PORTUGALIAE MATHEMATICA
Vol. 62 Fasc. 3 – 2005
Nova Série

CURVE SHORTENING AND THE FOUR-VERTEX THEOREM

Bernd Süssmann

Abstract: This paper shows how the four-vertex theorem, a famous theorem in differential geometry, can be proven by using curve shortening.

1 – Introduction

The four-vertex theorem, in its classical formulation, says, that every simple closed C^3 curve in the Euclidean plane E^2 has at least four vertices, i.e. points with $k_s = 0$, where k is the curvature and k_s the derivative of k by arclength s.

The theorem was proven first by S. Mukhopadhyaya in 1909 for convex and in 1912 by A. Kneser for nonconvex curves, see [Mu], [K].

Later on several interesting methods for proving the theorem were discovered, an overview can be found e.g. in [BF].

Recent publications on this topic deal with more-vertex theorems ([O]) or vertices of nonsimple curves ([P]). The reader may also have a look at the bibliographies of [BF] and [O].

S.B. Jackson extended in 1945 the four-vertex theorem to simple closed curves on simply connected surfaces M^2 of constant curvature K. His proof is based upon the four-vertex theorem for plane curves and a transformation $M^2 \to E^2$, that maps vertices on vertices, cf. [J].

We will prove this result in a somewhat weaker form:

Received: January 7, 2004; *Revised*: June 21, 2004.
AMS Subject Classification: 53A04, 53A35, 53C44.
Keywords: four-vertex theorem; curve shortening.

Theorem. *Let M^2 be a smooth, complete, simply connected surface with constant Gauss curvature K. Let C be a simple, closed, immersed C^3 curve in M^2.*

In the case $K < 0$ we additionally require for the geodesic curvature k of C in each point $k \geq \sqrt{-K}$ or $k \leq -\sqrt{-K}$.

Then C has at least four vertices.

The restriction on k in the hyperblic case has technical reasons, as we will see later.

Let us outline the proceeding in the following chapters:

We will construct a contradiction by asserting that C possesses only two vertices.

We will apply Curve shortening to C and consider the focal curve of $C(t)$ at an arbitrary time $t > 0$. The focal curve has the property that it possesses singularities or cusps at the same parameter values, where its source curve has vertices. We will then show that the focal curve encloses a domain with positive winding number, which expands during progressing time. This will yield a contradiction to the fact that the focal curve contracts to a geodesic segment (Lemma 5).

In E^2 the focal curve (or evolute, in this case) converges even to the same point as the curve itself. However, we do not have this result in the non-Euclidean case. This makes a somewhat more sophisticated analysis of the behaviour of that vertex necessary, which represents the curvature minimum (Lemma 3). This analysis requires a transition to the direction-preserving flow, that is essentially a parameter transformation to an angle parameter.

While in the case $K > 0$ the focal curve always exists, it needs to have strictly positive curvature in the Euclidean case, and for $K < 0$ it is required that $k > \sqrt{-K}$ or $k < -\sqrt{-K}$ holds.

Since there exists a moment $t_c < t_{\max}$ ($t_{\max}$ is the maximal lifespan of the evolving curve) for each nonconvex curve in E^2, at which the curve becomes convex and $k > 0$ is reached for $t > t_c$ (cf. [Gr1, §2, Main Theorem]), we are able to construct the focal curve (or evolute, respectively) for $t > t_c$, and so prove the Theorem also for nonconvex curves. For $K < 0$, however, it is not known, whether all the curves fulfill one of the curvatue restrictions mentioned earlier, before the evolution stops. So we have to require them a priori.

This work was part of the author's doctoral thesis.

2 – Preparations

For what follows, let C be a curve as described in the Theorem. We additionally assume for C that it has exactly two vertices.

For $K<0$ we consider only the case $k>\sqrt{-K}$, without loss of generality.

We take C as initial curve $C(0)=C$ of the initial value problem for curve shortening on M^2, let the solution have the parametrization X: $S^1\times[0,t_{\max})\to M^2$, $X=X(u,t)$.

According to [A2, Theorem 1.5], the number of vertices does not increase during the evolution. Since there must always be at least two points wih vanishing derivative of the curvature, each $C(t), 0\le t<t_{\max}$, has exactly two vertices.

With the assumption above it follows that C can have at most two inflection points (points with vanishing curvature), this amount cannot increase in time either, by [A2, Theorem 1.4]. Hence all $C(t), 0\le t<t_{\max}$, have at most two inflection points.

In the case $M^2=E^2$ we have by [Gr1, §2, Main Theorem] a $t_c<t_{\max}$, such that $C(t)$ possesses strictly positive curvature for all $t>t_c$. Without loss of generality we set $t_c=0$.

In the case $K<0$ we have $k(u,t)>\sqrt{-K}$ for all $(u,t)\in S^1\times(0,t_{\max})$. This follows by applying the strong maximum principle to $k-\sqrt{-K}$, see e.g. [Gr1, Lemma 1.8], and to the evolution equation of k.

Now we consider the *direction-preserving flow*, following [Gr2, section 2], in a slightly different manner.

Let $\theta=\theta(u,t)=\int_0^u kv\,du=\int_0^{s(u)} k\,ds$ with $v=\|X_u\|$ be the angle, $T(u,t)$ encloses with the from $X(0,t)$ to $X(u,t)$ parallel transported vector $T(0,t)$. For curves in the Euclidean plane with strictly positive curvature, θ can be used as a global curve parameter, since the curvature remains strictly positive during the evolution, cf. [GaH, §4].

Here we have, using $v_t=-k^2v$ ([Gr2, p. 74]) and $k_t=k_{ss}+k^3+Kk$ ([Gr2, Lemma 1.3])

$$\theta_t \;=\; \frac{\partial}{\partial t}\int_0^u kv\,du \;=\; k_s+K\theta\ . \tag{1}$$

Since we cannot use θ as angle parameter for $t>0$, we define a corrective function ϱ, the "angle density", comparable to the arc length density v. So we set

$$\varrho(u,t) \;:=\; \begin{cases} e^{Kt}\ , & k(u,t)>0\\ -e^{Kt}\ , & k(u,t)<0\ . \end{cases} \tag{2}$$

The new "angle parameter" will be the function

$$\varphi(u,t) := \frac{\theta(u,t)}{\varrho(u,t)} \ . \tag{3}$$

We have

$$\varphi_s = \frac{\theta_s}{\varrho} \qquad \text{and} \qquad \varphi_t = \frac{k_s}{\varrho} \tag{4}$$

with (1). We set $\tau(u,t) = \tau(t) := t$ as the new time parameter, and the functional determinant for the parameter transformation from (u,t) to $(\varphi(u,t), \tau(u,t))$ reads $\frac{\partial \varphi}{\partial u}\frac{\partial \tau}{\partial t} - \frac{\partial \varphi}{\partial t}\frac{\partial \tau}{\partial u} = \frac{v}{\varrho} k > 0$ for $k \neq 0$ and small t.

We investigate the behaviour of a function f under this parameter transformation: From

$$f(u,t) \ = \ f(\varphi(u,t), \tau(t))$$

follows with (4)

$$f_s(u,t) \ = \ f_\varphi(\varphi,\tau)\, \varphi_s(u,t) \ = \ \frac{k(u,t)}{\varrho(u,t)}\, f_\varphi(\varphi,\tau)$$

and so

$$f_\varphi(\varphi,\tau) \ = \ \frac{\varrho(u,t)}{k(u,t)}\, f_s(u,t) \ . \tag{5}$$

Also with (4) we obtain

$$f_t(u,t) \ = \ f_\varphi(\varphi,\tau)\, \varphi_t(u,t) + f_\tau(\varphi,\tau)\, \tau_t(t) \ = \ \frac{k_s(u,t)}{\varrho(u,t)}\, f_\varphi(\varphi,\tau) + f_\tau(\varphi,\tau) \ ,$$

and, eventually with (5)

$$f_\tau(\varphi,\tau) \ = \ f_t(u,t) - \frac{k_s(u,t)}{k(u,t)}\, f_s(u,t) \ . \tag{6}$$

(6) applied to φ yields with (4) (cf. also [Gr2, Lemma 2.1])

$$\varphi_\tau(\varphi,\tau) \ = \ \varphi_t(u,t) - \frac{k_s(u,t)}{k(u,t)}\, \varphi_s(u,t) \ \equiv \ 0 \ .$$

From this we see that φ is independent of τ.

The following calculations are similar to those in [GaH, Section 4.1] or [EGa, §3], there (in the Euclidean case) we always have $\varrho \equiv 1$.

