) implies $\theta \to \pi/2$. The function $\mathscr{R}(\theta)$ is therefore defined for all $\theta > \pi/2$, and, as proved by Hungerbühler and Smoczyk, $\mathscr{R}'(\theta) > 0$ for all $\theta > \pi/2$.

A.2. **Asymptotic expansion of $\mathscr{R}$.** We now show that $\mathscr{R}(\theta)$ has an asymptotic expansion of the form

$$(A.4) \qquad \mathscr{R}(\theta) = (2\theta)^{1/2}\left\{1 + \frac{c_1}{2\theta} + \frac{c_2}{(2\theta)^2} + \cdots + \frac{c_N}{(2\theta)^N} + \mathscr{O}(\theta^{-N-1})\right\} \qquad (\theta \to \infty)$$

for any $N \in \mathbb{N}$. These expansions can be differentiated any number of times. The coefficients c_j can be computed by substituting the expansions in (A.2) and recursively solving for c_j. In particular, one finds $c_1 = 0$, $c_2 = -1$, and $c_3 = \frac{11}{3}$ so that

$$(A.5) \qquad \mathscr{R}(\theta) = \sqrt{2\theta} - \frac{1}{(2\theta)^{3/2}} + \frac{11}{3}\frac{1}{(2\theta)^{5/2}} + \mathscr{O}(\theta^{-7/2}) \qquad (\theta \to \infty)$$

$$(A.6) \qquad \mathscr{R}'(\theta) = \frac{1}{\sqrt{2\theta}} + \frac{3}{(2\theta)^{5/2}} - \frac{55}{3}\frac{1}{(2\theta)^{7/2}} + \mathscr{O}(\theta^{-9/2}) \qquad (\theta \to \infty).$$

This implies that the quantities $R = R(t) = \mathscr{R}(-2t)$, $\varepsilon = 1/R$, and $R_\theta = R_\theta(t) = \mathscr{R}'(-2t)$, which we use in the construction of the cap, have expansions in powers of $\tau = -4t$,

given by

$$(A.7) \quad \begin{cases} R = \mathscr{R}(-2t) = \sqrt{\tau}\{1 - \tau^{-2} + \tfrac{11}{3}\tau^{-3} + \mathscr{O}(\tau^{-4})\} \\[2mm] \varepsilon = \dfrac{1}{R} = \dfrac{1}{\sqrt{\tau}}\{1 + \tau^{-2} - \tfrac{11}{3}\tau^{-3} + \mathscr{O}(\tau^{-4})\}. \\[2mm] R_\theta = \mathscr{R}'(-2t) = \dfrac{1}{\sqrt{\tau}}\{1 + 3\tau^{-2} - \tfrac{55}{3}\tau^{-3} + \mathscr{O}(\tau^{-4})\}. \end{cases}$$

Proof of (A.4) Consider the quantity $u(\theta) = \theta - \tfrac{1}{2}\mathscr{R}(\theta)^2$. Since $\mathscr{R}'(\theta) > 0$ for all θ it follows from (A.3) that

$$(A.8) \qquad\qquad 0 < u(\theta) < \frac{\pi}{2} \text{ for all } \theta > \frac{\pi}{2}.$$

Directly differentiating $u = \theta - \tfrac{1}{2}\mathscr{R}^2$ and using (A.3) we find an equation for u,

$$(A.9) \qquad\qquad u'(\theta) = 1 - 2(\theta - u(\theta))\tan u(\theta).$$

i.e.

$$u' + 2\theta u = F(\theta, u) \overset{\text{def}}{=} 1 + 2\theta(u - \tan u) + 2u \tan u$$

We use induction to show that $u(\theta)$ has an expansion of the form

$$(A.10) \qquad\qquad u(\theta) = \frac{u_1}{\theta} + \frac{u_2}{\theta^2} + \cdots + \frac{u_N}{\theta^N} + \mathscr{O}(\theta^{-N-1})$$

for any $N \in \mathbb{N}$.

