Conchoids in a Riemannian Setup

Meeting notes

Oliver Knill

Collaboration with:

Michael Teodorescu

HCRP project fall 2008

Contents

Contents **1**

1 Meeting of November 12, 2008 **2**
1.1 The exponential map 3
1.2 Conchoids 4

2 Meeting of November 19, 2008 **6**
2.1 Metric and geodesic Distance 6
2.2 Calculus of variations 7

3 Meeting of December 2, 2008 **9**
3.1 The connection 9
3.2 The geodesic flow 10

4 Meeting of December 9, 2008 **12**
4.1 The hyperbolic plane 12
4.2 Geodesics 13
4.3 Geodesic Polar coordinates 14

5 Meeting of December 16, 2008 **16**
5.1 Which curves are wave fronts? 16
5.2 Caustics for compact perturbations of the flat plane 17
5.3 Caustics for compact perturbations of the hyperbolic plane 18

Chapter 1

Meeting of November 12, 2008

During our few meetings, we will explore some basic differential geometry, some calculus of variations and some hyperbolic geometry. There were 5 meetings in the fall semester of 2008.

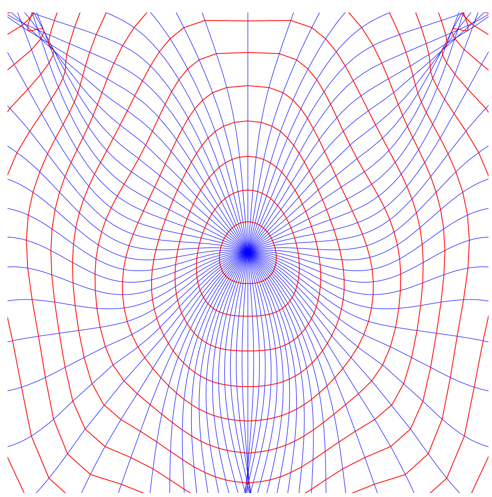


Figure: Geodesics and wave fronts in the plane.

1.1 The exponential map

For any Riemannian metric g on the plane R^2, a point $p \in R^2$ and a direction v, there is a geodesic $\gamma(t) = \exp_t(v)$, which starts at $p = \exp_0(v)$ and for which $\gamma'(t) = v$. By the existence theorem of differential equations, geodesics exist for a short time and by the Hopf-Rinow theorem, the curve exists for all times t but it does not mean that $\exp_t$ stays invertible. The set of all geodesics through p defines a vector field near the point and so a flow. Every point x near p gets transported along the geodesic flow.
Points in the plane, where the inverse of the exponential map is no more defined form a set called the **caustic** of the evolution problem. We have to pay attention to these caustics because it is there, where wave fronts and conchoids will change shape.

Assume a metric $g = g_0 + \epsilon h$ on M is a perturbation of the flat metric g_0 such that the support K of h is compact. If ϵ is small enough, some wave front W of p is a simple closed curve outside of K. The geodesics through W outside of C are straight lines perpendicular to C. For small enough ϵ, the curve W is a polar curve $r(t)\langle cos(t), sin(t)\rangle$ and the exponential map is given in an explicit form:

$$\exp_p(t,s) = (r(t) + sr'(t))\langle cos(t), sin(t)\rangle + sr(t)\langle -\sin(t), \cos(t)\rangle \ .$$

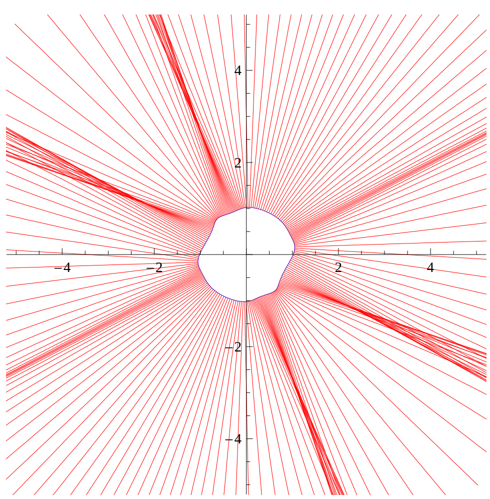


Figure: A Metric which is nonflat on a compact set K. It allows to study rays perpendicular to the wave front and to give rather concrete formulas for the exponential map outside K.

