Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 15–04

MULTIPLE SHORTEST PATHS ON CYLINDRICAL SURFACES IN PRE-HILBERT SPACES

EDUARDO MARQUES DE SÁ

Dedicated to Manuela Sobral on the occasion of her jubilee.

ABSTRACT: The concept of a 3D circular right cylinder is generalized to a real pre-Hilbert space. The surface $\mathscr{S}$ of such a cylinder has two pieces: the *wall* and the *septum* (or *cap*), whose intersection is the cylinder's *edge*. We study the intrinsic metric of $\mathscr{S}$ and discuss the existence, the nature and [non]uniqueness of shortest paths between two points in the cylinder's wall. The main problem addressed here is to determine the pairs of points in the wall for which there exist shortest paths going across the septum and shortest paths that do not cross the septum. We solve this problem in case one of the points lies in the edge, and show that this multiple shortest path problem is in essence a 3D riddle. Our methods involve the geometry of the traditional cycloid curve, its evolute and its negative pedal with respect to the cycloid cusp.

KEYWORDS: pre-Hilbert spaces, shortest paths, cycloid, negative pedal.
AMS SUBJECT CLASSIFICATION (2010): 46C05, 51M16, 51F, 14H50.

1. Introduction

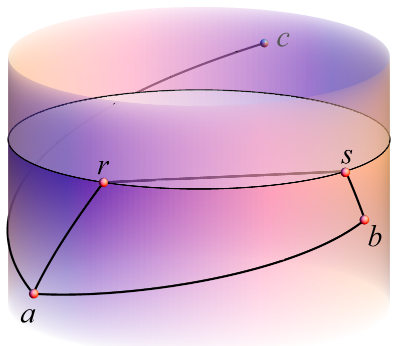

FIGURE 1

A cylinder in a real inner product space is a natural generalization of a three dimensional cylinder surface. Let us briefly discuss in the 3D case the problems to be addressed in this paper. Figure 1 represents a (right circular) 'cylinder surface' $\mathscr{S}$. This surface has two pieces: the 'wall', made up of the points at distance 1 from the axis, and the cylinder's 'septum', a circular unit disk orthogonal to the axis. The 'edge' of $\mathscr{S}$ is the relative boundary of the septum. We shall be concerned with the intrinsic metric of $\mathscr{S}$ in the sense of A. Aleksandrov (check [1, 2, 3]). Pick two points a, b in the cylinder's wall; the intrinsic distance between a and b, hereby denoted

Received February 11, 2015.
This work was partially supported by the Centre for Mathematics of the University of Coimbra (CMUC), funded by the European Regional Development Fund through the program COMPETE and by the Portuguese Government through the FCT under the project PEst-C/MAT/UI0324/2013.

$d_{\mathscr{S}}(a,b)$, turns out to be the length of a shortest path in $\mathscr{S}$ from a to b. There are two kinds of candidates to shortest paths in $\mathscr{S}$: a helicoidal shortest path across the wall, denoted $\widehat{ab}$, and the 'composite-paths', those that touch the septum, like the concatenation $\widehat{ar}[r,s]\widehat{sb}$ in figure 1, where $[r,s]$ is a straight line segment in the septum. Note that a helicoidal path (like $\widehat{ac}$) may meet the edge. If a and b are far enough from the cylinder's septum, then $\widehat{ab}$ is the unique shortest path; as a and b approach the septum, paths of the kind $\widehat{ar}[r,s]\widehat{sb}$ will eventually become shorter than $\widehat{ab}$. So, in certain critical positions of a,b, multiple shortest paths occur. The complete characterization of the pairs (a,b) for which we have multiple shortest paths seems to be a difficult problem that has not been solved yet, even in the 3D case.

In section 2 we treat the concept of a cylinder surface in a real pre-Hilbert space. The cylinder's axis is a subspace; we need completeness of the axis, or of its orthogonal space, to ensure the existence of orthogonal projections. Then we describe its intrinsic metric, and discuss the existence, the nature and [non]uniqueness of shortest paths between two points in the cylinder's wall, with special attention to composite-paths. Some results are within expectations, so this section is often sketchy. Section 3 is devoted to the crucial 3D case; our methods are classical in nature, going through the geometry of the traditional cycloid curve, its evolute and its negative pedal (with respect to the cycloid cusp), a curve which seems to have no name yet. For a fixed a in the cylinder's edge, we completely determine the locus of those points b in the wall from which we have multiple shortest paths to a. The case of a not in the edge is left open. In the last section we lift the three dimensional results to the real pre-Hilbert space case. The methods show that, in fact, the multiple shortest path problem is in essence a 3D riddle.

2. Cylinders in real pre-Hilbert spaces

Suppose V is a real vector space with inner product $\langle\cdot,\cdot\rangle$, and unit ball $\mathcal{B}=\{x : \|x\|\leqslant 1\}$. We assume V is a *pre-Hilbert space*, *i.e.*, the inner product is positive and non-degenerate. Fix a Hilbert subspace $\mathcal{H}\subset V$, such that $\dim\mathcal{H}^\perp > 1$. The *cylinder* with *axis* $\mathcal{H}$ is the set $(\mathcal{B}\cap\mathcal{H}^\perp)+\mathcal{H}$. The *septum* (or *cap*), the *edge* and the *wall* of the cylinder are the sets $\textsc{Cap}=\mathcal{B}\cap\mathcal{H}^\perp$, $\textsc{Edge}=\{x : \|x\|=1,\ x\perp\mathcal{H}\}$ and $\mathcal{W}=\mathcal{H}+\textsc{Edge}$, respectively. The *cylinder's surface* is the set $\mathscr{S}=\mathcal{W}\cup\textsc{Cap}$.

The orthogonal projections of V onto $\mathcal{H}$ and $\mathcal{H}^\perp$ are denoted by $\mathfrak{h}$ and $\mathfrak{e}$, respectively; these are continuous linear mappings such that any $a \in V$ is uniquely decomposed as $a = \mathfrak{h}(a) + \mathfrak{e}(a)$; it is well-known that $\mathfrak{h}(a)$ [$\mathfrak{e}(a)$] is the point of $\mathcal{H}$ [resp., $\mathcal{H}^\perp$] at a shortest distance from a (for the basics on these matters, see, *e.g.*, [4, ch.VI]). Note that if $a \in \mathcal{W}$, then $\mathfrak{e}(a) \in$ EDGE. For $a, b \in \mathcal{W}$, let $\omega(a, b)$ be the angle between $\mathfrak{e}(a)$ and $\mathfrak{e}(b)$:

$$\omega(a,b) = \arccos \langle \mathfrak{e}(a), \mathfrak{e}(b) \rangle = 2 \arcsin \tfrac{\|\mathfrak{e}(a-b)\|}{2}.$$

Given a set $\mathcal{U} \subseteq V$, a continuous injection $\gamma : [t_0, T] \to \mathcal{U}$, where $[t_0, T]$ is a real interval, is called *simple parametrized path in $\mathcal{U}$ from $\gamma(t_0)$ to $\gamma(T)$*. A *parametrized path* (shortened as *par-path*) is a concatenation of a finite number of (concatenable) simple par-paths. The concatenation of par-paths γ and ξ is denoted $\gamma\xi$. The image of a par-path is called *path*. If there is no danger of confusion, we use the same notation for a par-path and the corresponding path.