From

$$X(u,t) \;=\; X(\varphi(u,t),\tau(t))$$

we get

$$vT \;=\; X_u \;=\; \varphi_u\, X_\varphi \;=\; \frac{v}{\varrho}\, k\, X_\varphi$$

and thus

(7) $$X_\varphi = \frac{\varrho}{k}\, T \; .$$

Besides, we have $k = \theta_s = \varphi_s \theta_\varphi = \frac{k}{\varrho}\theta_\varphi$ and with that

(8) $$\theta_\varphi = \varrho \qquad \text{and} \qquad k\, ds = d\theta = \varrho\, d\varphi \; .$$

With $\frac{\partial}{\partial\theta} = \frac{1}{\varrho}\frac{\partial}{\partial\varphi}$ follows then

$$X_\theta \;=\; \frac{1}{k}\, T$$

as in the Euclidean case. From this we obtain (N is the unit normal vector of C)

$$kN \;=\; X_t \;=\; \varphi_t\, X_\varphi + X_\tau \;=\; \frac{k_s}{k}\, T + X_\tau \; .$$

From this we receive with

(9) $$k_s \;=\; \varphi_s\, k_\varphi \;=\; \frac{k}{\varrho}\, k_\varphi \;=\; k\, k_\theta$$

the new evolution equation for X:

(10) $$X_\tau \;=\; -\frac{k_\varphi}{\varrho}\, T + kN \;=\; -k_\theta T + kN \; .$$

The covariant derivatives $\nabla_\varphi = \nabla_{X_\varphi}$, $\nabla_\theta = \nabla_{X_\theta}$ and $\nabla_\tau = \nabla_{X_\tau}$ of T and N read

$$kN = \nabla_s T = \varphi_s \nabla_\varphi T = \frac{k}{\varrho}\, \nabla_\varphi T \quad \Longrightarrow \quad \nabla_\varphi T = \varrho N, \;\; \nabla_\varphi N = -\varrho T$$

as well as

(11) $$\nabla_\theta T = N\, , \qquad \nabla_\theta N = -T \; .$$

By $k_s N = \nabla_t T$ we get $k_s N = \nabla_t T = \varphi_t \nabla_\varphi T + \nabla_\tau T = k_s N + \nabla_\tau T$, therefore

(12) $$\nabla_\tau T = 0\, , \qquad \nabla_\tau N = 0 \; .$$

So equations (12) justify the name *direction-preserving flow.*

We calculate the new evolution equation of the curvature (cf. [Gr2, Lemma 2.7]):

We get $k_t = \varphi_t k_\varphi + k_\tau = k k_\theta^2 + k_\tau$ by (9). From $k_t = k_{ss} + k^3 + Kk$ ([Gr2, Lemma 1.3]) follows, using $k_{ss} = (k k_\theta)_s = k k_\theta^2 + k^2 k_{\theta\theta}$, the formula

$$k_\tau = k^2 k_{\theta\theta} + k^3 + Kk \ , \tag{13}$$

where θ and τ do not commute. Since ϱ does not depend on φ, the evolution equation (13) can also be written as

$$k_\tau = \varrho^{-2} k^2 k_{\varphi\varphi} + k^3 + Kk \ . \tag{14}$$

From now on, we consider the evolution only in the (φ, τ)-parameters.

If $C(t)$ has two inflection points, so let $\varphi \in I^-(\tau) \cup I^+(\tau)$, where $I^-(\tau), I^+(\tau)$ are to be understood as time-dependent open intervals with $k\Big|_{I^-(\tau)} < 0$ and $k\Big|_{I^+(\tau)} > 0$. If an inflection point vanishes at time $\tau_0 < \tau_{\max}$, so the other one must vanish at the same time, we have then $k > 0$ for $\tau_0 > 0$, and let $\varphi \in [0, \bar{\varphi})$ for $\tau > \tau_0$, $\bar{\varphi}$ kept fixed. For τ we have $\tau \in [0, \tau_{\max})$ with $\tau_{\max} = t_{\max}$, and we will always set $\tau_0 = 0$, without loss of generality, if τ_0 occurs.

3 – Evolute and focal curve

Let C, respectively $C(\tau)$, be given as in the previous section, i.e. especially $k > 0$ (in the case $K = 0$) or $k > \sqrt{-K}$ ($K < 0$) shall hold for $\tau > 0$.

Let in the Euclidean plane the *evolute* $\bar{C}(\tau) \subset E^2$ of $C(\tau)$ with the parametrization $\bar{X}$ be given by

$$\bar{X}(\varphi, \tau) := X(\varphi, \tau) + \frac{1}{k(\varphi, \tau)} N(\varphi, \tau) \ . \tag{15}$$

As a model for M^2 with $K < 0$ we use the Weierstrass model in the Lorentz space: $\mathbb{R}^3$ with the non-degenerated inner product $\langle x, y \rangle_{-1} = -x_1 y_1 + x_2 y_2 + x_3 y_3$ for $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3) \in \mathbb{R}^3$ is called the Lorentz space $\mathbb{R}^3_1$. Then the surface $H^2_K = \{x \in \mathbb{R}^3_1 \mid \langle x, x \rangle_{-1} = \frac{1}{K},\ x_1 > 0\}$ represents the Weierstrass model of the hyperbolic plane with curvature $K < 0$, cf. e.g. [C, p. 180].

We use the sphere $S^2_K = \{x \in E^3 \mid \langle x, x \rangle = x_1^2 + x_2^2 + x_3^2 = \frac{1}{K}\}$ in the Euclidean space E^3 as a model for M^2 with $K > 0$.

With this we can treat points and vectors in the case $K \neq 0$ in a way similar to the Euclidean case, without using the exponential map.

So for $K \neq 0$ we define a curve $\bar{C}(\tau)$ by

$$\bar{X}(\varphi,\tau) := \frac{k(\varphi,\tau)}{\sqrt{K+k^2(\varphi,\tau)}}\, X(\varphi,\tau) + \frac{1}{\sqrt{K+k^2(\varphi,\tau)}}\, N(\varphi,\tau)\ . \tag{16}$$

For $K<0$, $\bar{C}(\tau)$ is exactly the *focal curve,* and for $K>0$ one of the two possible focal curves of $C(\tau)$ (depending on the orientation of $C(\tau)$). For the definition of the focal curve in general cf. [C, Definition 4.5, p. 232], and for the derivation of the focal curve in the spherical case cf. e.g. [Ml, p. 18].

Since (16) also makes sense for $K=0$, we will work in the following without different cases.

Elementary calculations yield $\bar{v} = \|\bar{X}_\varphi\| = \frac{|k_\varphi|}{K+k^2}$ as well as unit tangent and normal vectors of $\bar{X}$ as

$$\bar{T} = \frac{K \operatorname{sign} k_\varphi}{\sqrt{K+k^2}}\, X - \frac{k \operatorname{sign} k_\varphi}{\sqrt{K+k^2}}\, N\ , \tag{17}$$

$$\bar{N} = \operatorname{sign} k_\varphi \cdot T\ , \tag{18}$$

only at points with $k_\varphi \neq 0$.

For the curvature $\bar{k}$ of $\bar{X}$ we obtain

$$\bar{k} = \frac{\varrho(K+k^2)^{\frac{3}{2}}}{k|k_\varphi|} = \frac{(K+k^2)^{\frac{3}{2}}}{|k|\,|k_\theta|}\ . \tag{19}$$

We determine the *(induced) evolution equation of the focal curve*:

Lemma 1. *If $C(\tau)$ evolves according to the equation $X_\tau = -\frac{k_\varphi}{\varrho}T + kN$, then for the focal curve $\bar{C}(\tau)$ of $C(\tau)$*

$$\bar{X}_\tau = \frac{k^2 k_{\varphi\varphi} \operatorname{sign} k_\varphi}{\varrho^2(K+k^2)}\, \bar{T} - \frac{k|k_\varphi|}{\varrho\sqrt{K+k^2}}\, \bar{N}\ , \tag{20}$$

is valid at any time $\tau>0$ at points where $k_\varphi \neq 0$.

Proof: Using (14) we get

$$\left(\frac{k}{\sqrt{K+k^2}}\right)_\tau = \frac{Kk^2k_{\varphi\varphi}}{\varrho^2(K+k^2)^{\frac{3}{2}}} + \frac{Kk}{\sqrt{K+k^2}}$$

and

$$\left(\frac{1}{\sqrt{K+k^2}}\right)_\tau = -\frac{k^3k_{\varphi\varphi}}{\varrho^2(K+k^2)^{\frac{3}{2}}} - \frac{k^2}{\sqrt{K+k^2}}\ .$$

$X_\tau = -\frac{k_\varphi}{\varrho}T + kN$ and $N_\tau = -KkX$ as well as (17), (18) lead to the assertion. ∎

4 – Convergence of the focal curve

By [Gr2, Theorem 0.1 and Corollary 3.4] we know: If $\tau_{\max}$ is finite, $C(\tau)$ converges for $\tau \to \tau_{\max}$ to a point $P \in M^2$ (in the Hausdorff metric). In the case $\tau_{\max} = \infty$ (which is only possible for $K > 0$), $C(\tau)$ converges to a large circle in the C^∞-sense.