1.2 Conchoids

Conchoid curves can be defined in any Riemannian setups, where it is the evolution of a curve under the exponential map. It can be defined in more general metric spaces where geodesics exist.

If a wave front is given as $\exp_p(t, s(t))$, then

$$s \to \exp_p(t, s(t) + c)$$

is called a **conchoid** of C with respect to the point p.

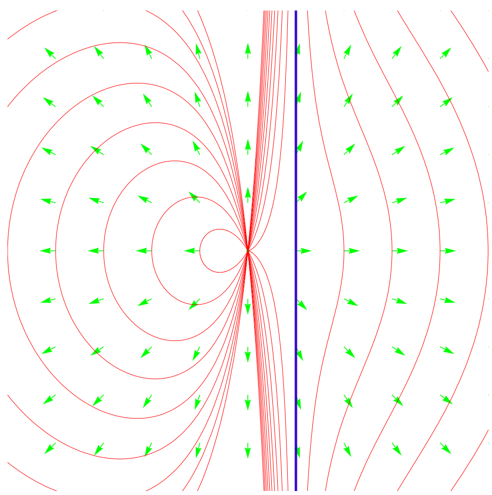


Figure: the figure shows the conchoid of Nicomachus. We look at generalizations of this geometric construction.

For the flat metric and if C is a line, the conchoid curve is called a **Conchoids of Nicomachus**. They are famous because with the help of such a curve, one can trisect angles.

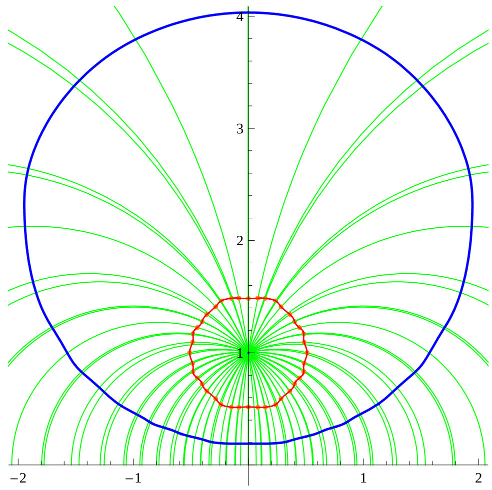


Figure: A conchoid in hyperbolic space is defined by a curve C given in hyperbolic polar coordinates as $r(t)$. The conchoid is in hyperbolic polar coordinates given by $R(t) = r(t) + c$.

Chapter 2

Meeting of November 19, 2008

If we replace the metric on the plane with a more general metric like the l^p metric what is the analogue of the conchoids? Do they become more interesting? How does one define geodesics in a general metric space? What happens in the perticular case of the l^p metric, which is no more Riemannian but an example of a Finsler metric?

2.1 Metric and geodesic Distance

Let (M, d) be a metric space. The **distance** $d(x, y)$ between two points x and y is often given by explicit formulas like $d(x, y) = ||x - y||_p = (\sum_i |x_i - y_i|^p)^{1/p}$.

Given a metric space (X, d), we can ask for **geodesics** between two points A, B, "paths" between A and B for which **arc length** is minimal. Independent of the fact that such minimal "curves" exist or not, one can define a new distance.

