# Variational Methods & Optimal Control

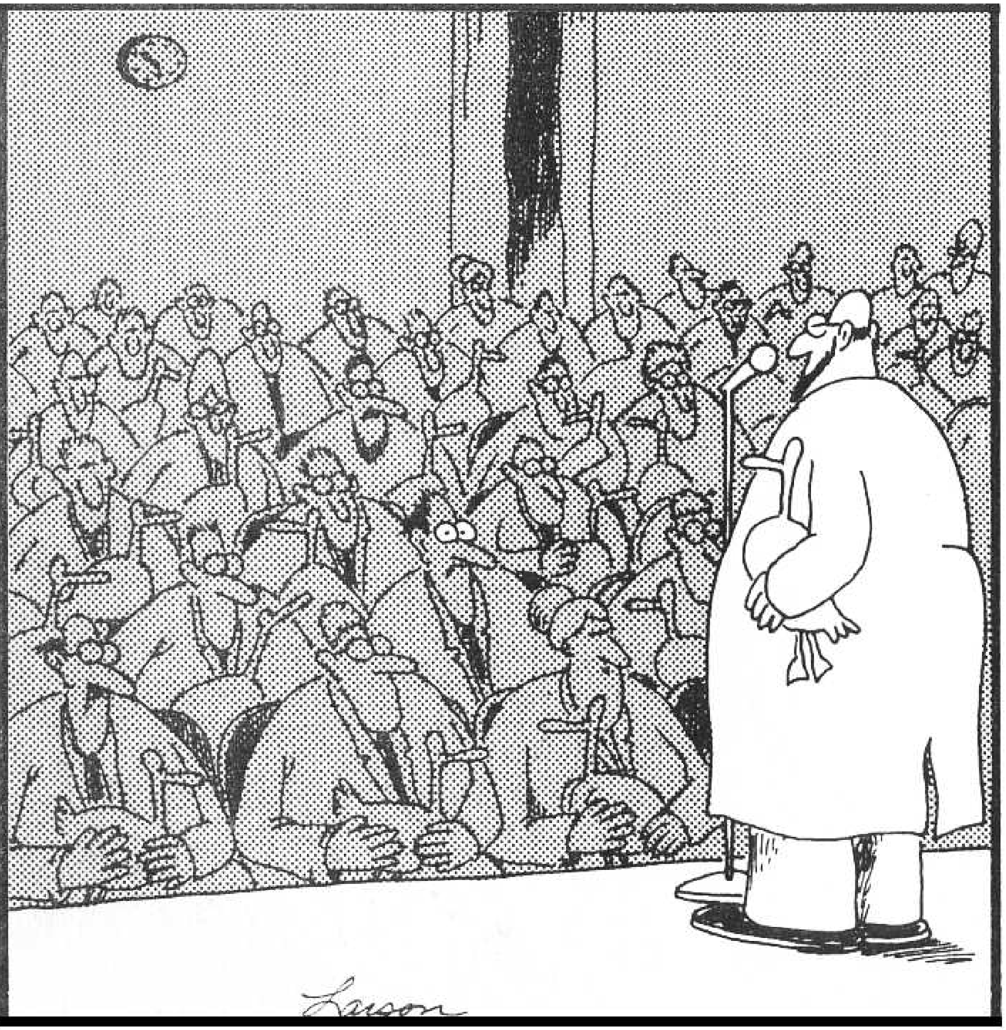
*lecture 01*

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

April 14, 2016

# Did you bring your duck?




Suddenly, Professor Liebowitz realizes he has come to the seminar without his duck.

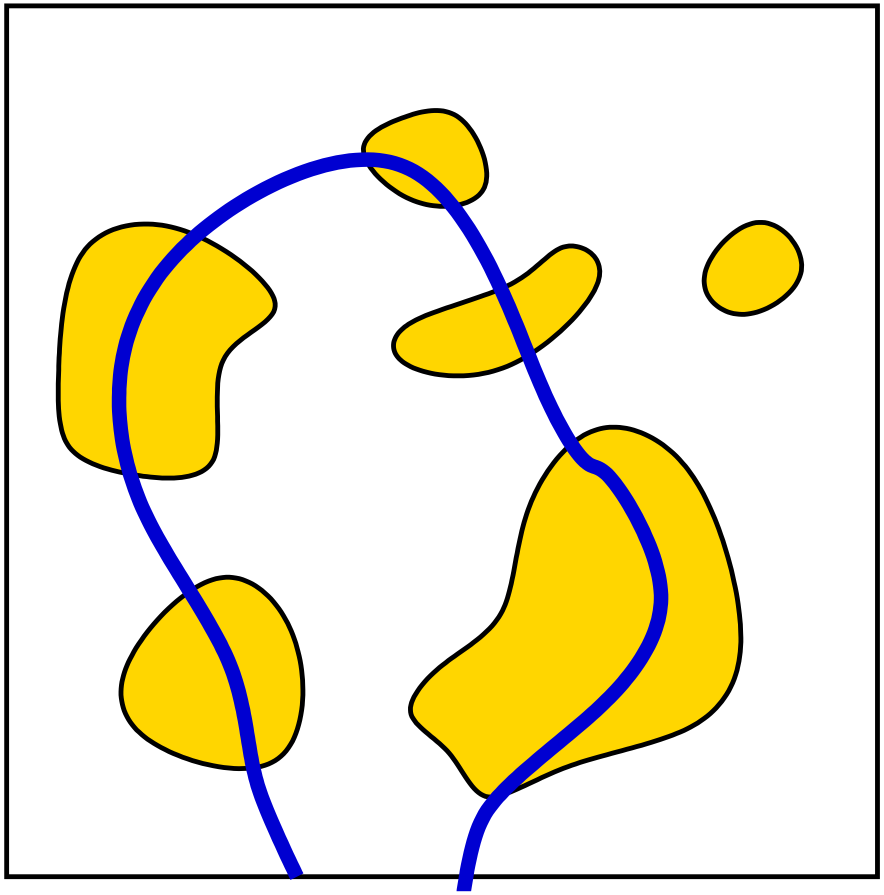
Larson, 1989

# Introduction

What is the point of this course?

# Motivation

- Imagine a field containing patches of gold.
- Collect the most gold
- We want to choose best **path**
- But the path length is limited.

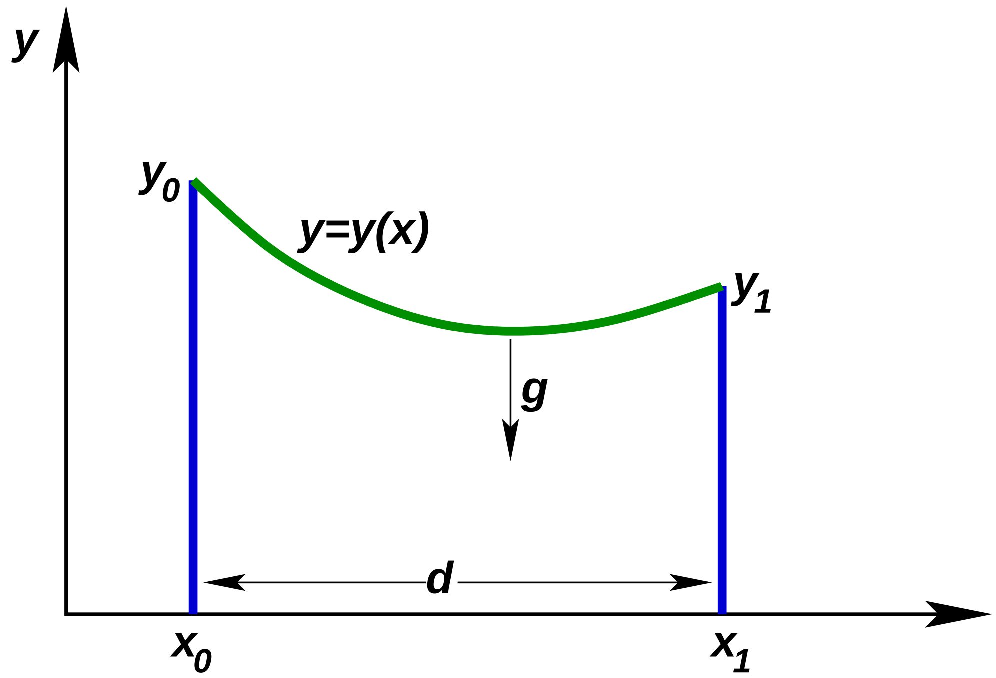


# Gold example (part ii)

- The gold collected on the path is the **integral** of the gold at each point.
- The length of the path is **fixed**.
- We are maximizing an integral over a path for **all possible paths**.
- Maximizing a function of a function (a **functional**).

# The catenary

Consider a thin, uniformly-heavy, flexible cable suspended from the top of two poles of height $y_0$ and $y_1$ spaced a distance $d$ apart. What is the shape of the cable between the two poles?




What is the difference if the cable is coiled at the base of the poles and is free to move up and down via a pulley?

# Brachystochrone problem

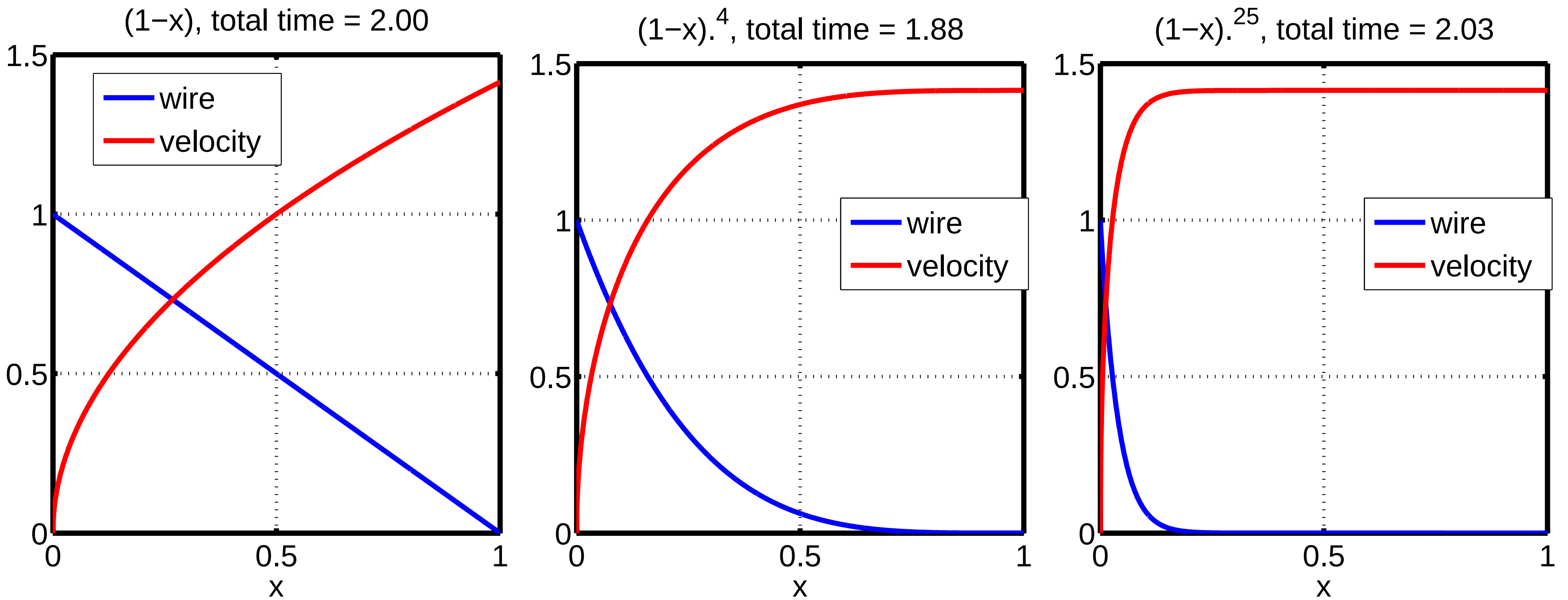
“Did Bernoulli sleep before he found the curves of quickest descent? ”,
Peter Parker, Spiderman II

Find the shape of a wire along which a bead, initially at rest, slides from one end to the other as quickly as possible under the influence of gravity.

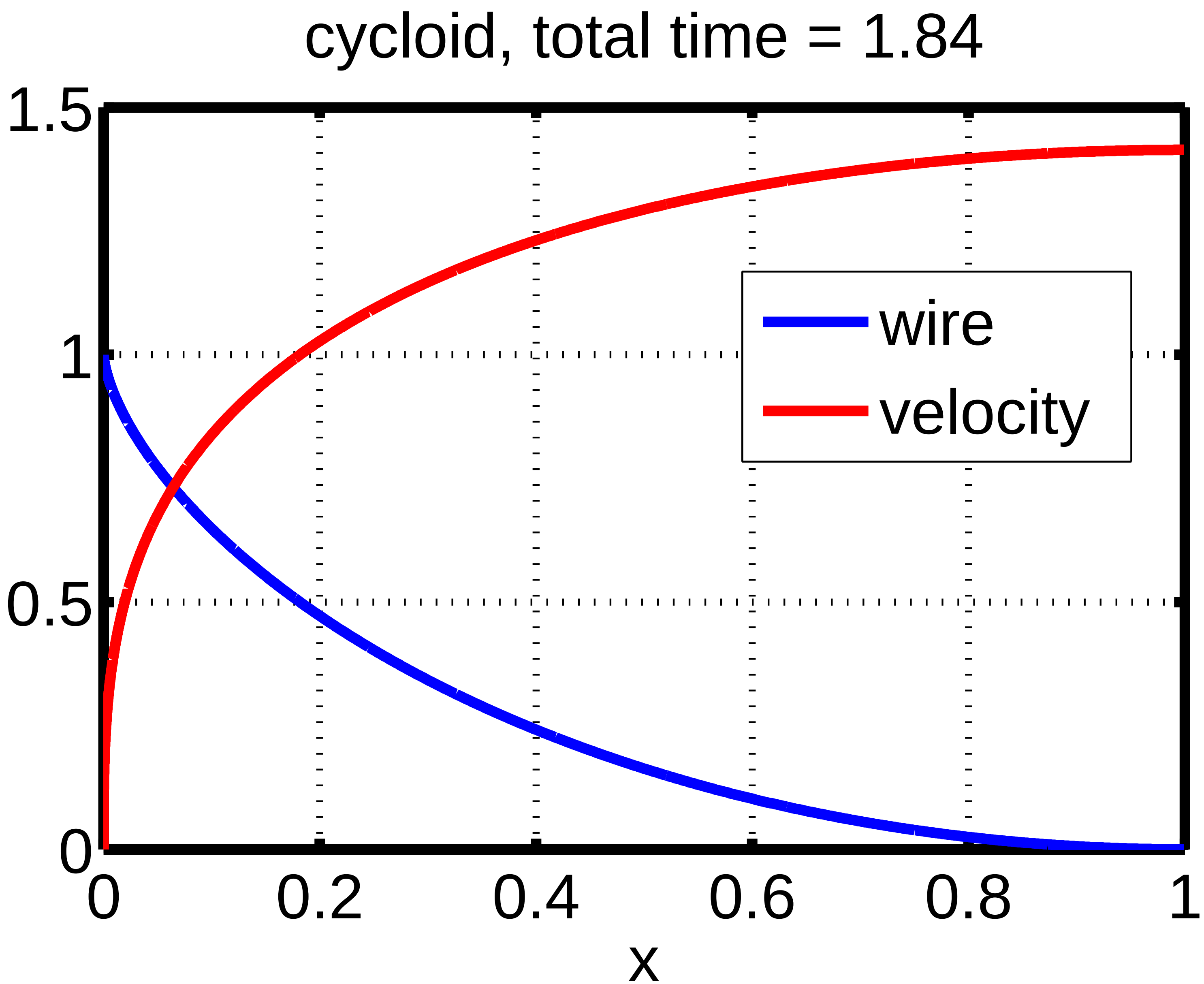
- endpoints are fixed
- motion is frictionless

Can think of as the “optimal slippery dip”

# Brachystochrone problem



# Brachystochrone solution



# Brachystochrone history

- problem posed by Johann Bernoulli (1696)
- Newton, Liebnitz, Huygens, Bernoulli's
- Euler developed method to solve it that was generalizable
- Jacob first to solve?
- Johann, "Ah, I recognize the paws of a lion"
- Christiaan Huygens discovered cycloid property

  A bead sliding down a cycloid generated by a circle of radius $\rho$ under gravity $g$ reaches the bottom after $\pi\sqrt{\rho/g}$ regardless of where the bead starts. Hence **cycloid = isochrone**