We will use the convention that $C(0)$ is positively oriented, i.e. that $\int_{C(0)} k\, ds \geq 0$ holds. With [Gr2, Section 1] we then have $\int_{C(\tau)} k\, ds \geq 0$ for all τ.

Lemma 2. *Let M^2 and C be as in the Theorem, additionally we assume that $C(0)$ has exactly two vertices and that $C(\tau)$ converges to a point for $\tau \to \tau_{\max}$.*

Then there exists a constant $k_0 > \infty$, such that $k(\varphi, \tau) \geq k_0$ is valid for all $(\varphi, \tau) \in (I^-(\tau) \cup I^+(\tau)) \times [0, \tau_{\max})$.

In the case $M^2 = E^2$ even $\lim_{\tau \to \tau_{\max}} \min_{\varphi \in [0, \bar{\varphi})} k(\varphi, \tau) = \infty$ holds.

Proof: The second assertion is known ([GaH, Corollary 5.6]). In order to prove the first assertion, we assume its contrary, i.e. a sequence $(\varphi_n, \tau_n)_{n \in \mathbf{N}}$ shall exist with $\tau_n \to \tau_{\max}$ and $k_n := k(\varphi_n, \tau_n) \to -\infty$ for $n \to \infty$. Following [Gr2, Lemma 5.2] (cf. also [Gr2, Theorem 5.1]) there exists to each τ_n and $k_n \leq 0$ a $\tilde{\tau} \in [\tau_n, \tau_{\max})$ and an interval $I(\tilde{\tau}_n) = \{\varphi \mid k(\varphi, \tilde{\tau}_n) < k_n\}$ with $\int_{I(\tilde{\tau}_n)} |\varrho|\, d\varphi = e^{K\tau_n} |I(\tilde{\tau}_n)| > \pi$ (cf. (2)).

The arc $C(I(\tilde{\tau}_n))$ belonging to $I(\tilde{\tau}_n)$ possesses, by the δ-*Whisker-Lemma* ([Gr2, Lemma 6.4]), a *whisker* with length $\delta > 0$ (i.e. geodesic segments of length δ, starting at $C(I(\tilde{\tau}_n))$, going into the domain enclosed by $C(\tilde{\tau}_n)$), which belong to a suitable foliation and are parallel to the respective tangents at the edges of $C(I(\tilde{\tau}_n))$), which does not intersect the rest of the curve. δ depends only on the initial curve $C(0)$. Since $C(\tilde{\tau}_n)$ converges to a point for $n \to \infty$, the whisker must intersect the curve at some time. This is a contradiction. ∎

Lemma 3. *Let M^2 and C be as in the Theorem, additionally we assume that $C(0)$ has exactly two vertices and that $C(\tau)$ converges to a point for $\tau \to \tau_{\max}$. The minimum of curvature shall be reached at $\varphi_0(\tau)$.*

Then there is a $\tau_0 < \tau_{\max}$, such that $\varphi_0(\tau)$ is continuous for $\tau_0 < \tau < \tau_{\max}$.

If the curvature minimum is additionally bounded by above, i.e. if there exists a $0 < k_1 < \infty$, such that $k(\varphi_0(\tau), \tau) < k_1$ holds for all $0 < \tau < \tau_{\max}$, then $\lim_{\tau \to \tau_{\max}} \varphi_0(\tau)$ and $\lim_{\tau \to \tau_{\max}} T(\varphi_0(\tau), \tau)$ exist as well.

Proof: φ_0 can have a discontinuity only, if at any time a further local cuvature minimum occurs, which is not possible; or if for a $\tau_0 < \tau_{\max}$ the respective vertex is at the same time a zero of the curvature, i.e. $k(\varphi_0(\tau_0),\tau_0) = k_\varphi(\varphi_0(\tau_0),\tau_0) = 0$. This (only) point with that property disappears immediately (see [A2, Theorem 1.3]), i.e. it is $k > 0$ for $\tau > \tau_0$, new such points do not occur, and so φ_0 is continuous for $\tau > \tau_0$ (after the adjustment of ϱ and φ at τ_0). If the vertex coincides never or as late as at $\tau_{\max}$ with an inflection point, we can set $\tau_0 = 0$.

Now let $k(\varphi_0(\tau),\tau) < k_1$ for all $0 < \tau < \tau_{\max}$ with k_1 as required.

In order to show that $\lim_{\tau\to\tau_{\max}} \varphi_0(\tau)$ exists, we consider two different cases. First we consider the situation, where $C(\tau)$ is or will become convex, i.e. where a $\tau_0 < \tau_{\max}$ exists, such that $k(\varphi_0(\tau),\tau) > 0$ is true for all $\tau > \tau_0$. Without loss of generality we set $\tau_0 = 0$. In the remaining case we then have $k(\varphi_0(\tau),\tau) < 0$ for all $\tau < \tau_{\max}$.

In the first case we assume that $\varphi_0(\tau)$ diverges for $\tau \to \tau_{\max}$.

Then there is a sequence $(\tau_n)_{n\in\mathbf{N}}$ with $\tau_n \to \tau_{\max}$, such that $\varphi_0(\tau_n)$ diverges. However, $(\varphi_0(\tau_n))_{n\in\mathbf{N}}$ is bounded and thus has an accumulation point φ_1, by the Bolzano–Weierstrass Theorem. φ_1 can not be the only accumulation point of $(\varphi_0(\tau_n))_{n\in\mathbf{N}}$, for then $\varphi_0(\tau_n)$ would have to converge to φ_1. Therefore each $(\varphi_0(\tau_n))_{n\in\mathbf{N}}$ has another accumulation point $\varphi_2 \neq \varphi_1$. But then every $\varphi \in [\varphi_1,\varphi_2]$ (without loss of generality the interval is of this form, it could also be $[\varphi_2,\varphi_1]$ or the entire parameter interval $[0,\bar\varphi)$) is an accumulation point of φ_0 due to the continuity of φ_0, in other words, φ_0 oscillates on $[\varphi_1,\varphi_2]$.

Now we set $\Delta := \min\left\{\frac{\varphi_2-\varphi_1}{100},\, \frac{1}{3}e^{-|K|\tau_{\max}}\right\}$ and $\varphi_3 := \varphi_1 + \Delta$, $\varphi_4 := \varphi_1 + 2\Delta$, $\varphi_5 := \varphi_1 + 3\Delta$, $\varphi_6 := \varphi_1 + 4\Delta$. Due to the closedness of $C(\tau)$ it is possible to admit also parameters $\varphi \bmod \bar\varphi$, $\varphi \in [0,\bar\varphi)$, for functions defined on $[0,\bar\varphi)$. Thus we set $\varphi_7 := \varphi_3 + \bar\varphi$, such that $\varphi_7 > \varphi_6$ holds, and consider $[0,\bar\varphi)$ also as $[0,\bar\varphi) = [\varphi_3,\varphi_6) \cup [\varphi_6,\varphi_7)$.

Then we have with (8) and the Gauss–Bonnet Theorem

$$\begin{aligned}
\theta(\varphi_7,\tau) - \theta(\varphi_6,\tau) &= \int_{\theta(\varphi_6,\tau)}^{\theta(\varphi_7,\tau)} d\theta = \int_{\theta(0,\tau)}^{\theta(\bar\varphi,\tau)} d\theta - \int_{\theta(\varphi_3,\tau)}^{\theta(\varphi_6,\tau)} d\theta \\
&= \int_0^{L(\tau)} k\,ds - \theta(\varphi_6,\tau) + \theta(\varphi_3,\tau) = 2\pi - KA(\tau) - e^{K\tau}(\varphi_6-\varphi_3) \\
&\geq 2\pi - KA(\tau) - 3\,\Delta e^{|K|\tau_{\max}} \geq 2\pi - KA(\tau) - 1\,,
\end{aligned}$$

where $L(\tau)$ is the length of $C(\tau)$ and $A(\tau)$ the area enclosed by $C(\tau)$. From $A'(\tau) = -\int_0^{L(\tau)} k\,ds$ (see e.g. [Ga, Lemma 1.3]) follows $A'(\tau) = KA(\tau) - 2\pi$

(Gauss–Bonnet) and hence the monotonicity of $A(\tau)$. By that and $\lim_{\tau\to\tau_{\max}} A(\tau) = 0$ we conclude, that a $\tau' < \tau_{\max}$ exists, such that

$$\theta(\varphi_7, \tau) - \theta(\varphi_6, \tau) > \pi \tag{21}$$

is true for all $\tau > \tau'$.