PATHS IN THE CYLINDER'S WALL. Given $a, b \in \mathcal{W}$ define

$$\eta(a,b) = \sqrt{\omega(a,b)^2 + \|\mathfrak{h}(a-b)\|^2}$$

We let $\mathscr{P}$ be the set of all partitions of $[t_0, T]$, where a *partition* is a finite tuple $P = (t_0, t_1, \ldots, t_n)$, with $t_0 < t_1 \cdots < t_n = T$. Given a par-path $\gamma : [t_0, T] \to \mathcal{W}$, from a to b, define $L(P, \gamma) = \sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\|$, and $L_\eta(P, \gamma) = \sum_{i=1}^n \eta(\gamma(t_i), \gamma(t_{i-1}))$; the *length* of γ is $\ell(\gamma) = \sup_{P \in \mathscr{P}} L(P, \gamma)$, and we let $\ell_\eta(\gamma) = \sup_{P \in \mathscr{P}} L_\eta(P, \gamma)$.

In case $0 < \omega(a, b) < \pi$, a natural candidate to shortest par-path is the par-path $\widehat{ab} : [0, \omega(a, b)] \to \mathcal{W}$ given by

$$\widehat{ab}\,(t) = \tfrac{\sin(\omega(a,b)-t)}{\sin \omega(a,b)}\, \mathfrak{e}(a) + \tfrac{\sin t}{\sin \omega(a,b)}\, \mathfrak{e}(b) + \mathfrak{h}\left(a + \tfrac{t}{\omega(a,b)}(b-a)\right). \tag{1}$$

Note that if v lies in the path $\widehat{ab}$, then $v = \widehat{ab}\,(\omega(a, v))$. If $\omega(a, b) = 0$, $[a, b]$ is a shortest path in $\mathscr{S}$; in such case we also denote $[a, b]$ by $\widehat{ab}$.

Theorem 2.1. *There exist shortest paths in $\mathcal{W}$ between any two points of $\mathcal{W}$. Moreover, η is the intrinsic metric on $\mathcal{W}$,* i.e., *$\eta = d_{\mathcal{W}}$.*

Proof. Firstly we show that for any par-path γ in $\mathcal{W}$, $\ell_\eta(\gamma) = \ell(\gamma)$. Clearly $\ell_\eta(\gamma) \geqslant \ell(\gamma)$. We may assume $0 < \ell(\gamma) < +\infty$. For $P \in \mathscr{P}$ as above, let

$\gamma_i = \gamma(t_i)$. As $\sqrt{x^2+H^2}-x$ is decreasing as a function of x, we get

$$L_\eta(P,\gamma) - L(P,\gamma) \leqslant \textstyle\sum_{i=1}^n \big(\omega(\gamma_i,\gamma_{i-1}) - \|\mathfrak{e}(\gamma_i - \gamma_{i-1})\|\big). \tag{2}$$

Now let $\varphi(x) = 2\arcsin\frac{x}{2} - x$. Clearly $\varphi(x) = x^3 g(x)$, with g analytic in $]-2,2[$; let $M = \max\{g(x) : 0 \leqslant x \leqslant 1\}$. Pick $\varepsilon \in]0,1[$, such that $\varepsilon M \ell(\gamma) < 1$. There exists $P \in \mathscr{P}$ such that $\|\gamma_i - \gamma_{i-1}\| < \varepsilon$ for all i. From (2) we get

$$L_\eta(P,\gamma) - L(P,\gamma) \leqslant \textstyle\sum_{i=1}^n \varphi(\|\mathfrak{e}(\gamma_i-\gamma_{i-1})\|) \leqslant M \sum_{i=1}^n \|\gamma_i - \gamma_{i-1}\|^3 < \varepsilon.$$

Therefore $\ell_\eta(\gamma) = \ell(\gamma)$, as desired. The existence of shortest paths from a to b in $\mathcal{W}$ is split into three cases. The case $\omega(a,b)=0$ is clear. In case $0 < \omega(a,b) < \pi$, the par-path (1) satisfies, for any $P \in \mathscr{P}$, $\omega(\widehat{ab}\,(t_i), \widehat{ab}\,(t_{i-1})) = t_i - t_{i-1}$ and $\mathfrak{h}(\widehat{ab}\,(t_i) - \widehat{ab}\,(t_{i-1})) = \frac{t_i - t_{i-1}}{\omega(a,b)}\,\mathfrak{h}(b-a)$. So $\ell(\widehat{ab}\,) = \eta(a,b)$. For any γ, $L_\eta(P,\gamma) = \sum_{i=1}^n \ell(\widehat{\gamma_{i-1}\,\gamma_i}) \geqslant \ell(\widehat{ab}\,)$. Therefore $\widehat{ab}$ is a shortest par-path.

Finally, the case $\omega(a,b) = \pi$. For any par-path γ from a to b, chose $c \in \operatorname{Im}\gamma$ such that $0 < \omega(a,c) < \pi$. Then γ is the concatenation of par-paths γ_{ac}, from a to c, and γ_{cb}, from c to b. We have $\ell(\gamma) = \ell(\gamma_{ac}) + \ell(\gamma_{cb}) \geqslant \eta(a,b)$. Now pick $p \in \textsc{Edge} \cap \mathfrak{e}(a)^\perp$, and let $m = p + \frac{1}{2}(a+b)$; the concatenation $\widehat{am}\,\widehat{mb}$ has length $\eta(a,b)$, and is therefore a shortest path. $\square$

Lemma 2.2. *Suppose $a,b,v \in \mathcal{W}$, $\omega(a,b) < \pi$, $\omega(a,v) < \pi$, $\omega(v,b) < \pi$. Then $\eta(a,b) = \eta(a,v) + \eta(v,b)$ implies that v is on the path $\widehat{ab}$.*

Proof. As the case $\omega(a,b) = 0$ is easy to handle, assume $\omega(a,b) > 0$. There exist $p,q \in \mathcal{H}^\perp$ such that $\|p\| = \omega(a,v)$, $\|q\| = \omega(v,b)$ and $\|p+q\| = \omega(a,b)$. Our assumption reads $\|p+q+\mathfrak{h}(a-b)\| = \|p+\mathfrak{h}(a-v)\| + \|q+\mathfrak{h}(v-b)\|$; this is equivalent to the existence of real nonnegative λ,μ, not both zero, such that $\lambda(p+\mathfrak{h}(a-v)) = \mu(q+\mathfrak{h}(v-b))$, *i.e.*: $\lambda p = \mu q$ and $\lambda\mathfrak{h}(a-v) = \mu\mathfrak{h}(v-b)$. Therefore, $\omega(a,b) = \omega(a,v) + \omega(v,b)$ and $\omega(a,v)\mathfrak{h}(v-b) = \omega(v,b)\mathfrak{h}(a-v)$. By 3D trigonometry applied to the spherical triangle with vertices $\mathfrak{e}(a), \mathfrak{e}(b), \mathfrak{e}(v)$, we get $\mathfrak{e}(v) \in \widehat{\mathfrak{e}(a)\mathfrak{e}(b)}$. Let $t^* = \omega(a,v)$; we have $\mathfrak{e}(v) = \widehat{\mathfrak{e}(a)\mathfrak{e}(b)}(t^*)$, and $\mathfrak{h}(v) = \mathfrak{h}\left(a + \frac{t^*}{\omega(a,b)}(b-a)\right)$. Therefore $v = \widehat{ab}\,(t^*)$. $\square$

Theorem 2.3.