Begin with the case $N = 0$. We know that $0 < u < \pi/2 < \theta$, so that $u < \tan u$, and hence

$$u'(\theta) = 1 - 2(\theta - u)\tan u \leqslant 1 - 2(\theta - u)u = 1 - 2\theta u + 2u^2$$

which implies

$$u' + 2\theta u \leqslant 1 + 2u^2 \leqslant 1 + \frac{\pi^2}{2} = C_0.$$

Multiply with e^{θ^2}, and integrate

$$0 < u(\theta) \leqslant u(\theta_0)e^{\theta_0^2 - \theta^2} + C_0 \int_{\theta_0}^{\theta} e^{\xi^2 - \theta^2}\, d\xi = O(\theta^{-1}) \quad (\theta \to \infty),$$

Thus the case $N = 0$ holds.

For the induction step we expand $\tan u$ in a Taylor series,

$$\tan u = u + \frac{u^3}{3} + \cdots = u + \sum_{k \geqslant 1} c_k u^{2k+1}$$

and rewrite the equation for u as

$$F(\theta, u) = 1 + 2u^2 + \sum_{k \geqslant 1} 2(\theta - u)c_k u^{2k+1}$$

Multiplying with e^{θ^2} and integrating from some fixed $\theta_0 > \pi/2$ we get

$$(A.11) \qquad\qquad u(\theta) = e^{\theta_0^2 - \theta^2}u(\theta_0) + \int_{\theta_0}^{\theta} e^{\xi^2 - \theta^2}F(\xi, u(\xi))\, d\xi$$

Repeated integration by parts leads to

$$(A.12) \quad \int_{\theta_0}^{\theta} e^{\xi^2 - \theta^2}\xi^{-k}\, d\xi =$$

$$\frac{1}{2}\theta^{-k-1} + \frac{k+1}{2}\theta^{-k-3} + \cdots + \frac{(k+1)\cdots(k+2m-1)}{2^m}\theta^{-k-2m-1} + \mathscr{O}(\theta^{-k-2m-3})$$

for all $k, m \in \mathbb{N}$. If we assume that u has an expansion up to $\mathcal{O}(\theta^{-(N-1)})$, then we also have expansions for u^2 and $F(\theta, u)$ up to $\mathcal{O}(\theta^{-(N-1)})$. Substitute these expansions in the integral equation (A.11) and use (A.12) to conclude that u has an expansion up to $\mathcal{O}(\theta^{-N})$, as claimed.

The expansion (A.10) for u, which we now have proved, implies $\mathcal{R}(\theta) = \sqrt{2(\theta - u(\theta))}$ also has an asymptotic expansion in powers of θ^{-1}.

Finally, while one cannot in general differentiate asymptotic expansions, one can integrate them. Thus if a function $f(\theta)$ and its derivative $f'(\theta)$ both have asymptotic expansions in powers of θ^{-1}, then by integrating the expansion of f' one should get the expansion for f, up to a constant: this implies that the expansion of f' can be found by differentiating the expansion for f. We therefore only have to show that all derivatives of $u(\theta)$ have expansions in powers of θ^{-1}, which will then imply that the expansions (A.10) can be differentiated.

To find expansions for $u^{(m)}$, note that if u has an expansion with remainder $\mathcal{O}(\theta^{-N-1})$, then simple substitution in the differential equation (A.9) leads to an expansion for $u'(\theta)$ with remainder $\mathcal{O}(\theta^{-N})$. Going further, one can differentiate (A.9) $m-1$ times and express $u^{(m)}(\theta)$ in terms of $u, u', u'', \ldots, u^{(m-1)}$. This implies that if one has an expansion in powers of θ^{-1} of the first $m-1$ derivatives of u, then one also has an expansion for $u^{(m)}$. By induction it follows that all derivatives of u have such expansions.

Similar arguments also apply to the expansions of $\mathcal{R}(\theta)$.