Define $d_\varepsilon(A, B)$ as the infimum of all possible sums $\sum_{i=1}^n d(P_i, P_{i+1})$, where P_i are points in M satisfying $P_0 = A, P_n = B$ and $d(P_i, P_{i+1}) \leq \varepsilon$. Define

$$d_{geo}(A, B) = \sup_{\varepsilon > 0} d_\varepsilon(A, B) \; .$$

This **geodesic distance** is by the triangle inequality always bigger or equal than the actual distance:

$$d_{geo}(A,B) \geq d(A,B) \; .$$

For the l^4 metric in the plane where $d(x,y) = ||x-y||_4$ the Euler equations in **calculus of variations** assure that the geodesics are straight lines. The geodesic distance is longer than the actual distance. This is an example of a **Finsler geometry**.
Example: On the sphere with the distance $d(A,B)$ as the Euclidean distance between the points A,B. The initial distance between the two poles N and S is $d(S,P) = 2$ while the geodesic distance is $d_{geo}(A,B) = \pi > 2$.
Example: Let M consist of the 4 vertices of the unit square where the distance is the distance of the underlying plane. The geodesic distance between two diagonally opposed points is 2, its actual distance is $\sqrt{2}$.

The geodesic distance of a geodesic distance is the geodesic distance itself, but only if the points are close enough. We can say in a general metric space that three points are **on a line** if the geodesic distance satisfies

$$d_{geo}(A,B) + d_{geo}(B,C) = d_{geo}(A,C) \; .$$

2.2 Calculus of variations

Some distances are given by local rules. Assume we have a function $F(x,y)$ so that $d(A,B) = \inf_\gamma \int_C F(x(t),x'(t)) \; dt$, where the infimum is taken over all smooth paths γ connecting A with B.

In order to compute the geodesic distance in a Riemannian plane, we have to find the shortest path. This is a classical topic and called calculus of variations. To derive the equation defining the shortest path, only multivariable calculus is needed:

Let $F(x,y)$ be a smooth function of 2 vector variables $x = (x_1,x_2), y = (y_1,y_2)$. Given two points A,B in the plane, we look at the problem to find the curve parametrized by the interval $[0,1]$ and connecting A with B for which the integral

$$I(r) = \int_0^1 F(r(t),r'(t)) \; dt$$

is minimal. From calculus we know that at a minimum, all directional derivatives must be zero. Because we do not have introduced a derivative in the infinite dimensional case yet, we find the curve r for which $D_h I = d/d\epsilon I(r + \epsilon h) = 0$ for all curves h for which $h(0) = h(1) = 0$. Since

$$I(r + \epsilon h) = \int_0^1 F(r(t) + \epsilon h(t), r'(t) + \epsilon h'(t)) \, dt \, ,$$

we get by the chain rule

$$\frac{d}{d\epsilon} I(r + \epsilon h) = \int_0^1 F_x(r + \epsilon h, r' + \epsilon h')h + F_y(r + \epsilon h, r' + \epsilon h')h' \, dt \, .$$

At $\epsilon = 0$, the right hand side is

$$\int_0^1 F_x(r, r')h + F_y(r, r')h' \, dt \, .$$

Integration by parts the second integral using $u = F_y(r), dv = h'dt$ gives

$$\int_0^1 F_x(r, r')h - \frac{d}{dt} F_y(r, r')h dt \, .$$

Since this is true for all h, we must have

$$\boxed{F_x(r, r') = \tfrac{d}{dt} F_y(r, r')}$$

These equations are called the **Euler equations** of the variational problem.

As in multi variable calculus, having a critical point does not necessarily mean that it is a minimum. But in many situations, which are physically relevant, this is the case.
Example: An important example from physics is the case $F(x, y) = V(x) + |y|^2/2$. Here $F_x = V'(x)$ and $F_y = y$. The Euler equations are $d/dt y = V'(x)$. If x is position and y is momentum, then these Euler equations are just the **Newton equations** in physics.

Chapter 3

Meeting of December 2, 2008

The question of the nature of conchoids becomes interesting already for surfaces with a general metric. If the exponential map is invertible, one has a good polar coordinate system and conchoids formally still look the same. Evenso we look especially at the case, when the surface is the graph of a function $f(x, y)$, we derive the equations for the geodesic flow for general parametrized surfaces.