(i) If $0 \leqslant \omega(a,b) < \pi$, $\widehat{ab}$ is the only shortest path in $\mathcal{W}$ from a to b.

(ii) If $\omega(a,b) = \pi$, for any $m \in \frac{1}{2}(a+b) + \textsc{Edge} \cap \mathfrak{e}(a)^\perp$, the concatenation $\widehat{am}\,\widehat{mb}$ is a shortest path in $\mathcal{W}$ from a to b, and all shortest paths in $\mathcal{W}$ from a to b are of this kind.

Proof. (i) is a trivial application of lemma 2.2. Concerning (ii), clearly $\widehat{am}\,\widehat{mb}$ is a shortest path. Let $\gamma : [t_0, T] \to \mathcal{W}$ be any shortest par-path from a to b; there exists $t^* \in]t_0, T[$ such that $\omega(a, \gamma(t^*)) = \frac{\pi}{2}$; let $m = \gamma(t^*)$. From (i), $\gamma([t_0, t^*]) = \operatorname{Im} \widehat{am}$ and $\gamma([t^*, T]) = \operatorname{Im} \widehat{mb}$; thus γ and $\widehat{am}\,\widehat{mb}$ determine the same path. It remains to prove that $\mathfrak{h}(m) = \frac{1}{2}(a+b)$. In fact, the identity $\eta(a, b) = \eta(a, m) + \eta(m, b)$ holds; expanding it and arguing as in the proof of lemma 2.2 we get $\mathfrak{h}(a-m) = \frac{1}{2}\mathfrak{h}(a-b)$. □

Composite-Paths. Given $a, b \in \mathcal{W}$, a *composite-path* from a to b is a path in $\mathcal{W} \cup \textsc{Cap}$ having at least one point in the cap. Concatenations of the kind $\widehat{ar}[r, s]\widehat{sb}$, with $r, s \in \textsc{Edge}$, are composite-paths, and a shortest composite-path must be a path like this. Note that

$$d_{\mathscr{S}}(a, b) = \inf\{\ell(\widehat{ab}), \inf_{r,s \in \textsc{Edge}} \ell(\widehat{ar}[r, s]\widehat{sb})\}.$$

Theorem 2.4. *Suppose that $\omega(a, b) > 0$, and $S := \|\mathfrak{h}(a)\| + \|\mathfrak{h}(b)\| > 0$. There exists a shortest composite-path from a to b and, if $\widehat{ar}[r, s]\widehat{sb}$ is a shortest composite-path, then $\operatorname{Span}(r, s, \mathfrak{e}(a), \mathfrak{e}(b))$ has dimension $\leqslant 2$. Moreover:*

- (i) *Case $\omega(a, b) < \pi$. There exist one or two shortest composite-paths from a to b. If there are two, one of them is of type $\Phi_1 = \widehat{ar}[r, s]\widehat{sb}$, with $r \neq s$, and the other is of type $\Phi_2 = \widehat{au}\widehat{ub}$, where $u \in \textsc{Edge}$ is uniquely determined by he conditions $\omega(a, u) = \frac{\|\mathfrak{h}(a)\|}{S}\omega(a, b)$ and $\omega(b, u) = \frac{\|\mathfrak{h}(b)\|}{S}\omega(a, b)$; if there is only one, then it is of one of the types Φ_1 or Φ_2 just described.*
- (ii) *Case $\omega(a, b) = \pi$. Let $H_\pi = \frac{\pi^2}{4} - 1$, and $\theta = \frac{\|\mathfrak{h}(a)\|}{S}\pi$. We have:*
  - $\alpha)$ *If $S < H_\pi$ there is a unique shortest composite-path from a to b, namely the polygonal line $[a, \mathfrak{e}(a), -\mathfrak{e}(a), b]$.*
  - $\beta)$ *If $S > H_\pi$ the shortest composite-paths are those of the form $\widehat{ar}\widehat{rb}$, where r runs over the set $\cos\theta\, \mathfrak{e}(a) + \sin\theta \big(\textsc{Edge} \cap \mathfrak{e}(a)^\perp\big)$.*
  - $\gamma)$ *If $S = H_\pi$ the shortest composite-paths are those in $\alpha)$-$\beta)$.*

Proof. Our task is to minimize $F(r, s) = \ell(\widehat{ar}[r, s]\widehat{sb})$, with $r, s \in \textsc{Edge}$. We have

$$F(r, s) = \alpha(\|\mathfrak{e}(a) - r\|) + \beta(\|\mathfrak{e}(b) - s\|) + \|s - r\|,$$

with $\alpha(x) = \sqrt{\left(2 \arcsin \frac{x}{2}\right)^2 + \|\mathfrak{h}(a)\|^2}$ and $\beta(x) = \sqrt{\left(2 \arcsin \frac{x}{2}\right)^2 + \|\mathfrak{h}(b)\|^2}$.

The introductory part of the theorem is proven separately in cases (i)-(ii).

Case (i). Let Σ be any finite dimensional subspace of $\mathcal{H}^\perp$ which contains $\mathrm{Span}(\mathfrak{e}(a),\mathfrak{e}(b))$. Let us firstly treat the case when one of $\mathfrak{h}(a),\mathfrak{h}(b)$ is 0; assume, *e.g.*, that $\mathfrak{h}(a)=0$, (*i.e.*, $a=\mathfrak{e}(a)$). By compactness, there exists $(\overline{r},\overline{s})$ minimizing $F(r,s)$ for $r,s\in\Sigma\cap\textsc{Edge}$. Then we must have $\overline{r}=\mathfrak{e}(a)$, and so $\overline{s}$ minimizes $\beta(\|s-\mathfrak{e}(b)\|)+\|s-\mathfrak{e}(a)\|$, for $s\in\Sigma$, under the constraint $\|s\|=1$. We obtain a contradiction from $\overline{s}\notin\mathrm{Span}(\mathfrak{e}(a),\mathfrak{e}(b))$. Introduce the functional $\Phi(s,\lambda)=F(\mathfrak{e}(a),s)-\lambda\|s\|$, where λ is a Lagrange multiplier; then $\overline{s}$ satisfies the first order conditions on Φ, namely

$$\tfrac{\beta'(\|\overline{s}-\mathfrak{e}(b)\|)}{\|\overline{s}-\mathfrak{e}(b)\|}\,(\overline{s}-\mathfrak{e}(b))+\tfrac{1}{\|\overline{s}-\mathfrak{e}(a)\|}\,(\overline{s}-\mathfrak{e}(a))-\lambda\overline{s}=0. \tag{3}$$

Our assumption $\overline{s}\notin\mathrm{Span}(\mathfrak{e}(a),\mathfrak{e}(b))$ implies $0<\|\overline{s}-\mathfrak{e}(a)\|<2$ and $0<\|\overline{s}-\mathfrak{e}(b)\|<2$; thus all terms of (3) are well defined, and (3) is a linear combination of $\overline{s},\mathfrak{e}(a),\mathfrak{e}(b)$, with nonzero coefficients in $\mathfrak{e}(a)$ and $\mathfrak{e}(b)$. This contradicts the linear independence of $\{\overline{s},\mathfrak{e}(a),\mathfrak{e}(b)\}$; therefore $\overline{s}\in\mathrm{Span}(\mathfrak{e}(a),\mathfrak{e}(b))$.