Also by (8) we obtain

$$\begin{aligned} \varphi_4 - \varphi_3 &= \int_{\varphi_3}^{\varphi_4} d\varphi = e^{-K\tau} \int_{\theta(\varphi_3,\tau)}^{\theta(\varphi_4,\tau)} d\theta = e^{-K\tau} \int_{s(\varphi_3,\tau)}^{s(\varphi_4,\tau)} k\, ds \\ &\leq e^{-K\tau} \max_{\varphi_3 \leq \varphi \leq \varphi_4} k(\varphi, \tau)\, [s(\varphi_4, \tau) - s(\varphi_3, \tau)] \\ &\leq e^{-K\tau} \max_{\varphi_3 \leq \varphi \leq \varphi_4} k(\varphi, \tau)\, L(\tau) \end{aligned}$$

and hence

$$\max_{\varphi_3 \leq \varphi \leq \varphi_4} k(\varphi, \tau) \geq \frac{e^{K\tau}}{L(\tau)} (\varphi_4 - \varphi_3) \geq \frac{e^{-|K|\tau_{\max}}}{L(\tau)} \Delta\,. \tag{22}$$

For $\max_{\varphi_5 \leq \varphi \leq \varphi_6} k(\varphi, \tau)$ the same estimation holds.

For the following we define a constant α as

$$\alpha := \left[\sin\left(\pi\, \frac{\varphi_2 - \Delta - \varphi_6}{\varphi_7 - \varphi_6}\right)\right]^{-1} = \left[\sin\left(\pi\, \frac{\varphi_2 - \varphi_1 - 5\Delta}{\bar{\varphi} - 3\Delta}\right)\right]^{-1} > 0\,. \tag{23}$$

Due to $L'(\tau) = -\int_0^{L(\tau)} k^2\, ds < 0$ ([Gr2, Section 1]), $L(\tau)$ is monotone decreasing, and $\lim_{\tau\to\tau_{\max}} L(\tau) = 0$ holds, there exists because of (22) a point in time $\tau'' < \tau_{\max}$, such that

$$\max_{\varphi_3 \leq \varphi \leq \varphi_4} k(\varphi, \tau) > \alpha\, k_1\, e^{|K|\tau_{\max}} \quad \text{and} \quad \max_{\varphi_5 \leq \varphi \leq \varphi_6} k(\varphi, \tau) > \alpha\, k_1\, e^{|K|\tau_{\max}}$$

is true for all $\tau > \tau''$. We choose a $\tilde{\tau} > \max\{\tau', \tau''\}$, $\tilde{\tau} < \tau_{\max}$, such that $\varphi_0(\tilde{\tau})$ lies within (φ_4, φ_5) (possible, since φ_0 oscillates on $[\varphi_1, \varphi_2]$).

But then $k(\varphi, \tilde{\tau}) > \alpha k_1 e^{|K|\tau_{\max}}$ is true for all $\varphi \in [\varphi_6, \varphi_7]$, otherwise $k(\varphi, \tilde{\tau})$ must have a local minimum in $[\varphi_6, \varphi_7]$; so together with $\varphi_0(\tilde{\tau}) \in (\varphi_4, \varphi_5)$ at least two different ones, which is impossible.

We define a comparison function f, similar as in the proof of Lemma 5.4 in [Gr2, p. 98]:

$$f(\varphi, \tau) := \alpha\, k_1\, e^{|K|(\tau_{\max} - \tau)} \sin\left(\pi\, \frac{\varphi - \varphi_6}{\varphi_7 - \varphi_6}\right), \quad \varphi_6 \leq \varphi \leq \varphi_7, \quad \tilde{\tau} \leq \tau \leq \tau_{\max}\,. \tag{24}$$

By this $k(\varphi,\tilde\tau) > \alpha k_1 e^{|K|\tau_{\max}} > f(\varphi,\tilde\tau)$ holds for all $\varphi \in [\varphi_6,\varphi_7]$.

We calculate the derivatives of f:

$$f_{\varphi\varphi} = -\left(\frac{\pi}{\varphi_7-\varphi_6}\right)^2 f\,, \qquad f_\tau = -|K|\,f\ . \tag{25}$$

Since $k(\varphi_0(\tau),\tau) > 0$ is always true, the graphs of k and f can never meet at the edges $\varphi_6,\ \varphi_7$.

If the graph of k touches the graph of f at a time $\bar\tau > \tilde\tau,\ \bar\tau < \tau_{\max}$, for the first time, at a point $\varphi_8 \in (\varphi_6,\varphi_7)$, then we have there

$$k(\varphi_8,\bar\tau) = f(\varphi_8,\bar\tau) \quad \text{and} \quad k_\varphi(\varphi_8,\bar\tau) = f_\varphi(\varphi_8,\bar\tau)\ . \tag{26}$$

By the maximum principle follows that

$$k_{\varphi\varphi}(\varphi_8,\bar\tau) \ \geq\ f_{\varphi\varphi}(\varphi_8,\bar\tau)\ . \tag{27}$$

With (14), (2), (27), (26) and (25) we get

$$\begin{aligned} k_\tau(\varphi_8,\bar\tau) &= e^{-2K\bar\tau}\, k^2(\varphi_8,\bar\tau)\, k_{\varphi\varphi}(\varphi_8,\bar\tau) + k^3(\varphi_8,\bar\tau) + Kk(\varphi_8,\bar\tau) \\ &\geq e^{-2K\bar\tau}\, f^2(\varphi_8,\bar\tau)\, f_{\varphi\varphi}(\varphi_8,\bar\tau) + f^3(\varphi_8,\bar\tau) + Kf(\varphi_8,\bar\tau) \\ &\geq \left[1 - e^{-2K\bar\tau}\left(\frac{\pi}{\varphi_7-\varphi_6}\right)^2\right] f^3(\varphi_8,\bar\tau) + Kf(\varphi_8,\bar\tau)\ . \end{aligned}$$

Now we see

$$e^{-2K\bar\tau}\left(\frac{\pi}{\varphi_7-\varphi_6}\right)^2 = \left(\frac{\pi}{e^{K\bar\tau}(\varphi_7-\varphi_6)}\right)^2 = \left(\frac{\pi}{\theta(\varphi_7,\bar\tau)-\theta(\varphi_6,\bar\tau)}\right)^2 <\ 1$$

because of (21) and $\bar\tau > \tau'$.

Hence we have

$$k_\tau(\varphi_8,\bar\tau) \ >\ Kf(\varphi_8,\bar\tau) \ \geq\ -|K|\,f(\varphi_8,\bar\tau) \ =\ f_\tau(\varphi_8,\bar\tau)\ .$$

This means $k(\varphi_8,\tau) > f(\varphi_8,\tau)$ for $\tau > \bar\tau,\ \tau$ close $\bar\tau$, and, altogether,

$$k(\varphi,\tau) \ \geq\ f(\varphi,\tau) \quad \text{for} \quad \varphi_6 \leq \varphi \leq \varphi_7,\quad \tilde\tau < \tau < \tau_{\max}\ ,$$

i.e. the graph of k cannot cross the graph of f. From this follows with (23)

$$\begin{aligned} k(\varphi_2-\Delta,\tau) \ \geq\ f(\varphi_2-\Delta,\tau) \ &=\ \alpha\, k_1\, e^{|K|(\tau_{\max}-\tau)}\, \sin\!\left(\pi\ \frac{\varphi_2-\Delta-\varphi_6}{\varphi_7-\varphi_6}\right) \\ &=\ k_1\, e^{|K|(\tau_{\max}-\tau)} \ \geq\ k_1 \quad \text{for} \quad \tilde\tau \leq \tau \leq \tau_{\max}\ . \end{aligned}$$

Since φ_0 oscillates continuously on $[\varphi_1, \varphi_2]$, there is a $\tau \geq \tilde{\tau}$ with $\varphi_0(\tau) = \varphi_2 - \Delta$ and thus $k(\varphi_0(\tau), \tau) = k(\varphi_2 - \Delta, \tau) \geq k_1$, in contradiction to the assumption $k(\varphi_0(\tau), \tau) < k_1$ for all τ.

With this the assumption, that $\varphi_0(\tau)$ diverges for $\tau \to \tau_{\max}$, must have been wrong.

Now we treat the remaining case, where $k(\varphi_0(\tau), \tau) < 0$ holds for all $\tau < \tau_{\max}$.

Let $I^-(\tau) = (\varphi^-(\tau), \varphi^+(\tau))$, e.g. the two inflection points occur at $\varphi^-(\tau)$ and $\varphi^+(\tau)$. By this $\varphi^-(\tau) < \varphi_0(\tau) < \varphi^+(\tau)$ is true for all $\tau < \tau_{\max}$.