3.1 The connection

For a parametrized surface $\vec{r}(u, v)$ in space, the **tangent vectors** are $\vec{r}_u, \vec{r}_v$. Write $dx = \vec{r}_u du$, $dy = \vec{r}_v dv$. The **distance element** $ds = \sqrt{dx \cdot dx + dy \cdot dy}$ satisfies

$$ds^2 = (r_u du + r_v dv)^2 = r_u \cdot r_u du du + r_u \cdot r_v du dv + r_v \cdot r_u dv du + r_v \cdot r_v dv dv \; .$$

With $g = \begin{bmatrix} r_u \cdot r_u & r_u \cdot r_v \\ r_v \cdot r_u & r_v \cdot r_v \end{bmatrix}$, this becomes $ds^2 = (du, dv) \cdot g(du, dv)$. We measure arc length of a curve $(u(t), v(t))$ surface with

$$\int_a^b ds = \int_a^b \sqrt{g_{11} r_u \dot{u} r_u \dot{u} + g_{12} r_u \dot{u} r_v \dot{v} + g_{21} r_v \dot{v} r_u \dot{u} + g_{22} r_v \dot{v} r_v \dot{v}} \; dt \; .$$

In the special case of a graph $\vec{r}(u,v) = (u, v, f(u,v))$, the metric is

$$g(u,v) = \begin{bmatrix} \vec{r}_u \cdot \vec{r}_u & \vec{r}_u \cdot \vec{r}_v \\ \vec{r}_v \cdot \vec{r}_u & \vec{r}_v \cdot \vec{r}_v \end{bmatrix} = \begin{bmatrix} 1 + f_u^2 & f_u f_v \\ f_u f_v & 1 + f_v^2 \end{bmatrix} .$$

Define the **Christoffel symbols**

$$\Gamma_{ijk} = \frac{1}{2}[\frac{\partial}{\partial x^i} g_{jk}(x) + \frac{\partial}{\partial x^j} g_{ik}(x) - \frac{\partial}{\partial x^k} g_{ij}(x)] .$$

This is

$\Gamma_{111} = r_{uu} \cdot r_u$, $\Gamma_{112} = r_{uu} \cdot r_v$ $\quad$ $\Gamma_{211} = r_{vu} \cdot r_u$, $\Gamma_{212} = r_{vu} \cdot r_v$
$\Gamma_{121} = r_{uv} \cdot r_u$, $\Gamma_{122} = r_{uv} \cdot r_v$ $\quad$ $\Gamma_{221} = r_{vv} \cdot r_u$, $\Gamma_{222} = r_{vv} \cdot r_v$

3.2 The geodesic flow

The equations for the geodesic flow are

$$\ddot{x}^k + \Gamma^k_{ij} \dot{x}^i \dot{x}^j .$$

These differential equations can be derived in calculus of variations by finding the curves minimizing arc length.

In the graph case, these differential equations simplify to

$$x'' = \frac{-f_x(f_{xx}x'^2 + 2f_{xy}x'y' + f_{yy}y'^2)}{(1 + f_x^2 + f_y^2)} , y'' = \frac{-f_x(f_{xx}x'^2 + 2f_{xy}x'y' + f_{yy}y'^2)}{(1 + f_x^2 + f_y^2)} .$$

For a **conformal metric** $g = e^f g_0$, the equations for the geodesic flow are

$$x'' = -(x'y'f_y + (x'^2 - y'^2)f_x/2) , y'' = -(x'y'f_x + (y'^2 - x'^2)f_y/2) .$$

Locally in a small neighborhood of a point (u_0, v_0), we can find **geodesic coordinates**, where $ds^2 = du^2 + G(u,v)dv^2$. In other words, locally, any metric is a conformal metric.
To get a Geodesic coordinate system, we draw a geodesic curve through $r(u_0, v_0)$ in the direction $(1, 0)$. Then draw on each point on that curve

an other geodesic curve perpendicular to the curve (where perpendicularity is with respect to the dot product at the point). Now draw curves parallel to the first geodesic and perpendicular to the perpendicular geodesic field.

Example. Consider the sphere and a point P on the equator. The equator itself is a geodesic. The meridians (lines of longitudes) are geodesics too which hit the equator perpendicularly. The circles of latitude are also called loxodromes. The geodesic coordinate system works except at the north and south pole.