# Cycloid generation

# Cycloid generation

# Cycloid generation

# Cycloid generation

# Cycloid generation

# Cycloid generation

# Cycloid generation

# Cycloid generation

# Cycloid generation

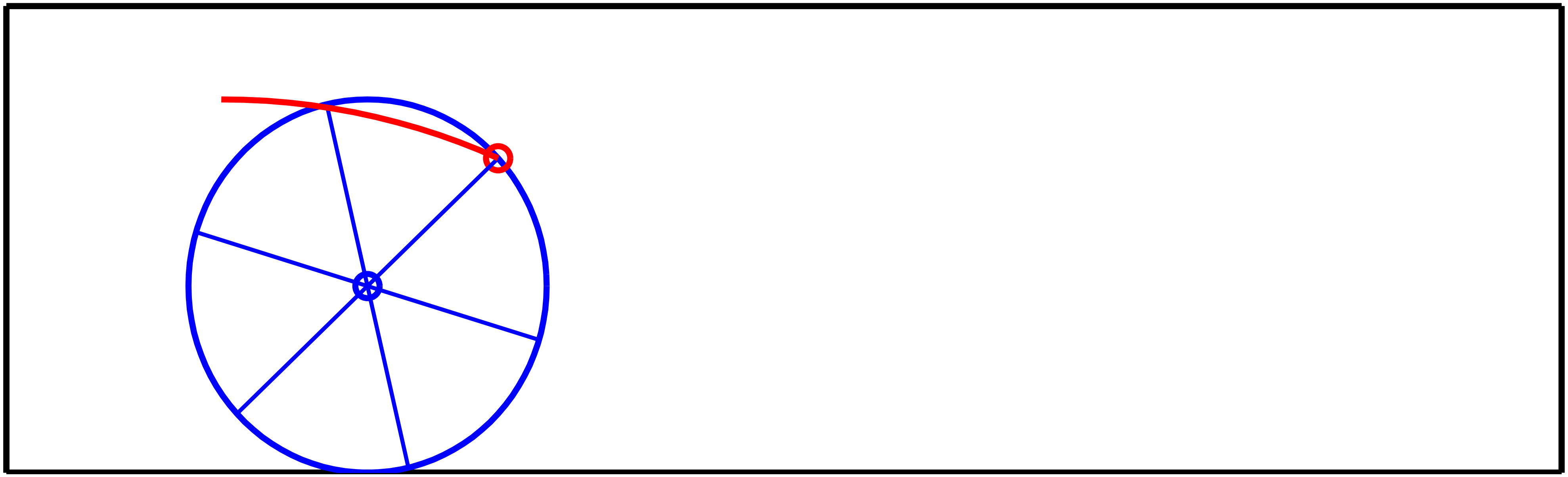
# Cycloid generation

# Cycloid generation

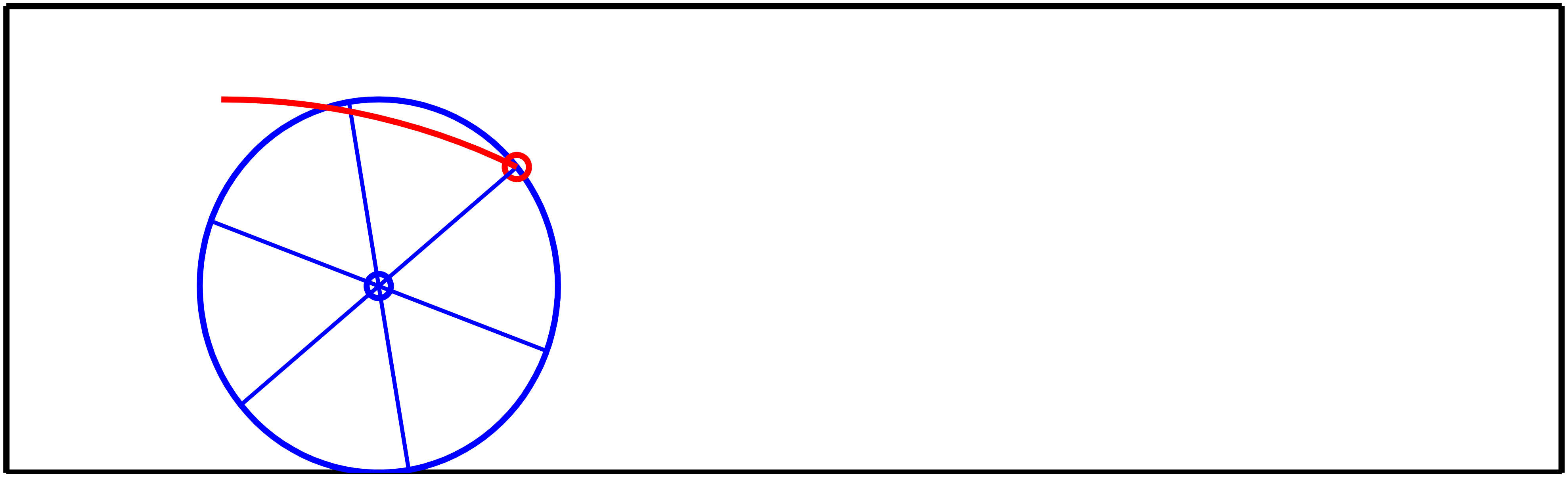
# Cycloid generation

# Cycloid generation

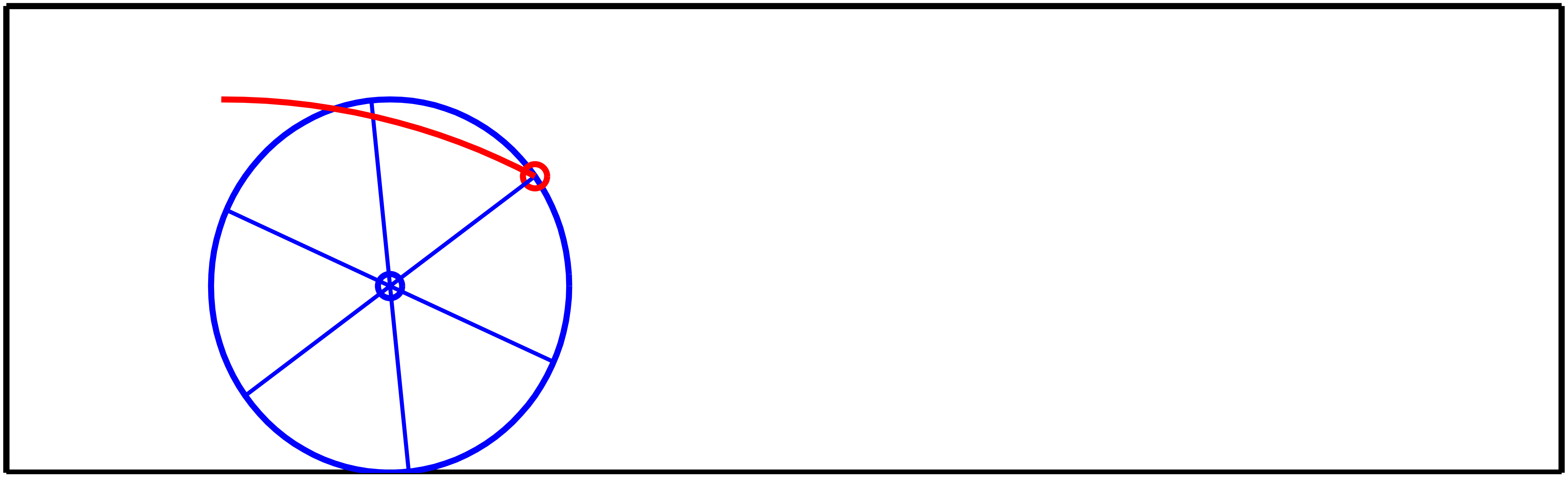
# Cycloid generation



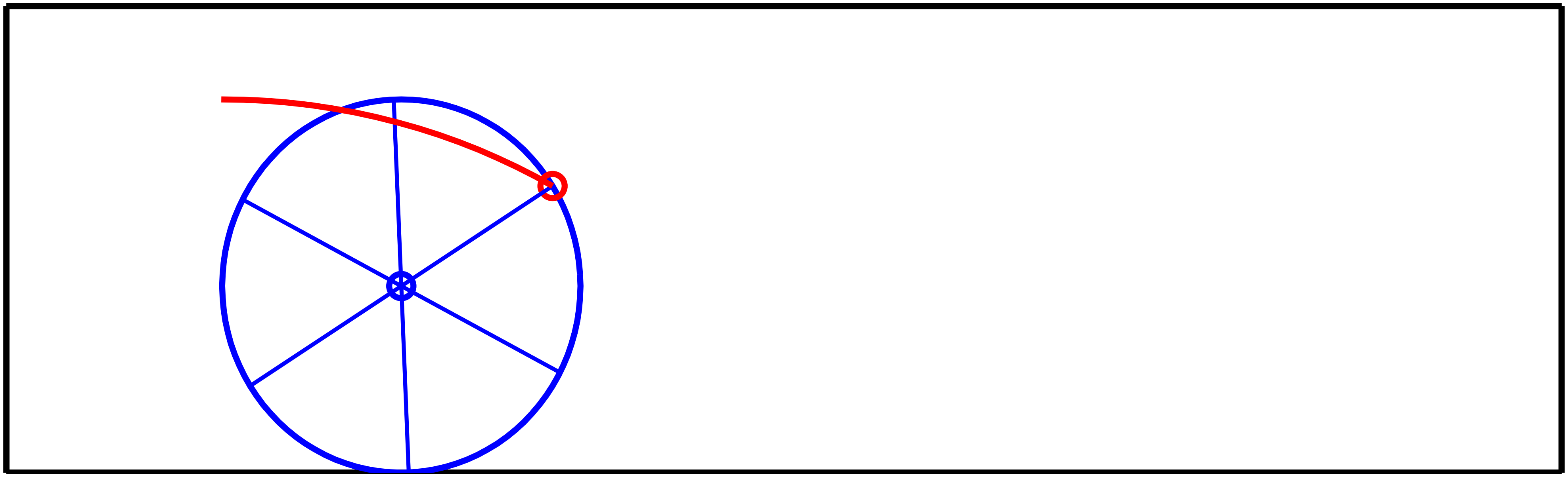
# Cycloid generation



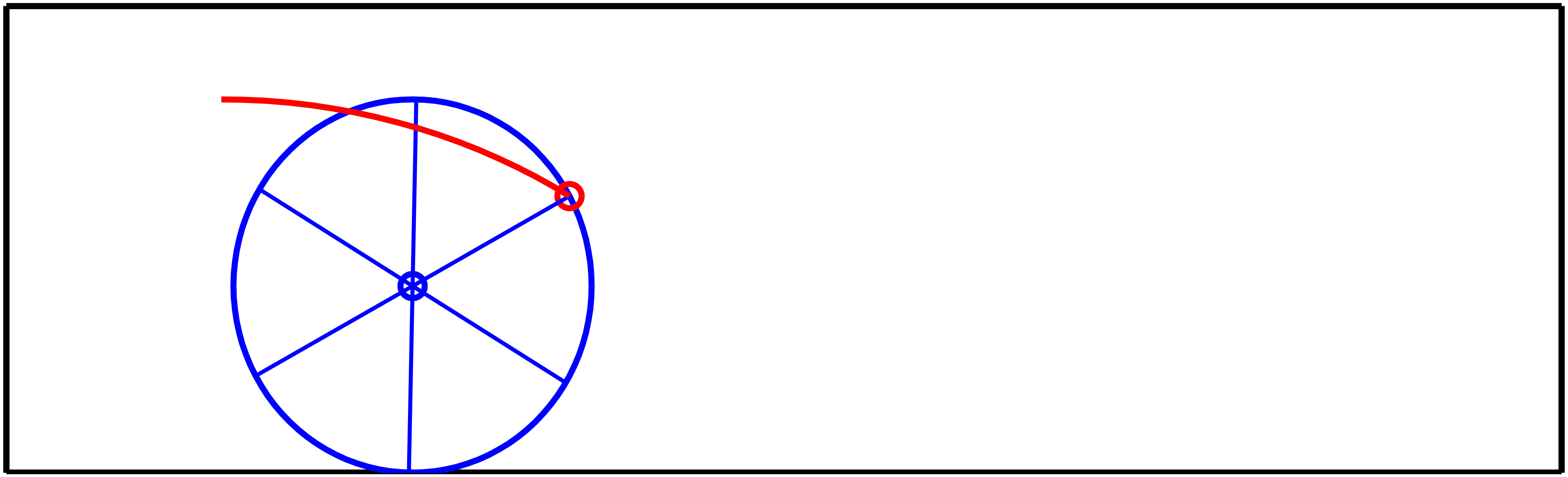
# Cycloid generation



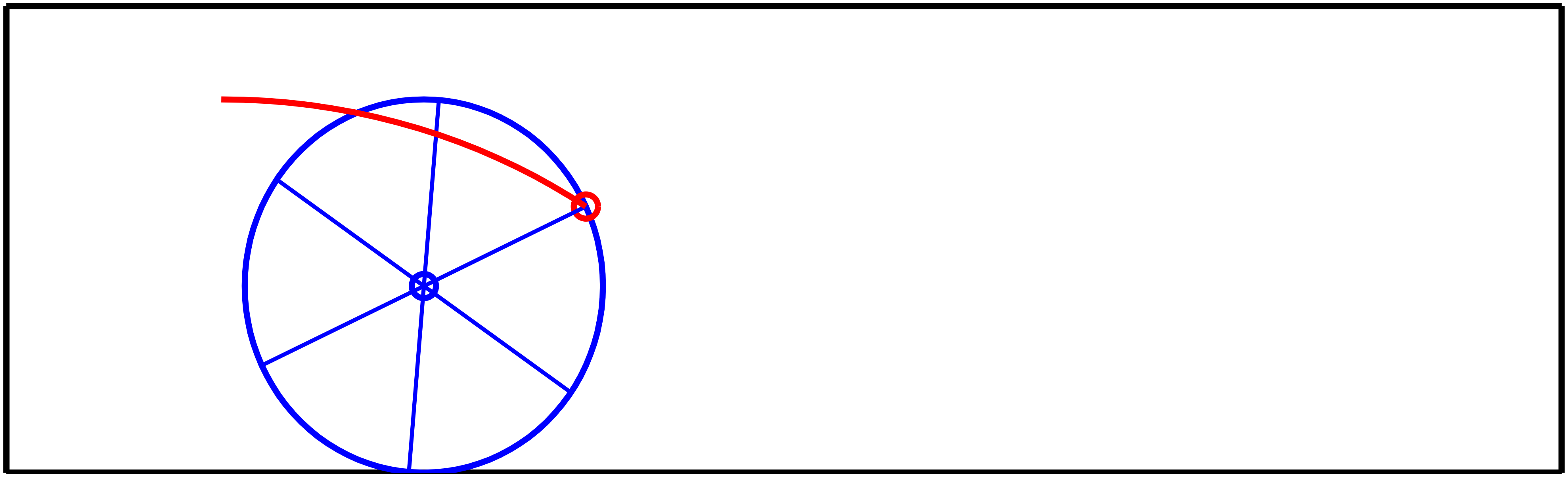
# Cycloid generation



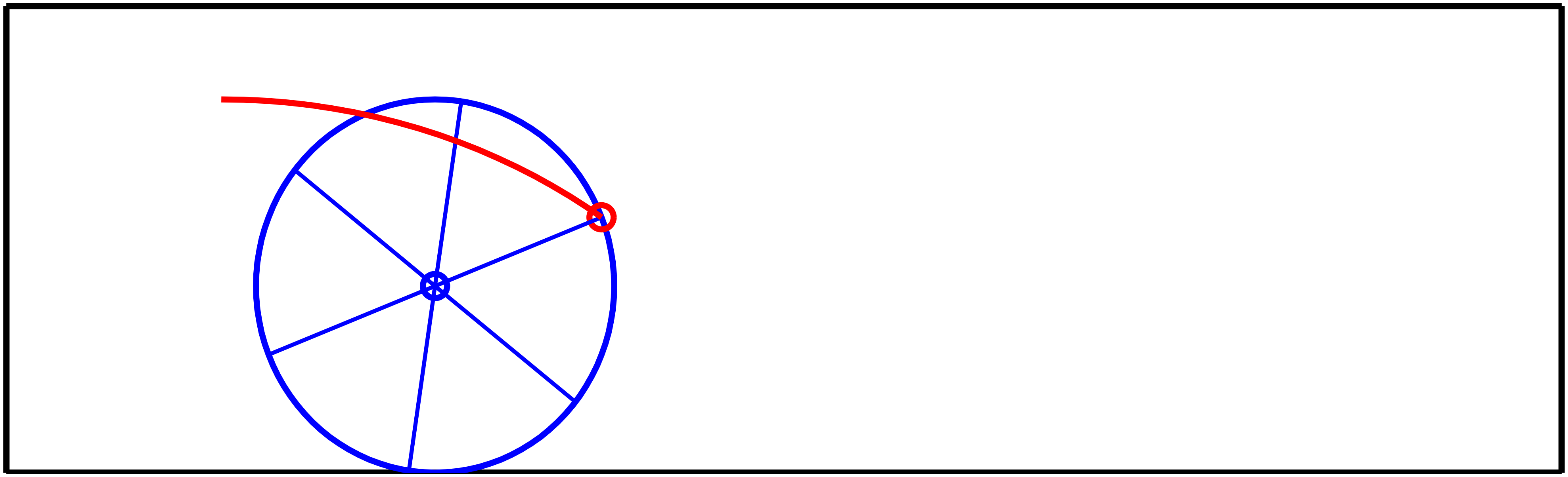
# Cycloid generation



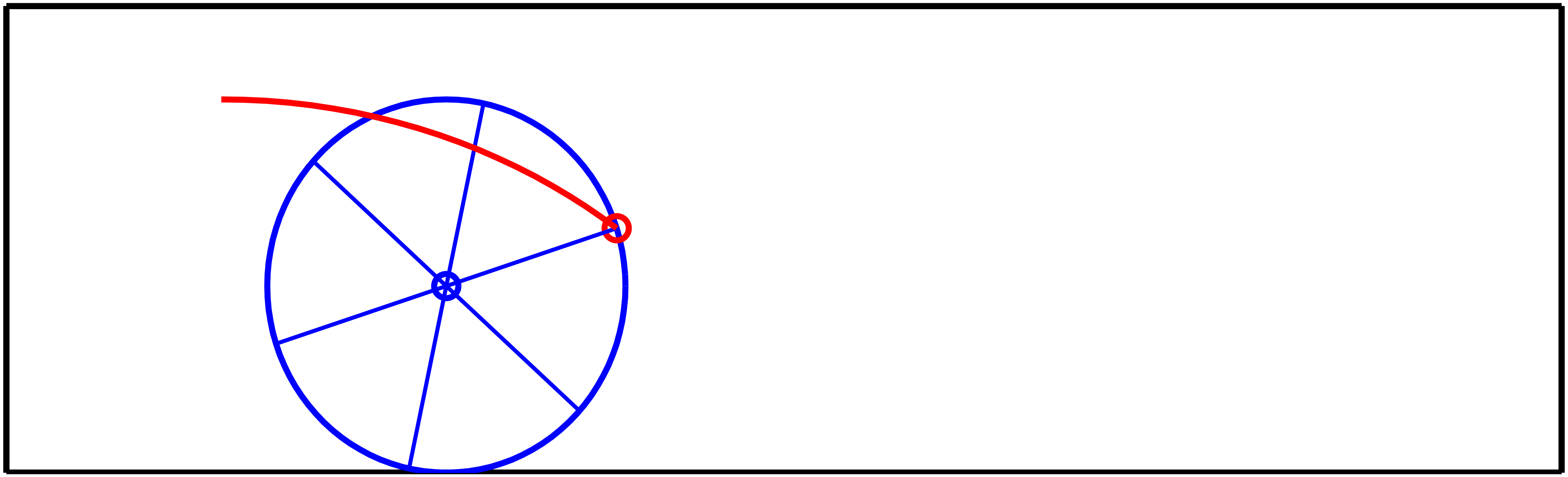
# Cycloid generation



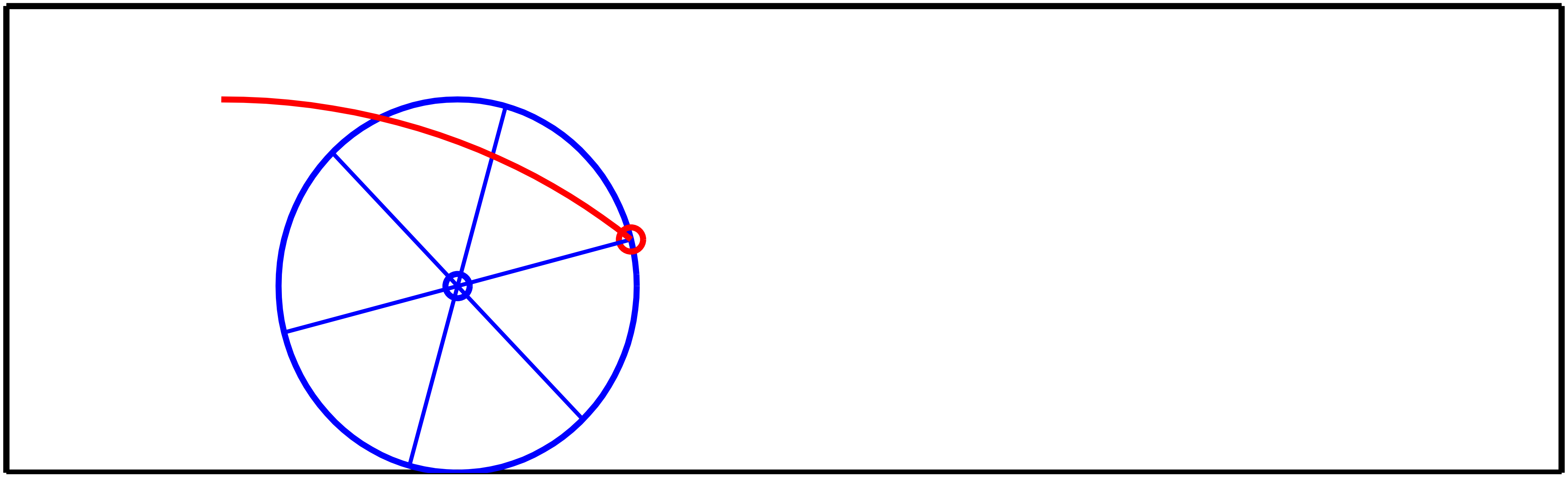
# Cycloid generation



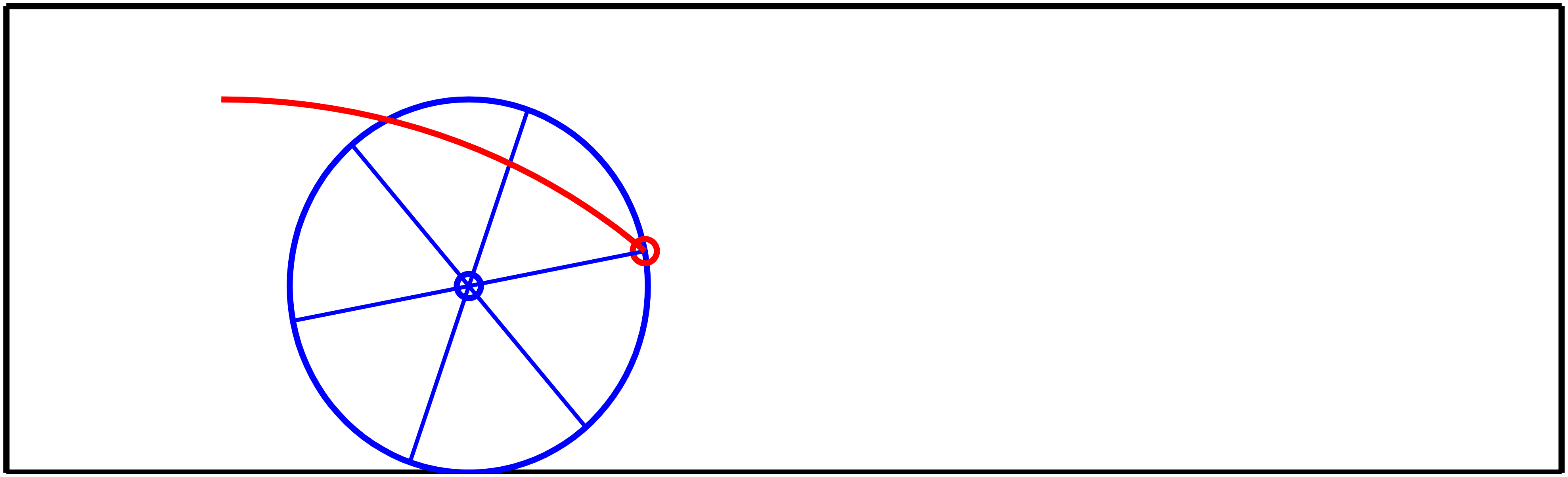
# Cycloid generation



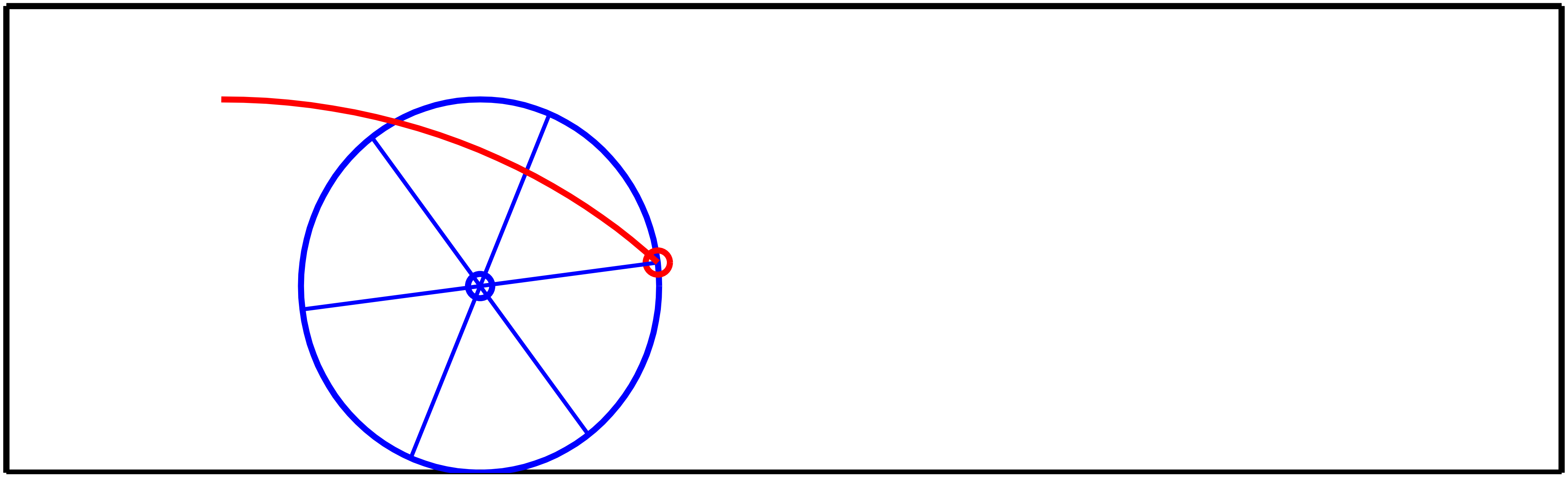
# Cycloid generation



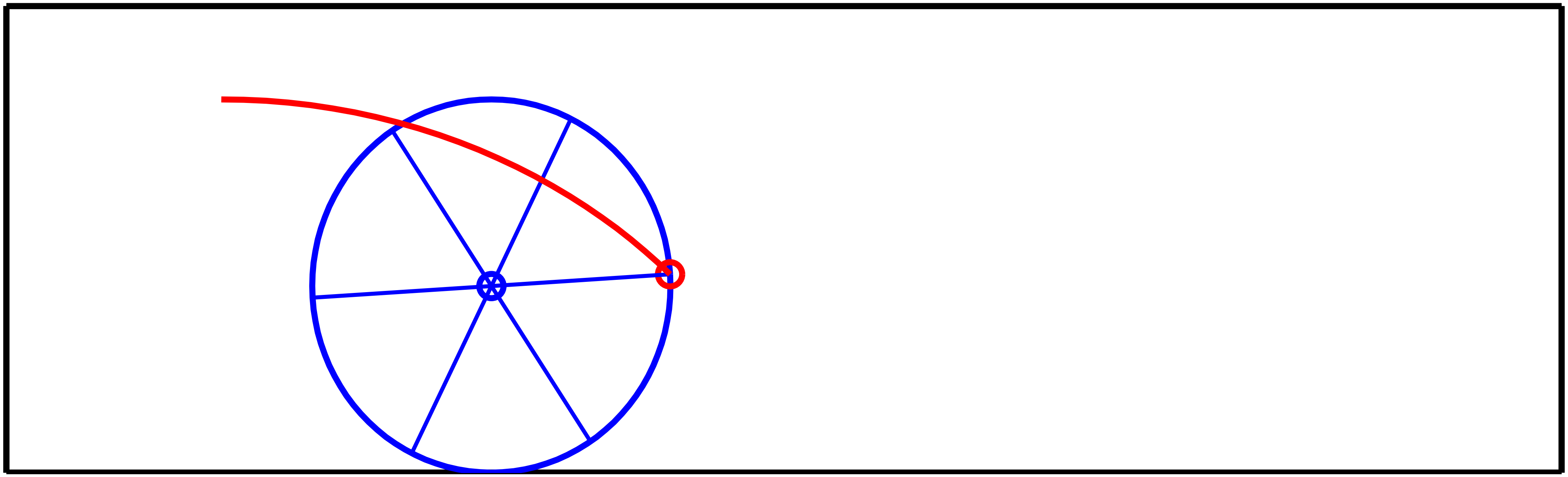
# Cycloid generation



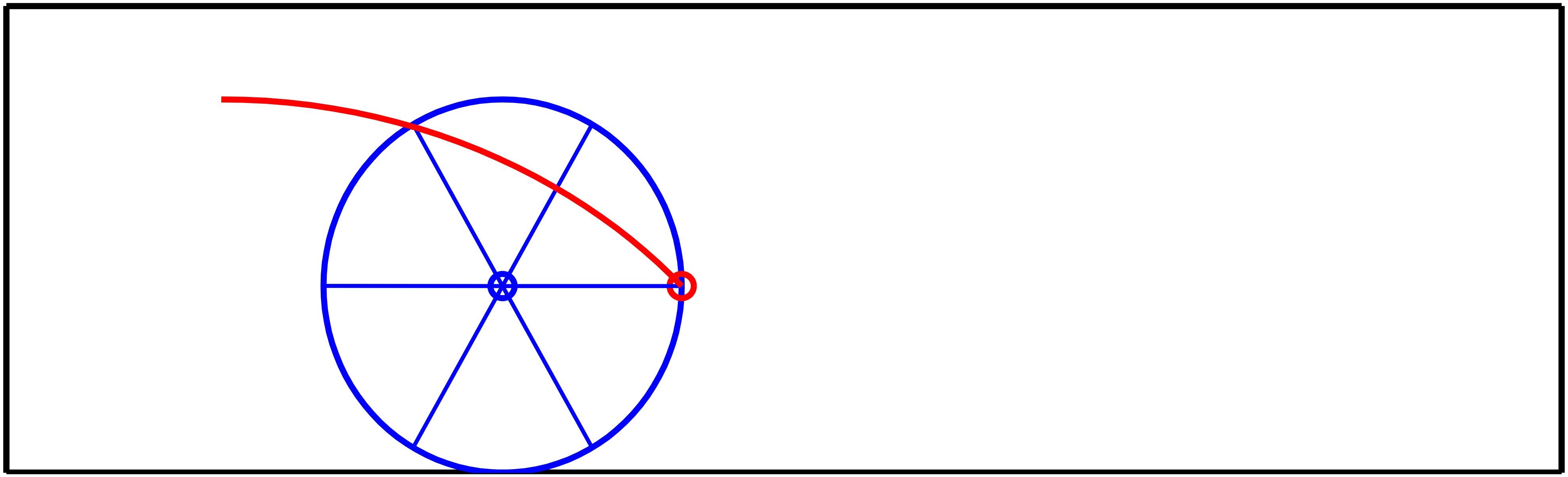
# Cycloid generation



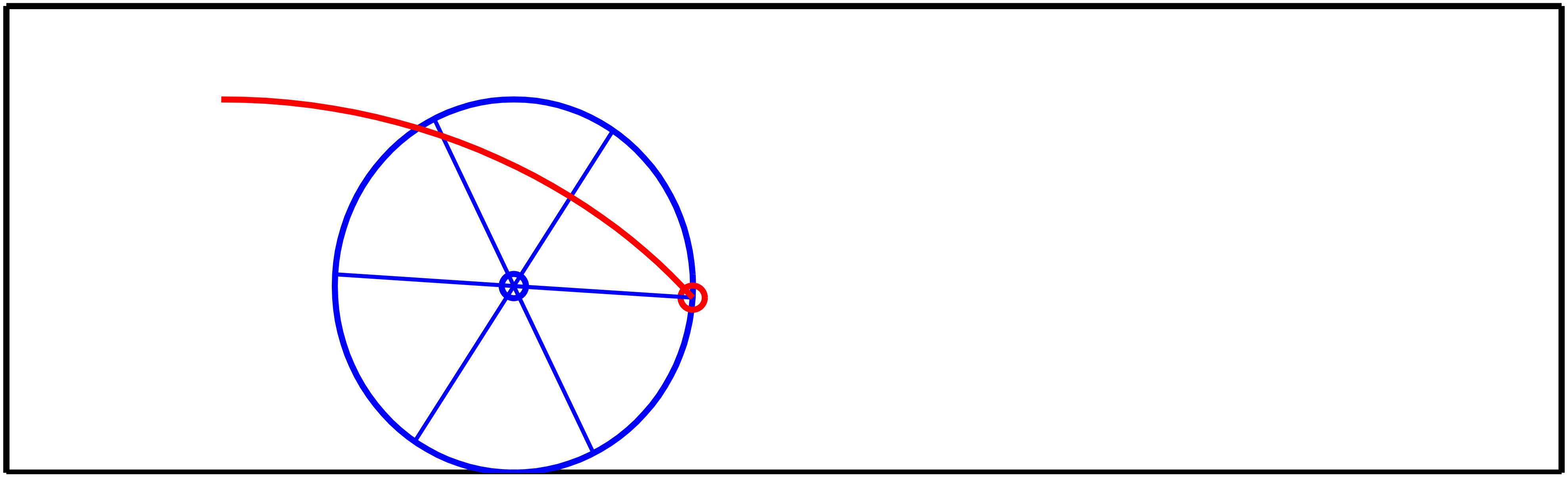
# Cycloid generation



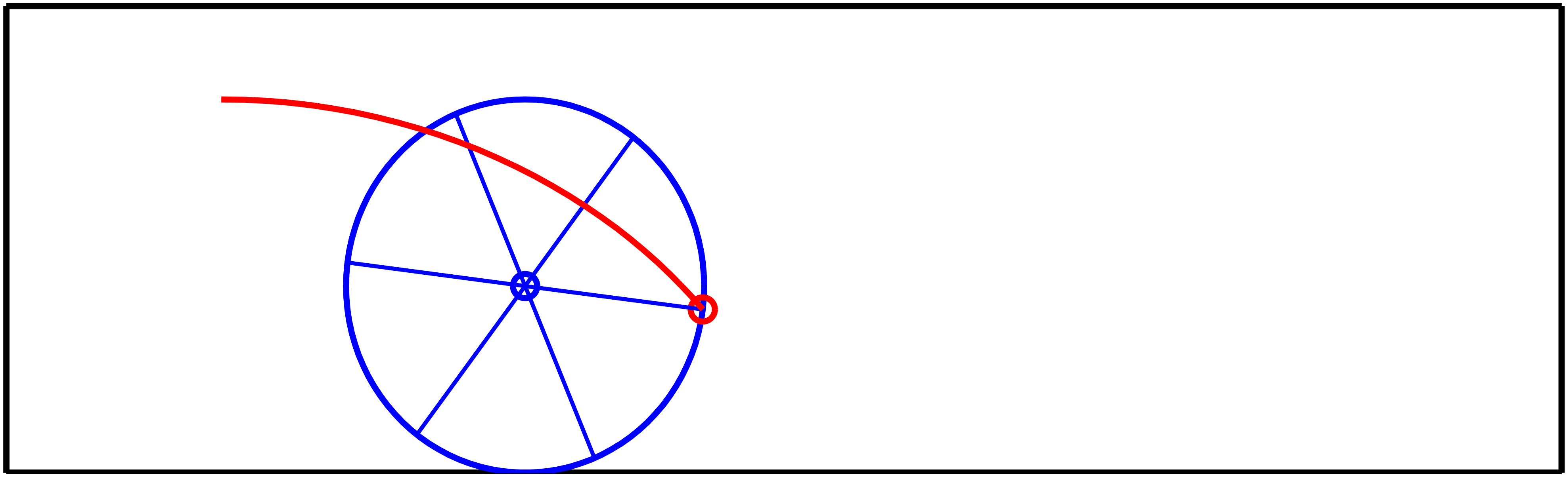
# Cycloid generation



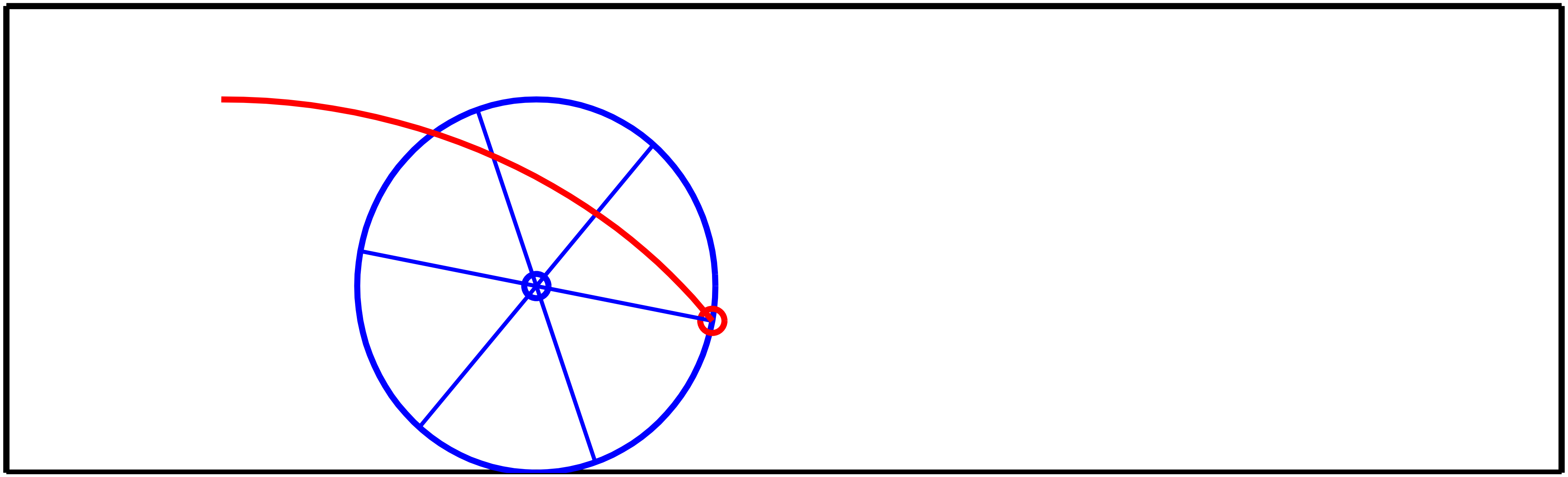
# Cycloid generation



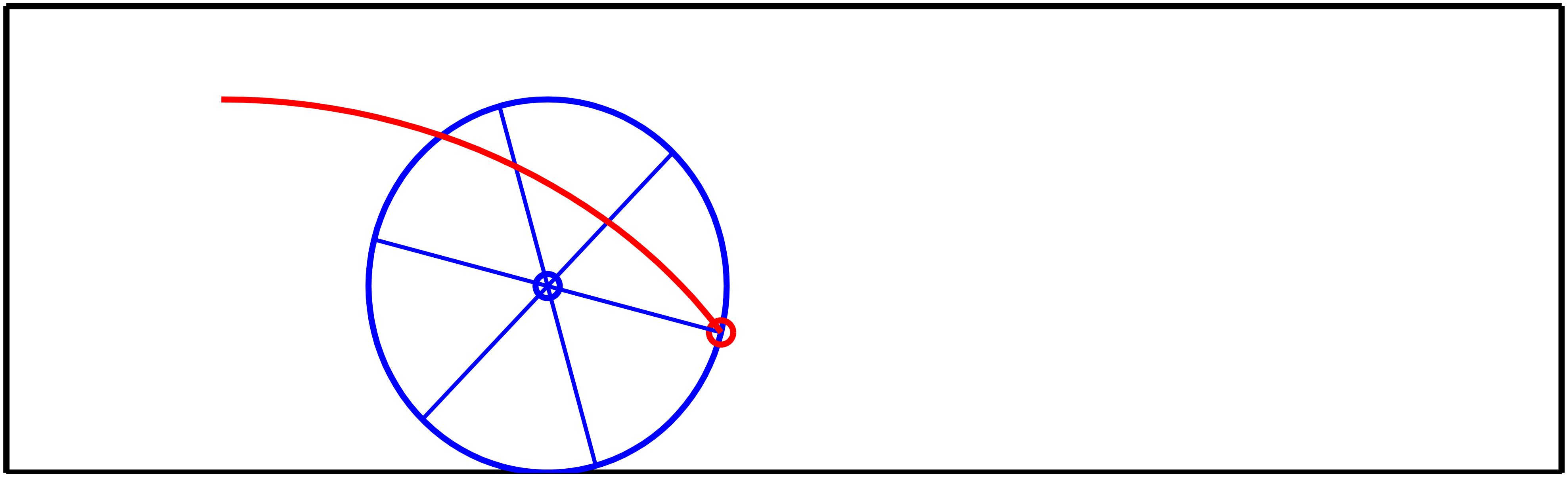
# Cycloid generation



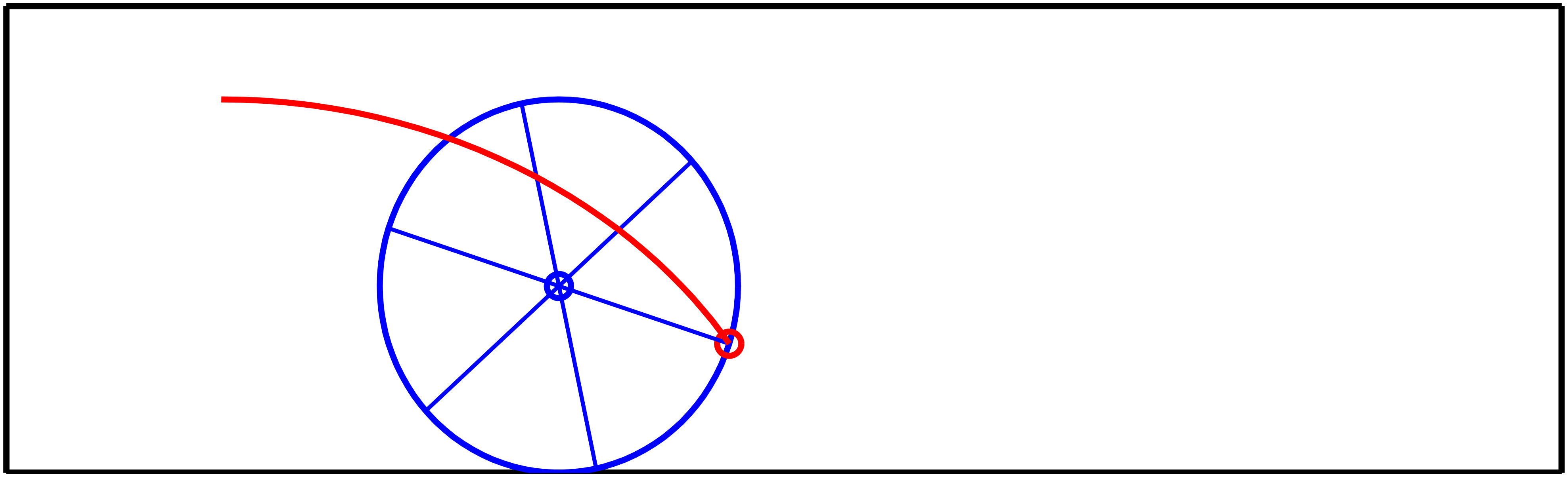
# Cycloid generation



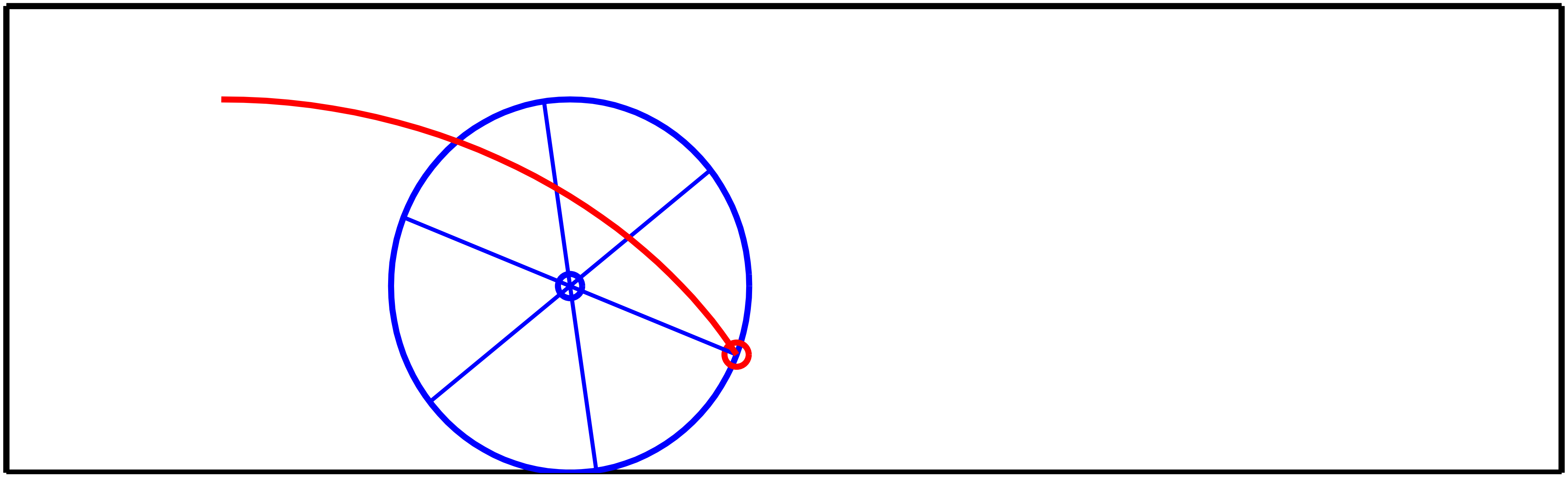
# Cycloid generation



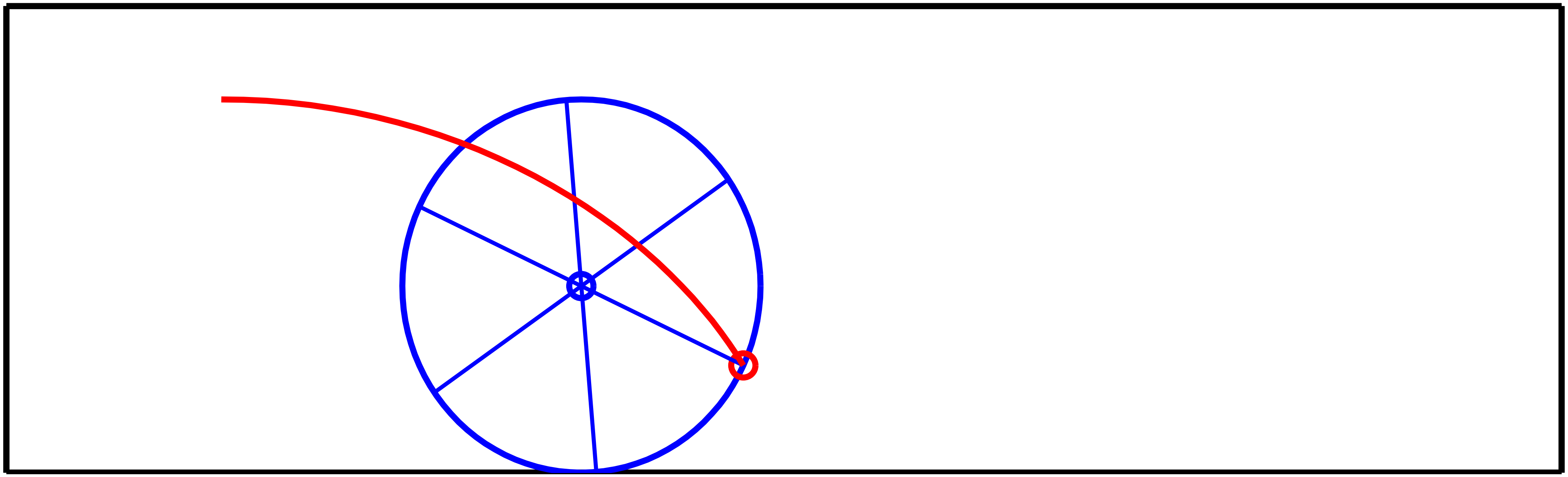
# Cycloid generation



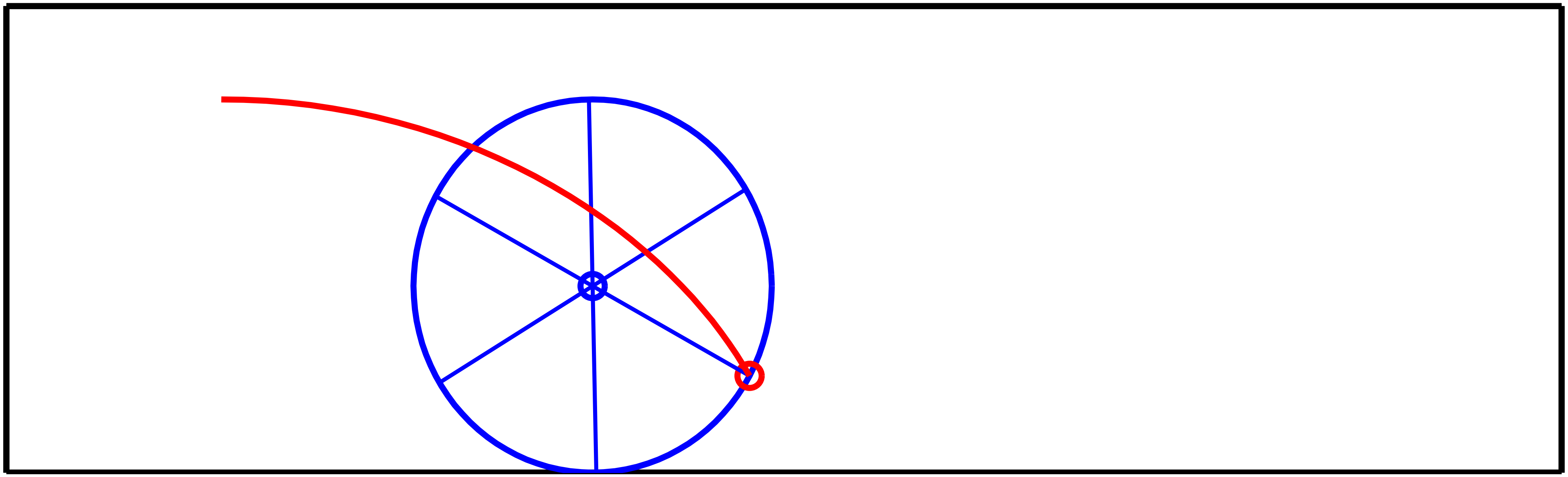
# Cycloid generation



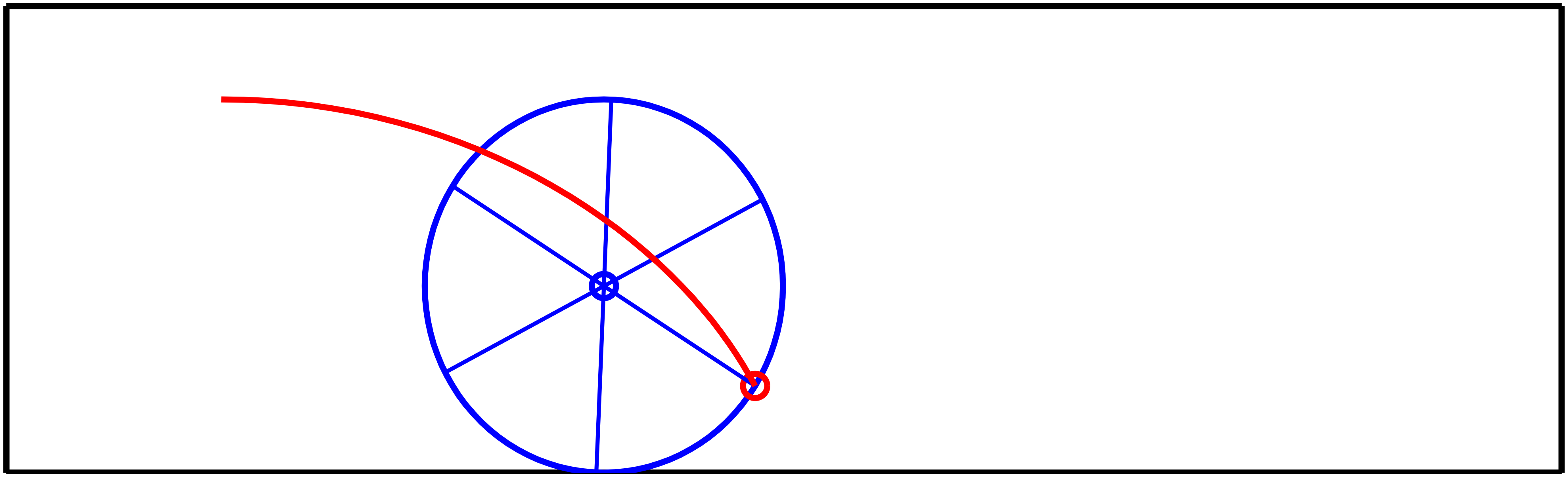
# Cycloid generation



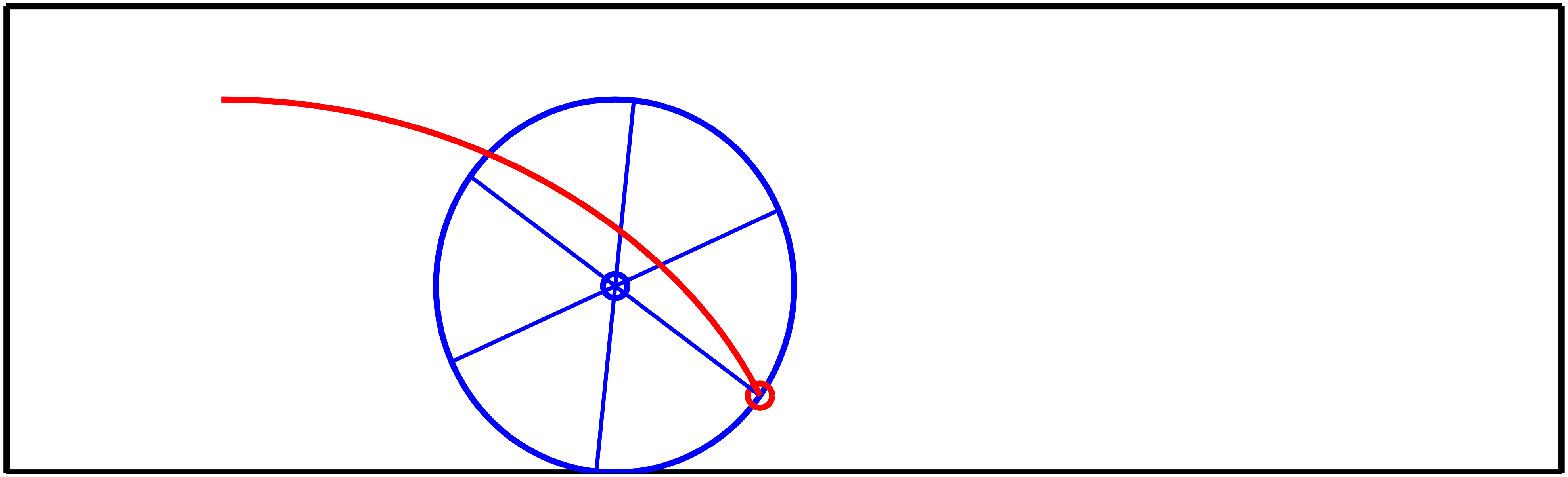
# Cycloid generation



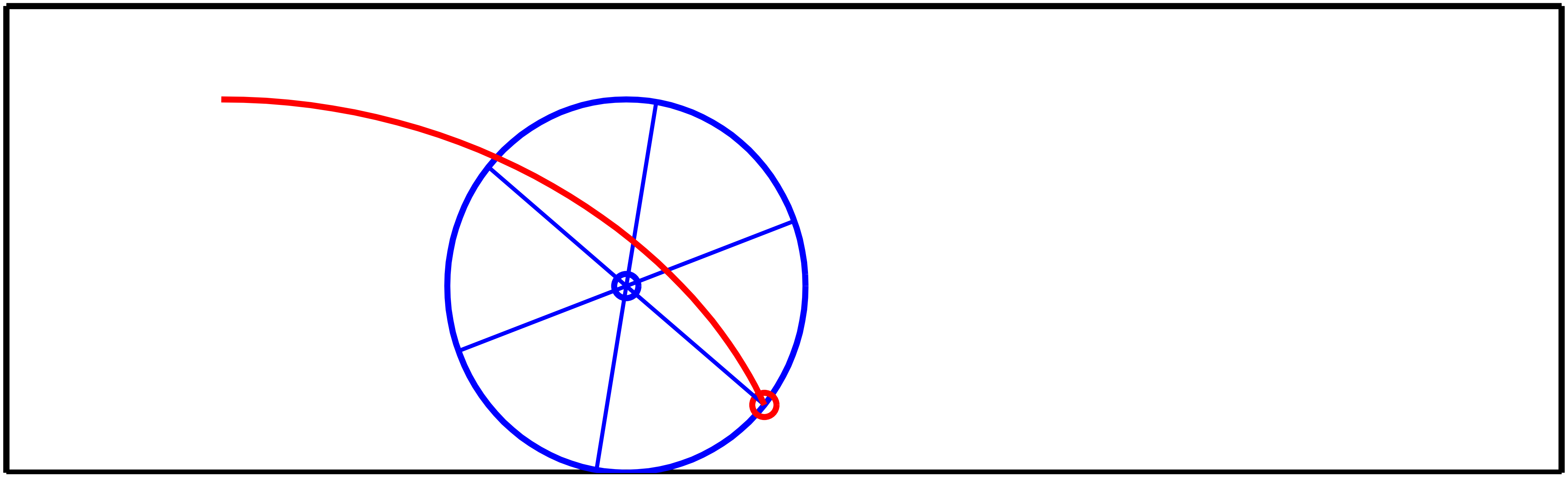
# Cycloid generation



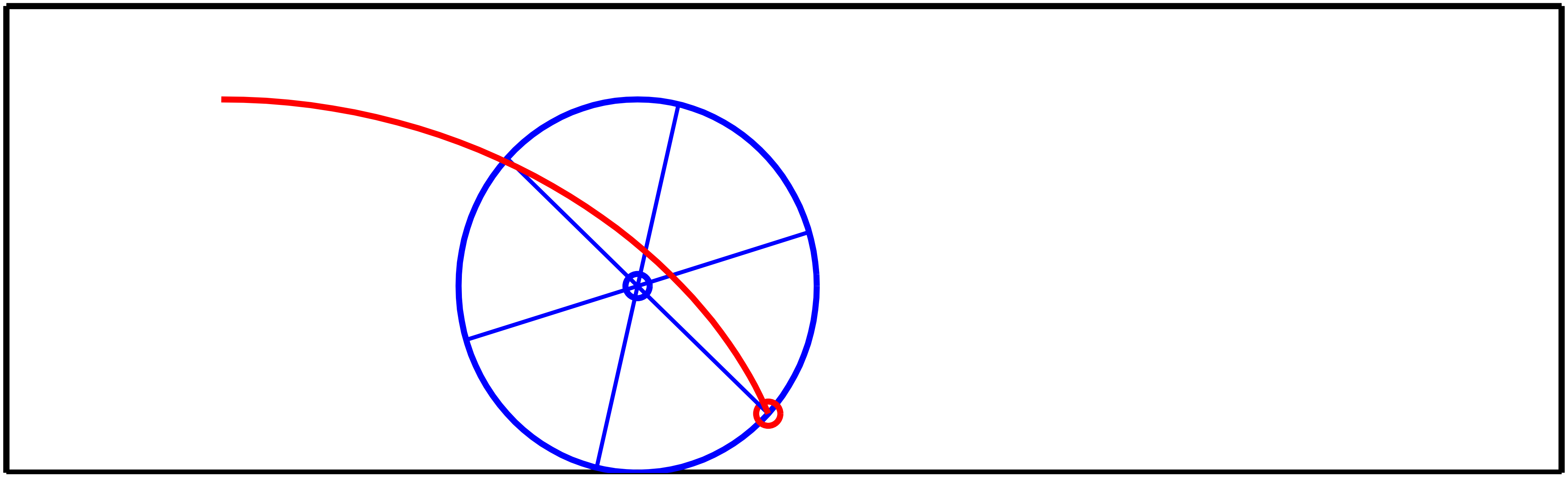
# Cycloid generation



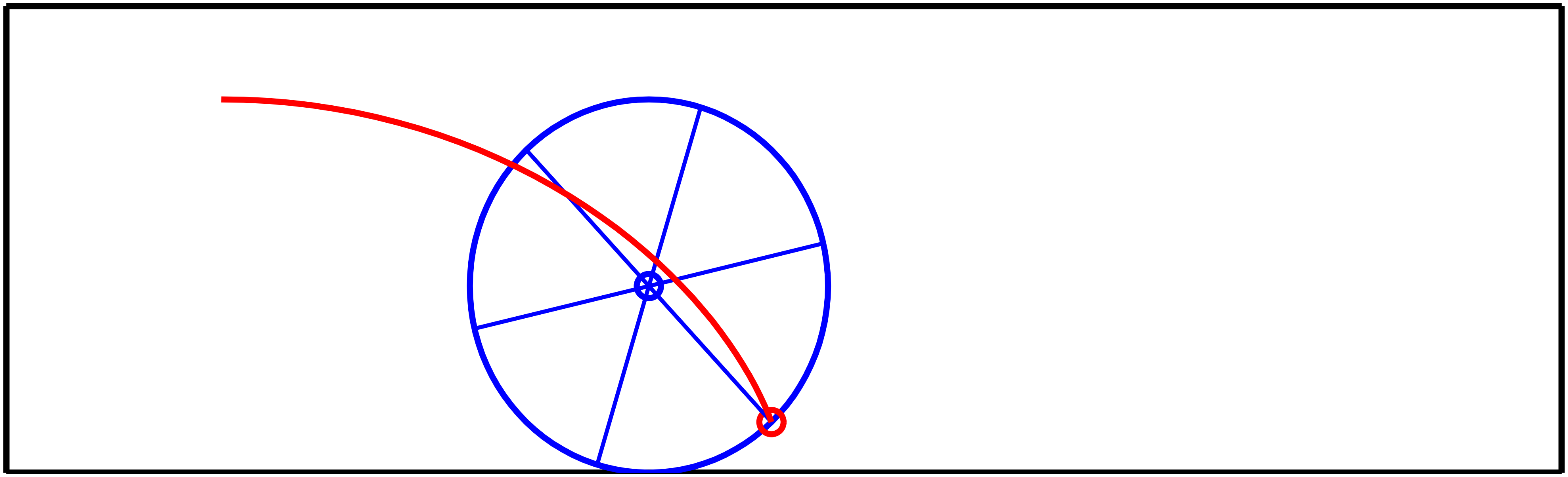
# Cycloid generation



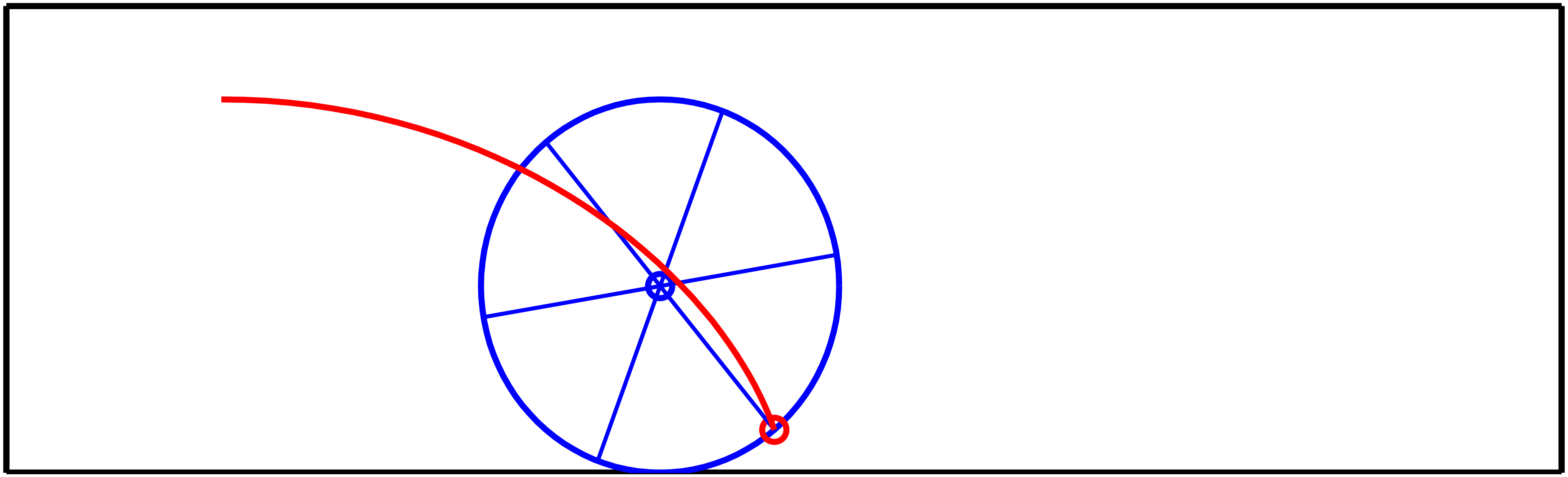
# Cycloid generation



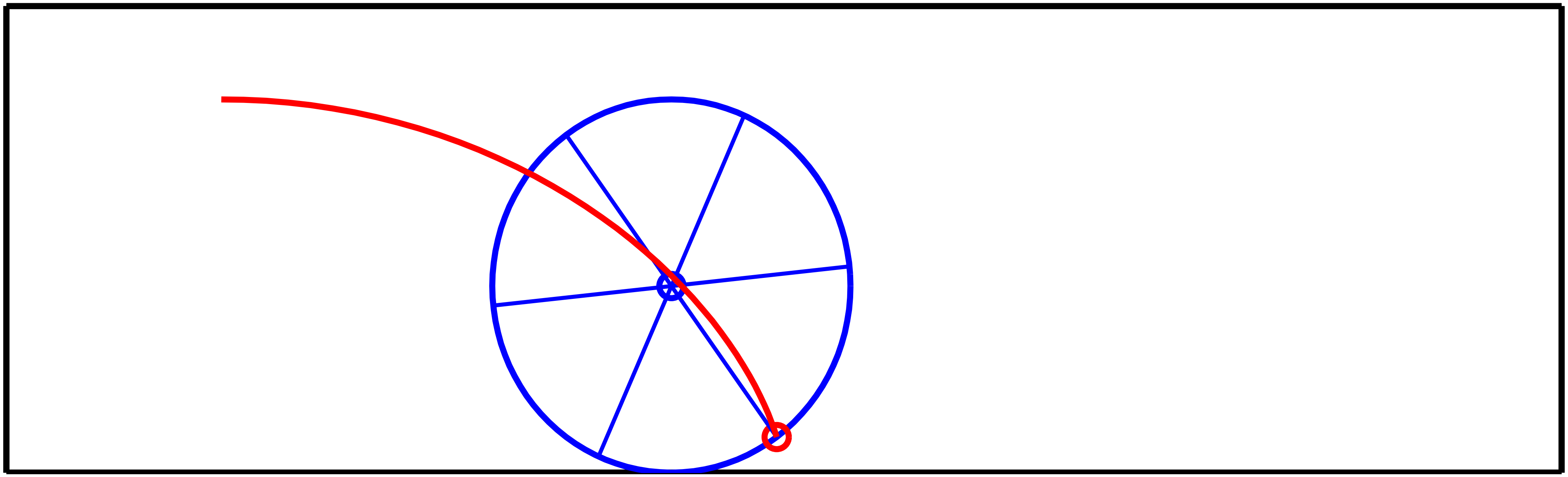
# Cycloid generation



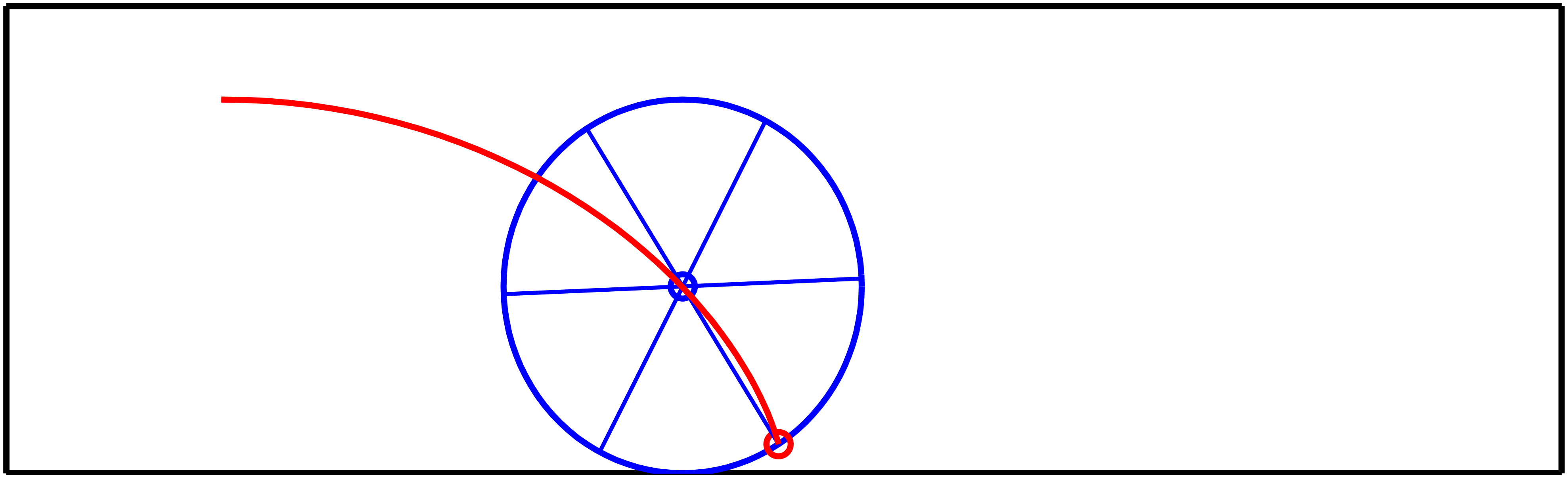