Then we have

$$\begin{aligned} |\varphi^+(\tau) - \varphi^-(\tau)| &= e^{-K\tau}\Big|\theta(\varphi^+(\tau), \tau) - \theta(\varphi^-(\tau), \tau)\Big| = e^{-K\tau}\left|\int\limits_{s(\varphi^-(\tau),\tau)}^{s(\varphi^+(\tau),\tau)} k\, ds\right| \\ &\leq e^{-K\tau} \int\limits_{s(\varphi^-(\tau),\tau)}^{s(\varphi^+(\tau),\tau)} |k|\, ds \leq e^{-K\tau}|k_0|\, L(\tau) \end{aligned}$$

by Lemma 2. This means $\lim_{\tau\to\tau_{\max}} |\varphi^+(\tau) - \varphi^-(\tau)| = 0$. By [Gr2, Corollary 2.6] φ^- and φ^+ cannot oscillate, thus $\lim_{\tau\to\tau_{\max}} \varphi^-(\tau)$, $\lim_{\tau\to\tau_{\max}} \varphi^+(\tau)$ exist, and hence also $\lim_{\tau\to\tau_{\max}} \varphi_0(\tau)$.

So in both cases $\lim_{\tau\to\tau_{\max}} \theta(\varphi_0(\tau), \tau)$ and therefore also $\lim_{\tau\to\tau_{\max}} T(\varphi_0(\tau), \tau)$ exist.

We set $\varphi_0(\tau_{\max}) := \lim_{\tau\to\tau_{\max}} \varphi_0(\tau)$, if this limit exists. ∎

Lemma 4. *Let M^2 and C be as in the Theorem, additionally we assume that $C(0)$ possesses exactly two vertices and that $C(\tau)$ converges to a point for $\tau \to \tau_{\max}$. Then $\lim_{\tau\to\tau_{\max}} k(\varphi_0(\tau), \tau) \in (-\infty, \infty]$ exists.*

Are furthermore in the case $k(\varphi_0(\tau), \tau) < k_1$ for all $0 < \tau < \tau_{\max}$ with $0 < k_1 < \infty$ a $\delta > 0$ and a sequence $(\varphi_n, \tau_n)_{n\in\mathbb{N}}$ with $\lim_{n\to\infty} \tau_n = \tau_{\max}$ and $|\varphi_0(\tau_{\max}) - \varphi_n| \geq \delta$ for all $n \in \mathbb{N}$ given, then for this sequence $\lim_{n\to\infty} k(\varphi_n, \tau_n) = \infty$ holds.

Proof: We first show the differentiability of φ_0 for almost all $\tau \in (0, \tau_{\max})$: By (14) follows that

$$\begin{aligned} k_{\varphi\tau} = k_{\tau\varphi} &= \varrho^{-2}k^2 k_{\varphi\varphi\varphi} + (\varrho^{-2}k^2)_\varphi k_{\varphi\varphi} + 3\,k^2 k_\varphi + K k_\varphi \\ &= \varrho^{-2}k^2 k_{\varphi\varphi\varphi} + 2\,\varrho^{-2}k k_\varphi k_{\varphi\varphi} + 3\,k^2 k_\varphi + K k_\varphi \end{aligned}$$

with $(\varrho^{-2}k^2)_\varphi = 2\,\varrho^{-2}k k_\varphi$ because of $\varrho_\varphi \equiv 0$.

For $\kappa := k_\varphi$ we obtain by this the evolution equation

$$\kappa_\tau = \varrho^{-2} k^2 \kappa_{\varphi\varphi} + 2\,\varrho^{-2} k\kappa\kappa_\varphi + 3\,k^2\kappa + Kk\ . \tag{28}$$

(28) is of the same type as (14), hence Proposition 1.2 of [A2] can be applied to (28). From this we get: If a time $\hat\tau > 0$ with $\kappa(\varphi_0(\hat\tau),\hat\tau)) = \kappa_\varphi(\varphi_0(\hat\tau),\hat\tau) = 0$ occurs, $\kappa_\varphi(\varphi_0(\tau),\tau) \neq 0$ must hold for any small $\tau > \hat\tau$.

This means that $k_{\varphi\varphi}(\varphi_0(\tau),\tau) = \kappa_\varphi(\varphi_0(\tau),\tau) = 0$ can only occur on a discrete subset of $(0,\tau_{\max})$. On its complement we have $k_\varphi(\varphi_0(\tau),\tau)=0$, $k_{\varphi\varphi}(\varphi_0(\tau),\tau)\neq 0$, there the Theorem on implicit functions yields the differentiability of φ_0.
We consider the same two cases as in the proof before.

So let first be $k(\varphi_0(\tau),\tau) > 0$ for all $\tau > 0$. Then by (14), $k_\varphi(\varphi_0(\tau),\tau) = 0$, $k_{\varphi\varphi}(\varphi_0(\tau),\tau) \geq 0$ and $k(\varphi_0(\tau),\tau) > \sqrt{-K}$ (in the case $K < 0$)

$$\begin{aligned}
\frac{d}{d\tau}k(\varphi_0(\tau),\tau) &= k_\varphi(\varphi_0(\tau),\tau)\,\varphi_0'(\tau) + k_\tau(\varphi_0(\tau),\tau)\\
&= e^{-2K\tau}k^2(\varphi_0(\tau),\tau)\,k_{\varphi\varphi}(\varphi_0(\tau),\tau) + k^3(\varphi_0(\tau),\tau) + Kk(\varphi_0(\tau),\tau)\\
&\geq k(\varphi_0(\tau),\tau)\,\Big(k^2(\varphi_0(\tau),\tau) + K\Big) > 0
\end{aligned}$$

for almost all $\tau\in(0,\tau_{\max})$. φ_0 is continuous for $0<\tau<\tau_{\max}$ (following Lemma 3) and thus also $k(\varphi_0(\tau),\tau)$ for $0 < \tau < \tau_{\max}$. With the previous estimation we conclude, that $k(\varphi_0(\tau),\tau)$ is monotone increasing for $0 < \tau < \tau_{\max}$.

If $k(\varphi_0(\tau),\tau)$ is limited above by k_1, $\lim_{\tau\to\tau_{\max}} k(\varphi_0(\tau),\tau) \leq k_1 < \infty$ exists.

If $k(\varphi_0(\tau),\tau)$ does not have an upper limit, $\lim_{\tau\to\tau_{\max}} k(\varphi_0(\tau),\tau) = \infty$ follows.

In the case $k(\varphi_0(\tau),\tau) < 0$ for all $\tau < \tau_{\max}$, which can only occur for $K > 0$, we have, analogue to above with Lemma 2

$$\frac{d}{d\tau}k(\varphi_0(\tau),\tau) \geq k_0^3 + Kk_0 > -\infty$$

for almost all $\tau\in(0,\tau_{\max})$. φ_0 and so $k(\varphi_0(\tau),\tau)$ are continuous for $0<\tau<\tau_{\max}$; $k(\varphi_0(\tau),\tau)$ is bounded by k_0 and 0, and cannot oscillate, because then $k_\varphi(\varphi_0(\tau),\tau)$ would have to be bonded below and above. Hence also in this case $\lim_{\tau\to\tau_{\max}} k(\varphi_0(\tau),\tau) \leq 0$ exists.

For the proof of the second assertion let $k(\varphi_0(\tau),\tau)$ be bounded above by k_1. Then $\varphi_0(\tau_{\max}) = \lim_{\tau\to\tau_{\max}} \varphi_0(\tau)$ exists by Lemma 3. Additionally, let $\delta > 0$ and $(\varphi_n,\tau_n)_{n\in\mathbf{N}}$ be as mentioned. If $\varphi^-(\tau)$ and $\varphi^+(\tau)$ occur for all $\tau < \tau_{\max}$, a $\tau_\delta < \tau_{\max}$ exists due to $\lim_{\tau\to\tau_{\max}} \varphi^-(\tau) = \lim_{\tau\to\tau_{\max}} \varphi^+(\tau) = \varphi_0(\tau_{\max})$, such that $\varphi^-(\tau),\varphi_0(\tau),\varphi^+(\tau) \in (\varphi_0(\tau_{\max}) - \frac{\delta}{2}, \varphi_0(\tau_{\max}) + \frac{\delta}{2})$ holds for all $\tau > \tau_\delta$. So $\varphi^-(\tau),\varphi_0(\tau),\varphi^+(\tau) \notin [\varphi_n - \frac{\delta}{2}, \varphi_n + \frac{\delta}{2}]$ (or only $\varphi_0(\tau) \notin [\varphi_n - \frac{\delta}{2}, \varphi_n + \frac{\delta}{2}]$, respectively) for all $\tau > \tau_\delta$ and each $n \in \mathbf{N}$ due to the assumption $|\varphi_0(\tau_{\max}) - \varphi_n| \geq \delta$ for all $n \in \mathbf{N}$.