Now suppose $\mathfrak{h}(a),\mathfrak{h}(b)$ are both nonzero. There exists $\overline{r}$ which minimizes $F(r,r)$ for $r\in\Sigma\cap\textsc{Edge}$. We obtain a contradiction from $\overline{r}\notin\mathrm{Span}(\mathfrak{e}(a),\mathfrak{e}(b))$; arguing as above, and omitting details, $\overline{r}$ satisfies the first order conditions on the functional $\Psi(r,\lambda)=F(r,r)-\lambda\|r\|$, namely

$$\tfrac{\alpha'(\|\overline{r}-\mathfrak{e}(a)\|)}{\|\overline{r}-\mathfrak{e}(a)\|}\,(\overline{r}-\mathfrak{e}(a))+\tfrac{\beta'(\|\overline{r}-\mathfrak{e}(b)\|)}{\|\overline{r}-\mathfrak{e}(b)\|}\,(\overline{r}-\mathfrak{e}(b))-\lambda\overline{r}=0.$$

As this goes against the linear independence of $\{\overline{r},\mathfrak{e}(a),\mathfrak{e}(b)\}$, we must have $\overline{r}\in\mathrm{Span}(\mathfrak{e}(a),\mathfrak{e}(b))$. A similar argument shows that all minimizers of $F(r,-r)$, for $r\in\Sigma\cap\textsc{Edge}$, lie in $\mathrm{Span}(\mathfrak{e}(a),\mathfrak{e}(b))$.

Now assume (r^*,s^*) minimizes $F(r,s)$, for $r,s\in\Sigma\cap\textsc{Edge}$. If $r^*=\pm s^*$ the previous argument shows that $r^*,s^*\in\mathrm{Span}(\mathfrak{e}(a),\mathfrak{e}(b))$; if $r^*\neq\pm s^*$, the argument produced about (3) shows that $r^*\in\mathrm{Span}(\mathfrak{e}(a),s^*)$ and $s^*\in\mathrm{Span}(r^*,\mathfrak{e}(b))$; therefore $r^*,s^*\in\mathrm{Span}(\mathfrak{e}(a),\mathfrak{e}(b))$.

This establishes, in case (i), the existence of a global minimum of $F(r,s)$ in $\textsc{Edge}^2$, and that all minimizing r,s lie in the unit circle $\textsc{Edge}\cap\mathrm{Span}(\mathfrak{e}(a),\mathfrak{e}(b))$. An elementary 2D argument shows that a minimizing (r,s) must satisfy $r,s\in\widehat{\mathfrak{e}(a)\mathfrak{e}(b)}$ and $r\in\widehat{\mathfrak{e}(a)\,s}$; in other words, $\omega(a,b)=\omega(a,r)+\omega(r,s)+\omega(s,b)$. Using the notation $x=\omega(a,r)$, $y=\omega(s,b)$, $z=\omega(r,s)$, our problem is minimizing

$$f(x,y,z)=\sqrt{x^2+\|\mathfrak{h}(a)\|^2}+\sqrt{y^2+\|\mathfrak{h}(b)\|^2}+2\sin\frac{z}{2}, \tag{4}$$

subject to $x+y+z=\omega(a,b)$. Without loss of generality we assume $\mathfrak{h}(a)\neq 0$. A minimizing (x,y,z) satisfies $y\|\mathfrak{h}(a)\|=x\|\mathfrak{h}(b)\|$; thus (4) transforms into

$$f_1(x)=\tfrac{S}{\|\mathfrak{h}(a)\|}\sqrt{x^2+\|\mathfrak{h}(a)\|^2}+2\sin\tfrac{1}{2}\left(\omega(a,b)-\tfrac{S}{\|\mathfrak{h}(a)\|}\,x\right). \tag{5}$$

For $x=0$ or $x=\omega(a,b)\frac{\|\mathfrak{h}(a)\|}{S}$, the derivative of f_1 is negative; for x in the open interval $]0,\omega(a,b)\frac{\|\mathfrak{h}(a)\|}{S}[$ the equation $f_1'(x)=0$ is equivalent to

$$x=\|\mathfrak{h}(a)\|\cot\tfrac{1}{2}\left(\omega(a,b)-\tfrac{S}{\|\mathfrak{h}(a)\|}x\right); \tag{6}$$

due to the strict convexity of $\cot\theta$ in $]0,\frac{\pi}{2}]$, $f_1'(x)=0$ has at most two solutions in $[0,\omega(a,b)\frac{\|\mathfrak{h}(a)\|}{S}]$. So the minimum of $f_1(x)$ is either $f_1\left(\omega(a,b)\frac{\|\mathfrak{h}(a)\|}{S}\right)$, or $f(x_0)$ for a uniquely determined $x_0\in]0,\omega(a,b)\frac{\|\mathfrak{h}(a)\|}{S}[$; the former case corresponds to the triple $(x,y,z)=\left(\omega(a,b)\frac{\|\mathfrak{h}(a)\|}{S},\omega(a,b)\frac{\|\mathfrak{h}(b)\|}{S},0\right)$, and the latter to a triple (x_0,y_0,z_0) with $z_0>0$. So (i) follows at once.

Case (ii). Note that $\mathfrak{e}(b)=-\mathfrak{e}(a)$. The proof of (i) may be easily adapted to prove that if (r',s') minimizes $F(r,s)$ for $r,s\in\Sigma\cap\textsc{Edge}$, then $\{r',s',\mathfrak{e}(a)\}$ is linearly dependent. Now take any 2-space $\Sigma\subseteq\mathcal{H}^\perp$ containing $\mathfrak{e}(a)$; among the composite-paths $\widehat{ar}[r,s]\widehat{sb}$ whose $\mathfrak{e}$-projected images fall inside Σ, shortest composite-paths exist; obviously the lengths of these paths do not depend on Σ; therefore shortest composite-paths exist.