# Cycloid generation



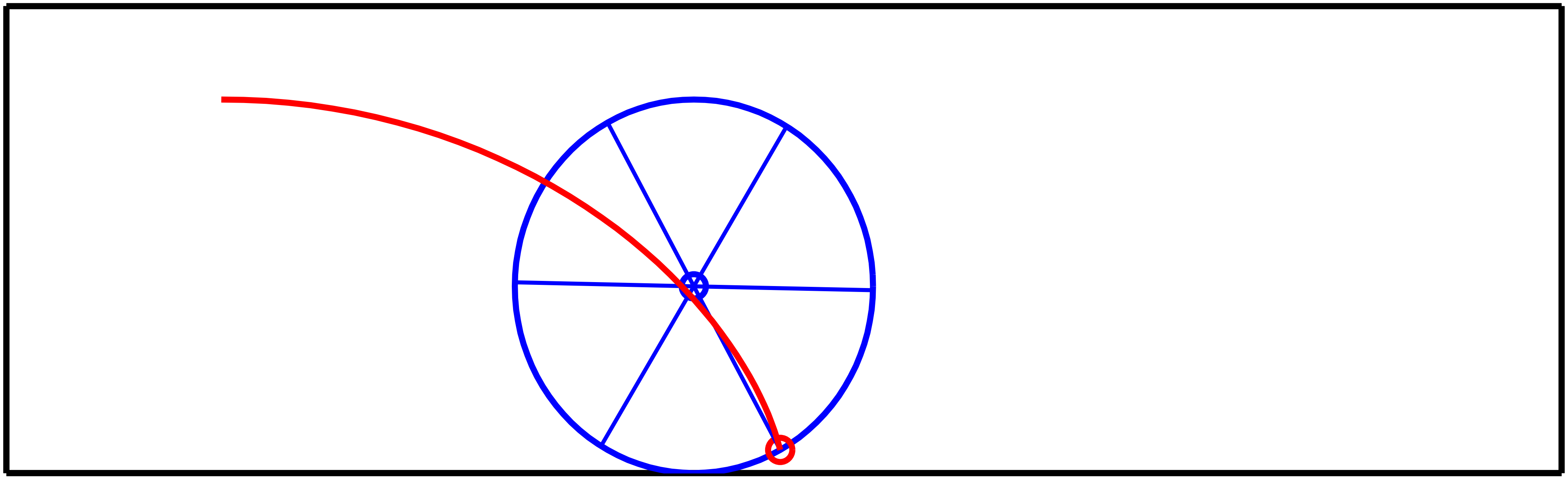
# Cycloid generation



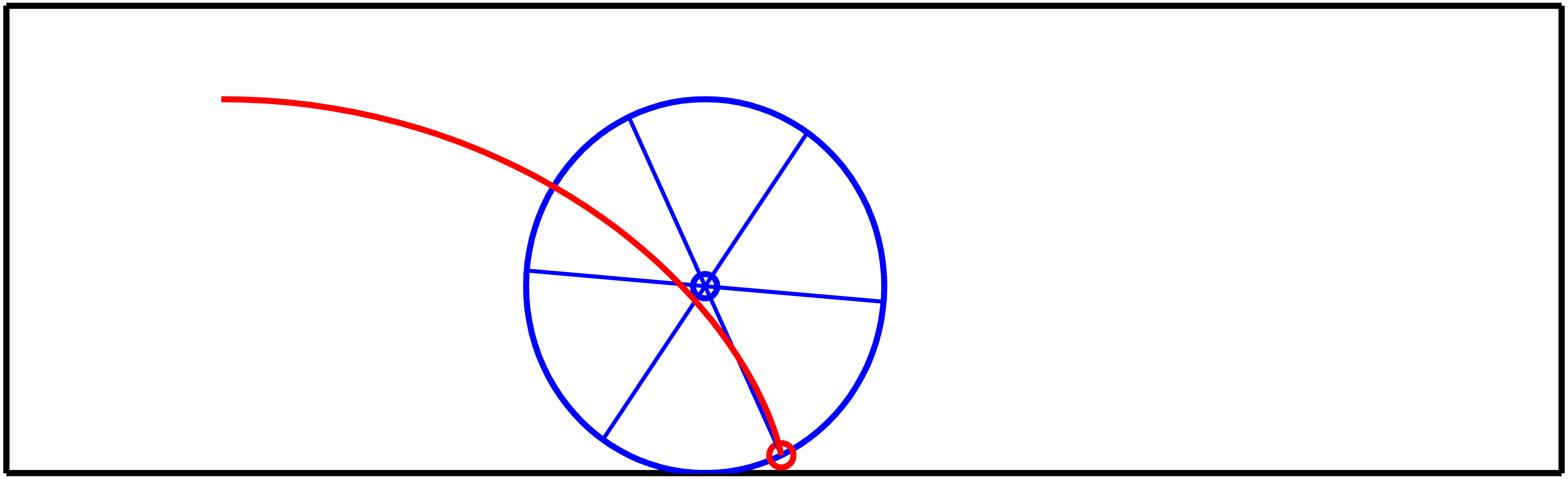
# Cycloid generation



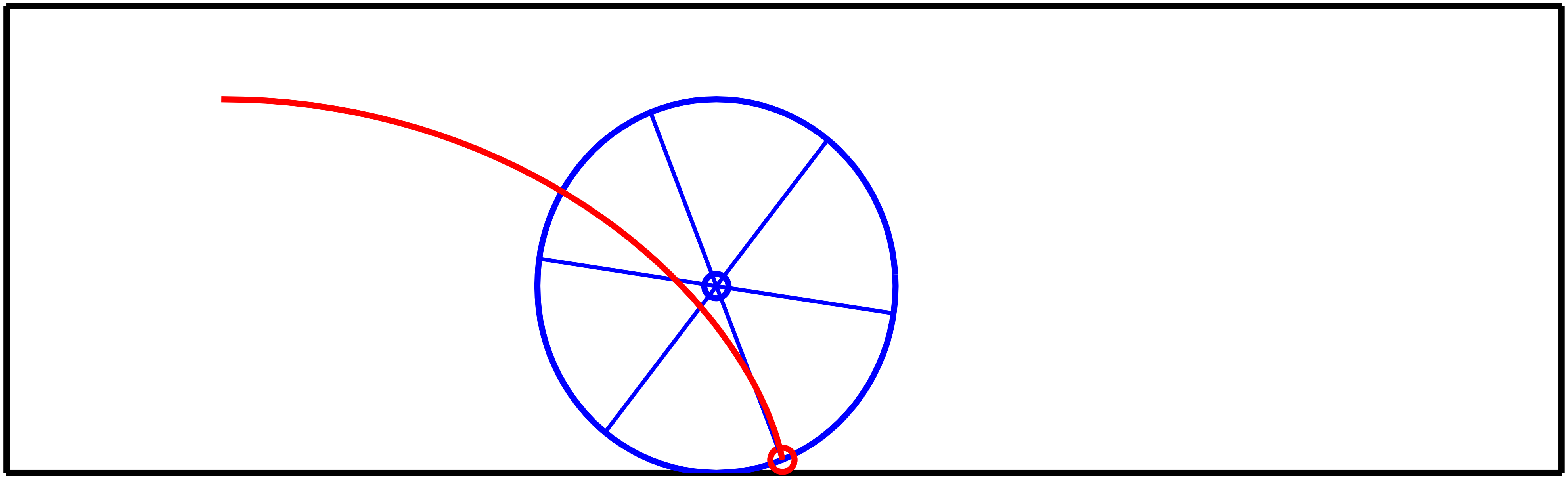
# Cycloid generation



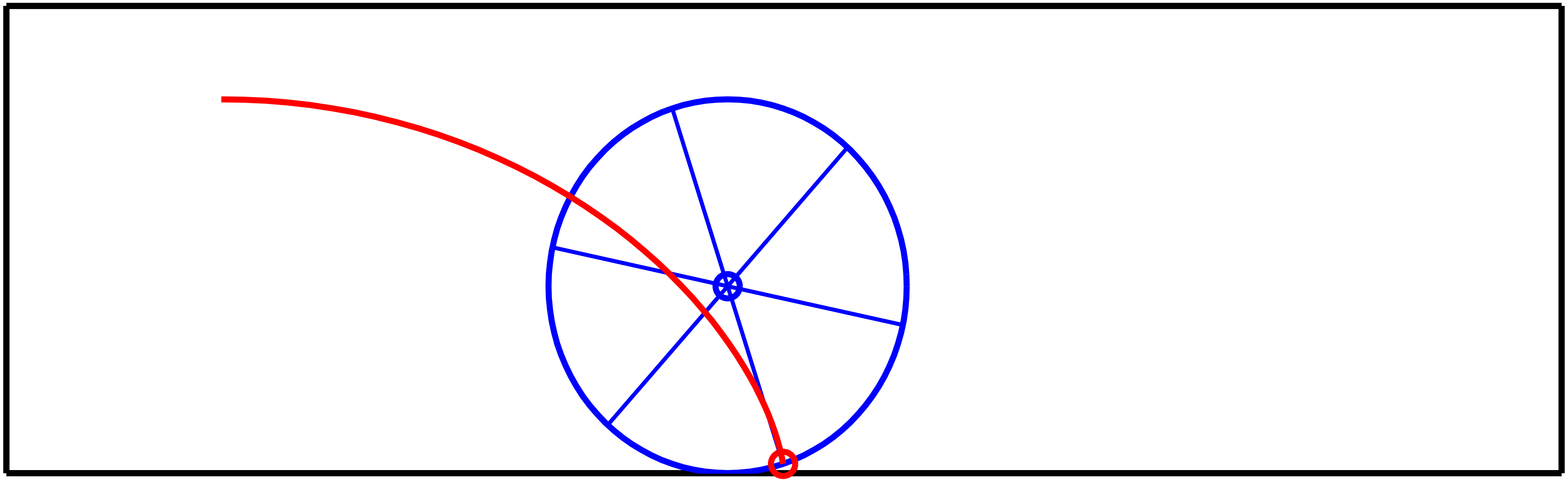
# Cycloid generation



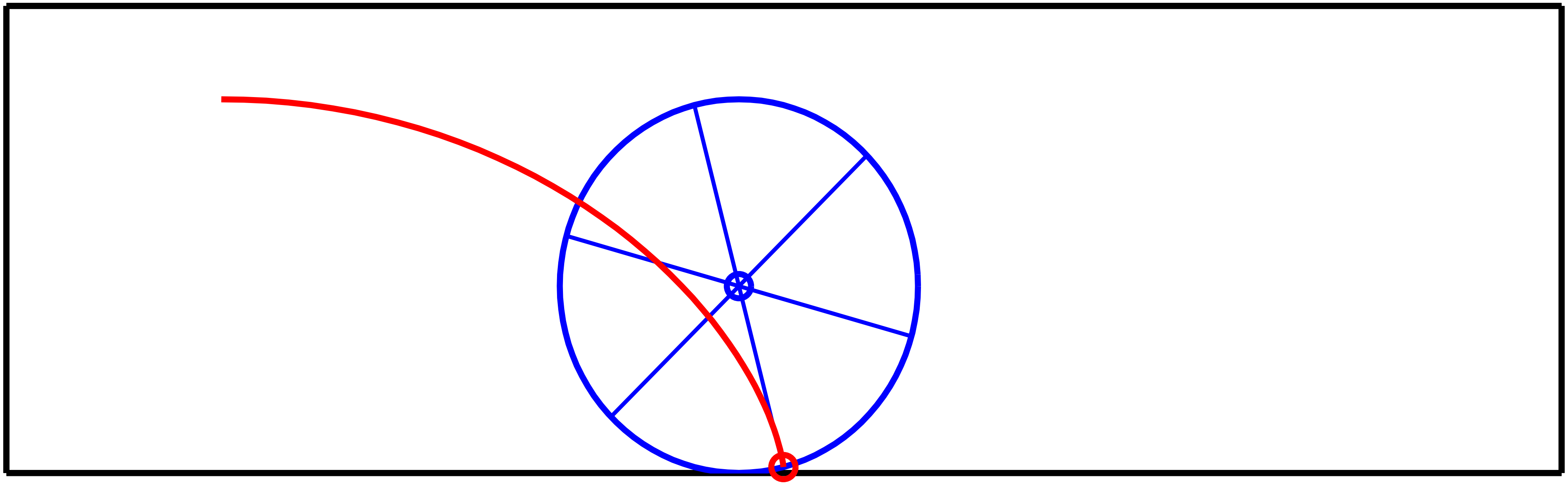
# Cycloid generation



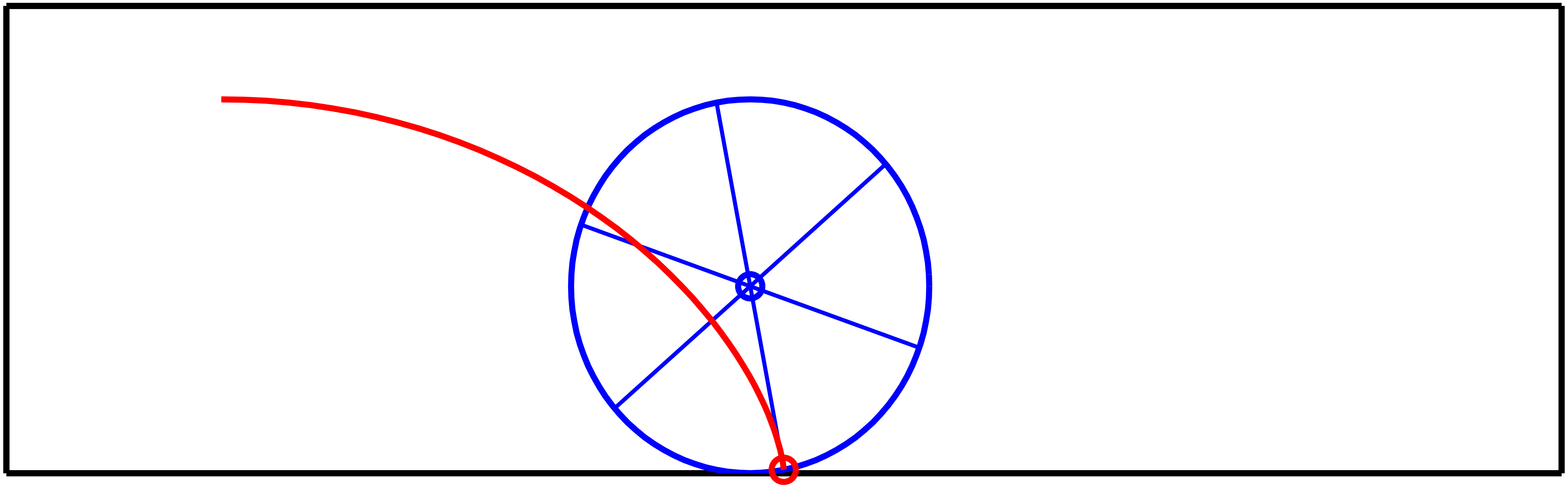
# Cycloid generation



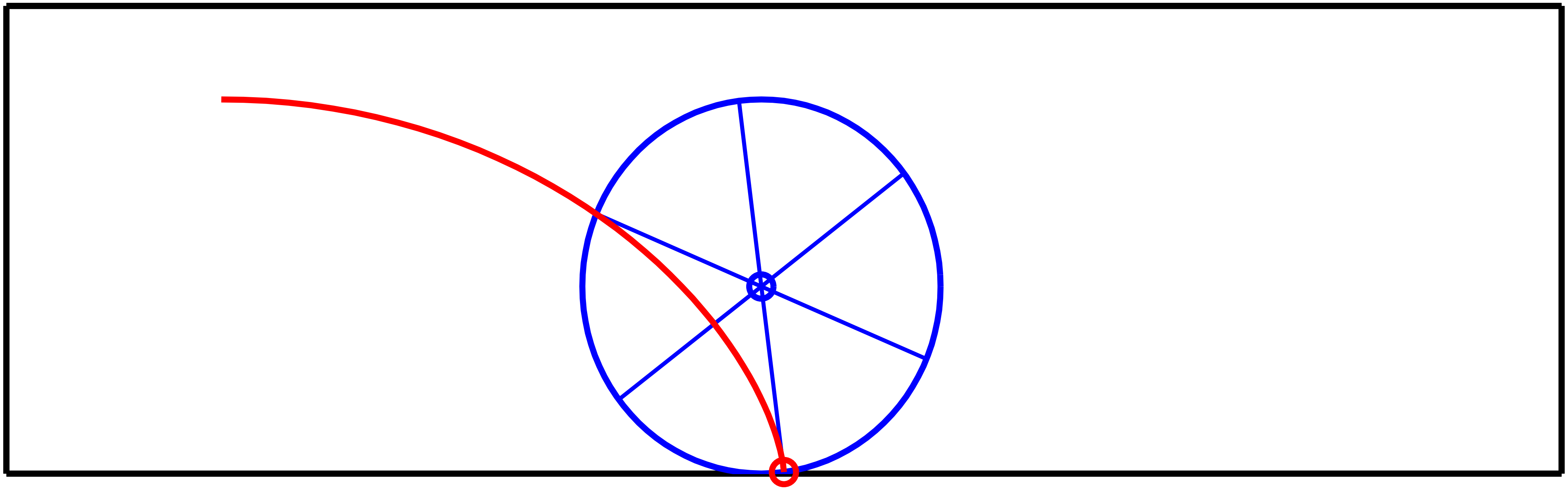
# Cycloid generation



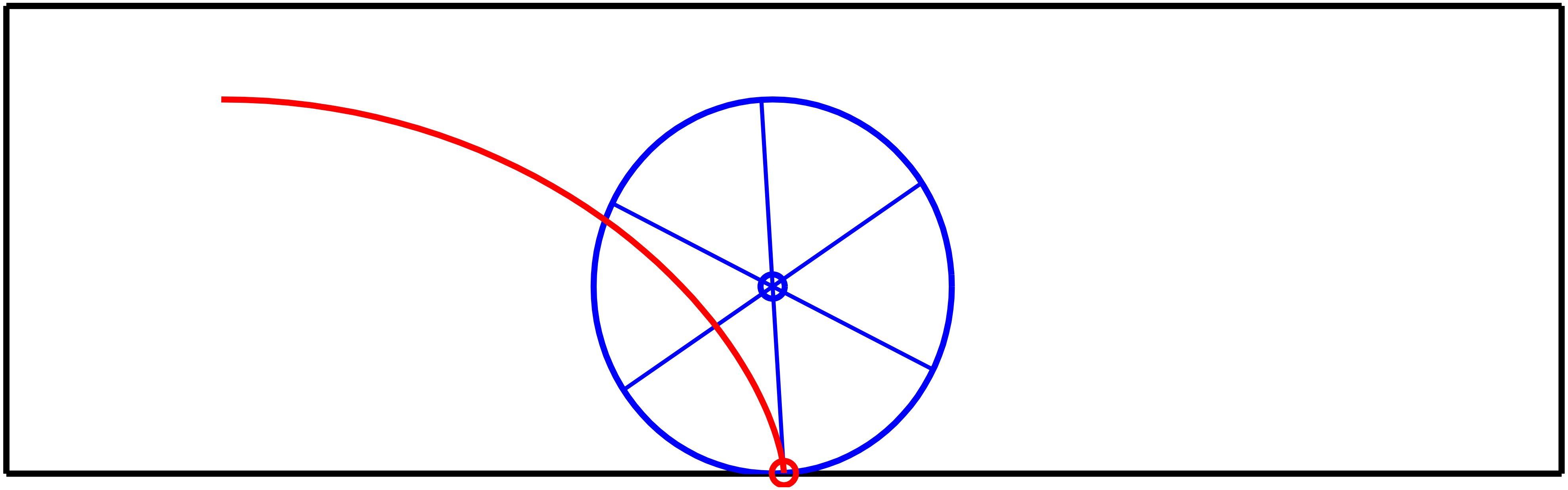
# Cycloid generation



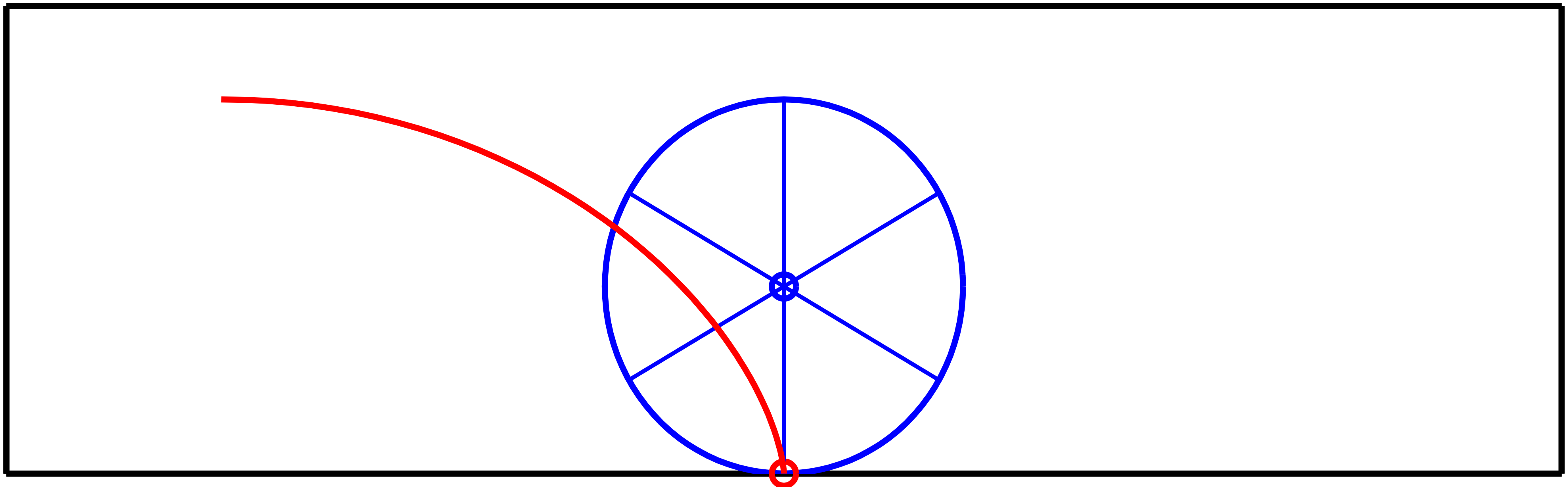
# Cycloid generation



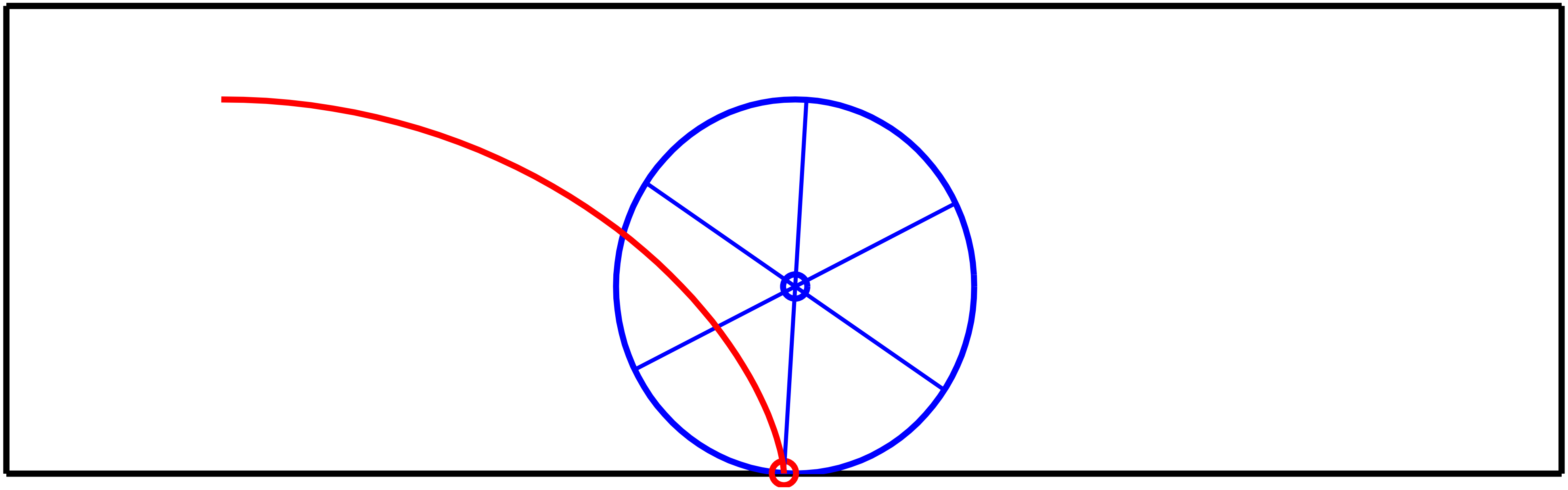
# Cycloid generation



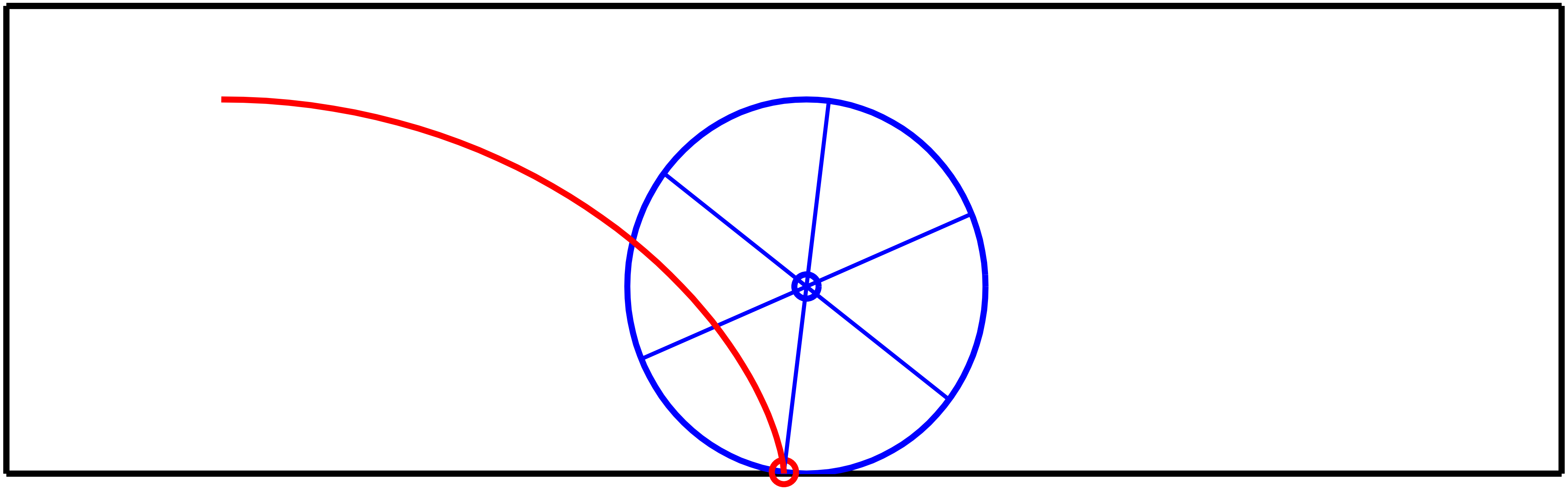
# Cycloid generation



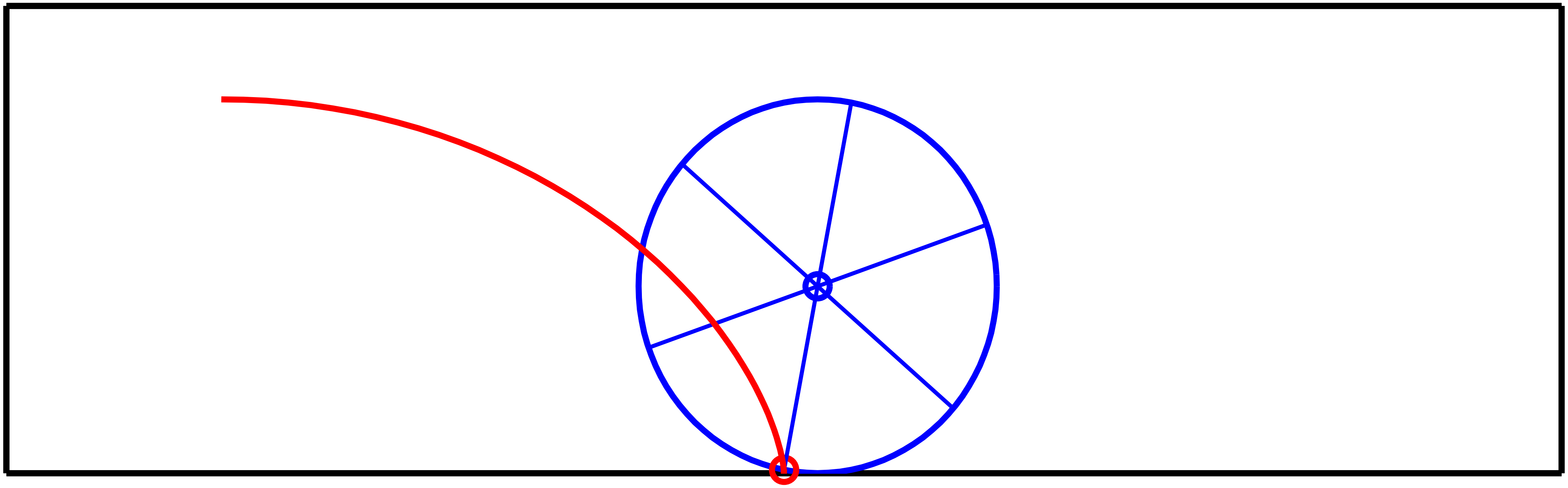
# Cycloid generation



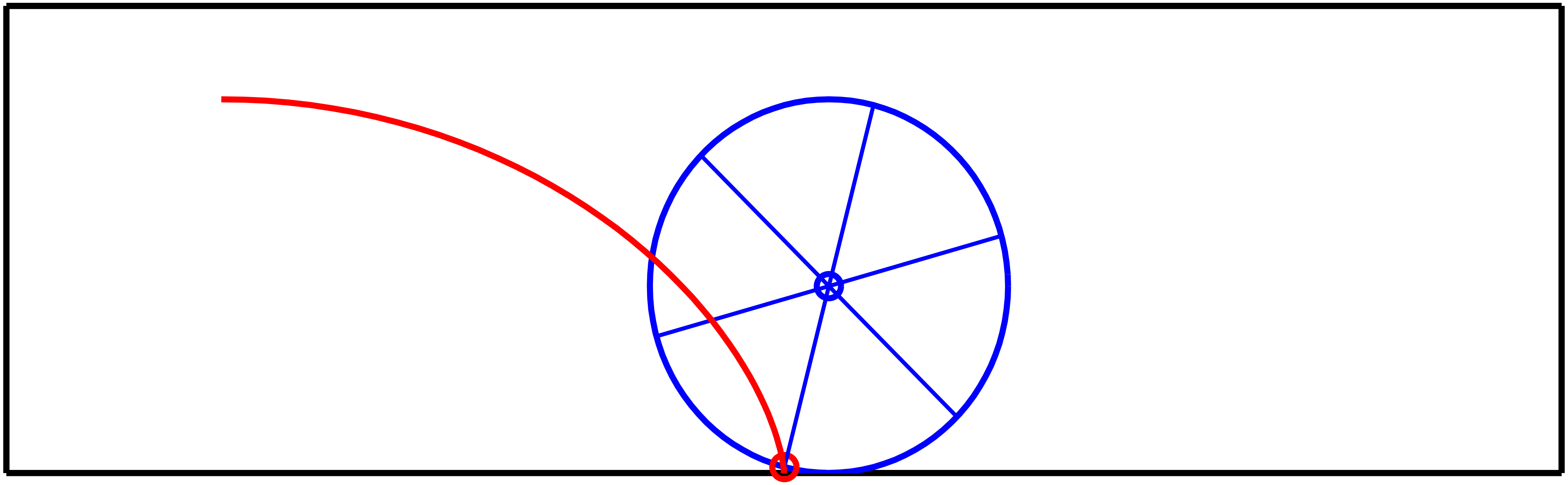
# Cycloid generation



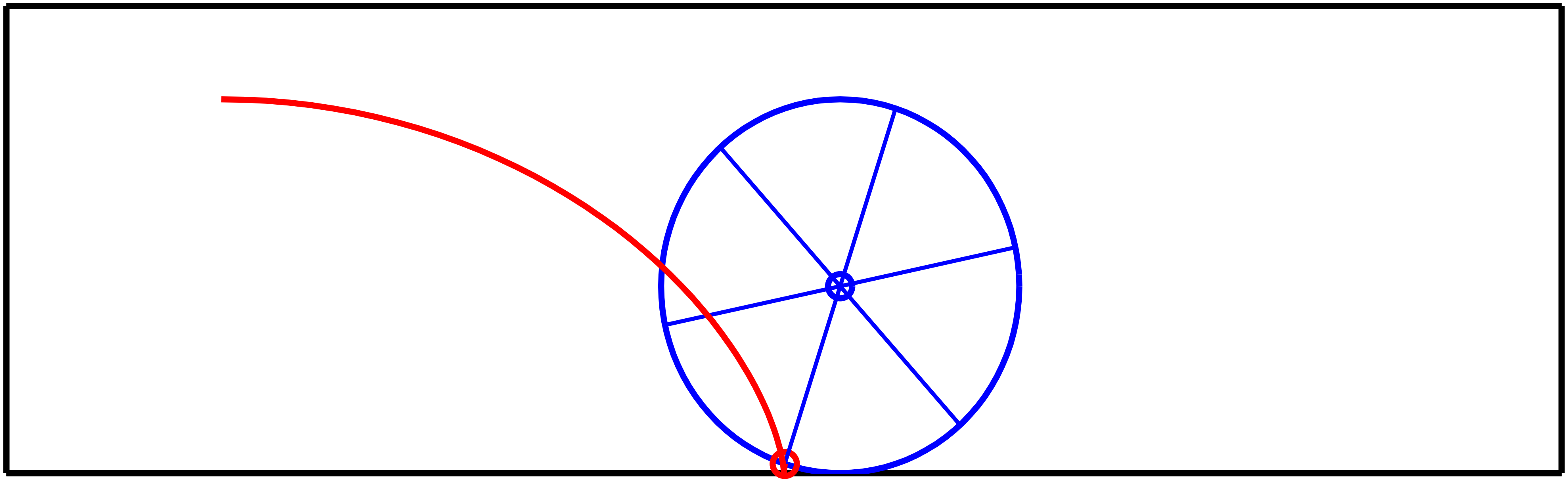
# Cycloid generation



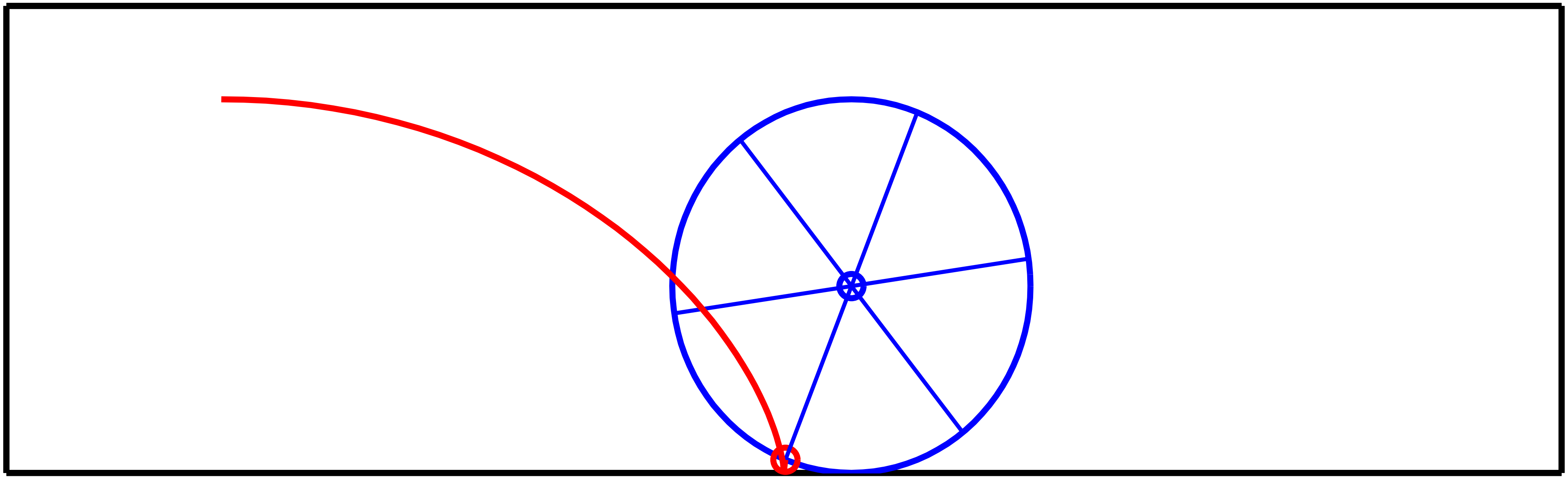
# Cycloid generation



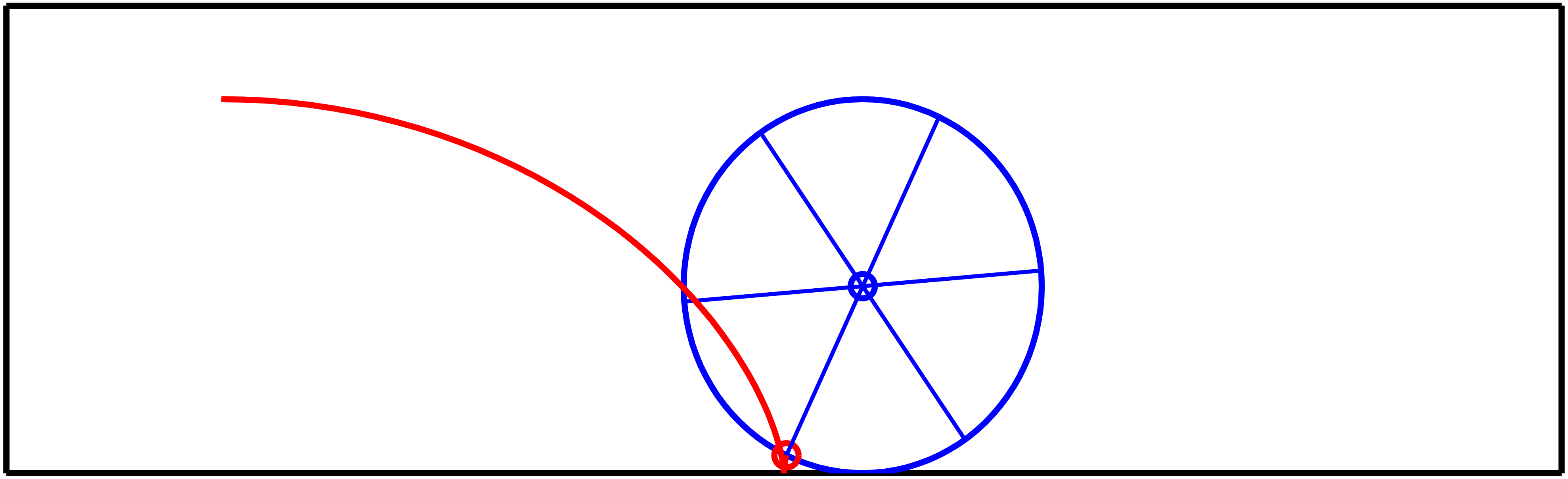
# Cycloid generation



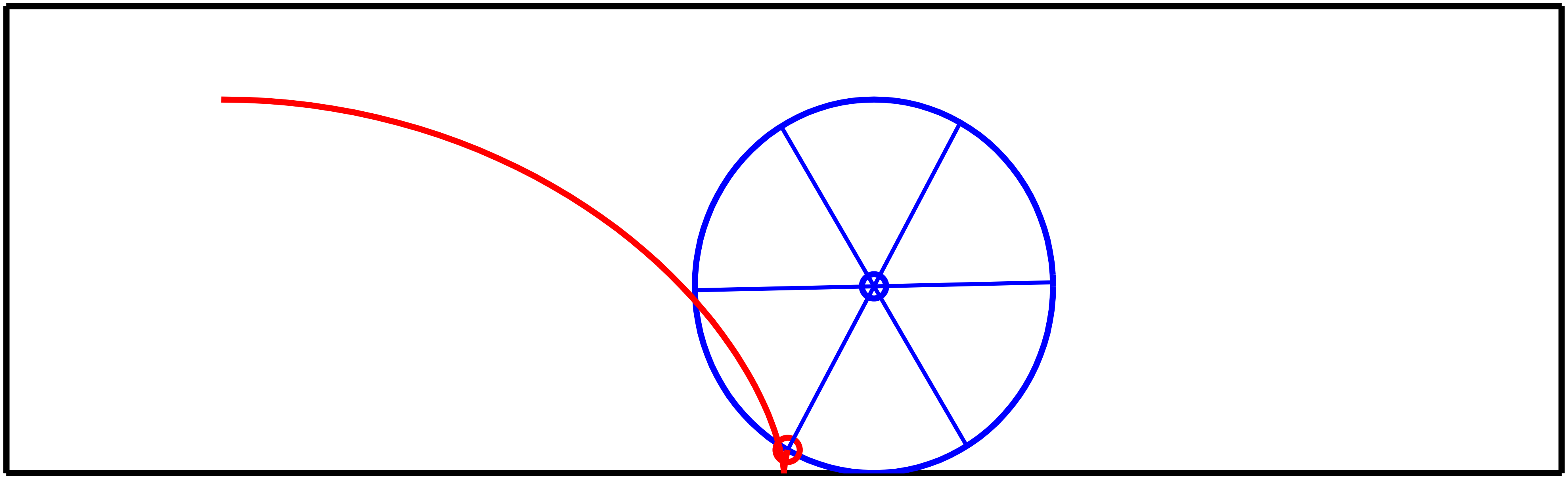
# Cycloid generation



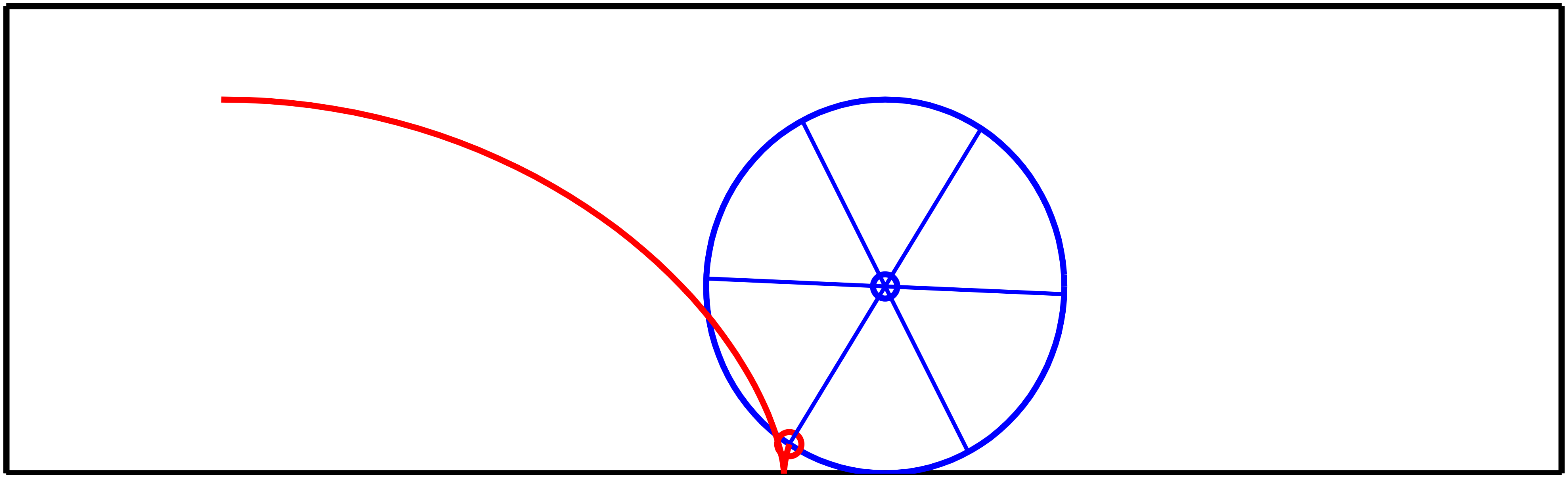
# Cycloid generation



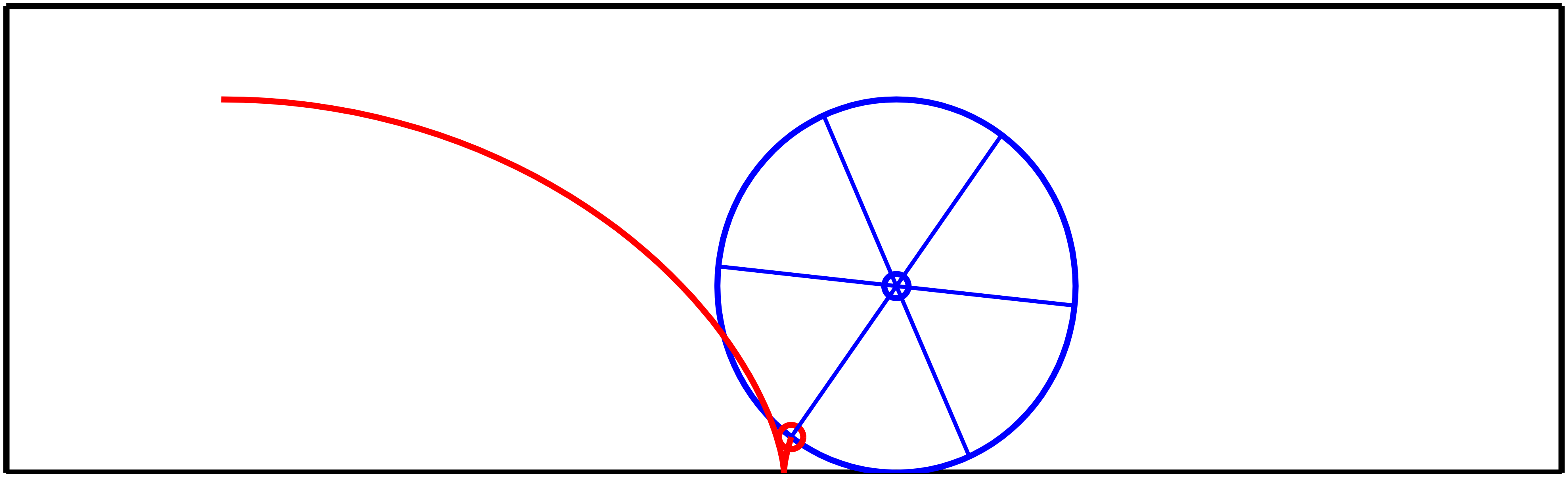
# Cycloid generation



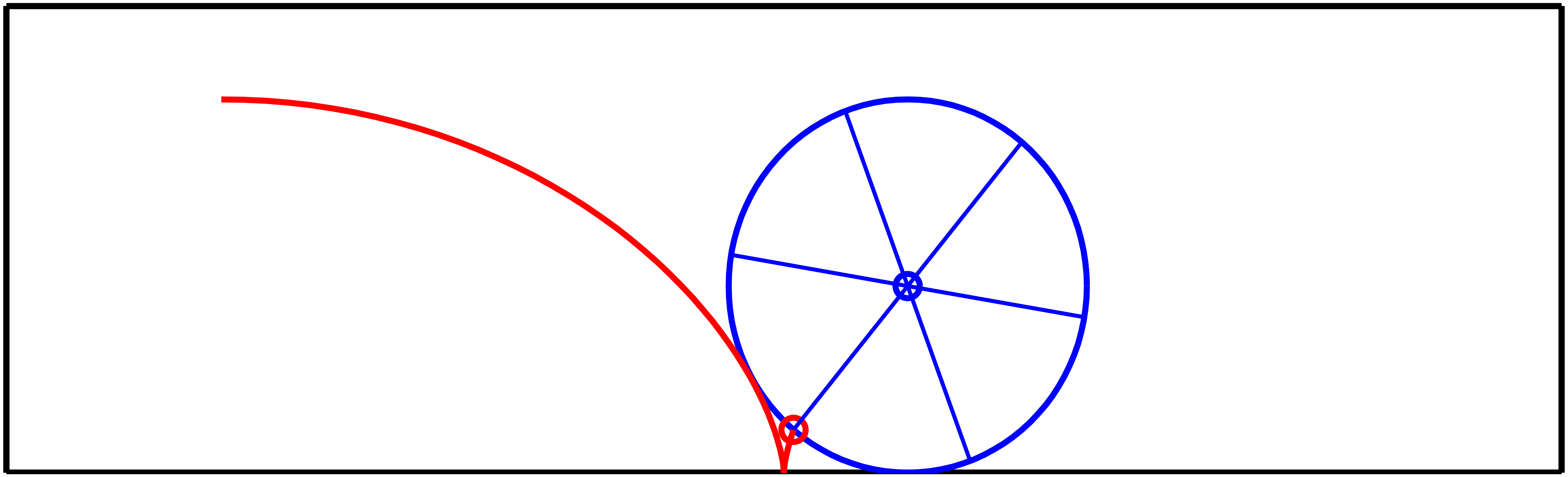
# Cycloid generation



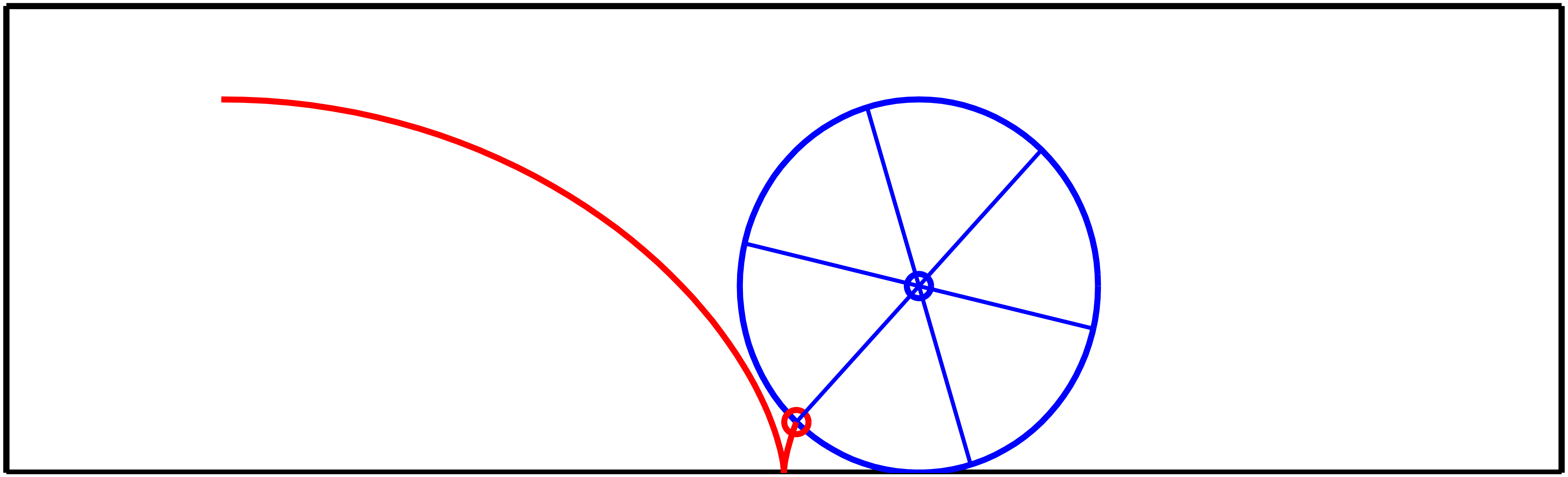
# Cycloid generation



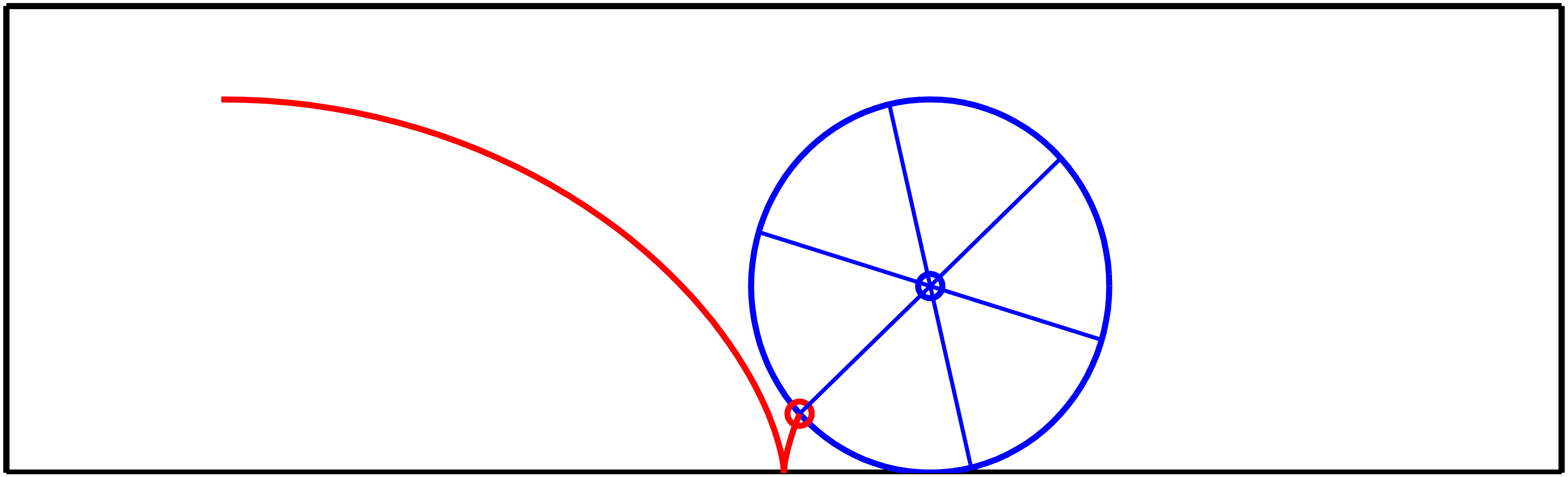
# Cycloid generation



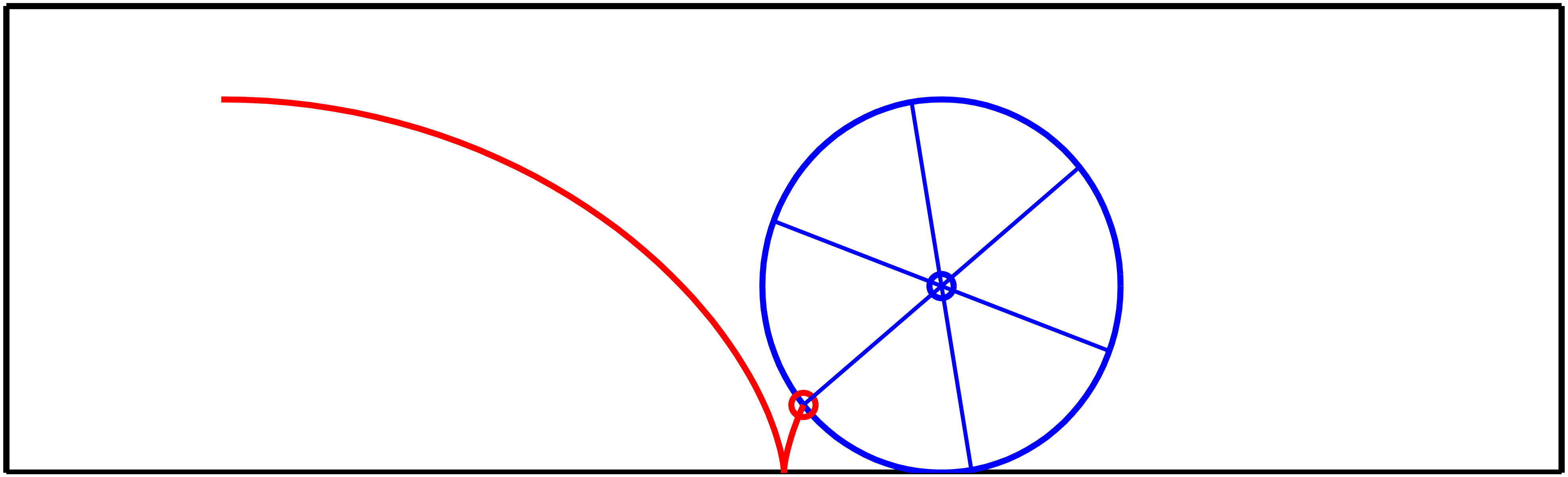
# Cycloid generation



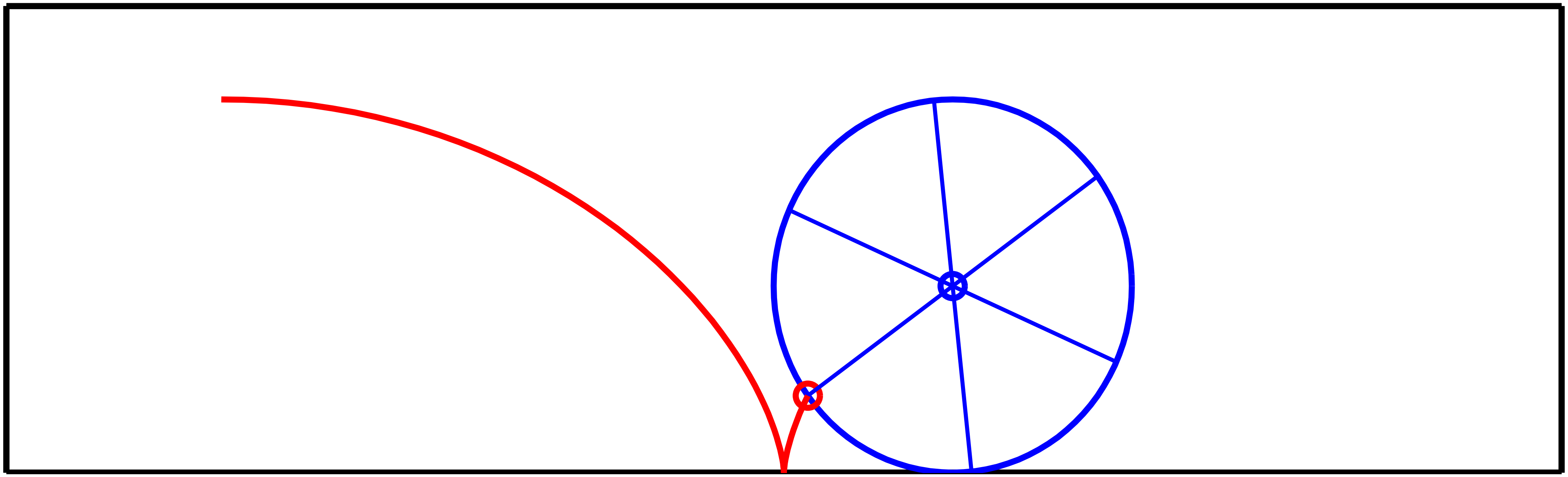
# Cycloid generation



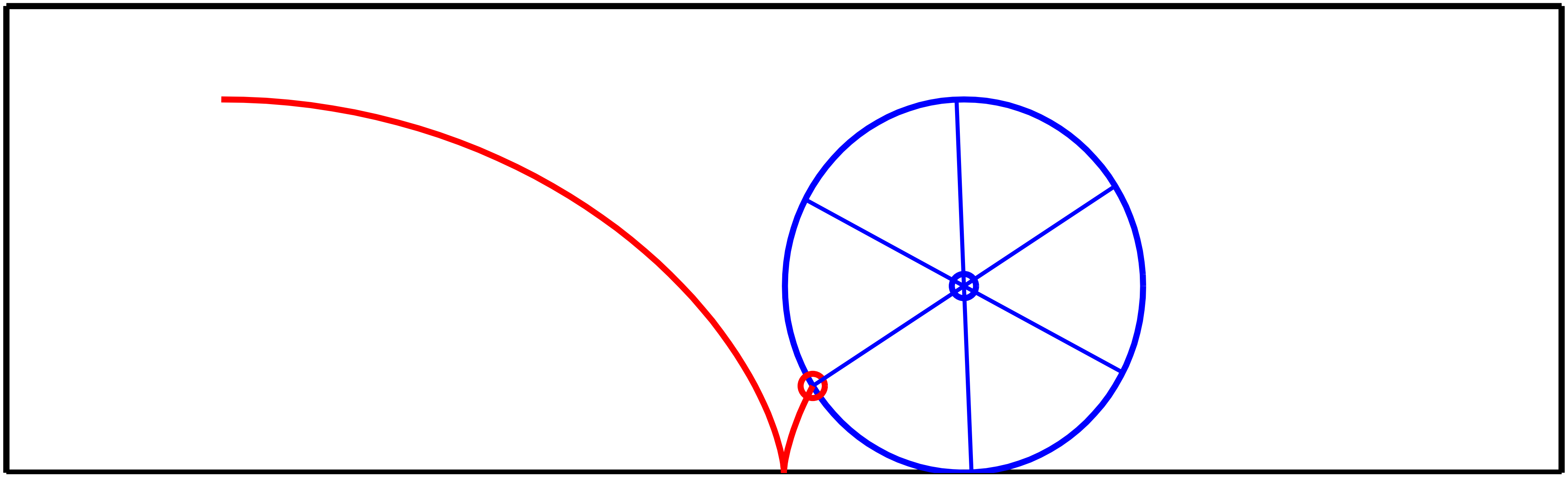
# Cycloid generation



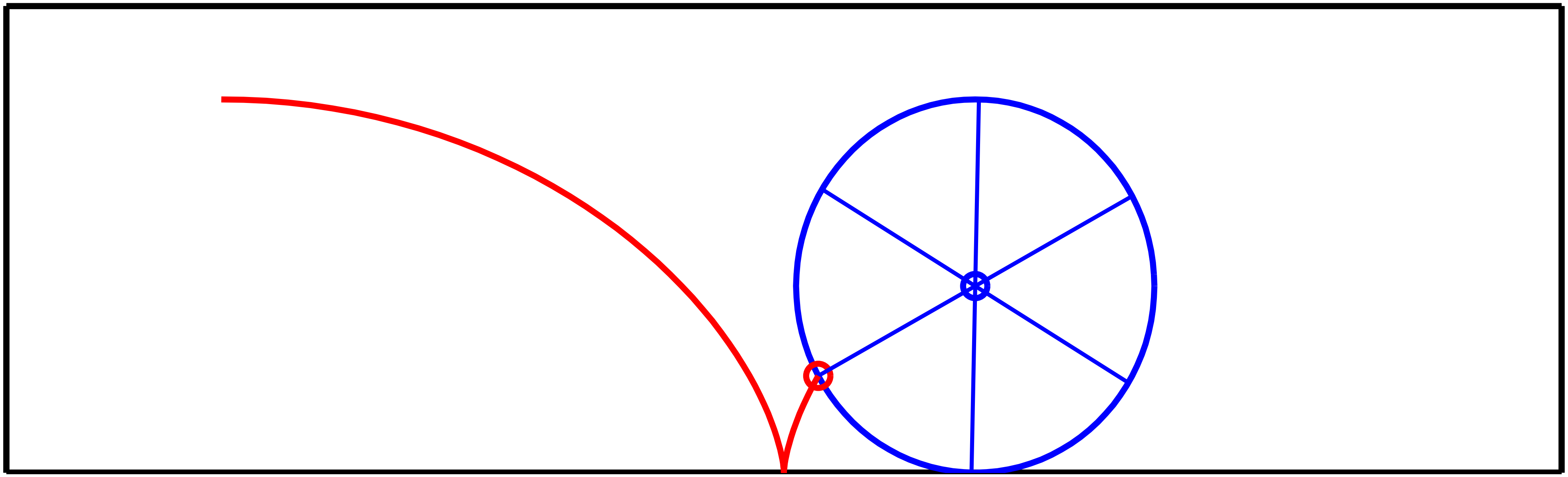
# Cycloid generation



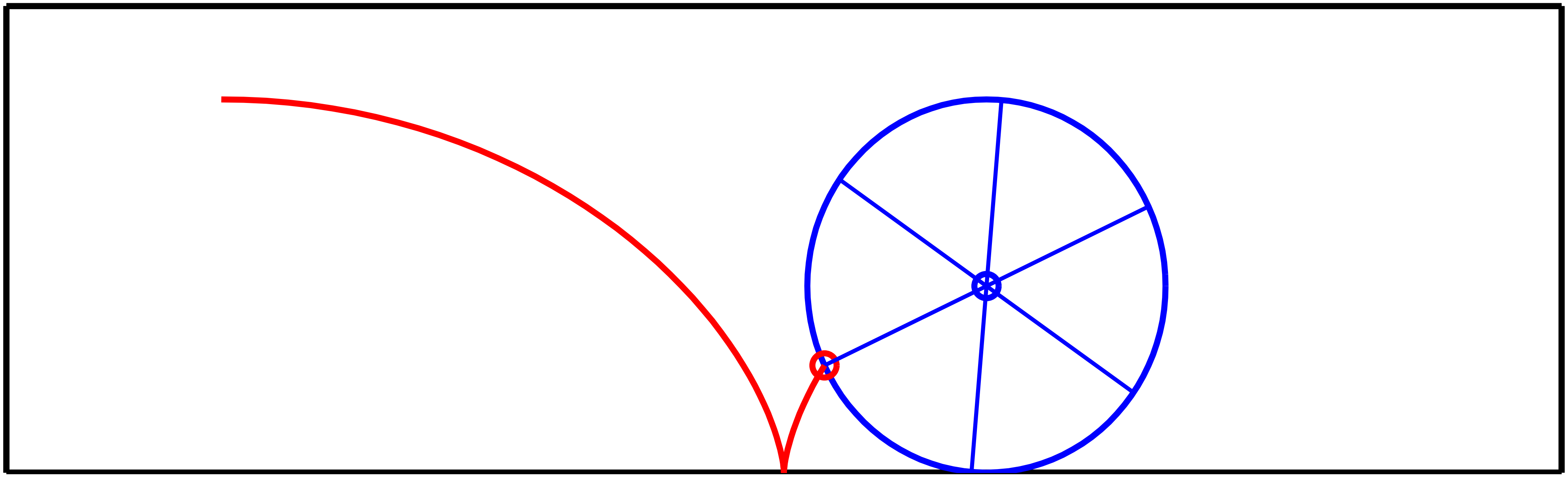
# Cycloid generation



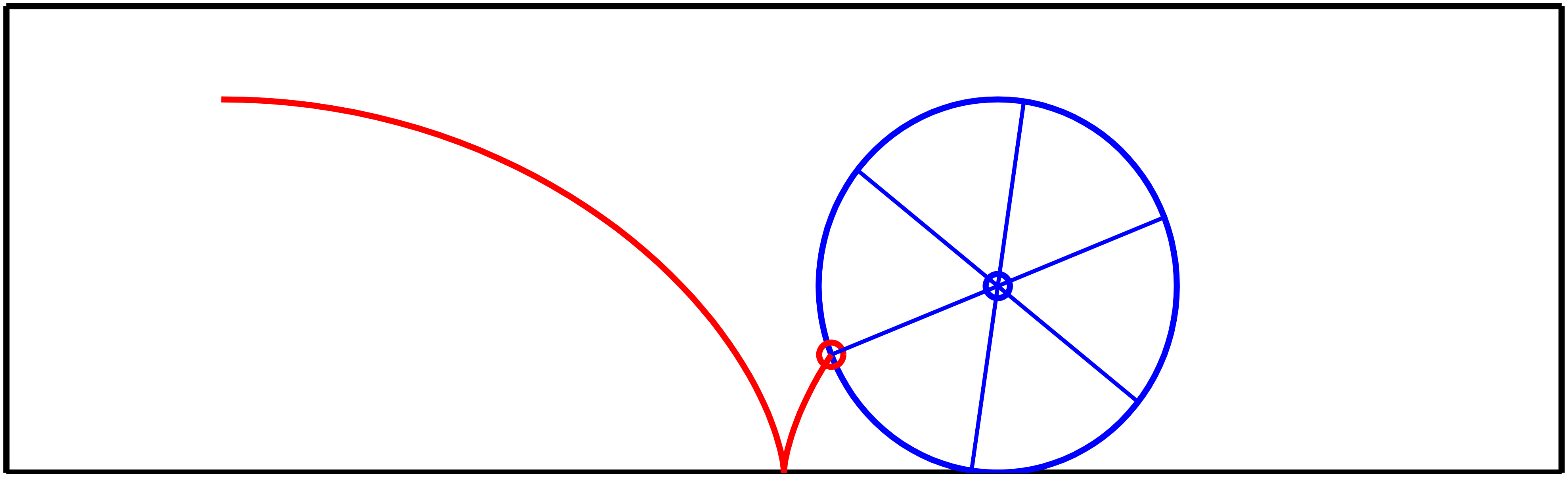
# Cycloid generation



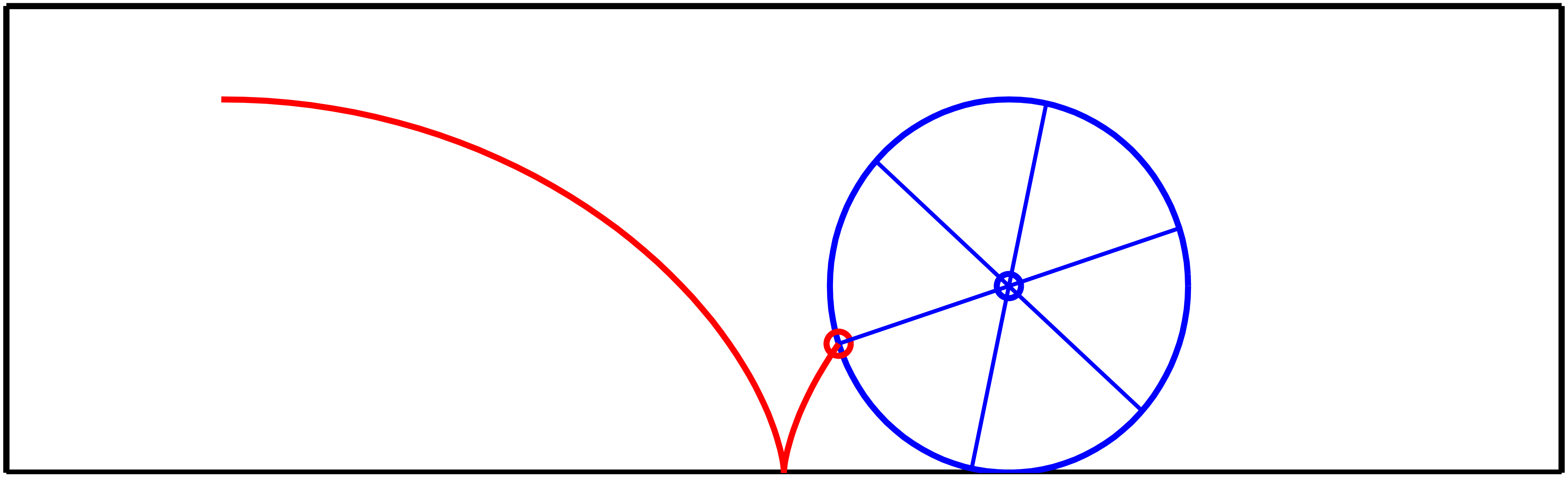
# Cycloid generation



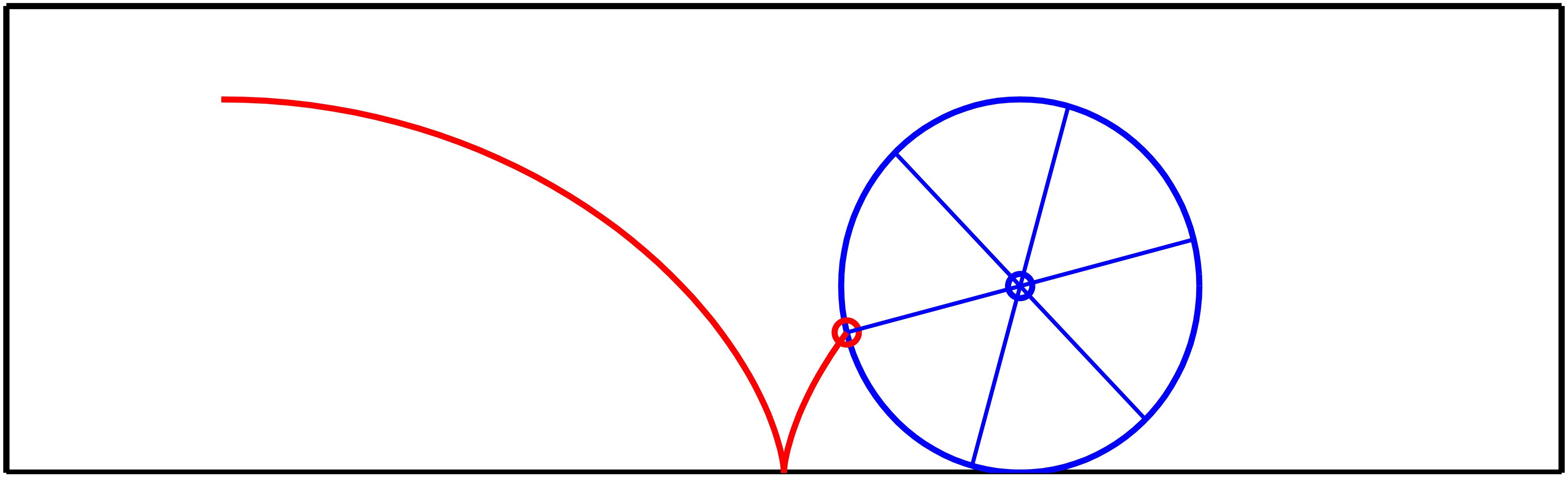
# Cycloid generation



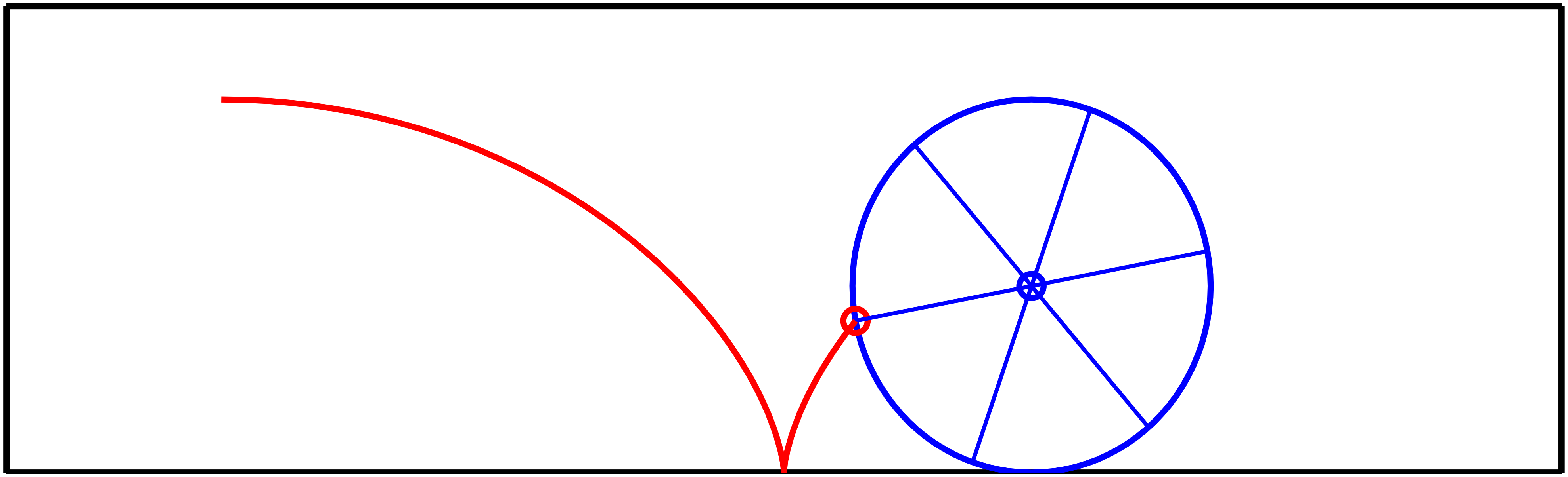
# Cycloid generation



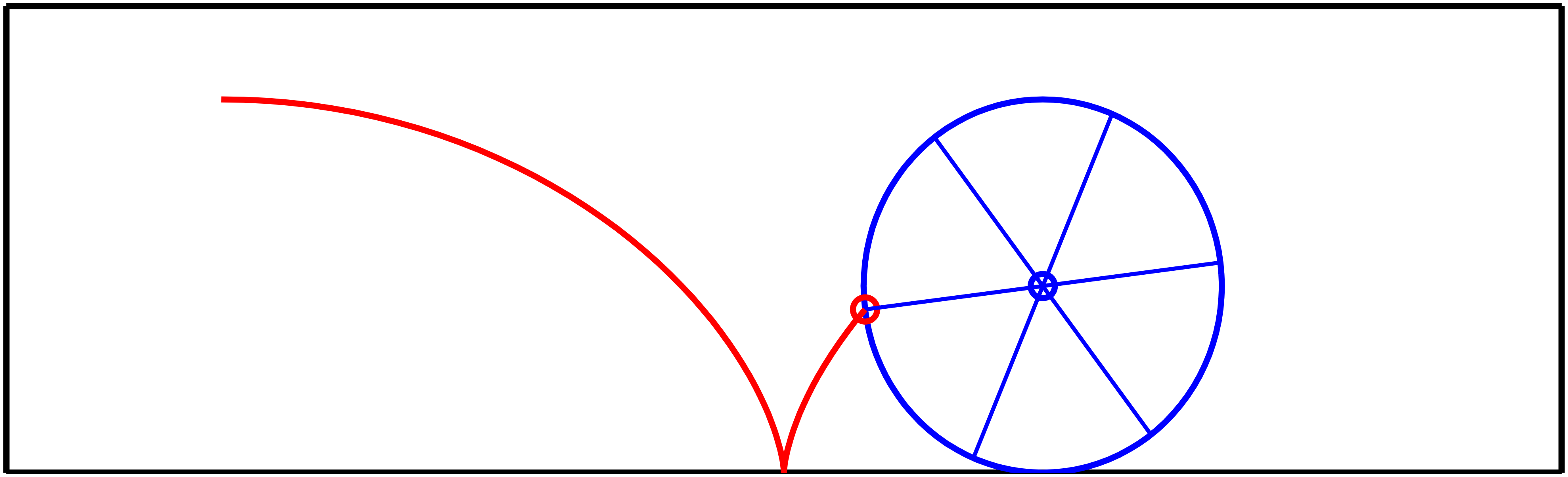