A.3. **Inversion of the expansion of $\mathcal{R}$.** The expansion (A.4), which expresses $\mathcal{R}$ as a function of θ, implies that one can invert the function $\theta \mapsto \mathcal{R}(\theta)$, and that the inverse has an asymptotic expansion. It follows from (A.4) that

$$\mathcal{R}(\theta)^2 = (2\theta)\left\{1 + \frac{\bar{c}_1}{2\theta} + \frac{\bar{c}_2}{(2\theta)^2} + \cdots + \frac{\bar{c}_N}{(2\theta)^N} + \mathcal{O}(\theta^{-N-1})\right\}$$

and hence

$$2\theta = \mathcal{R}^2\left\{1 + \frac{\bar{c}_1}{2\theta} + \frac{\bar{c}_2}{(2\theta)^2} + \cdots + \frac{\bar{c}_N}{(2\theta)^N} + \mathcal{O}(\theta^{-N-1})\right\}^{-1}.$$

Repeated substitution of this expansion in itself allows one to convert all powers of (2θ) on the left into powers of $\mathcal{R}^2$, so that we have an expansion

$$2\theta = \mathcal{R}^2\left\{1 + \frac{\widetilde{c}_1}{\mathcal{R}^2} + \frac{\widetilde{c}_2}{\mathcal{R}^4} + \cdots + \frac{\widetilde{c}_N}{\mathcal{R}^{2N}} + \mathcal{O}(\mathcal{R}^{-2N-2})\right\}$$

for certain coefficients $\widetilde{c}_i$.

References

[1] Dylan J. Altschuler, Steven J. Altschuler, Sigurd B. Angenent, and Lani F. Wu. The zoo of solitons for curve shortening in $\mathbb{R}^n$. *Nonlinearity*, 26(5):1189–1226, 2013.

[2] Steven J. Altschuler and Matthew A. Grayson. Shortening space curves and flow through singularities. *J. Differential Geom.*, 35(2):283–298, 1992.

[3] Sigurd Angenent and Qian You. Ancient solutions to curve shortening with finite total curvature. *Trans. Amer. Math. Soc.*, 374(2):863–880, 2021.

[4] Jumageldi Charyyev. A compact non-convex ancient curve shortening flow, 2022. Preprint: arXiv:2204.05978.

[5] Panagiota Daskalopoulos, Richard Hamilton, and Natasa Sesum. Classification of compact ancient solutions to the curve shortening flow. *J. Differential Geom.*, 84(3):455–464, 2010.

[6] Matthew A. Grayson. The heat equation shrinks embedded plane curves to round points. *J. Differential Geom.*, 26:284–314, 1986.

[7] Hoeskuldur P. Halldorsson. Self-similar solutions to the curve shortening flow. *Trans. Amer. Math. Soc.*, 364(10):5285–5309, 2012.

[8] N. Hungerbühler and K. Smoczyk. Soliton solutions for the mean curvature flow. *Differential Integral Equations*, 13(10-12):1321–1345, 2000.

[9] Wikipedia contributors. Fermat's spiral — Wikipedia, the free encyclopedia, 2021. https://en.wikipedia.org/w/index.php?title=Fermat%27s_spiral&oldid=1025901425 [accessed 2-July-2021].

[10] Qian You. *Some Ancient Solutions of Curve Shortening*. PhD thesis, UW–Madison, December 2014.

YONGZHE ZHANG
Email address: yongzhe@caltech.edu

CONNOR OLSON
Email address: cjo5325@psu.edu

ILYAS KHAN, UNIVERSITY OF OXFORD, MATHEMATICAL INSTITUTE, ANDREW WILES BUILDING, WOODSTOCK RD, OXFORD, UK, OX2 6GG
Email address: Ilyas.Khan@maths.ox.ac.uk

SIGURD ANGENENT, UNIVERSITY OF WISCONSIN–MADISON, VAN VLECK HALL, MADISON, WI53706, USA
Email address: angenent@wisc.edu