Pick $p\in\textsc{Edge}\cap\mathfrak{e}(a)^\perp$, let $\Sigma_p=\mathrm{Span}(p,\mathfrak{e}(a))$ and seek for shortest composite-paths $\widehat{ar}[r,s]\widehat{sb}$ with $r,s\in\Sigma_p\cap\textsc{Edge}$. Obviously, at a minimum of $F(r,s)$, $\langle r,p\rangle$ and $\langle s,p\rangle$ cannot have opposite signs; we choose the case where these numbers are both nonnegative (the non-positive case is handled with p replaced by $-p$). Moreover, we must have $\omega(a,r)+\omega(r,s)+\omega(s,b)=\pi$. So we may argue as in (i) till the point of minimizing $f_1(x)$, as in (5), in the interval $[0,\theta]$. In the current case, the minimum of $f_1(x)$ is the minimum of $f_1(0)=S+2$ and $f_1(\theta)=\sqrt{\pi^2+S^2}$. Note that $x=0$ corresponds to the polygonal line $[a,\mathfrak{e}(a),-\mathfrak{e}(a),b]$, and $x=\theta$ corresponds to the composite-path $\widehat{ar}\,\widehat{rb}$, with $r=\cos\theta\;\mathfrak{e}(a)+\sin\theta\;p$. The rest is obvious. $\square$

We present a nice, expected geometrical characteristic of a shortest composite-path $\widehat{ar}[r,s]\widehat{sb}$, in case $\widehat{ar}\in]0,\pi[$, $r\neq s$ and $\mathfrak{e}(a),\mathfrak{e}(b)$ are not zero. After introducing a Lagrange multiplier, the first order conditions of (4), subject to $x+y+z=\omega(a,b)$, are

$$\frac{x}{\sqrt{x^2+\|\mathfrak{h}(a)\|^2}}=\frac{y}{\sqrt{y^2+\|\mathfrak{h}(b)\|^2}}=\cos\frac{z}{2}. \tag{7}$$

As $x=\omega(a,r)$, it is easy to see that $u=\cot x\ r-\csc x\ \mathfrak{e}(a)$ is a unit tangent vector to the unit circle $U=\operatorname{Span}(\mathfrak{e}(a),r)\cap$ EDGE, at r, oriented in the direction from a to r; from (1) one easily gets that $\frac{d}{dt}\widehat{ar}(t)$ is, for $t=x=\omega(a,r)$: $w=u-\frac{1}{x}\ \mathfrak{h}(a)$. Hence the left hand side of (7) equals $\cos\alpha$, with α the angle between u and w, *i.e.*, the angle the path $\widehat{ar}$ makes with the circle U at r. If β denotes the angle $\widehat{sb}$ makes with U at s, then (7) reads: $\alpha=\beta=\frac{z}{2}$.

In figure 2 we give a planar drawing of a composite-path; points $a,b,r,s,\mathfrak{e}(a),\mathfrak{e}(b)$ have planar images labeled, respectively, A,B,R,S,A^*,B^*; moreover, EDGE$\cap$Span$(\mathfrak{e}(a),\mathfrak{e}(b))$ is represented by a unit circle where R,S lie at an angular distance z; then A^* is plotted so that $[A^*R]$ has length x and is tangent to the circle at R; then A is such that $\measuredangle AA^*R$ is a right angle and $\overline{AA^*}=\|\mathfrak{h}(a)\|$; points B^*,B are planted in a similar manner. Clearly $\ell(\widehat{ar})=\overline{AR}$, $\ell(\widehat{sb})=\overline{SB}$, and $\ell(\widehat{rs})=\overline{RS}$; moreover, $\alpha=\measuredangle ARA^*$ and $\beta=\measuredangle BSB^*$. So the polygonal line $[ARSB]$ is a nice image of the composite path $\ell(\widehat{ar}[r,s]\widehat{sb})$. Note that the cylinder's wall is not unfolded as usual, but rather in a deformed manner to show in the same "cubist view" the contacts septum-wall in both R and S. The punch line to this is that the first order conditions, $\alpha=\beta=\frac{z}{2}$, mean that A,R,S,B are colinear.

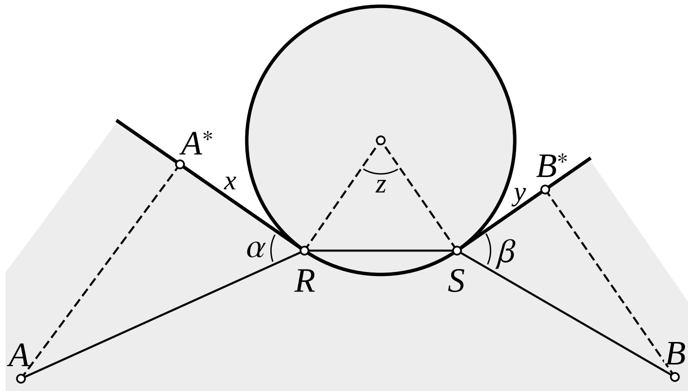

FIGURE 2

3. A critical 3D problem

In case (i) of theorem 2.4 some more information can be extracted from the proof. For $S>2$, equation (6) has no roots (in the interval $]0,\omega(a,b)\frac{\|\mathfrak{h}(a)\|}{S}[$); for $S\leqslant 2$, a unique, double root exists iff $\omega(a,b)=2\arcsin\sqrt{\frac{S}{2}+\sqrt{S(2-S)}}$; in all these cases, a unique shortest composite-path exists and it is of type Φ_2. So shortest composite-paths of type Φ_1 may only exist in case $S<2$

and $\omega(a,b) > 2\arcsin\sqrt{\frac{S}{2}} + \sqrt{S(2-S)}$; in this case there exist two distinct solutions of $f_1'(x) = 0$, say $x_0 < x_1$; then f_1 has a local minimum at x_0, and a maximum at x_1. The critical situation where two shortest composite-paths exist is characterized by the following three conditions:

$$f_1'(x) = 0, \quad f_1''(x) > 0, \quad f_1(x) = f_1\left(\omega(a,b)\frac{\|\mathfrak{h}(a)\|}{S}\right),$$

a system numerically tractable, but from which interesting theoretical information seems to be difficult to extract. However, in case one of a, b lies in Edge, a geometrical approach may be used to handle the critical situation.

The problem will be treated in the 3D case, on a conspicuous cylinder surface as in figure 1. As we are going to fix a in te cylinder's edge, we may eliminate the upper part of the wall, so that the circular septum is now more like a cap.

Points of the cylinder will be identified by cylindrical coordinates. Denote by c the center of the cap, so that $c = (0,0,0)$; the z-axis is the cylinder's axis with top-to-down orientation. To measure angles we fix a point o in the edge. Any point v in the cylinder's wall has an orthogonal projection v^* on the cap's plane; clearly $v^* = \mathfrak{e}(v)$. The distance of v to the cap is denoted by h_v (so $h_v = \|\mathfrak{h}(v)\|$), and the *polar angle* of v, denoted by ω_v, is the angle the half-line $\dot{c}v^*$ makes with $\dot{c}o$, counted positively counter-clockwise, when the upper cap is viewed from top-to-down. Note that $0 \leqslant \omega_v < 2\pi$, and $\omega_v = \omega(o,v)$ if $\omega_v \leqslant \pi$.

3.1. The cycloid approach. From now on, a is in the edge, and we choose $o = a$. So $h_a = \omega_a = 0$. We consider points b such that $h_b > 0$ and $\omega_b \in]0,\pi]$. Clearly shortest paths from a to b have the form $[a,s]\widehat{sb}$, with $s \in \widehat{ab^*}$. For each choice of s, we unfold the cylinder in a traditional way as figure 3 shows; the 2D-image of a 3D point is labeled by the same capitalized letter; point O, the planar image of $o = a$, is the origin of our 2D reference system. The cylinder's edge is mapped into a horizontal Ox-axis where angles 'ω' are marked, oriented from left to right; the Oy-axis has down-to-top orientation; the planar image of b is $B = (\omega_b, -h_b)$ also denoted by $B = (\omega_B, -h_B)$. The edge point s has planar image $S_t = (t,0)$, where t is the polar angle of s. The cylinder's cap is represented by a unit circle touching the Ox axis at S_t.