# Cycloid generation



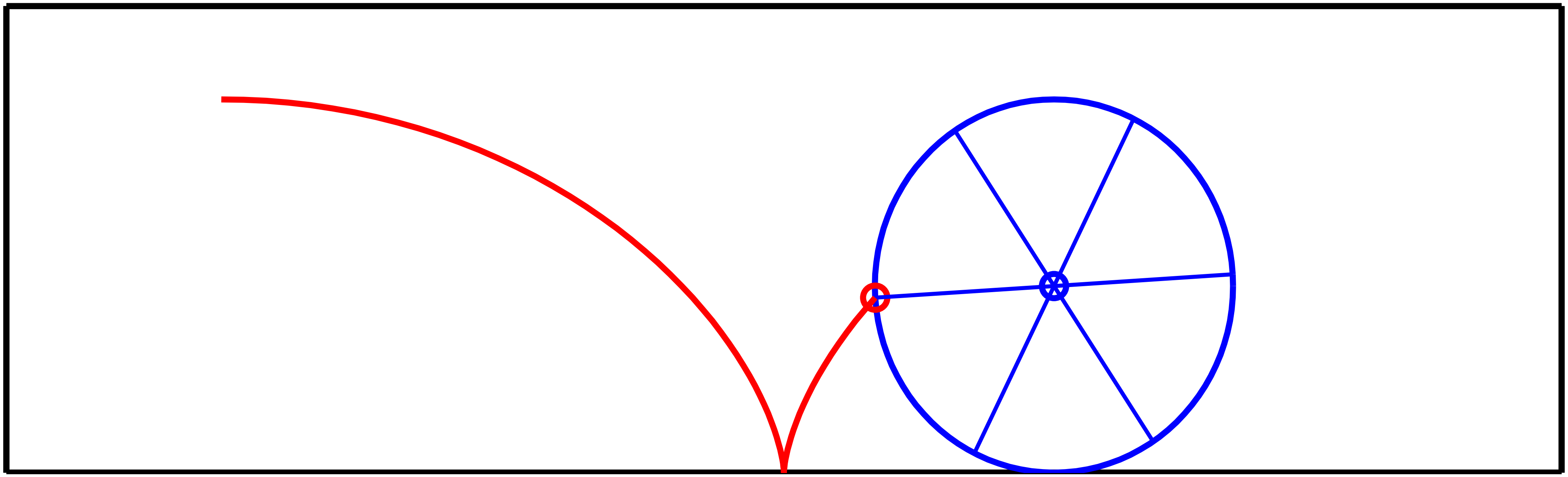
# Cycloid generation



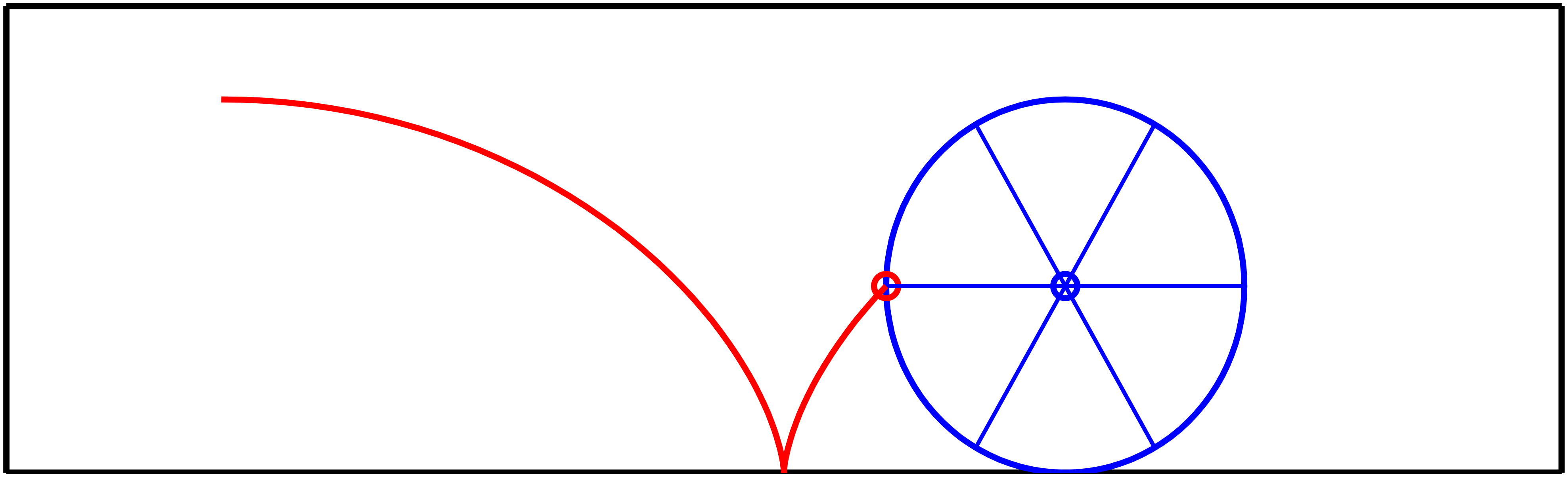
# Cycloid generation



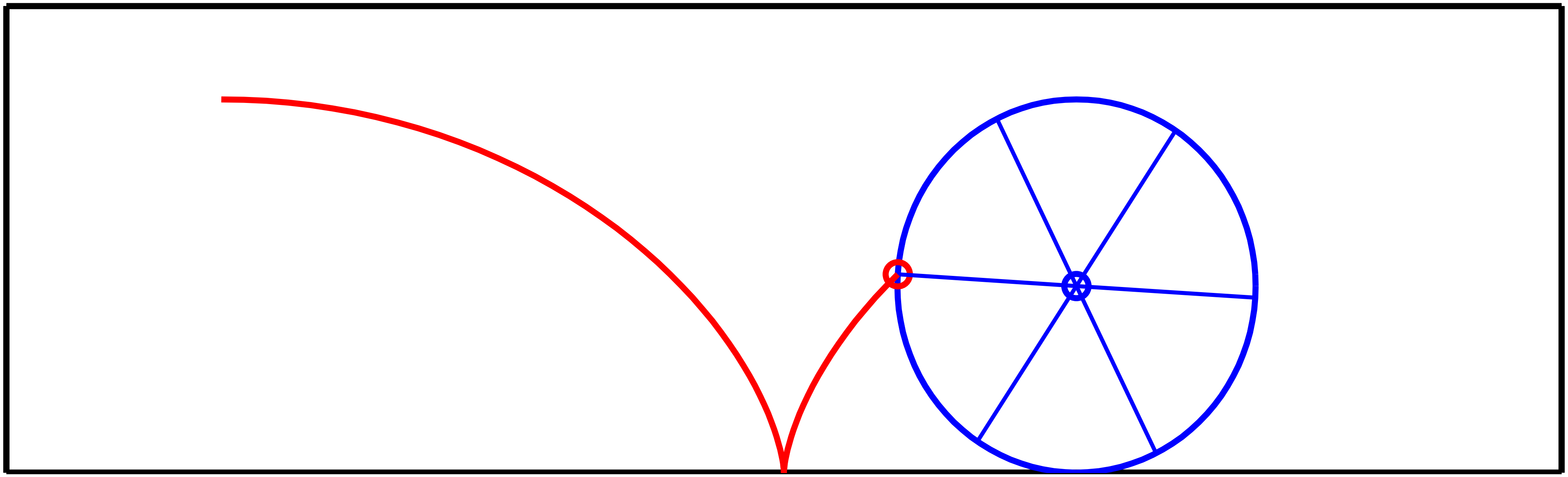
# Cycloid generation



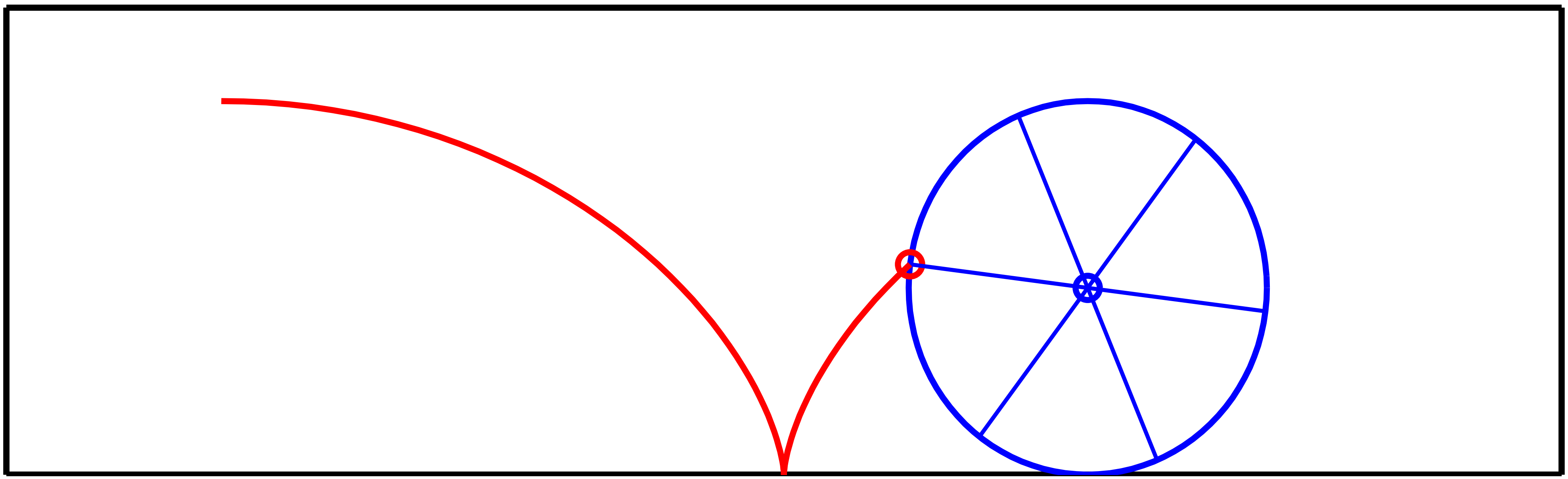
# Cycloid generation



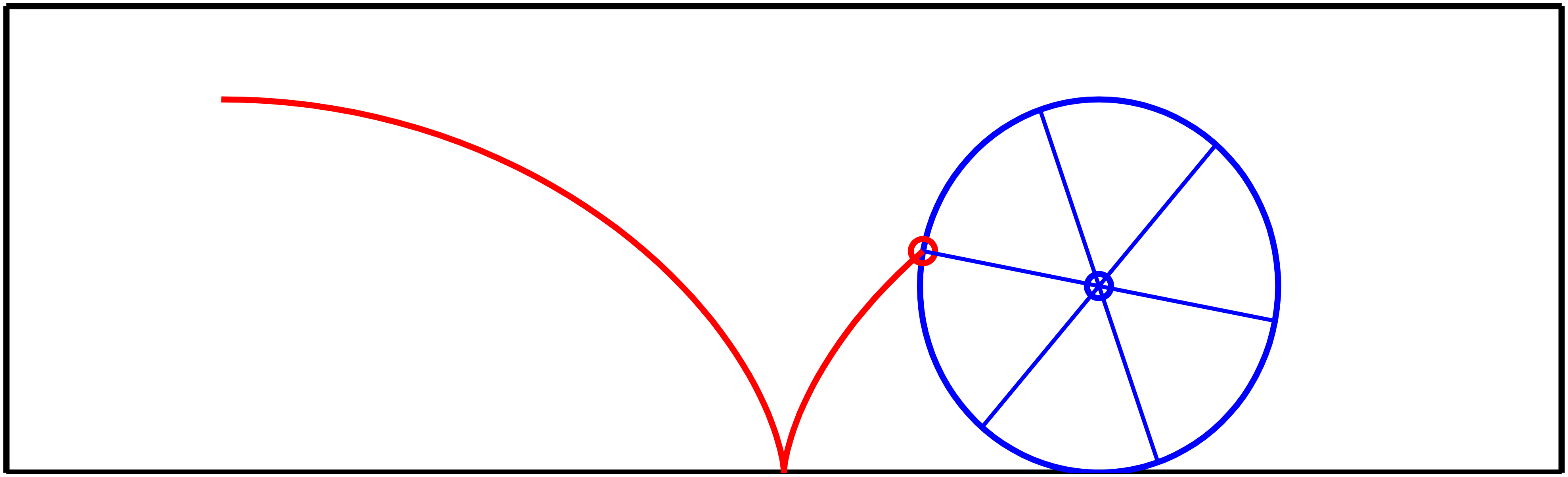
# Cycloid generation



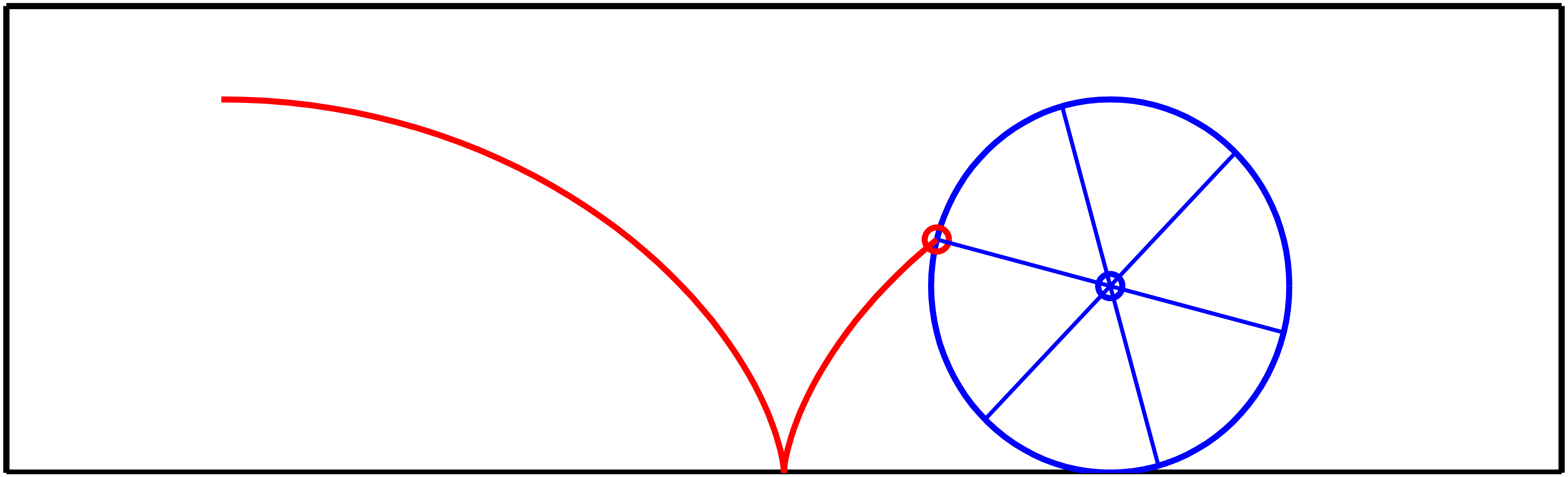
# Cycloid generation



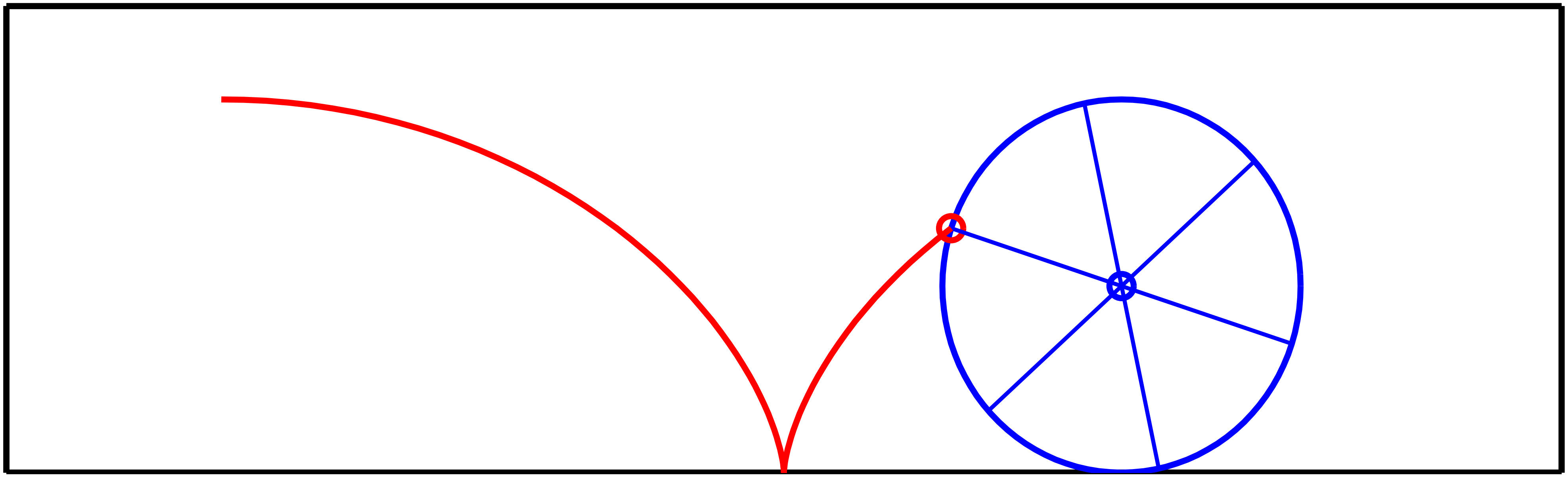
# Cycloid generation



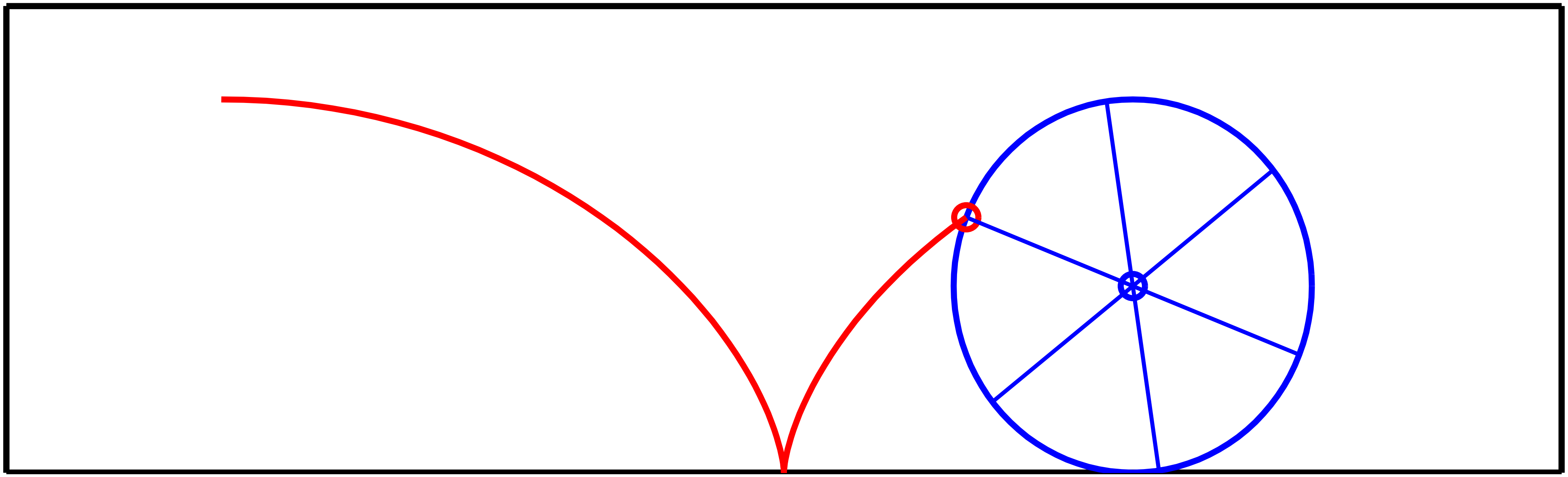
# Cycloid generation



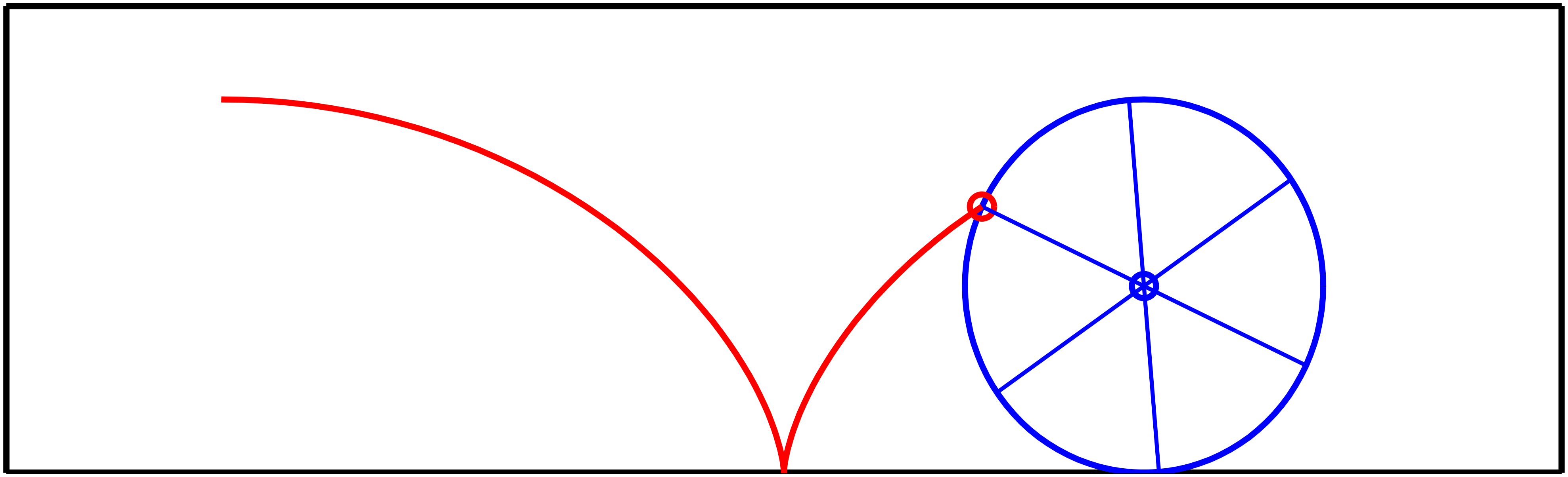
# Cycloid generation



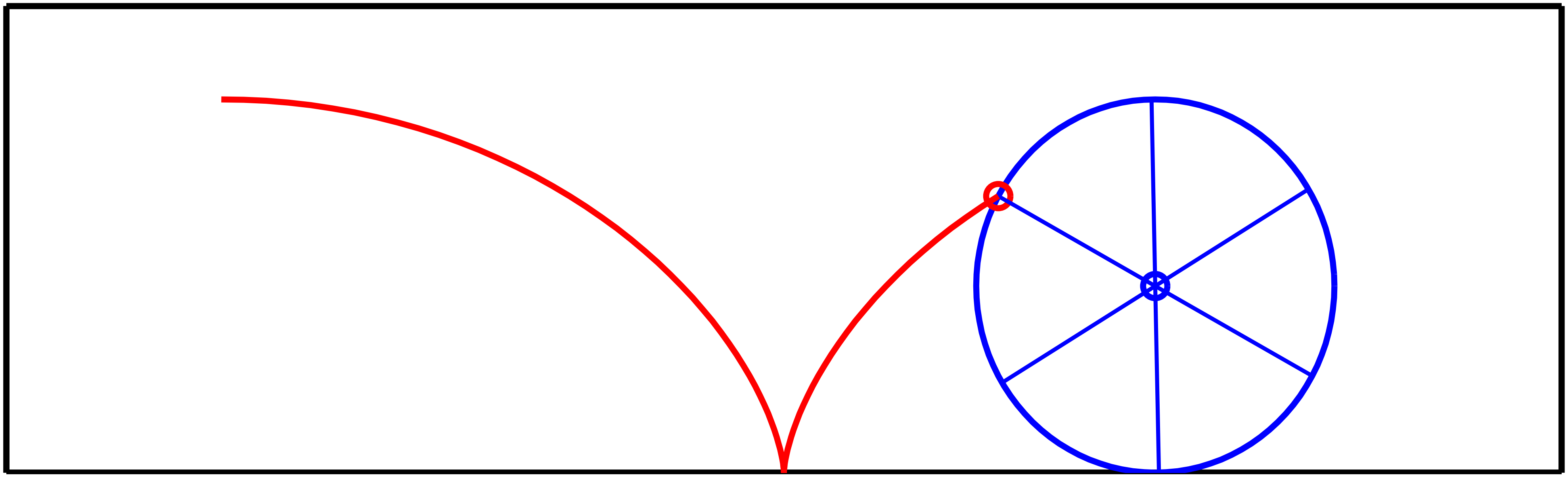
# Cycloid generation



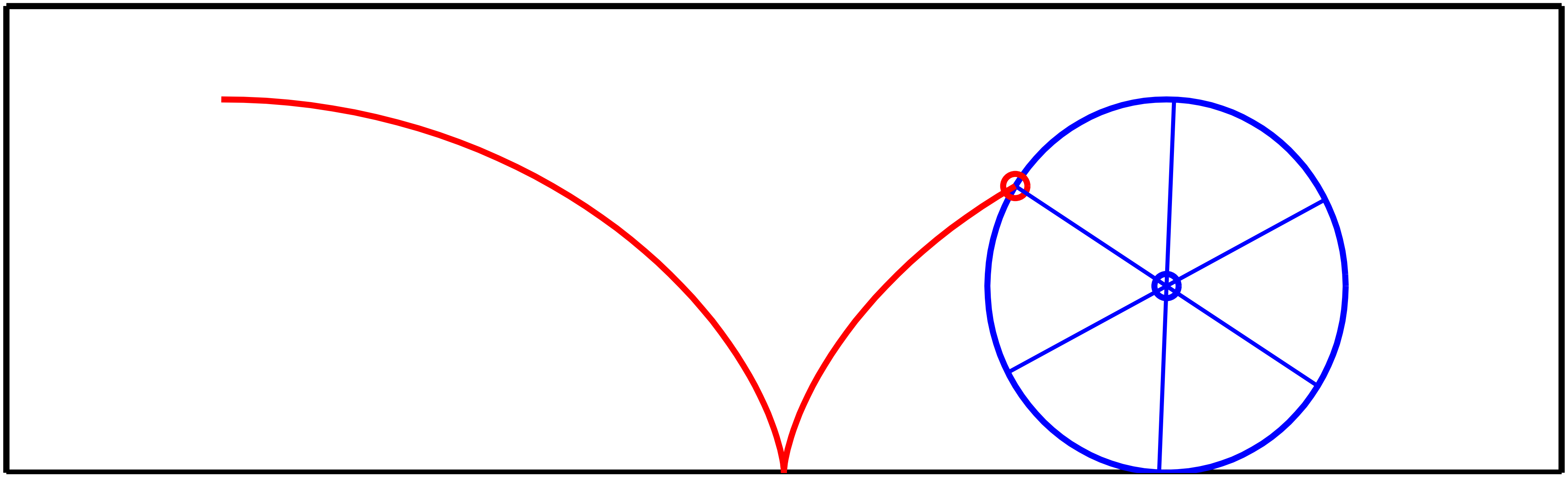
# Cycloid generation



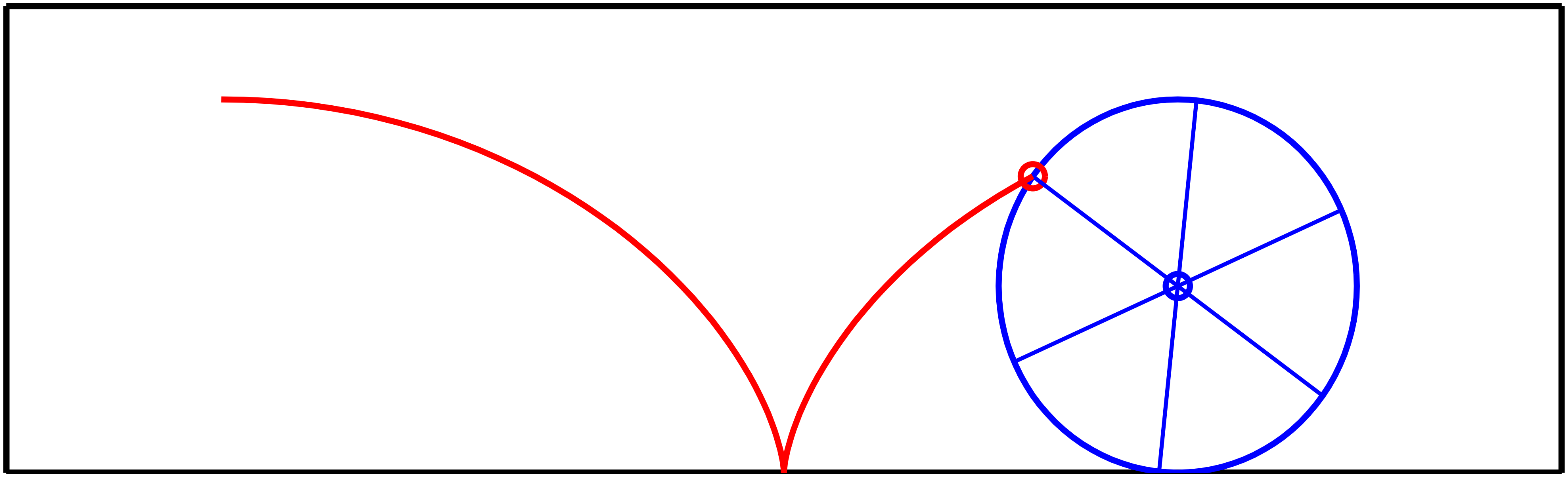
# Cycloid generation



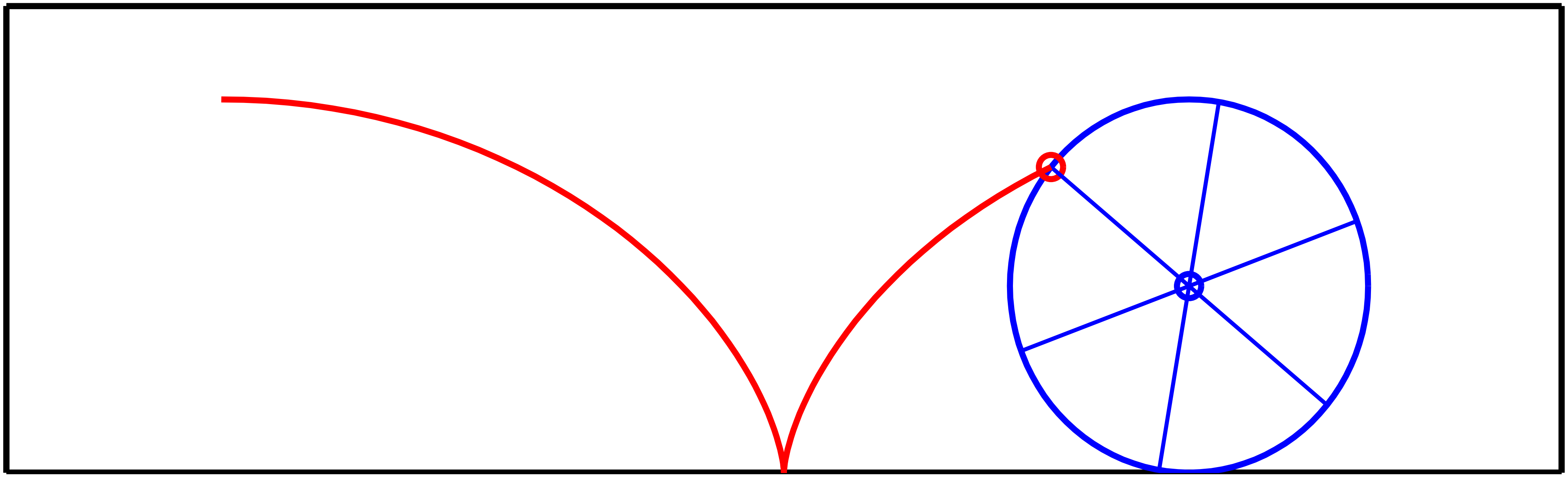
# Cycloid generation



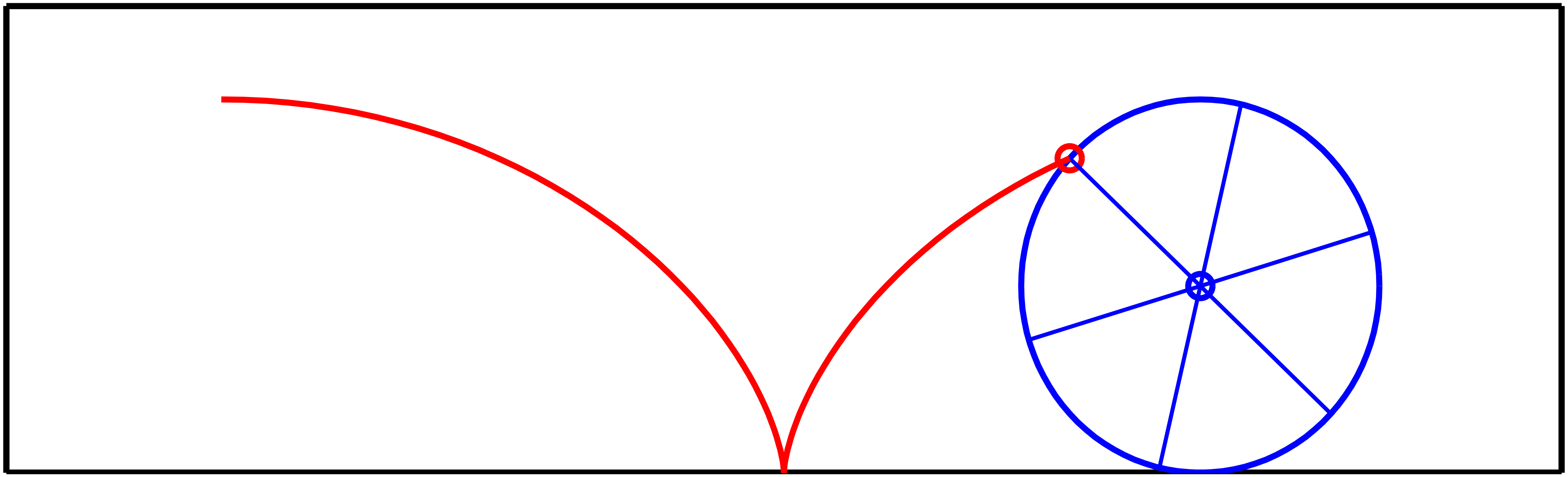
# Cycloid generation



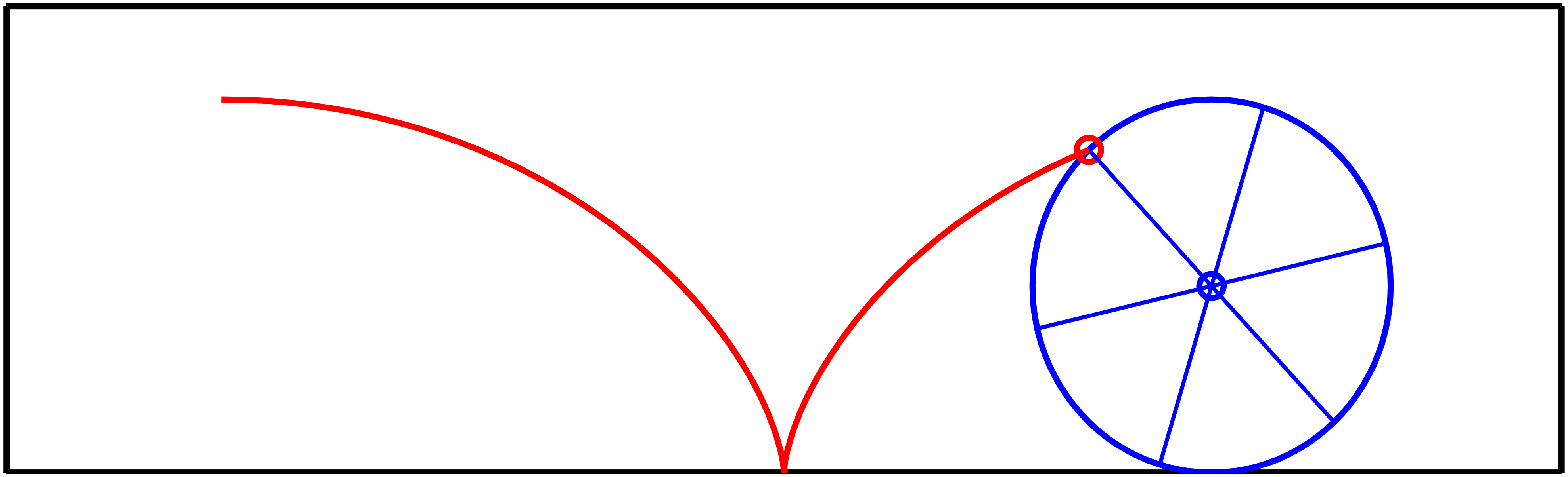
# Cycloid generation



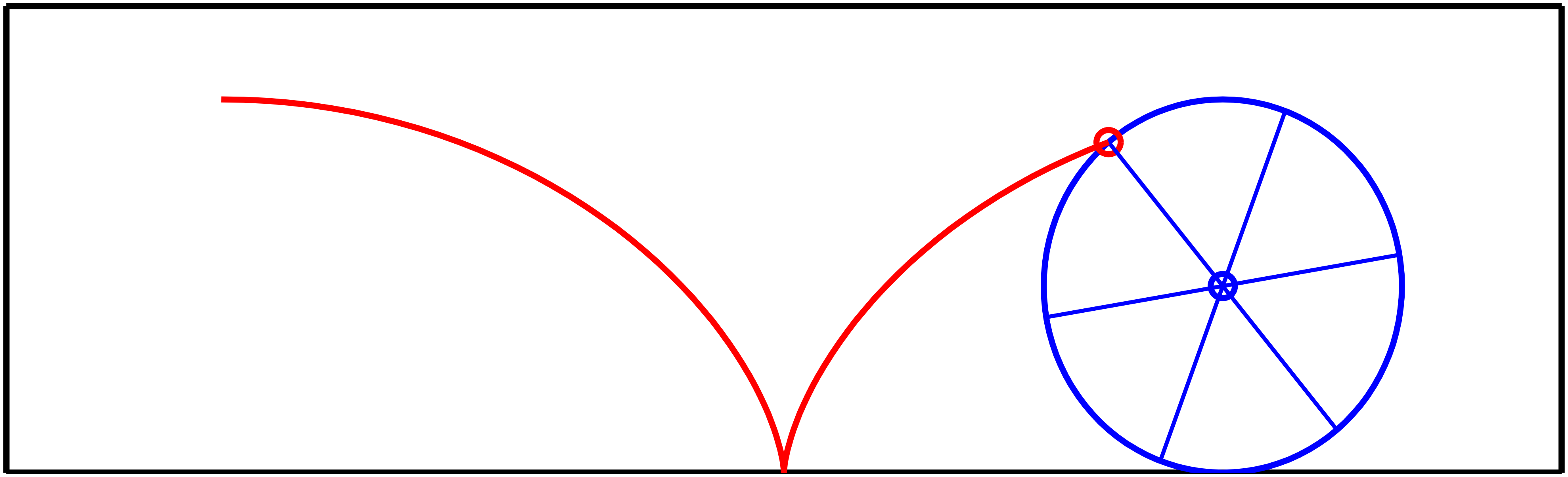
# Cycloid generation



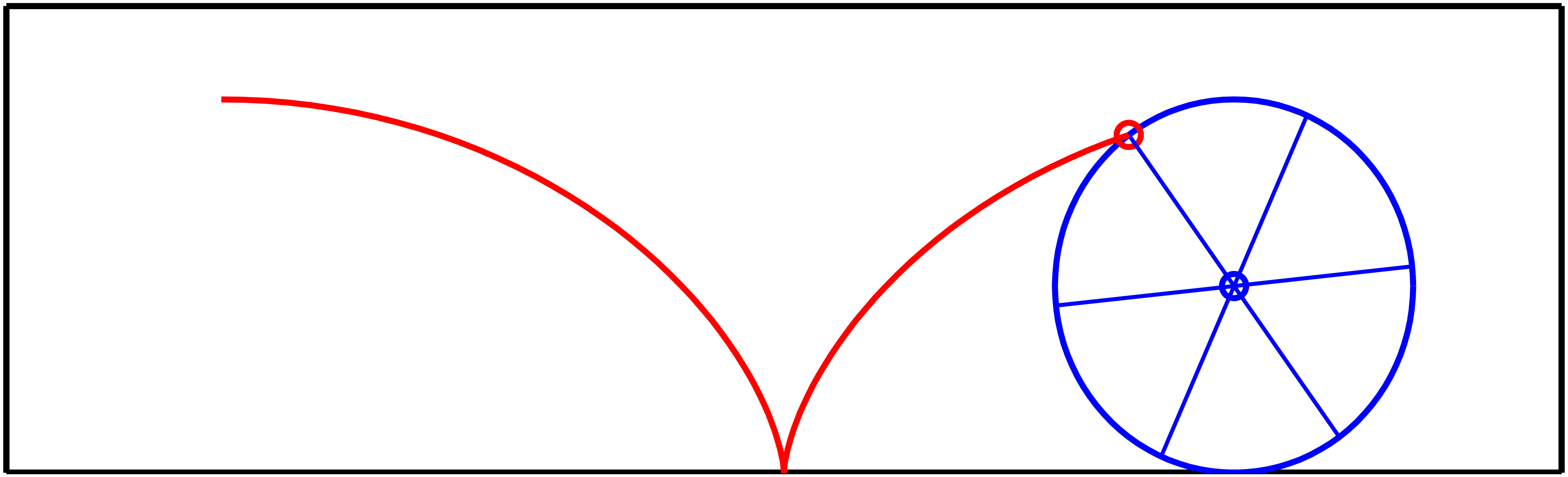
# Cycloid generation



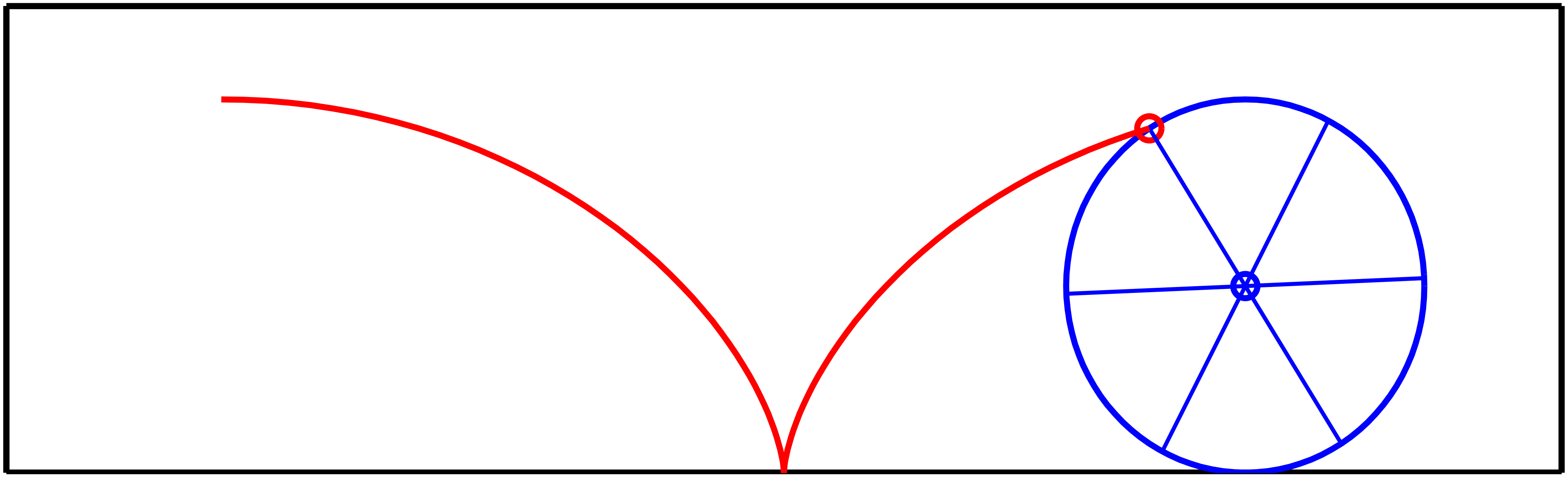
# Cycloid generation



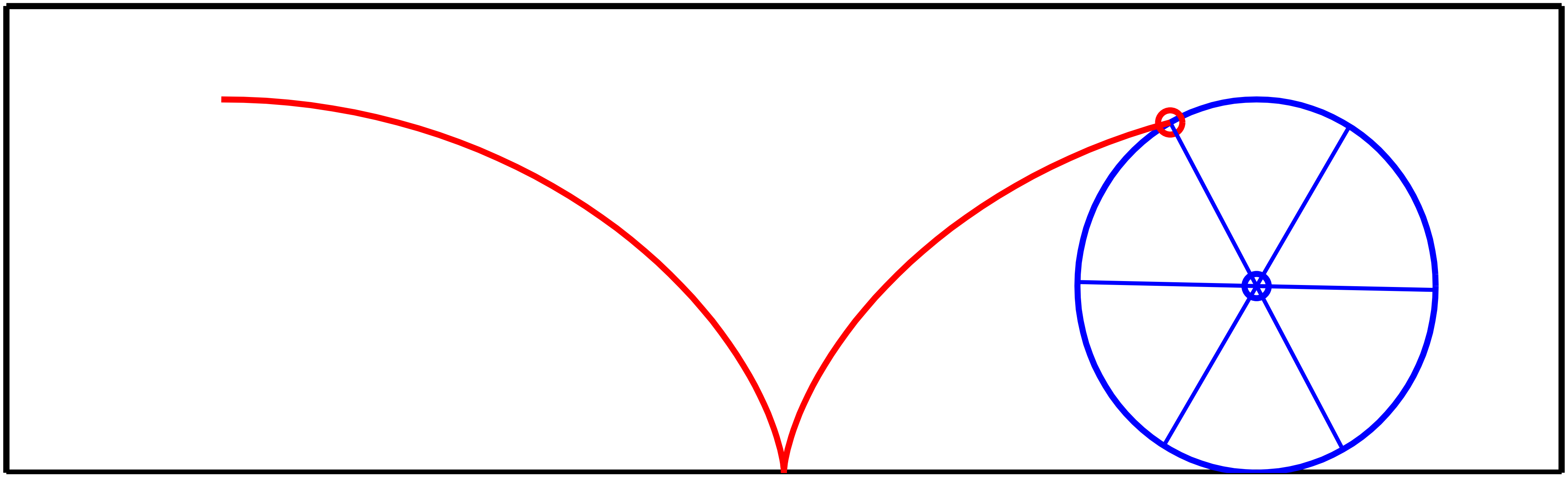
# Cycloid generation



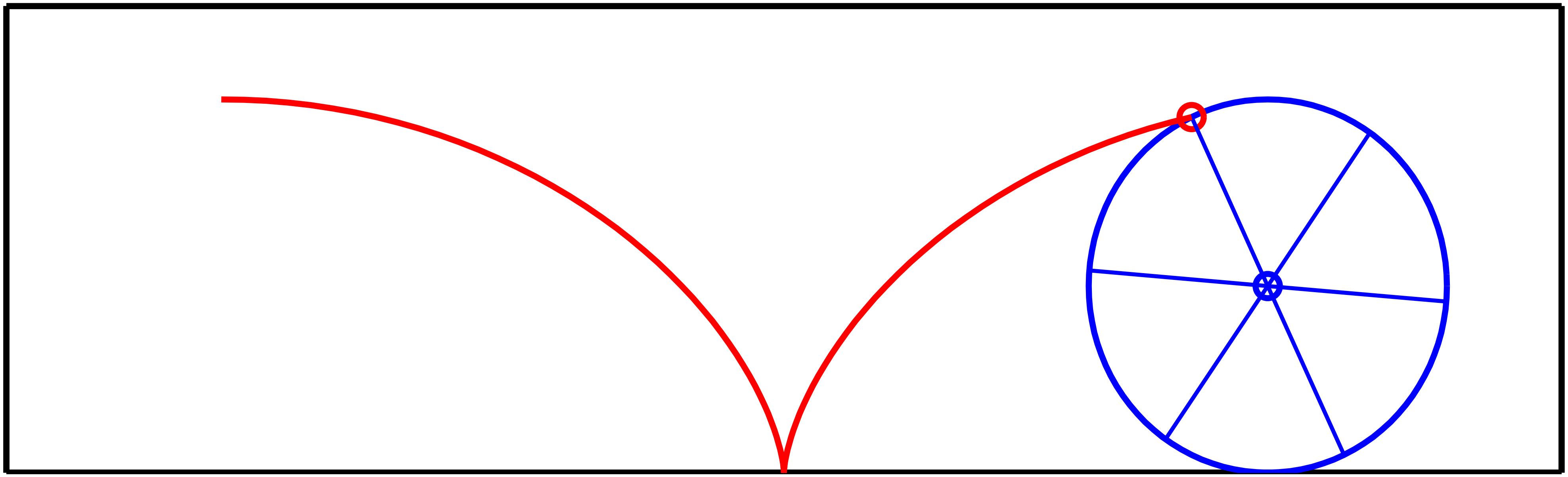
# Cycloid generation



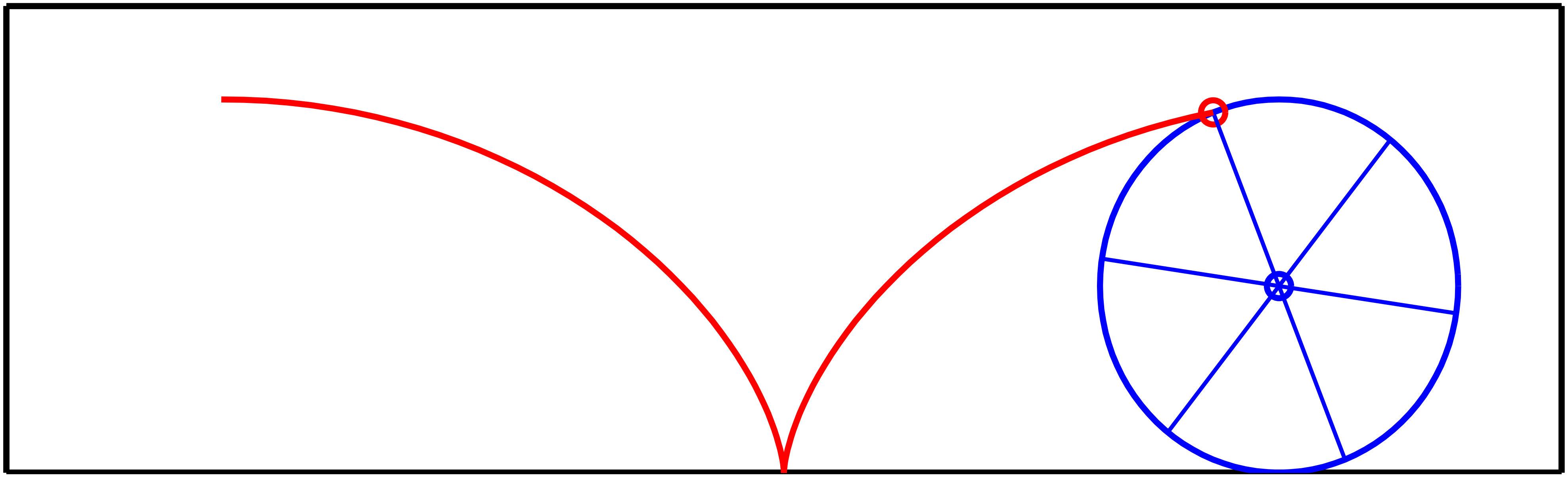
# Cycloid generation



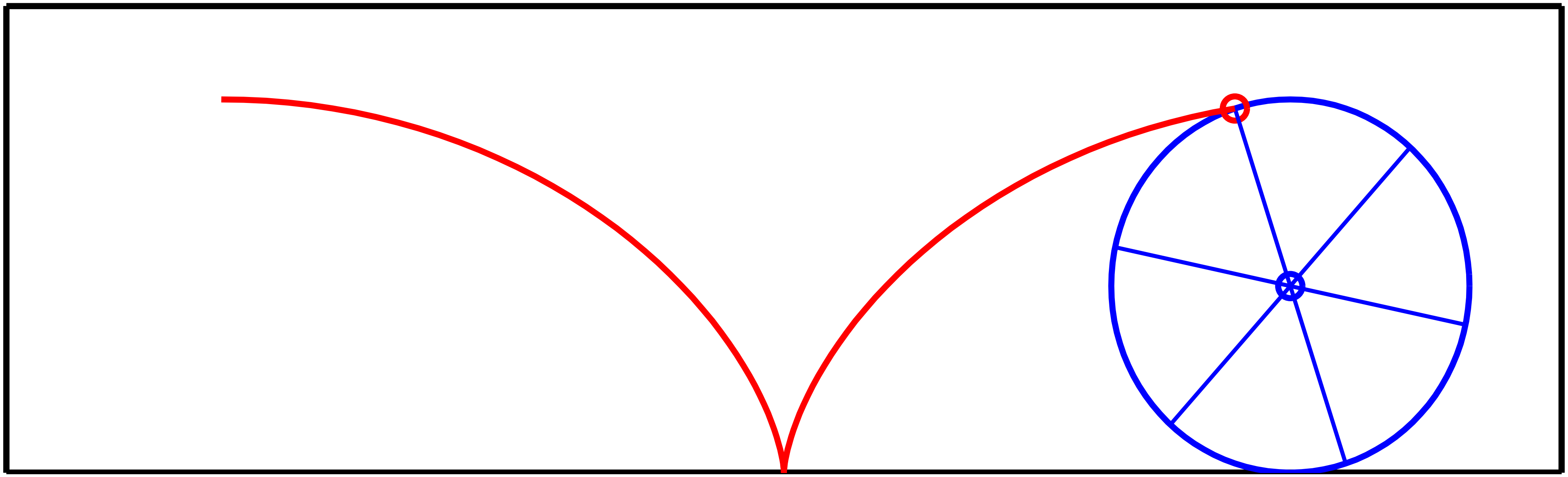
# Cycloid generation



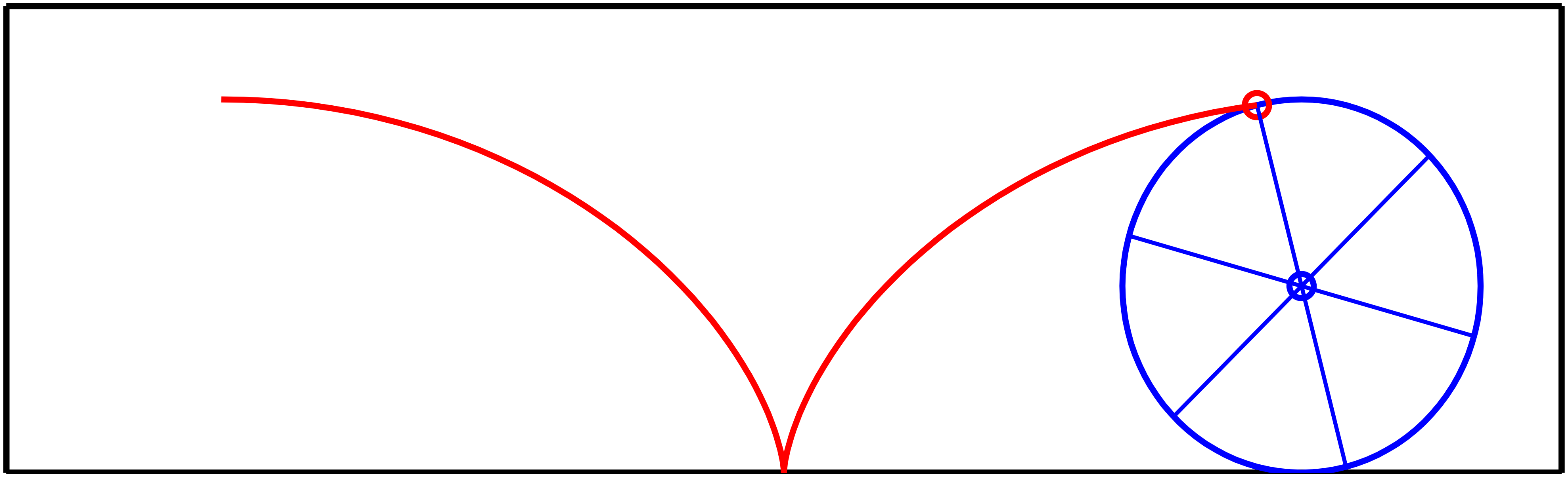
# Cycloid generation



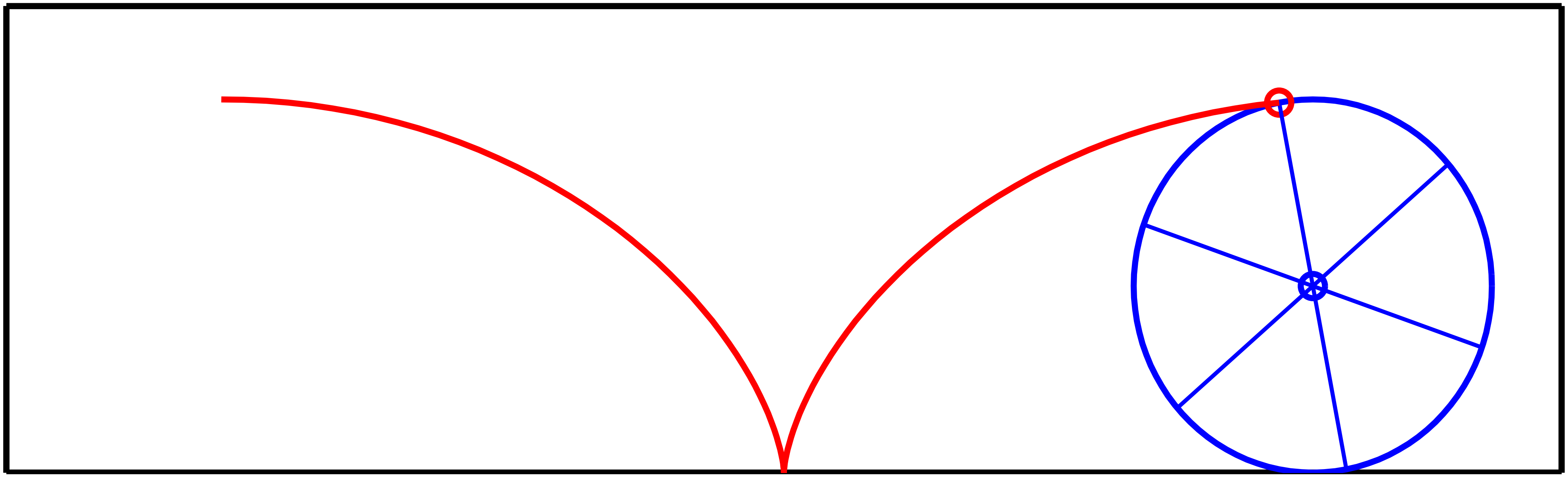
# Cycloid generation



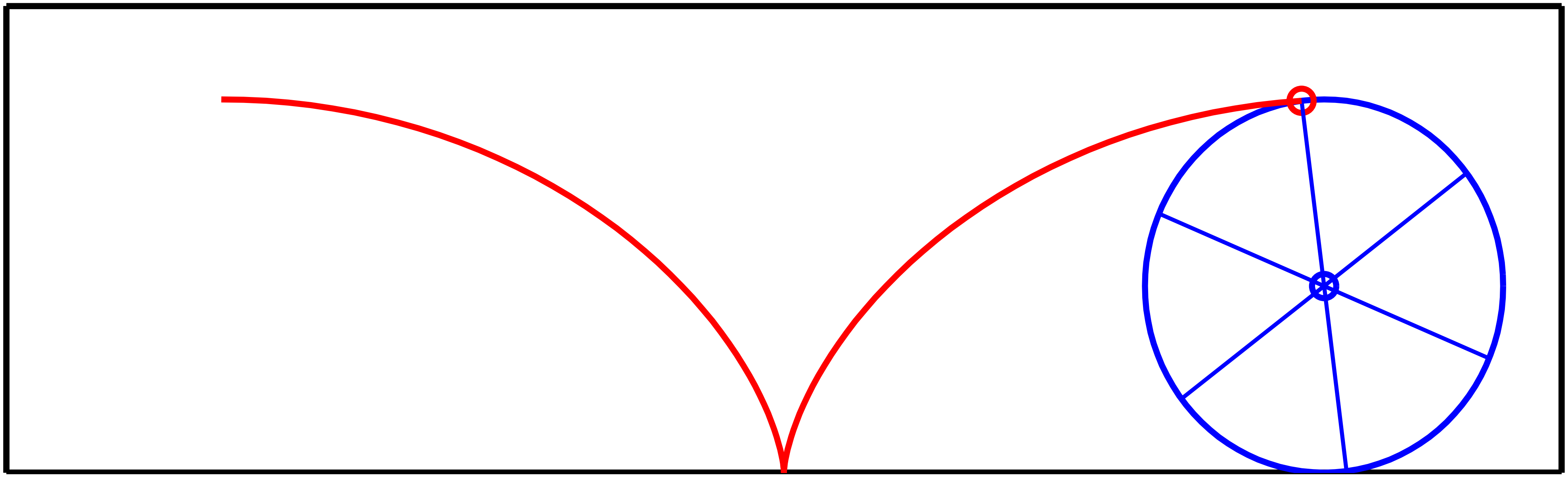
# Cycloid generation



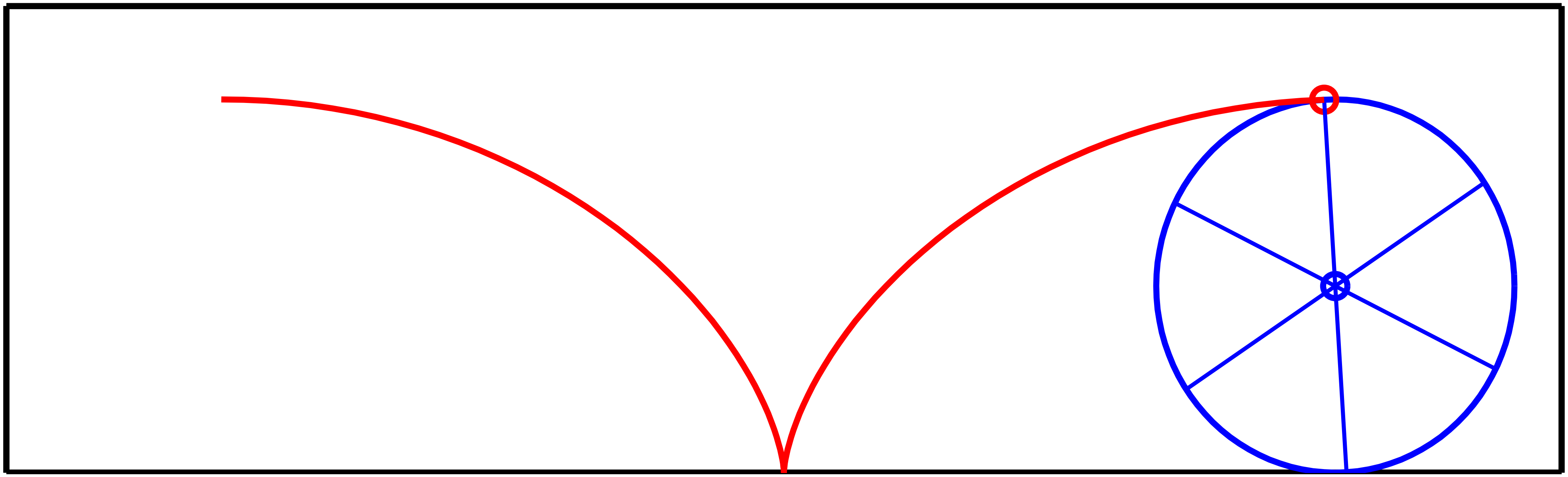
# Cycloid generation



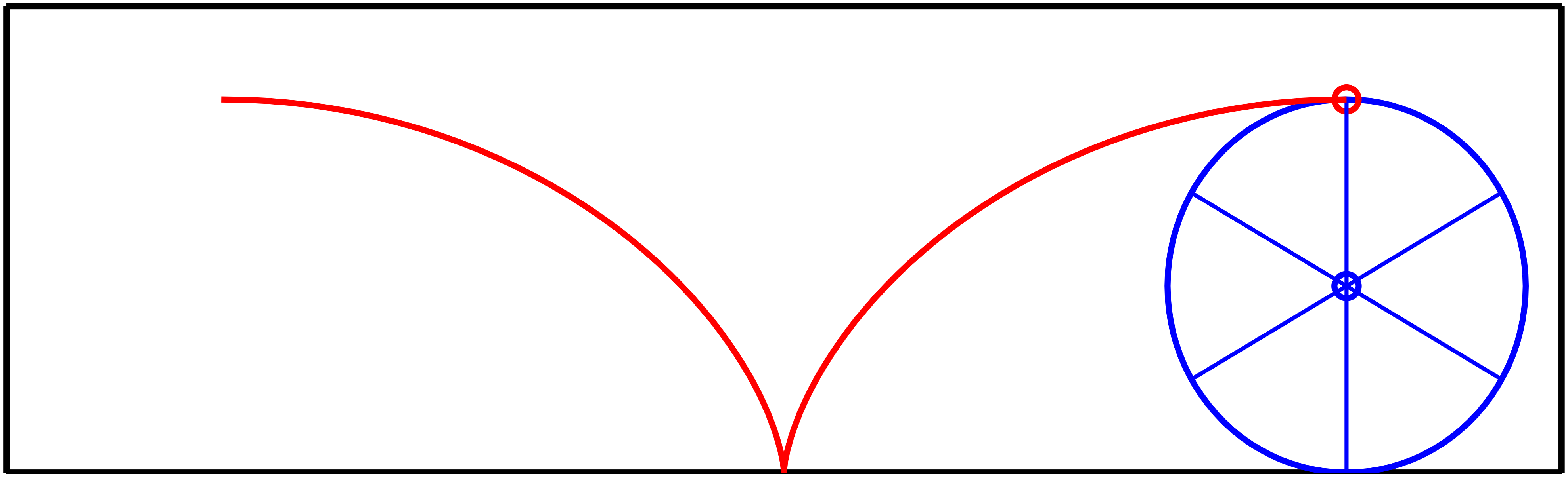
# Cycloid generation



# Cycloid generation



# Cycloid generation



# Geodesics

Geodesic = shortest path

- shortest path between two points on a plane
- shortest path between two points on a sphere

- shortest path on an arbitrary manifold on $\mathbb{R}^n$