Then $k\Big|_{[\varphi_n - \frac{\delta}{2}, \varphi_n + \frac{\delta}{2}]} > 0$ for all $\tau > \tau_\delta$ and each n; and analogue to (22)

$$\max_{\varphi_n - \frac{\delta}{2} \leq \varphi \leq \varphi_n - \frac{\delta}{4}} k(\varphi, \tau) \geq \frac{\delta}{4} \cdot \frac{e^{K\tau}}{L(\tau)}, \qquad \max_{\varphi_n + \frac{\delta}{4} \leq \varphi \leq \varphi_n + \frac{\delta}{2}} k(\varphi, \tau) \geq \frac{\delta}{4} \cdot \frac{e^{K\tau}}{L(\tau)}$$

holds for all $\tau > \tau_\delta$ and each n. Hence $k(\varphi_n, \tau) \geq \delta e^{-|K|\tau_{\max}}/4L(\tau)$ follows for all $\tau > \tau_\delta$ and each n, since there cannot lie any further local minimum of k in $[\varphi_n - \frac{\delta}{4}, \varphi_n + \frac{\delta}{4}]$. Thus also $k(\varphi_n, \tau_n) \geq \delta e^{-|K|\tau_{\max}}/4L(\tau_n)$ for all n with $\tau_n > \tau_\delta$; and eventually $\lim_{n\to\infty} k(\varphi_n, \tau_n) = \infty$ because of $\lim_{n\to\infty} \tau_n = \tau_{\max}$, $\lim_{\tau\to\tau_{\max}} L(\tau) = 0$ and the continuity of L. ∎

Lemma 5. *Let M^2 and C be as in the Theorem, additionally we assume that $C(0)$ has exactly two vertices.*

Then for each $\varepsilon > 0$ there exists a $\tau_\varepsilon < \tau_{\max}$, such that $\bar{C}(\tau)$ lies for all $\tau > \tau_\varepsilon$ in the ε-neighbourhood of a point or a geodesic segment on M^2.

Furthermore, in the case $K > 0$, there exists a $\tau_+ < \tau_{\max}$, such that $\bar{C}(\tau)$ lies in a hemisphere S_K^{2+} for all $\tau > \tau_+$.

Proof: We consider first the case, where $C(\tau)$ converges to a point for $\tau \to \tau_{\max}$. $\hat{k} := \lim_{\tau\to\tau_{\max}} k(\varphi_0(\tau), \tau) \in (-\infty, \infty]$ exists by Lemma 4.

If $k(\varphi_0(\tau), \tau)$ has the upper bound k_1, $\hat{k} < \infty$ holds and also $\hat{N} := \lim_{\tau\to\tau_{\max}} N(\varphi_0(\tau), \tau)$ exists by Lemma 3.

We further set

$$\begin{aligned}
P &:= \lim_{\tau\to\tau_{\max}} X(\varphi, \tau) = \lim_{\tau\to\tau_{\max}} C(\tau)\ , \\
S &:= \lim_{\tau\to\tau_{\max}} \bar{X}(\varphi_0(\tau), \tau) = \\
&= \lim_{\tau\to\tau_{\max}} \left[\frac{k(\varphi_0(\tau), \tau)}{\sqrt{K + k^2(\varphi_0(\tau), \tau)}}\, X(\varphi_0(\tau), \tau) + \frac{1}{\sqrt{K + k^2(\varphi_0(\tau), \tau)}}\, N(\varphi_0(\tau), \tau) \right] \\
&= \begin{cases} P, & \text{if } \hat{k} = \infty\,, \\ \frac{\hat{k}}{\sqrt{K + \hat{k}^2}}\, P + \frac{1}{\sqrt{K + \hat{k}^2}}\, \hat{N}, & \text{if } \hat{k} < \infty\ . \end{cases}
\end{aligned}$$

By PS we mean the geodesic segment or the shortest connection between P and S on M^2, respectively (in the case $K > 0$ PS lies in an open hemisphere of S_K^2), and by $U_\varepsilon(PS)$ the ε-neighbourhood of PS.

Now we treat the subcase, where $k(\varphi_0(\tau), \tau)$ has the upper bound k_1. By Lemma 3 $\varphi_0(\tau_{\max}) = \lim_{\tau\to\tau_{\max}} \varphi_0(\tau)$ exists.

Let $\varepsilon > 0$ be fixed. We assume, there is no τ_ε as mentioned. Then there exists a sequence $(\tau_n)_{n\in\mathbb{N}}$ with $\lim_{n\to\infty} \tau_n = \tau_{\max}$ and $\bar{C}(\tau_n) \nsubseteq U_\varepsilon(PS)$ for all n, i.e. we also find a sequence $(\varphi_n)_{n\in\mathbf{N}}$ such that $\bar{X}(\varphi_n, \tau_n) \notin U_\varepsilon(PS)$ for all $n \in \mathbf{N}$ is true.

If $(\varphi_n)_{n\in\mathbf{N}}$ has a subsequence $(\varphi_{n_m})_{m\in\mathbf{N}}$ with $\lim_{m\to\infty} \varphi_{n_m} = \varphi_0(\tau_{\max})$, then, due to the continuity of X and N, $\lim_{m\to\infty} X(\varphi_{n_m}, \tau_{n_m}) = P$ and $\lim_{m\to\infty} N(\varphi_{n_m}, \tau_{n_m}) = \hat{N}$. In general, however, $\lim_{m\to\infty} k(\varphi_{n_m}, \tau_{n_m}) = \hat{k}$ is not true, and so neither $\lim_{m\to\infty} \bar{X}(\varphi_{n_m}, \tau_{n_m}) = S$, since the limit function of k at $\varphi_0(\tau_{\max})$ does not have to be continuous.

But $k(\varphi_{n_m}, \tau_{n_m}) \geq k(\varphi_0(\tau_{n_m}), \tau_{n_m})$ holds for all m; and so

$$\begin{aligned} \lim\inf_{m\to\infty} k(\varphi_{n_m}, \tau_{n_m}) \;&\geq\; \lim\inf_{m\to\infty} k(\varphi_0(\tau_{n_m}), \tau_{n_m}) \\ &= \lim_{m\to\infty} k(\varphi_0(\tau_{n_m}), \tau_{n_m}) \;=\; \hat{k} \;. \end{aligned}$$

For each λ, $\hat{k} \leq \lambda \leq \infty$,

$$Y(\lambda) := \frac{\lambda}{\sqrt{K+\lambda^2}}\, P + \frac{1}{\sqrt{K+\lambda^2}}\, \hat{N}$$

is a point of the segment PS with $Y(\hat{k}) = S$ and $\lim_{\lambda\to\infty} Y(\lambda) = P$.

Now we set $\lambda := \lim\inf_{m\to\infty} k(\varphi_{n_m}, \tau_{n_m}) \in [\hat{k}, \infty]$. There exists a subsequence $(\varphi_{n_{m_l}}, \tau_{n_{m_l}})_{l\in\mathbf{N}}$ of $(\varphi_{n_m}, \tau_{n_m})_{m\in\mathbf{N}}$, with $\lim_{l\to\infty} k(\varphi_{n_{m_l}}, \tau_{n_{m_l}}) = \lambda$. By this we have

$$\begin{aligned} &\lim_{l\to\infty} \bar{X}(\varphi_{n_{m_l}}, \tau_{n_{m_l}}) \;= \\ &= \lim_{l\to\infty}\left[\frac{k(\varphi_{n_{m_l}}, \tau_{n_{m_l}})}{\sqrt{K+k^2(\varphi_{n_{m_l}}, \tau_{n_{m_l}})}}\, X(\varphi_{n_{m_l}}, \tau_{n_{m_l}}) + \frac{1}{\sqrt{K+k^2(\varphi_{n_{m_l}}, \tau_{n_{m_l}})}}\, N(\varphi_{n_{m_l}}, \tau_{n_{m_l}})\right] \\ &= \frac{\lambda}{\sqrt{K+\lambda^2}}\, P + \frac{1}{\sqrt{K+\lambda^2}}\, \hat{N} \;=\; Y(\lambda) \;\in\; PS \;, \end{aligned}$$

in contradiction to $\bar{X}(\varphi_{n_{m_l}}, \tau_{n_{m_l}}) \notin U_\varepsilon(PS)$ for all $l \in \mathbf{N}$.

So $(\varphi_n)_{n\in\mathbf{N}}$ cannot possess a subsequence as stated.

Hence a $\delta > 0$ and a $n_\delta \in \mathbf{N}$ exist, such that $\varphi_n \notin U_\delta(\varphi_0(\tau_{\max}))$ holds for all $n \geq n_\delta$. But then $\lim_{n\to\infty} k(\varphi_n,\tau_n) = \infty$ by Lemma 4, and thus $\lim_{n\to\infty} \bar{X}(\varphi_n,\tau_n) = P$, in contradiction to $\bar{X}(\varphi_n, \tau_n) \notin U_\varepsilon(PS)$ for all $n \in \mathbf{N}$.
So a τ_ε exists as required, and the first part of the Lemma is proven for this subcase.

In the second subcase, where $k(\varphi_0(\tau), \tau)$ does not have an upper bound, we have by Lemma 4 $\hat{k} = \lim_{\tau\to\tau_{\max}} k(\varphi_0(\tau), \tau) = \infty$.