The unfolding depends on the choice of s, *i.e.*, on its polar angle t, and therefore the unfolded image of any point of the cylinder's cap depends on t;

accordingly, the images of c and a are denoted with a subindex t. As t varies,

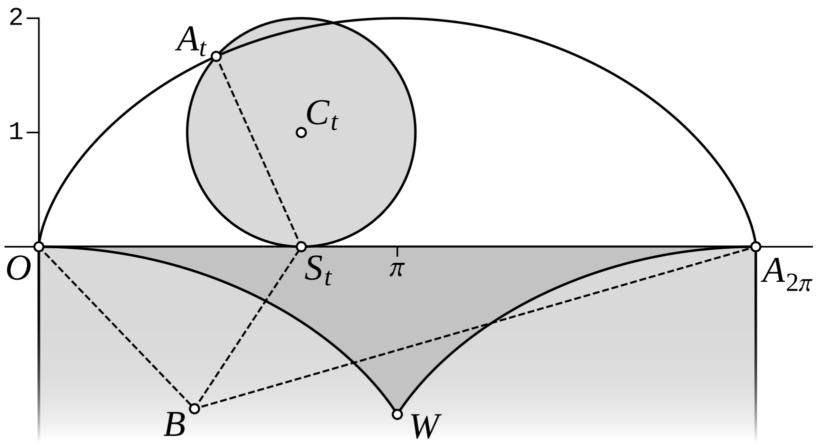

FIGURE 3. The bat-shaped curve is the critical curve.

the circle of figure 3 rolls on the Ox axis without slipping, and A_t describes the cycloid arc $\mathfrak{A}$ with parametric representation

$$\mathfrak{A}: \quad A_t = \big(t - \sin t, 1 - \cos t\big), \quad \text{with } 0 \leqslant t \leqslant 2\pi. \tag{8}$$

This arc has three prominent points: the *left cusp* $A = O$, the *right cusp* $A_{2\pi} = (2\pi, 0)$, and the *top point* $(\pi, 2)$.

Definition 3.1. For each $t \in [0, 2\pi]$, define $\mathfrak{p}_t$ as the composite-path from $a = o$ to b whose planar image is the polygonal line $[A_t S_t B]$. For a point u in the cylinder's wall and its planar image U, denote by Z_u or Z_U the circle with center U and radius $\operatorname{dist}(U, \mathfrak{A})$.

Note that $\mathfrak{p}_0$ and $\mathfrak{p}_{2\pi}$ are helicoidal paths on the wall, whose planar images are the segments $[BO]$ and $[BA_{2\pi}]$ (dashed in figure 3).

Theorem 3.2. *The distance from b to a in the intrinsic metric of $\mathscr{S}$ is the Euclidean distance from B to the cycloid $\mathfrak{A}$. The shortest paths from a to b are the $\mathfrak{p}_\mu$ such that A_μ is a point of contact of Z_B with $\mathfrak{A}$.*

Proof. For any $t \in [0, 2\pi]$, we have $\ell(\mathfrak{p}_t) = \ell[A_t S_t B]$. Therefore $\ell(\mathfrak{p}_t) \geqslant \|B - A_t\|$, and so $d_{\mathscr{S}}(B, A) \geqslant \operatorname{dist}(B, \mathfrak{A})$.

Now let $A_\mu \in \mathfrak{A} \cap Z_B$. Then $A_\mu - B$ is normal to $\mathfrak{A}$ at A_μ; by a well-known property of the cycloid (due to R. Descartes, *cf.* [6, p. 135]), $A_\mu - S_\mu$ is also normal to $\mathfrak{A}$ at A_μ, and so S_μ lies in $[BA_\mu]$. (All this trivially holds if A_μ is a cusp, because then $S_\mu = A_\mu$, and any vector is orthogonal to $\mathfrak{A}$ at a cusp.)

This proves $d_{\mathscr{S}}(b,a) = \mathrm{dist}(B, \mathfrak{A})$, and also that $\mathfrak{p}_\mu$ is a shortest path. The proof that $\mathfrak{p}_t$ is a shortest path implies $A_t \in \mathfrak{A} \cap Z_B$ is now obvious. $\square$

Part of the following theorem can be obtained from theorem 2.4. However, we give an independent treatment based on the functional

$$f(t,\omega,h) = (t - \sin t - \omega)^2 + (1 - \cos t + h)^2 - \omega^2 - h^2. \tag{9}$$

Clearly, $f(t, \omega_B, h_B) = 0$ means that A_t and the left cusp of $\mathfrak{A}$ are equidistant from the point B.

Theorem 3.3. *$Z_B \cap \mathfrak{A}$ consists of one, two or three points. More specifically, there exists a function $H :]0, \pi] \to \mathbb{R}$, such that $H(\pi) = \frac{\pi^2}{4} - 1$ and:*

- (α) *Case $0 < \omega_b < \pi$:*
  - α_1) *If $h_b < H(\omega_b)$, then $Z_B \cap \mathfrak{A}$ is a singleton, namely a point A_{t_0}, such that $0 < t_0 < \pi$.*
  - α_2) *If $h_b > H(\omega_b)$, then $Z_B \cap \mathfrak{A}$ is a singleton, namely $\mathfrak{A}$'s left cusp.*
  - α_3) *If $h_b = H(\omega_b)$, then Z_B has a double contact with $\mathfrak{A}$, namely the left cusp of $\mathfrak{A}$, and a regular point as in α_1).*
- (β) *Case $\omega_b = \pi$:*
  - β_1) *If $h_b < H(\pi)$, $Z_B \cap \mathfrak{A}$ is a singleton, namely the top point of $\mathfrak{A}$;*
  - β_2) *If $h_b > H(\pi)$, $Z_B \cap \mathfrak{A}$ is a doubleton, namely the two cusps of $\mathfrak{A}$;*
  - β_3) *If $h_b = H(\pi)$, $Z_B \cap \mathfrak{A}$ is a tripleton, namely the points in β_1)-β_2).*

Proof. The derivative $\partial_t f$ may be given the form

$$\partial_t f(t,\omega,h) = 2h(1 - \cos t)\left[\tfrac{t-\omega}{h} + \cot\tfrac{t}{2}\right]. \tag{10}$$

(α) Assume $A_\xi \in Z_B$. We cannot have $\xi > \pi$, otherwise the point $A_{2\pi-\xi}$, symmetrically located with respect to the symmetry axis of $\mathfrak{A}$, would lie strictly inside Z_B contradicting the definition of Z_B; so $\xi \leqslant \pi$. On the other hand, $f(t, \omega_b, h_b)$ has a minimum at $t = \xi$, and so $\partial_t f(\xi, \omega_b, h_b) = 0$; this implies $\xi \neq \pi$. Therefore, $0 \leqslant \xi < \pi$.