# Dido’s problem

Isoperimetric problem: what shaped curve encompasses the largest area given a fixed perimeter.

- 200 B.C. proof by Zendorus (but flawed)
- Steiner proved that “if it exists” its a circle
- Weierstraβ proved using **Calculus of Variations**

# Control problems

Control of systems is critical in modern life

- Mech.Eng: Design of active suspension
- Medicine: Drug delivery to minimize harmful side-effects
- Aerospace: optimize rocket thrust (to minimize fuel consumption)
- Economics: maximize utility of consumption (vs savings)
- Environment: optimal harvesting (say of fish)
- Minimizing cost of A/C

Optimal control is the best (cheapest, fastest, smoothest, ...) we can do.

# Other examples

- Design of vehicle profile that minimizes drag
- Finding shapes of soap bubbles

# Revision

Extrema of functions of one variable.

“Nothing takes place in the world whose meaning is not that of some maximum or minimum.”

L.Euler

# Revision

Calculus of variations is concerned with maximization (minimization)

We are going to maximize (minimize) functionals, not functions

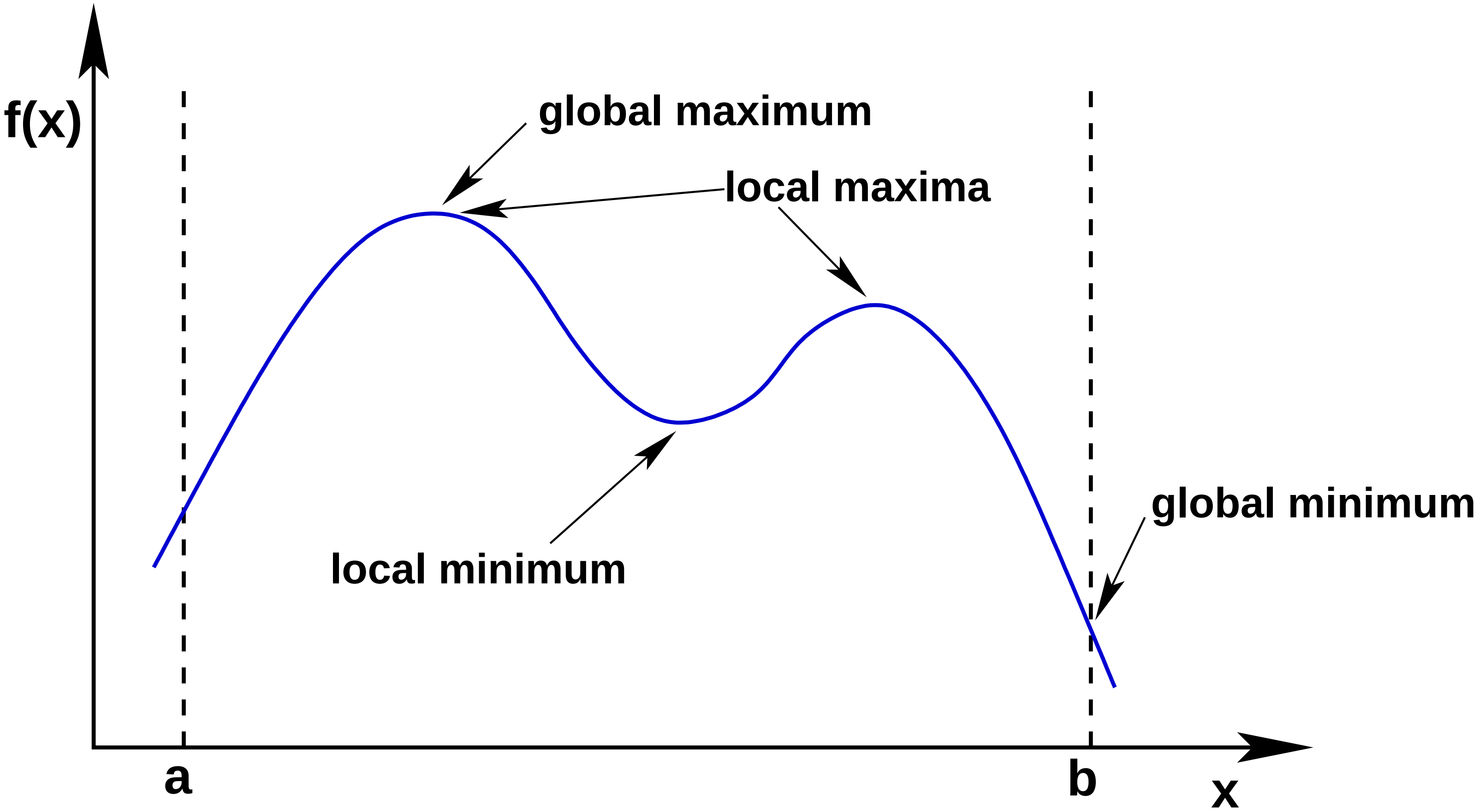
Let us first revise maximization (minimization) of function

# Maxima and minima

Functions of one variable:

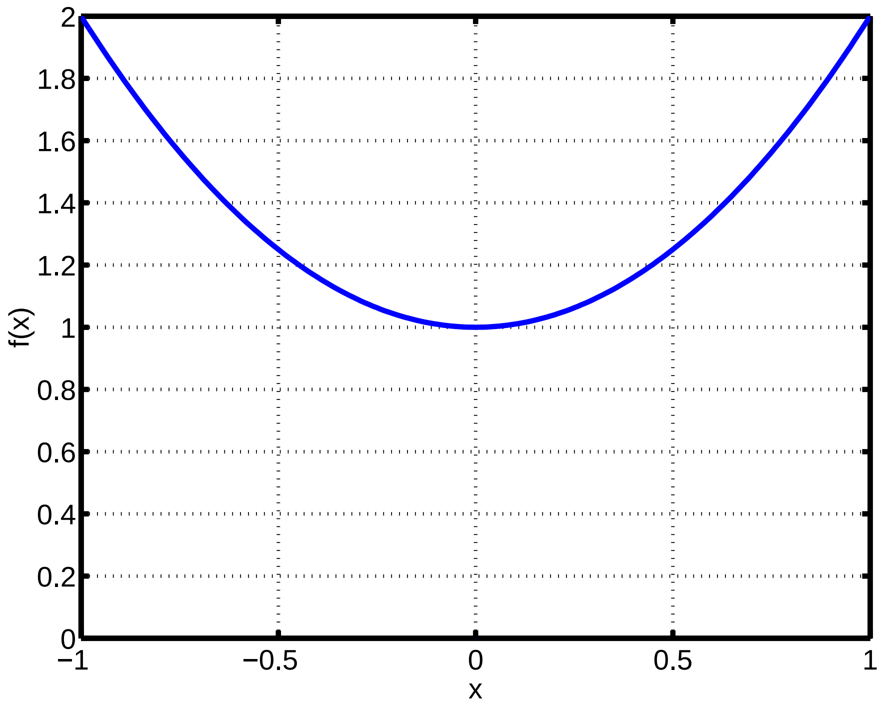
- Let $x \in [a,b]$ and $f(x) : [a,b] \to \mathbb{R}$
- If there is a point $x_{\min}$ such that $f(x_{\min}) \leq f(x)$ for all $x \in [a,b]$, then $x_{\min}$ is called a **global minima** of $f(x)$ in $[a,b]$.
- The set of points $x$ such that $f(x) = f(x_{\min})$ is called the **minimal set**.
- If there is an interior point $x \in (a,b)$ such that there exists a $\delta > 0$ with $f(x) \leq f(\hat{x})$ for all $\hat{x} \in (x-\delta, x+\delta)$, then $x$ is called a **local minimum** of $f(\cdot)$.
- similar definitions apply for maxima, note maxima of $f(x)$ are the minima of $-f(x)$

# Maxima and minima: example 1



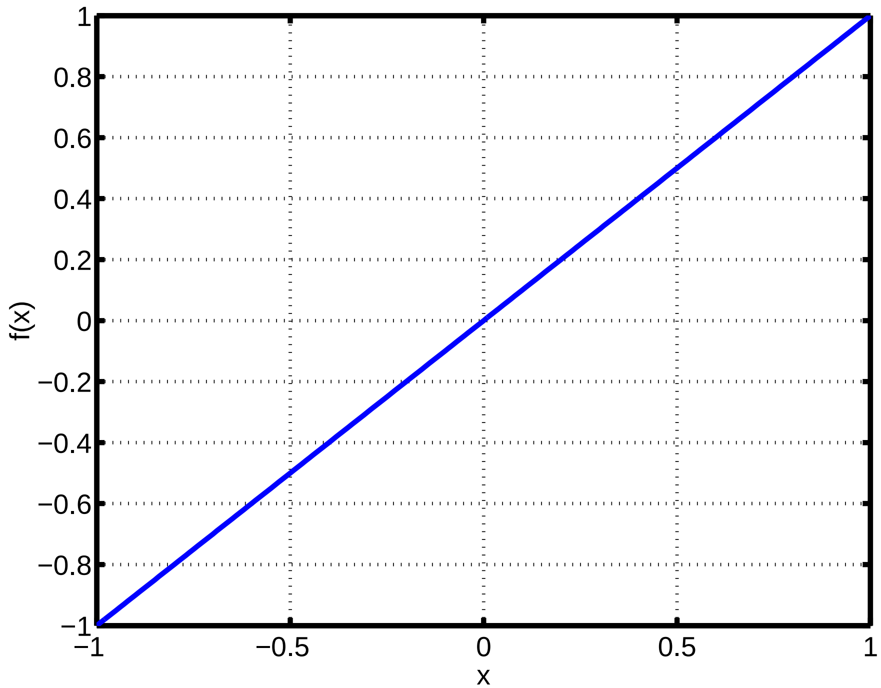
# Maxima and minima: example 2

- $f(x) = 1 + x^2$ on $[-1, 1]$
- global minimum at $x = 0$
- local minimum at $x = 0$