Chapter 4

Meeting of December 9, 2008

Today, we look at the geodesic evolution in the hyperbolic plane. Our goal is to understand conchoids in hyperbolic space.

4.1 The hyperbolic plane

Arc length in the **Poincaré plane** $\{(x, y) \mid y > 0\}$ is defined as

$$\int_a^b \frac{|r'(t)|}{y}\, dt\ ,$$

To find the distance between two points A, B in this hyperbolic space, we need to know about **geodesics** in this space, curves which minimize arc-length.

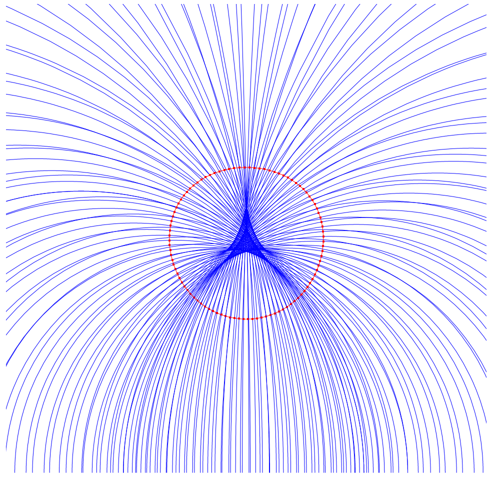


Figure: geodesics perpendicular to a given curve in the hyperbolic plane.

4.2 Geodesics

A geodesic connection between two points $A = (a_1, a_2), B = (b_1, b_2)$ is a circular half circle through A, B with center $(M, 0) = (a_1^2 + a_2^2 - b_1^2 - b_2^2, 0)/(2a_1 - 2b_1)$ and radius $r = |MA|$.
In the special case when $a_1 = b_1$, the geodesic connection is the line segement connecting the two points.

One can reduce the special case to the general case using **Möbius transformations**:
Let $z = x + iy$ denote a point on the upper half plane. In complex coordinates, the metric is $ds^2 = dz\overline{dz}/y^2$. Transformations $z \to az + b/(cz + d)$ with real a, b, c, d satisfying $ad - bc = 1$ form the subgroup of all **Möbius transformations**. Every M̈bius transformations leave the upper half plane invariant.

Real Möbius transformations T are isometries of the hyperbolic metric. If A, B are two points, then $d(A, B) = d(T(A), T(B))$.

Proof: Let $z' = x' + iy'$ the image point of z under the Möbius transformation T. A direct computation shows that $y' = y/|cz + d|^2$ and

$dz' = dz/(cz+d)^2$. The second identity gives

$$\frac{dz'\overline{dz}'}{y'^2} = \frac{dz\overline{dz}}{|cz+d|^4 y'^2}$$

Multiply the right hand side $1 = \frac{|cz+d|^4 y'^2}{y^2}$ to get

$$\frac{dz'\overline{dz}'}{y'^2} = \frac{dz\overline{dz}}{y^2} .$$

If A, B are two points have the same real value, then the geodesic is the straight line connecting them. In general, we can find an explicit Möbius transformation T such that $T(A) = i, T(B) = hi$ for some h. This shows that $d(A,B) = \log(h)$ and that the geodesic line is the image of the positive imaginary axes under T. M̈bius transformations map half circles perpendicular to the real axes into half circles perpendicular to the real axes.