The function $\cot\frac{t}{2}$ is strictly convex and strictly decreasing in $]0, \pi[$, and goes to $+\infty$ as $t \to 0^+$. As a consequence, there exists a positive ε such that $\partial_t f(t, \omega_b, h_b) > 0$ for $0 < t < \varepsilon$. Moreover, the equation

$$\tfrac{t-\omega_b}{h_b} + \cot\tfrac{t}{2} = 0 \tag{11}$$

represents the intersection of the straight line L with equation $y = \frac{\omega_b - t}{h_b}$, with the graph G of $\cot\frac{t}{2}$; therefore (11), and a fortiori $\partial_t f(t, \omega_b, h_b) = 0$, has at most two roots in the interval $]0, \pi[$. The line L passes in $(\omega_b, 0)$; when the slope of L is such that L is tangent to G (*i.e.*, (11) has a double root) easy computations show that the value θ of t in the contact point of L and G is implicitly given by $\theta + \sin\theta = \omega_b$; there is a unique such θ, which is positive and independent of h_b, and we have

$$\text{When (11) has two roots in }]0,\pi[, \text{ the largest root is } > \theta. \tag{12}$$

Suppose $f(t, \omega_B, h_B)$ has a minimum at $t = \xi > 0$; then it must have a local maximum at a positive $t' < \xi$; therefore, ξ *is the largest root of* (11) *in* $]0,\pi[$.

For a given $0 < \omega \leqslant \pi$, define $H(\omega)$ as the infimum of the set

$$\{h \geqslant 0 : f(t,\omega,h) \geqslant 0, \quad \text{for all } t \in]0,\pi]\}. \tag{13}$$

Note that, for $\omega = \omega_B$, h_B lies in (13) iff the left cusp of $\mathfrak{A}$ lies in Z_B. As

$$\partial_t f(t,\omega,0) = 4(t-\omega)\sin^2 \tfrac{t}{2}$$

is negative for $0 < t < \omega$, we have $H(\omega) > 0$. The set (13) is obviously closed; to see it is nonempty, note that the derivative of $\cot\frac{t}{2}$ is $-\frac{1}{2}$ at $t = 0$; so, if $h \geqslant 2$, no $t \in]0,\pi[$ zeroes out $\partial_t f(t,\omega,h)$; therefore any $h \geqslant 2$ lies in (13).

As $\partial_h f(t,\omega,h) = 2(1-\cos t)$, $f(t,\omega,h)$ is strictly increasing with h; from this fact, α_1) and α_2) follow at once.

To prove α_3), let $h_B = H(\omega_B)$, and consider a sequence of points (B_k) such that $h_{B_k} < H(\omega_B)$, $\omega_{B_k} = \omega_B$ and (h_{B_k}) converges to $H(\omega_B)$. By α_1), let A_{t_k} be the unique point of $Z_{B_k} \cap \mathfrak{A}$. By continuity (A_{t_k}) converges to the point A_τ of $Z_B \cap \mathfrak{A}$, where $\tau = \lim_k t_k$; by (12), $t_k > \theta$ for all k; this implies $\tau > 0$, and so A_τ is not the left cusp.

(β) Let $\omega_B = \pi$. The equation $\partial_t f(t,\pi,h_B) = 0$ has 5 roots in $[0, 2\pi]$, symmetrically placed with respect to π, namely, in increasing order, $0, t_1, \pi, t_2, 2\pi$, where $t_1 + t_2 = 2\pi$. Arguing as in (α), we now get: if $f(t,\pi,h_B)$ has a minimum at $\xi \in]0,2\pi[$, then $\xi = \pi$. So the two cusps and the top point of $\mathfrak{A}$ are the only candidates to points of contact of Z_B with $\mathfrak{A}$. The value $H(\pi) = \frac{\pi^2}{4} - 1$ arises naturally, because the circle with center $(\pi, -H(\pi))$ and radius $H(\pi) + 2$ is the only one passing in the cusps and the top point of $\mathfrak{A}$. The conclusion is left to the reader. □

3.2. The critical curve. The graph of the function $\omega \mapsto -H(\omega)$, called the *critical curve*, is the unfolded representation of the set of points b in the cylinder's wall having multiple shortest paths to $a = o$. That curve arises from the solutions of the following non-linear system

$$\begin{cases} f(t,\omega,h) = 0 \\ \partial_t f(t,\omega,h) = 0 \qquad 0 < t,\omega \leqslant \pi \text{ and } h \geqslant 0, \end{cases} \tag{14}$$

where $f(t,\omega,h)$ is the functional introduced in (9). As a matter of fact, a solution (t,ω,h) of (14), with positive t,ω,h and $\omega < \pi$, has the following interpretation; the equation $f(t,\omega,h) = 0$ says that the circle Z centered at $(\omega,-h)$ and passing in the left cusp of $\mathfrak{A}$ passes in A_t; on the other hand, $\partial_t f(t,\omega,h) = 0$ asserts that the circle Z is tangent to $\mathfrak{A}$ at A_t. Therefore $H(\omega) = h$, which means that $(\omega,-h)$ is a point of the critical curve.

Theorem 3.3 has a symmetric counterpart with respect to the symmetry axis of $\mathfrak{A}$. That twin theorem handles the case of points B with $\pi \leqslant \omega_B < 2\pi$; these points are, so to speak, under the jurisdiction of the right hand side cusp of $\mathfrak{A}$; in fact, if $\omega_B > \pi$, B is closer to the right cusp than to the left cusp. The role of the function $H :]0,\pi] \to \mathbb{R}$ is then played by a reflected twin $H^r : [\pi, 2\pi[\to \mathbb{R}$ given by $H^r(\omega) = H(2\pi - \omega)$.

In figure 3 the join of these two critical curves is represented with junction point $W = (\pi, -H_\pi)$; the bat-shaped darkened area is made up of those B for which some regular point of $\mathfrak{A}$ is a closest point of $\mathfrak{A}$ from B.

Theorem 3.4. *The function $\omega \mapsto H(\omega)$ is analytic at any $\omega \in]0,\pi]$.*

Proof. The functional $f(t,\omega,h)$ is analytic, so we only need to prove that its Jacobian determinant with respect to the variables t,h is nonzero at any solution (t,ω,h), with positive t,ω,h and $\omega \leqslant \pi$ (check, *e.g.*, [4, X§2]). A straightforward calculation leads to the value of the Jacobian

$$J(t,\omega,h) := \begin{vmatrix} \partial_t f & \partial_h f \\ \partial_t\partial_t f & \partial_h\partial_t f \end{vmatrix} = 8(h + \cos t - 1)\sin^2\frac{t}{2}.$$

The case $\omega = \pi$ is easy to handle, because the (unique) solution is explicitly known: $(t,\omega,h) = (\pi,\pi,\frac{\pi^2}{4} - 1)$.