- maximal set $\{-1, 1\}$



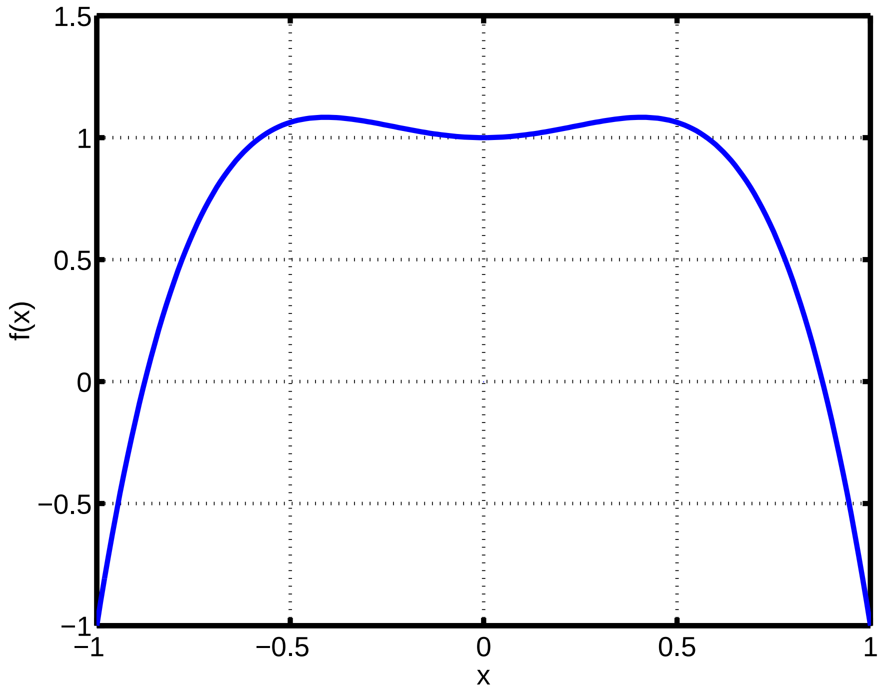
# Maxima and minima: example 3

- $f(x) = x$ on $[-1, 1]$
- global minimum at $x = -1$
- not a local min. because not an interior point
- global maximum at $x = 1$



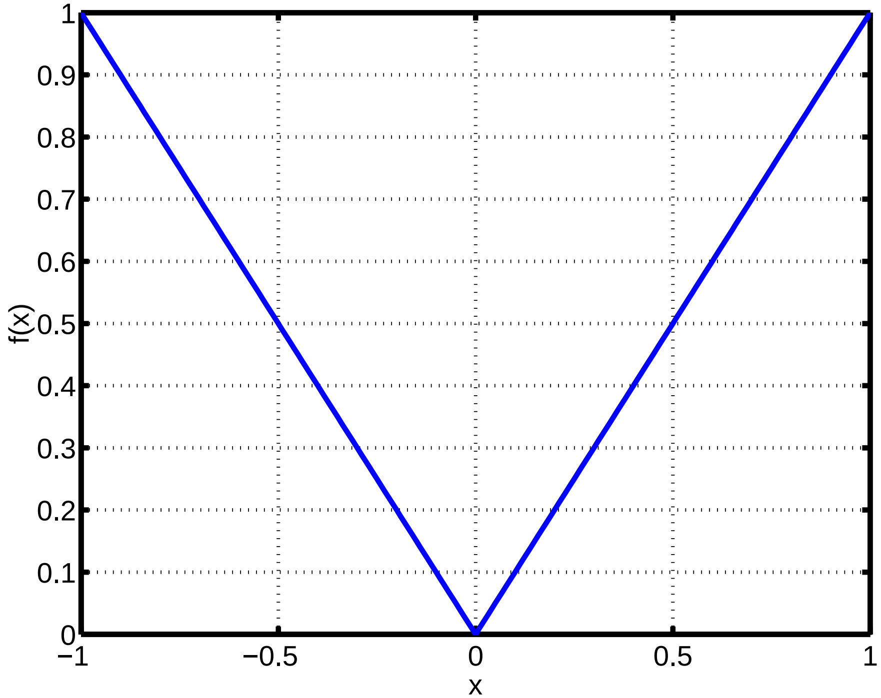
# Maxima and minima: example 4

- $f(x) = 1 + x^2 - x^4$ on $[-1, 1]$
- global minimum at $x = -1$ and $x = 1$
- local minimum at $x = 0$.



# Maxima and minima: example 5

- $f(x) = |x|$ on $[-1,1]$
- global minimum at $x = 0$
- local minimum at $x = 0$



# How to find maxima and minima

**Theorem 1**: Let $f(x) : [a,b] \to I\!R$ be differentiable in $(a,b)$. If $f(\cdot)$ has a local extrema at $x$ then

$$\frac{df}{dx} = f'(x) = 0$$

**Proof**: The derivative is given by

$$f'(x) = \lim_{\hat{x} \to x} \frac{f(\hat{x}) - f(x)}{\hat{x} - x}$$

Suppose $x$ is a local minima, then $\exists \delta > 0$ such that $\hat{x} \in (x - \delta, x + \delta) \Rightarrow f(\hat{x}) > f(x)$, hence the numerator $> 0$. The denominator changes sign at $\hat{x} = x$. Differentiability implies the left and right hand limits exist and are equal, and hence $f'(x) = 0$.

# Sufficient conditions

**Theorem 2**: Let $f(x) : [a,b] \to I\!R$ be twice differentiable in $(a,b)$.
Sufficient conditions for a local minimum at $x$ are

$$f'(x) = 0 \quad \text{and} \quad f''(x) > 0$$

**Proof**: see following.

# Some useful theorems

- **Mean Value Theorem:** Let $x_0 < x_1$, and $f(\cdot)$ be a continuous function in $[x_0, x_1]$, and differentiable in $(x_0, x_1)$, then $\exists \xi \in (x_0, x_1)$ such that
$$f(x_1) = f(x_0) + (x_1 - x_0) f'(\xi)$$
- **Taylor's theorem:** Let $f(\cdot)$ be a function whose first $n$ derivatives exist and are continuous in the interval $[x_0, x_1]$, and $f^{(n+1)}(x)$ exists for all $x \in (x_0, x_1)$, then $\exists \xi \in (x_0, x_1)$
$$f(x_1) = f(x_0) + (x_1 - x_0) f'(x_0) + \frac{(x_1 - x_0)^2}{2} f''(x_0) + \cdots$$
$$+ \frac{(x_1 - x_0)^n}{n!} f^{(n)}(x_0) + \frac{(x_1 - x_0)^{n+1}}{(n+1)!} f^{(n+1)}(\xi)$$

