Notes about Hyperbolic Space

Zhijian Lai
https://galvinlai.github.io/

October 9, 2023

This note is modified from [Bou23, § 7.6] and [CYRL19].

Contents

Chapter 1

Basic

1.1 Calculus

We write $F : A \to B$ to designate a map F whose domain is all of A. If C is a subset of A, we write $F|_C : C \to B$ to designate the restriction of F to the domain C, so that $F|_C(x) = F(x)$ for all $x \in C$.

Let U, V be open sets in two linear spaces $\mathcal{E}, \mathcal{F}$. A map $F : U \to V$ is smooth if it is infinitely differentiable (class C^∞) on its domain. We also say that F is smooth at a point $x \in U$ if there exists a neighborhood U' of x such that $F|_{U'}$ is smooth. Accordingly, F is smooth if it is smooth at all points in its domain.

If $F : U \to V$ is smooth at x, the differential of F at x is the linear map $\mathrm{D}F(x) : \mathcal{E} \to \mathcal{F}$ defined by

$$\mathrm{D}F(x)[u] = \frac{\mathrm{d}}{\mathrm{d}t} F(x + tu)\bigg|_{t=0} = \lim_{t\to 0} \frac{F(x + tu) - F(x)}{t}. \tag{1.1}$$

For a curve $c : \mathbb{R} \to \mathcal{E}$, we write c' to denote its velocity: $c'(t) = \frac{\mathrm{d}}{\mathrm{d}t} c(t)$.

For a smooth function $f : \mathcal{E} \to \mathbb{R}$ defined on a Euclidean space $\mathcal{E}$, the (Euclidean) gradient of f is the map $\operatorname{grad} f : \mathcal{E} \to \mathcal{E}$ defined by the following property:

$$\forall x, v \in \mathcal{E}, \quad \langle \operatorname{grad} f(x), v\rangle = \mathrm{D}f(x)[v]. \tag{1.2}$$

The (Euclidean) Hessian of f at x is the linear map $\operatorname{Hess} f(x) : \mathcal{E} \to \mathcal{E}$ defined by

$$\operatorname{Hess} f(x)[v] = \mathrm{D}(\operatorname{grad} f)(x)[v] = \lim_{t\to 0} \frac{\operatorname{grad} f(x + tv) - \operatorname{grad} f(x)}{t}. \tag{1.3}$$

1.2 Hyperbolic functions and their inverses

In mathematics, hyperbolic functions are analogues of the ordinary trigonometric functions, but defined using the hyperbola rather than the circle. Just as the points $(\cos t, \sin t)$ form a circle with a unit radius, the points $(\cosh t, \sinh t)$ form the right half of the unit hyperbola. Also, similarly to how the derivatives of $\sin(t)$ and $\cos(t)$ are $\cos(t)$ and $-\sin(t)$ respectively, the derivatives of $\sinh(t)$ and $\cosh(t)$ are $\cosh(t)$ and $+\sinh(t)$ respectively.

1.2.1 Hyperbolic functions: Definitions in terms of exponentiation

There are various equivalent ways to define the hyperbolic functions. Here, we use the definitions in terms of the exponential function.

- Hyperbolic sine: the odd part of the exponential function, that is,

$$\sinh x = \frac{e^x - e^{-x}}{2} = \frac{e^{2x} - 1}{2e^x} = \frac{1 - e^{-2x}}{2e^{-x}}. \tag{1.4}$$

- Hyperbolic cosine: the even part of the exponential function, that is,

$$\cosh x = \frac{e^x + e^{-x}}{2} = \frac{e^{2x} + 1}{2e^x} = \frac{1 + e^{-2x}}{2e^{-x}}. \tag{1.5}$$

- Hyperbolic tangent:

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x} - 1}{e^{2x} + 1}. \tag{1.6}$$

- Hyperbolic cotangent: for $x \neq 0$,

$$\coth x = \frac{\cosh x}{\sinh x} = \frac{e^x + e^{-x}}{e^x - e^{-x}} = \frac{e^{2x} + 1}{e^{2x} - 1}. \tag{1.7}$$

- Hyperbolic secant:

$$\operatorname{sech} x = \frac{1}{\cosh x} = \frac{2}{e^x + e^{-x}} = \frac{2e^x}{e^{2x} + 1}. \tag{1.8}$$

- Hyperbolic cosecant: for $x \neq 0$,

$$\operatorname{csch} x = \frac{1}{\sinh x} = \frac{2}{e^x - e^{-x}} = \frac{2e^x}{e^{2x} - 1}. \tag{1.9}$$

1.2.2 Inverse hyperbolic functions: Definitions in terms of logarithms

Since the hyperbolic functions are quadratic rational functions of the exponential function $\exp x$, they may be solved using the quadratic formula and then written in terms of the natural logarithm.