Remark: The shape of geodesics can also be derived from **calculus of variations**. The Euler equations for the functional $F(x_1, x_2, y_1, y_2) = \frac{\sqrt{y_1^2+y_2^2}}{x_2}$ are

$$\frac{d}{dt}F_{y_1} = 0, \frac{d}{dt}F_{y_2} = -F_{x_2} .$$

This gives

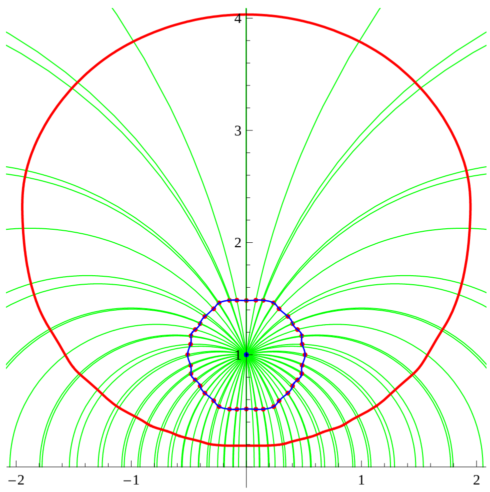
$$\frac{d}{dt}\frac{y_1}{x_2\sqrt{y_1^2+y_2^2}} = 0, \quad \frac{d}{dt}\frac{y_2}{x_2\sqrt{y_1^2+y_2^2}} = \frac{\sqrt{y_1^2+y_2^2}}{x_2^2} .$$

The first equation defines a direction field for which the solution curves are half circles.

4.3 Geodesic Polar coordinates

Geodesic polar coordinates in the hyperbolic plane are defined as in the Euclidean plane. Choose a center O and look at all half circles with center on the x which pass through O. A point has the coordinates (r, θ), where r is the distance to O and θ is the limiting angle of the half circle hitting O.

Conchoids in the hyperbolic plane. If O is the center and C is a curve defined in polar coordinates as $r = f(\theta)$, then the conchoid C_c has the coordinates $r = f(\theta) + c$.



Chapter 5

Meeting of December 16, 2008

5.1 Which curves are wave fronts?

Given a Riemannian metric g on the plane R^2 and a point $p \in R^2$, we can look at **wave fronts** $S(r) = \{x \in R^2 \mid d(x,p) = r\}$. At least for small r, they are **geodesic circles** for the geodesic distance $d_g(x,y)$. They are also level curves of a solution to the eiconal partial differential equation $|\nabla_g f|^2 = 1$, where $|\nabla_g f|^2 = g_{11}f_x^2 + 2g_{12}f_x f_y + g_{22}f_y^2$. For any closed curve C diffeomorphic to the circle and containing the point p in its interior, there is a metric g in the plane so that C is the unit circle in this Riemannian metric space, There is a conformal diffeomorphism φ which maps the interior of C into the unit disc. Using φ one can pull back the flat metric to the surface. It follows for any diffeomorphism T in the plane with fixed point p, there is a metric g such that $S(r) = \{g(T(x,y)) = r\}$. In other words, given an arbitrary smooth metric $d(x,y)$ in the plane, there exists a Riemannian metric tensor g such that the Riemannian distance d_g satisfies $d_g(p,x) = d(p,x)$.

We know therefore that the problem of a wavefront emanating from a given curve can be considered as a special case of the exponential map. Also the problem of wave fronts emanating from a time dependent point can be seen as a restriction in 2 dimensions. Short: all simple curves can be a wave fronts of some point and some metric.

5.2 Caustics for compact perturbations of the flat plane

A metric defines a smooth map $T(tv) = \exp_t(v)$ from R^2 to R^2, the **exponential map** from the origin. Define $\rho(x) = \det(DT(x))$. The image $C = T(Y)$ of the zero set $Y = \{\rho = 0\}$ is called the **caustic** of the point p with respect to the metric g. What is the nature of this set? For which metrics g does C have compact support?

Example: Let (M, g) be a compact perturbation of the flat plane. Consider a wave front