# Sufficient conditions

**Theorem 3**: Let $f(x) : [a,b] \to I\!R$ have derivatives of all orders, then a necessary and sufficient condition for a local minima is that for some $n$

$$f'(x) = f''(x) = \cdots = f^{(2n-1)}(x) = 0 \quad \text{and} \quad f^{(2n)}(x) > 0$$

**Proof**: Taylor's theorem, where $\hat{x} - x = \varepsilon$

$$f(\hat{x}) = f(x) + \varepsilon f'(x) + \cdots + \frac{\varepsilon^{2n-1}}{(2n-1)!} f^{(2n-1)}(x) + \frac{\varepsilon^{2n}}{(2n)!} f^{(2n)}(x) + O(\varepsilon^{2n+1})$$

Then

$$\begin{aligned} f(\hat{x}) - f(x) &= \frac{\varepsilon^{2n}}{(2n)!} f^{(2n)}(x) + O(\varepsilon^{2n+1}) \\ &> 0 \quad \text{for small enough } \varepsilon \end{aligned}$$

# Classifying extrema

Assume that $f'(x) = 0$

- local maxima $f''(x) < 0$
- local minima $f''(x) > 0$
- turning point $f''(x) = 0$, and $f^{(3)}(x) \neq 0$
- $+$ a lot of higher order conditions

Call all points with $f'(x) = 0$ the set of **stationary** points

# Conclusion

We have looked at 1D local maxima and minima
We need to generalize this

- next lecture, to functions of $N$ variables
- then, to functions of functions ($\infty$ variables)

# Extra bits

Some notation and definitions

# Notation

- $[a,b]$ is the closed interval, i.e. the set $\{x \in \mathbb{R} | a \leq x \leq b\}$
- $(a,b)$ is the open interval, i.e. the set $\{x \in \mathbb{R} | a < x < b\}$
- $(a,b]$ is the set $\{x \in \mathbb{R} | a < x \leq b\}$
- $f(x) : [a,b] \to \mathbb{R}$ denotes a function that maps the set $[a,b]$ to a real number.
- $\frac{d^n f}{dx^n} = f^{(n)}(x)$ denotes the $n$th derivative of $f(x)$.

# Synonyms

- the global minimum is sometimes called a strong minimum
- a local minimum is sometimes called a weak minimum
- the local extrema are the collection of local minima and maxima
  We sometimes abuse notation to include stationary points in the set of extrema.

# Useful Definitions: continuity

- a function $f(x)$ is **continuous** at $x_0$ iff the left and right limits at $x_0$ exist and are equal, i.e.,

$$\lim_{x \to x_0^-} f(x) = \lim_{x \to x_0^+} f(x)$$

  otherwise it is said to have a **discontinuity**.