We now consider the case $\omega < \pi$. We have to show that, if (t,ω,h) is a solution of (14), then $h \neq 1 - \cos t$. For this purpose we go back to what has been said about (11)-(12). As (t,ω,h) is a solution, t is the largest root of $\frac{t-\omega}{h} + \cot\frac{t}{2} = 0$ in the interval $]0,\pi[$. We now get a contradiction from $h = 1 - \cos t$; in fact, this hypothesis combined with the previous equation gives us $\frac{t-\omega}{1-\cos t} + \cot\frac{t}{2} = 0$, which may be transformed into

$$\tfrac{1}{2}(t + \sin t - \omega)\csc^2\tfrac{t}{2} = 0.$$

Thus we get $t+\sin t=\omega$. However, this contradicts what (12) says, namely, that $t>\theta$, where θ is the unique such that $\theta+\sin\theta=\omega$. $\square$

The fact that H is analytic at π means, obviously, that H can be analytically extended to an open neighborhood of π. Note, however, that we don't need that definition of $H(\omega)$ for $\omega>\pi$; the planar images of cylinder points with ω beyond π fall into the jurisdiction of the right cusp of $\mathfrak{A}$.

We now sketch an alternative approach that will produce an explicit parametrization of the critical curve. Define $\mathcal{P}=\{\frac{1}{2}A_t : 0<t\leqslant\pi\}$; this is the left half of the cycloid arc $\mathfrak{A}$ reduced by a scale factor $1/2$, with the origin O dropped out. For each point $P\in\mathcal{P}$, let L_P be the straight line through P orthogonal to OP; we denote by $\mathcal{N}$ the envelope of the family $\{L_P : P\in\mathcal{P}\}$. In the literature (for general curves), $\mathcal{P}$ is known as the *pedal* of $\mathcal{N}$, and $\mathcal{N}$ as the *negative pedal* (or *orthocaustic*) of $\mathcal{P}$ with respect to O (*cf.* [6, vol. III, pp. 250-252], [5, pp.157-9], [7, pp.153,182]).

Theorem 3.5. *$\mathcal{N}$ is the critical curve.*

Proof. The argument is sketched in figure 4, showing the main cycloid where A_t runs, the halved cycloid arc $\mathcal{P}$ where $P_t=\frac{1}{2}A_t$ runs, and the negative pedal $\mathcal{N}$. P_tN_t is the mediatrix of $[OA_t]$, and is tangent to $\mathcal{N}$ at N_t. The point $Q_t:=N_t-P_t$ completes a rectangle $[OP_tN_tQ_t]$; while P_t describes the pedal of $\mathcal{N}$, Q_t describes the so-called *contrapedal* of $\mathcal{N}$ (with respect to O). Theorem II of [7, p. 151] says that P_tQ_t is normal to $\mathcal{P}$ at P_t, and so N_tA_t is normal to $\mathfrak{A}$ at A_t. Therefore the circle centered at N_t and passing through O is tangent to $\mathfrak{A}$ at A_t. This is known in a much more general setting. The peculiarity of the cycloid, and punch line of the proof is that the referred circle centered at N_t is the Z_{N_t} of definition 3.1; the details to show this are more or less the same as expanded around (11)-(12). So N_t lies in the critical curve. $\square$

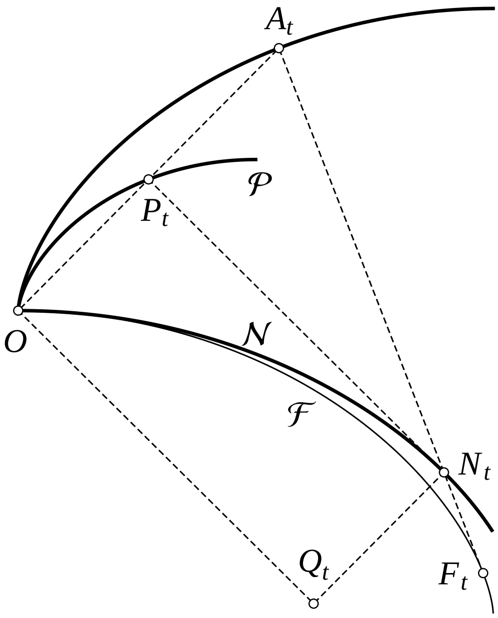

FIGURE 4

In figure 4, the thin curve $\mathcal{F}$ is an arc of the evolute of $\mathfrak{A}$, which is a translate of $\mathfrak{A}$ (see, *e.g.*, [6, vol. II, p. 135]). The figure shows that $\mathcal{F}$ lies bellow $\mathcal{N}$, more precisely, $N_t\in]A_tF_t[$, where F_t is the center of curvature of

$\mathfrak{A}$ at A_t (recall that A_tF_t is tangent to $\mathcal{F}$ at F_t). A short proof of $N_t \in]A_tF_t[$ may go as follows: $\|A_t - F_t\|$, the radius of curvature at A_t, equals the length of the cycloid arc $\widehat{OF_t}$; so O lies strictly inside the osculating circle of $\mathfrak{A}$ at A_t; the radius of Z_{N_t} is then strictly less than the radius of curvature at A_t.

To parametrize $\mathcal{N}$, write the equation of the line P_tN_t in the form $g(X, Y, t) = 0$, more explicitly

$$\left[Y - \tfrac{1}{2}(1 - \cos t)\right](1 - \cos t) + \left[X - \tfrac{1}{2}(t - \sin t)\right](t - \sin t) = 0;$$

then solve the system $\{g(X, Y, t) = 0, \partial_t g(X, Y, t) = 0\}$, to determine X and Y as functions of the parameter t. After some computations we get the required parametric equations that have been used in some of our figures:

$$\mathcal{N}: \quad \begin{cases} X(t) = \frac{4t(1-\cos t)-(2+t^2)\sin t+\sin 2t}{4(1-\cos t)-2t\sin t} \\ Y(t) = \frac{(1-\cos t)(2-t^2-2\cos t)}{4(1-\cos t)-2t\sin t}. \end{cases}$$

Acknowledgment. I thank my colleague A. Guedes de Oliveira for many helpful discussions and references on these matters.

References

[1] A. D. Aleksandrov, *Die Innere Geometrie der konvexen Flächen*, Akademie-Verlag, Berlin, 1955.

[2] D. A. Aleksandrov and V. A. Zalgaller: *Intrinsic geometry of surfaces*, Translations of Mathematical Monographs **15**, American Mathematical Society (AMS), Providence, (1967).

[3] D. Burago, Y. Burago and S. Ivanov, *A course in metric geometry*, Graduate studies in mathematics 33, AMS, Providence, RI, 2001.

[4] J. Dieudonné, *Foundations of modern analysis*, Pure and Applied Mathematics 10, Academic Press, New York, 1960.

[5] E. Lockwood, *A book of curves*, Cambridge University Press, 1961 (1976 edition).

[6] F. Gomes Teixeira, *Traité des Courbes Spéciales Remarcables, Planes et Gauches, Tome II*, Chelsea Publ. Comp., New York, 1971.

[7] C. Zwikker, *The advanced geometry of plane curves and their applications*, Dover Publications, Inc., New York, 1963.

Eduardo Marques de Sá
CMUC, Department of Mathematics, University of Coimbra, Apartado 3008, EC Santa Cruz 3001-501 Coimbra, Portugal

E-mail address: emsa@mat.uc.pt