In the following we will prove that $\bar{C}$ converges uniformly to P. We obtain

$$\begin{aligned}
\sup_\varphi \|P-\bar{X}\| &= \sup_\varphi \left\| P - \frac{k}{\sqrt{K+k^2}}\, X - \frac{1}{\sqrt{K+k^2}}\, N \right\| \\
&\le \sup_\varphi \left\| P - \frac{k}{\sqrt{K+k^2}}\, X \right\| + \sup_\varphi \frac{1}{\sqrt{K+k^2}} \\
\le \sup_\varphi \left\| P - \frac{k}{\sqrt{K+k^2}}\, P \right\| &+ \sup_\varphi \left\| \frac{k}{\sqrt{K+k^2}}\, P - \frac{k}{\sqrt{K+k^2}}\, X \right\| + \sup_\varphi \frac{1}{\sqrt{K+k^2}} \\
\le \|P\| \sup_\varphi \left| 1 - \frac{k}{\sqrt{K+k^2}} \right| &+ \sup_\varphi \frac{|k|}{\sqrt{K+k^2}} \cdot \sup_\varphi \|P-X\| + \sup_\varphi \frac{1}{\sqrt{K+k^2}}.
\end{aligned}$$

In the case $K>0$ we have $-1 < k/\sqrt{K+k^2} < 1$ for $-\infty < k < \infty$, and $k/\sqrt{K+k^2}$ is monotone increasing in k.

Thus

$$\sup_\varphi \left| 1 - \frac{k}{\sqrt{K+k^2}} \right| = 1 - \inf_\varphi \frac{k}{\sqrt{K+k^2}} \le 1 - \frac{\inf_\varphi k}{\sqrt{K+(\inf_\varphi k)^2}} \ ,$$

and we have $\lim_{\tau\to\tau_{\max}} \sup_\varphi \left|1 - k/\sqrt{K+k^2}\right| = 0$ by $\lim_{\tau\to\tau_{\max}} \inf_\varphi k(\varphi,\tau) = \lim_{\tau\to\tau_{\max}} k(\varphi_0(\tau),\tau) = \infty$. Additionally,

$$\lim_{\tau\to\tau_{\max}} \sup_\varphi \frac{|k|}{\sqrt{K+k^2}} \le 1$$

and

$$\lim_{\tau\to\tau_{\max}} \sup_\varphi \frac{1}{\sqrt{K+k^2}} \le \lim_{\tau\to\tau_{\max}} \frac{1}{\sqrt{K+(\inf_\varphi |k|)^2}} = 0 \ .$$

In the case $K<0$ we have $k/\sqrt{K+k^2} > 1$ for $k > \sqrt{-K}$ and $k/\sqrt{K+k^2}$ is monotone decreasing in k.

Here we see

$$\sup_\varphi \left| 1 - \frac{k}{\sqrt{K+k^2}} \right| = \sup_\varphi \frac{k}{\sqrt{K+k^2}} - 1 \le \frac{\inf_\varphi k}{\sqrt{K+(\inf_\varphi k)^2}} - 1 \ ,$$

and we obtain

$$\lim_{\tau\to\tau_{\max}} \sup_\varphi \left| 1 - \frac{k}{\sqrt{K+k^2}} \right| = 0$$

as above; as well as

$$\lim_{\tau\to\tau_{\max}} \sup_\varphi \frac{|k|}{\sqrt{K+k^2}} \le \lim_{\tau\to\tau_{\max}} \frac{\inf_\varphi |k|}{\sqrt{K+(\inf_\varphi |k|)^2}} = 1$$

and $\lim_{\tau\to\tau_{\max}} \sup_\varphi 1/\sqrt{K+k^2} = 0$.

For $K = 0$ we have

$$\sup_{\varphi}\left|1 - \frac{k}{\sqrt{K+k^2}}\right| \equiv 0\,, \qquad \sup_{\varphi}\frac{|k|}{\sqrt{K+k^2}} \equiv 1$$

and $\lim_{\tau\to\tau_{\max}} \sup_{\varphi} 1/\sqrt{K+k^2} = 0$.

Besides, in all cases follows $\lim_{\tau\to\tau_{\max}} \sup_{\varphi} \|P - X\| = 0$ from the convergence of $C(\tau)$ to P in the Hausdorff metric for $\tau \to \tau_{\max}$.

Altogether we now obtain

$$\lim_{\tau\to\tau_{\max}} \sup_{\varphi} \|P - \bar{X}\| \;=\; 0$$

and hence

$$\lim_{\tau\to\tau_{\max}} \mathrm{d}(P, \bar{C}(\tau)) \;=\; \lim_{\tau\to\tau_{\max}} \sup_{\varphi}\; \mathrm{d}(P, \bar{X}(\tau)) \;=\; 0\;.$$

This yields us the existence of $\tau_\varepsilon < \tau_{\max}$ for given $\varepsilon > 0$.

In order to find the wanted τ_+ for $K > 0$, we calculate the angle $\angle(P,S)$ between P and S on S^2_K as $\cos\angle(P,S) = \hat{k}/\sqrt{K+\hat{k}^2}$ and obtain from the montonicity of the occuring function and $-\infty < k_0 \le \hat{k} \le \infty$ (by Lemma 2) $-1 < k_0/\sqrt{K+k_0^2} \le \hat{k}/\sqrt{K+\hat{k}^2} \le 1$. Thus $0 \le \angle(P,S) < \pi$, hence PS lies in an open hemisphere of S^2_K, therefore also $U_\varepsilon(PS)$ for small ε and by the first part of the Lemma eventually also $\bar{C}(\tau)$ for $\tau > \tau_+ = \tau_\varepsilon$.

Now we treat the remaining case, where $\tau_{\max} = \infty$ holds and $C(\tau)$ converges to a large circle on S^2_K for $\tau \to \tau_{\max}$. Each large circle is the intersection of a plane in E^3 with S^2_K, the normal vector of the plane is then (with suitable orientation) the unit normal vector $\hat{N}$ along the large circle.

We will show the uniform convergence of $\bar{X}$ to $\hat{N}/\sqrt{K} \in S^2_K$: From

$$\begin{aligned}
\sup_{\varphi}\left\|\frac{\hat{N}}{\sqrt{K}} - \bar{X}\right\| \;&=\; \sup_{\varphi}\left\|\frac{\hat{N}}{\sqrt{K}} - \frac{k}{\sqrt{K+k^2}}\,X - \frac{1}{\sqrt{K+k^2}}\,N\right\| \\
&\le\; \sup_{\varphi}\frac{|k|}{\sqrt{K+k^2}}\,\|X\| + \sup_{\varphi}\left\|\frac{\hat{N}}{\sqrt{K}} - \frac{N}{\sqrt{K+k^2}}\right\| \\
\le\;\; & \frac{\sup_{\varphi}|k|}{\sqrt{K+(\sup_{\varphi}|k|)^2}}\cdot\frac{1}{\sqrt{K}} + \sup_{\varphi}\left\|\frac{\hat{N}}{\sqrt{K}} - \frac{N}{\sqrt{K}}\right\| + \sup_{\varphi}\left\|\frac{N}{\sqrt{K}} - \frac{N}{\sqrt{K+k^2}}\right\| \\
\le\;\; & \frac{\sup_{\varphi}|k|}{\sqrt{K}\sqrt{K+(\sup_{\varphi}|k|)^2}} + \frac{1}{\sqrt{K}}\sup_{\varphi}\left\|\hat{N} - N\right\| + \frac{1}{\sqrt{K}} - \frac{1}{\sqrt{K+(\sup_{\varphi}|k|)^2}}
\end{aligned}$$

results with $\lim_{\tau\to\infty} \sup_\varphi |k| = 0$ and $\lim_{\tau\to\infty} \sup_\varphi \|\hat{N} - N\| = 0$ eventually

$$\lim_{\tau\to\infty} \sup_\varphi \left\| \frac{\hat{N}}{\sqrt{K}} - \bar{X} \right\| = \frac{1}{\sqrt{K}} - \frac{1}{\sqrt{K}} = 0$$

and so

$$\lim_{\tau\to\infty} \mathrm{d}\left(\frac{\hat{N}}{\sqrt{K}},\ \bar{C}(\tau) \right) := \lim_{\tau\to\infty} \sup_\varphi \ \mathrm{d}\left(\frac{\hat{N}}{\sqrt{K}},\ \bar{X} \right) = 0\ .$$

By this method we obtain the wanted $\tau_\varepsilon < \infty$ and $\tau_+ = \tau_\varepsilon$ for small ε. ∎

5 – The proof of the Theorem

We assume that $C(0) = C$ has exactly two vertices, and we will bring this assumption to a contradiction.