- We say a function is continuous on an interval if it is continuous at every point inside the interval and the limits exist at the boundaries.
- A function is **piecewise continuous** on an interval if it has at most finite number of discontinuities.

# Useful Definitions: differentiability

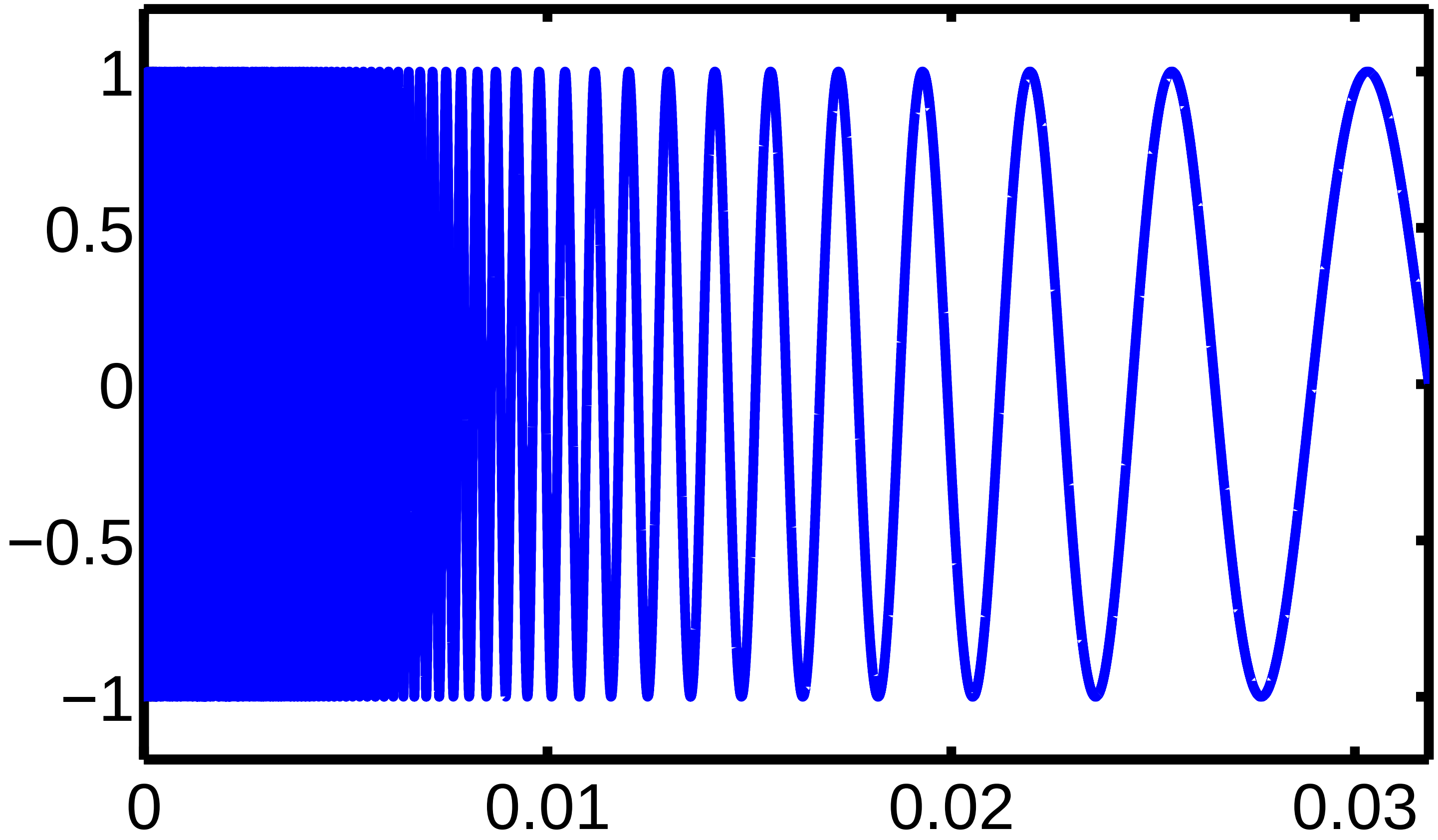
- A function is **differentiable** at $x_0$ if its derivative exists, and is continuous at $x_0$, i.e., the following limit exists and is the same from both directions

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

- We say a function is differentiable on an interval if it is differentiable at every point inside the interval and the limits exist at the boundaries.
- A function is **piecewise differentiable** if the derivative has at most a finite number of discontinuities.
- A function is **twice differentiable** if its second derivative exists and is continuous.

# Useful Definitions

- We also eliminate from consideration functions whose derivative changes sign an infinite number of times in a finite interval.
  - e.g. $\sin(1/x)$



# Notation

We define the **del** or **grad** operator by

$$\nabla = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$$

So, given a scalar function $\phi(x, y, z)$, then $\nabla\phi$ is a vector function

$$\nabla\phi = \left( \frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z} \right)$$

Given a vector function $\mathbf{f}(x, y, z) = (f_1, f_2, f_3)$ then we define the **div** operator div $\mathbf{f} = \nabla \cdot \mathbf{f}$, e.g.

$$\nabla \cdot \mathbf{f} = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \cdot (f_1, f_2, f_3) = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}$$

# Notation

We can also use del to define the **curl** operator using a cross-product curl $=$ del$\times$, e.g.

$$\text{curl}\,\mathbf{f} = \nabla \times \mathbf{f}$$

The **Laplacian operator**, or del-squared operator of a scalar function (of $(x, y, z)$) is defined by

$$\nabla^2\phi = \nabla \cdot (\nabla\phi) = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}$$