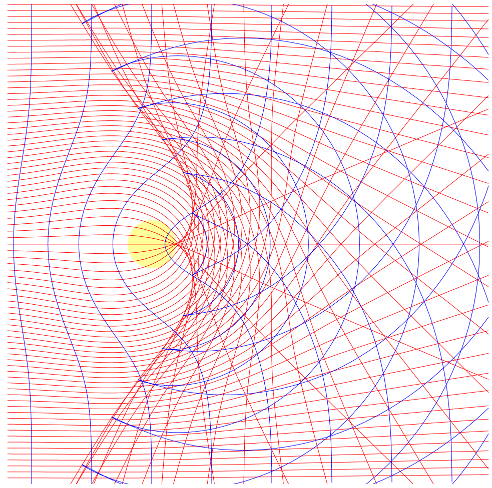
$$W_p = \{\exp_p(y) \mid |y| = 1\}$$

which has K in its interior. Assume $W_p(1)$ is a smooth curve, parametrized by arc length. Denote by $\kappa(t) = x'y'' - y'x''$ the **signed curvature** of the curve. Let $v(t)$ the unit vector in T_pM such that $\exp_p(v) = r(t)$. We have $T(t,s) = \exp_p(sv(t)) = r(t) + sn(t) = (x(t) - sy'(t), y(t) + sx'(t))$. The caustic is the set

$$C(p) = \{r(t) + sn(t) \mid \det(dT(t)) = \det(\begin{bmatrix} x' - sy'' & -y' \\ y' + sx'' & x' \end{bmatrix}) = 1 + \kappa s = 0\}.$$

If the curvature of the curve κ is positive everywhere, that is if the curve C encloses a strictly convex region, then the caustic $C(p)$ is empty. Note that we consider only the geodesic half ray with $s > 0$. For $s < 0$, caustics of course occur and they can leave the interior of the curve as the ellipse shows.

If κ changes sign, then $C(p)$ is nonempty and unbounded. The number of unbounded pieces is equal to the number of **inflection points** $\kappa(t) = 0$ of the wave front.



5.3 Caustics for compact perturbations of the hyperbolic plane

What is the topological nature of caustics in the hyperbolic plane? We expected that under a deformation $g = g_0 + th$ and any point p, the caustic $C_p(t)$ will remain compact. However this is not the case: here is the computation:

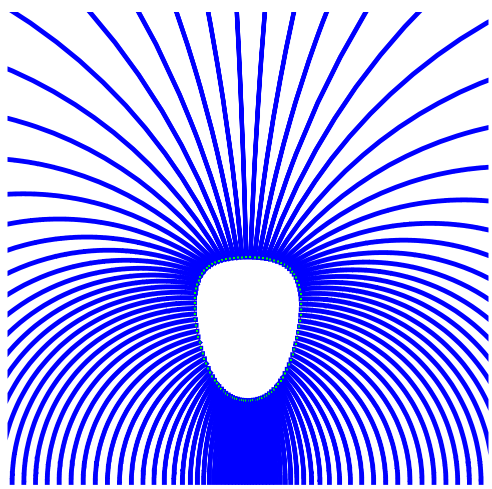


Figure: For small perturbations of the hyperbolic metric, there are no caustics according to Hadamard's theorem.

Lets compute the caustic from a curve C with parametrization $R(t) = (x(t), 1 + y(t)) = (x, y)$. The geodesic is a half circle with center $(m, 0)$ and radius r. Because $(x - m, y) \cdot (-y', x') == -xy' + my' + x' = 0$, we get $m = x - (x'/y')$ and $r = \sqrt{(x-m)^2 + y^2}$. The caustic is the set of (t, s) for which the transformation $T(t, s) = (m + r\cos(s), r\sin(s))$ has the property that the Jacobian is zero:

$$\det(dT(t,s)) = \det \begin{bmatrix} m' + r'\cos(s) & -r\sin(s) \\ -r'\sin(s) & r\cos(s) \end{bmatrix} = r(m'\cos(s) + r') = 0\ .$$

Since $r > 0$, this is equivalent to $\partial_t X(t, s) = (m\cos(s) + r)' = 0$ where $X(t, s)$ gives the x coordinate of the point of the geodesics through $(x(t), y(t))$ with coordinates s.

If the initial curve $(x(t), y(t))$ is close enough to a circle, then $m'\cos(s) + r'$ is never zero. There are no caustics confirming Hadamard's theorem. Unlike we thought first that caustics remain compact from some time, caustics appear to be noncompact always even after small perturbations. Noncompact caustics always hit the x axes perpendicularly.