As mentioned earlier (p. 271), all $C(\tau)$, $\tau > 0$, have exactly two vertices, which correspond to the two local extrema of k. Since k_φ changes sign at the extrema, the corresponding points of $\bar{C}$ are singularities, i.e. cusps in the sense, that $\bar{T}$ jumps by $\pm\pi$ (cf. (17)) and $\bar{C}$ does not have a unique tangent vector there. We call these singularities $\bar{S}_1 = \bar{S}_1(\tau)$ and $\bar{S}_2 = \bar{S}_2(\tau)$. Except for these points $\bar{C}(\tau)$ is smooth by [A1, Theorem 3.1], also at the inflection points, if these are not at the same time curvature extrema, cf. (17) and (9).

For the following let $0 < \tau < \tau_{\max}$ for $K \le 0$ or $\tau_+ < \tau < \tau_{\max}$ for $K > 0$, respectively (such that $\bar{C}(\tau)$ lies in a hemisphere S^{2+}_K by Lemma 5), arbitrary, but kept fixed.
We will investigate two cases: $\bar{S}_1 \ne \bar{S}_2$ and $\bar{S}_1 = \bar{S}_2$.

In the first case, $\bar{C}$ cannot be contained completely in the line F, which connects $\bar{S}_1$ and $\bar{S}_2$, because then we would have $\bar{k} \equiv 0$, which is not possible by (19).

Now we consider another line H, which shall not intersect F (for $K \le 0$), or shall have the same intersection points with the boundary of S^{2+}_K as F. If we move H towards F in a way, that the conditions above still hold, then H must touch the focal curve $\bar{C}$ in a nonsingular point (at least at one side of F), because otherwise nonsingualr points of $\bar{C}$ must exist outside of F.

In the second case $\bar{X} \equiv \bar{S}_1 = \bar{S}_2$ cannot hold, because this would imply $\bar{X}_\varphi \equiv 0$ and so $k_\varphi \equiv 0$. By this C would have infinitely many vertices, what is not possible. Thus there must also exist a line H, which touches $\bar{C}$ in a nonsingular point.

For both cases, let $\bar{Y}$ be the first point of contact of $\bar{C}$ with H, i.e. the first point of $\bar{C}$, H reaches (if there are more points with this property, we choose one of them).

$\bar{C}$ lies completely on one side of H; and due to $\bar{k} > 0$ (cf. (19)) the unit normal vector $\bar{N}(\bar{Y})$ of $\bar{C}$ at $\bar{Y}$ points on this side. Now we consider points $\bar{Z}_\delta := \exp_{\bar{Y}} \delta\bar{N}(\bar{Y})$ for $\delta > 0$. Then there must exist a δ, such that the winding number $w(\bar{Z}_\delta)$ of $\bar{Z}_\delta$ with respect to $\bar{C}$ is strictly positive, otherwise there must be for each δ a subarc of $\bar{C}$ between $\bar{Z}_\delta$ and $\bar{Y}$, which is traversed in the opposite direction as the subarc, on which $\bar{Y}$ lies, what is impossible due to $\bar{k} > 0$ and the first contact of H in $\bar{Y}$.

Hence there exists a nonempty domain $\bar{G} = \bar{G}(\tau)$ with $w|_{\bar{G}} \geq 1$ (w is taken with respect to $\bar{C}$) and $\bar{A} := \operatorname{area}(\bar{G}) = \int\int_{\bar{G}} dA > 0$.

Along the boundary $\partial\bar{G}$ of $\bar{G}$, $\bar{N}$ points in direction to $\bar{G}$ because of $\bar{k} > 0$ along $\partial\bar{G}$ and the increase of the winding number at crossing $\partial\bar{G}$ in direction to $\bar{G}$. By considering the Taylor expansion of $\bar{X} = \bar{X}(\varphi, \tau)$, for $\bar{X} \neq \bar{S}_i$, $(i = 1, 2)$, we see with (20) and (19), that $\bar{X}(\varphi, \tau + \varsigma)$, for $\varsigma > 0$ small (and dependent of φ), lies outside of $\operatorname{cl}\bar{G}(\tau) = \partial\bar{G}(\tau) \cup \bar{G}(\tau)$.

This is also true for the edges of $\partial\bar{G}$, which are at the same time self-intersections of $\bar{C}$.

At the singularities $\bar{S}_1, \bar{S}_2$ k_φ disappears, and the Taylor expansion of $\bar{X}$ has no part anymore in $\bar{N}$-direction. But under consideration of the continuity of $\bar{X}(\varphi, \tau)$ in both variables and the convexity of the curve arcs bordering on $\bar{S}_1, \bar{S}_2$ one can see, that $\bar{S}_i(\tau + \varsigma)$ $(i = 1, 2)$, for $\varsigma > 0$ small, cannot lie inside $\bar{G}(\tau)$.

By this $\bar{G}(\tau) \subseteq \bar{G}(\tau + \varsigma)$ follows for $\varsigma > 0$ small, and hence $\bar{A}(\tau) \leq \bar{A}(\tau + \varsigma)$. This means especially $\bar{A}(\tau_{\max}) \geq \bar{A}(\tau) > 0$.

But by Lemma 5 there is for each $\varepsilon > 0$ a $\tau_\varepsilon < \tau_{\max}$, such that $\bar{C}(\tau)$ for $\tau > \tau_\varepsilon$ and so also $\bar{G}(\tau)$ lie inside the ε-neighbourhood of a fixed segment (or a point, respectively). So we have $\lim_{\tau \to \tau_{\max}} \bar{A}(\tau) = 0$, in contradiction to $\bar{A}(\tau_{\max}) > 0$.

The assertion was wrong, therefore $C(0)$ must have at least three vertices. If this third vertex is only a saddle of the curvature, i.e. if k_φ does not change sign there, then this saddle disappears immediately, and $C(\tau)$ has only two vertices for $\tau > 0$. But then we get with our proof again a contradiction.

Hence the curvature of $C(0)$ has another local extremum; but since two local extrema of the same type cannot consecute, there must be another local extremum, which represents the fourth vertex. ∎

Remark. For surfaces of variable curvature one can in general not expect a four-vertex theorem: Each distance circle in a sufficiently small neighbourhood of a point with non-stationary Gauss curvature has exactly two vertices (see [J, Chapter 6]).

However, G. Thorbergsson proved for surfaces of nonpositive curvature a four-vertex theorem using a more general definition of a vertex (see [T]).

REFERENCES

[A1] Angenent, S. – Parabolic equations for curves on surfaces. Part I. Curves with p-integrable curvature, *Ann. Math.,* 132 (1990), 451–483.

[A2] Angenent, S. – Parabolic equations for curves on surfaces. Part II. Intersections, blow-up and generalized solutions, *Ann. Math.,* 133 (1991), 171–215.

[BF] Barner, M. and Flohr, F. – Der Vierscheitelsatz und seine Verallgemeinerungen, *Mathematikunterricht,* 4 (1958), 43–73.

[C] Carmo, M.P. do – *Riemannian Geometry*, Birkhäuser, Boston, MA, 1992.

[EGa] Epstein, C.L. and Gage, M.E. – The curve shortening flow. Wave motion: Theory, modelling, and computation, Proc. Conf. Hon. 60th birthday P.D. Lax, Publ., *Math. Sci. Res. Inst.,* 7 (1987), 15–59.

[Ga] Gage, M.E. – Curve shortening on surfaces, *Ann. Sci. Éc. Norm. Supér., IV. Sér.,* 23 (1990), 229–256.

[GaH] Gage, M.E. and Hamilon, R.S. – The heat equation shrinking convex plane curves, *J. Diff. Geom.,* 23 (1986), 69–96.

[Gr1] Grayson, M.A. – The heat equation shrinks embedded plane curves to round points, *J. Diff. Geom.,* 26 (1987), 285–314.

[Gr2] Grayson, M.A. – Shortening embedded curves, *Ann. Math.,* 129 (1989), 71–111.

[J] Jackson, S.B. – The four-vertex-theorem for surfaces of constant curvature, *Amer. J. Math.,* 67 (1945), 563–582.

[K] Kneser, A. – *Bemerkungen über die Anzahl der Extrema der Krümmung auf geschlossenen Kurven und über verwandte Fragen in einer nichteuklidischen Geometrie*, Festschrift H. Webers 70. Geb., Teubner 1912, 170–180.

[Ml] Müller, H.R. – *Sphärische Kinematik*, VEB, Berlin, 1962.

[Mu] Mukhopadhyaya, S. – New methods in the geometry of a plane arc, *Bull. Calcutta Math. Soc.,* 1 (1909), 31–37.

[O] Osserman, R. – The four or more vertex theorem, *Amer. Math. Monthly,* 92 (1985), 332–337.

[P] Pinkall, U. – On the four-vertex theorem, *Aequationes Math.,* 34 (1987), 221–230.

[T] Thorbergsson, G. – Vierscheitelsatz auf Flächen nichtpositiver Krümmung, *Math. Z.,* 149 (1976), 47–56.

Bernd Süssmann,
Robert Bosch GmbH,
Postfach 30 02 20, 70442 Stuttgart – GERMANY