$$\operatorname{arsinh}(x) = \ln\left(x + \sqrt{x^2 + 1}\right) \qquad x \in \mathbb{R} \tag{1.10}$$

$$\operatorname{arcosh}(x) = \ln\left(x + \sqrt{x^2 - 1}\right) \qquad x \geqslant 1 \tag{1.11}$$

$$\operatorname{artanh}(x) = \frac{1}{2}\ln\left(\frac{1 + x}{1 - x}\right) \qquad |x| < 1 \tag{1.12}$$

$$\operatorname{arcoth}(x) = \frac{1}{2}\ln\left(\frac{x + 1}{x - 1}\right) \qquad |x| > 1 \tag{1.13}$$

$$\operatorname{arsech}(x) = \ln\left(\frac{1}{x} + \sqrt{\frac{1}{x^2} - 1}\right) = \ln\left(\frac{1 + \sqrt{1 - x^2}}{x}\right) \qquad 0 < x \leqslant 1 \tag{1.14}$$

$$\operatorname{arcsch}(x) = \ln\left(\frac{1}{x} + \sqrt{\frac{1}{x^2} + 1}\right) \qquad x \neq 0 \tag{1.15}$$

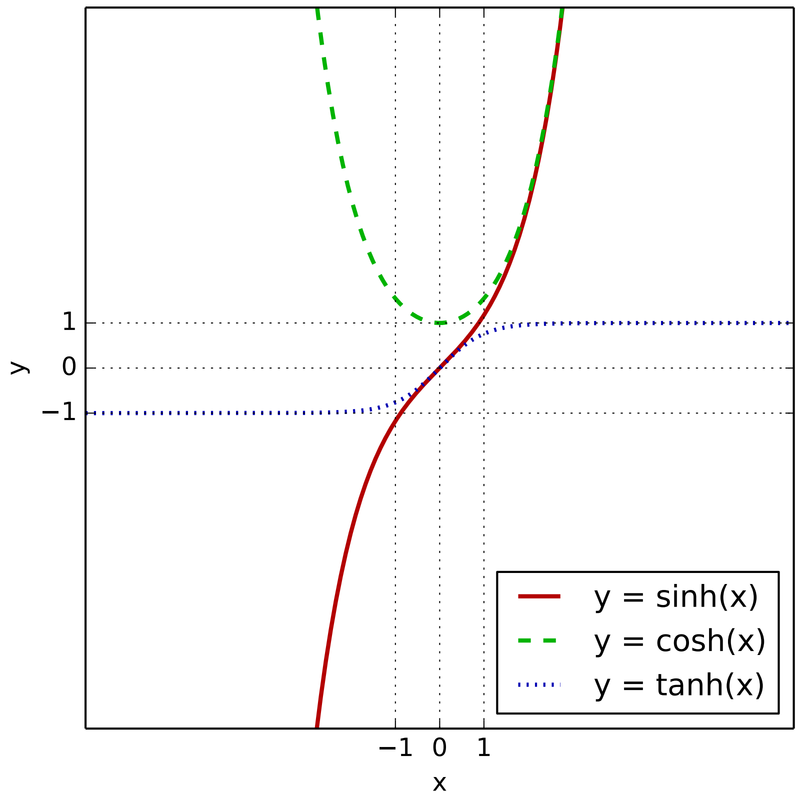

Figure 1.1: sinh, cosh and tanh

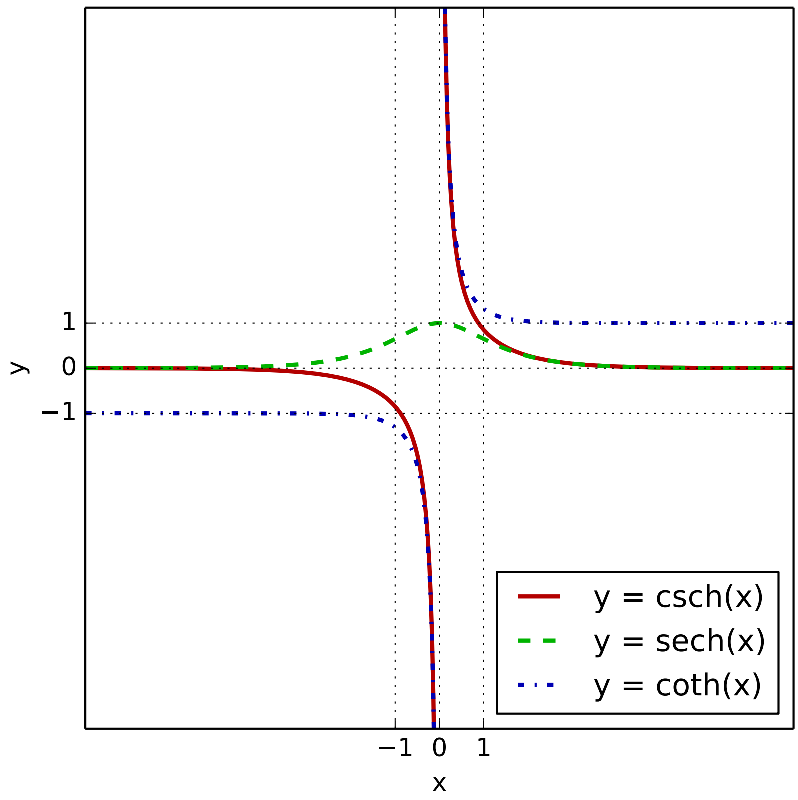

Figure 1.2: csch, sech and coth

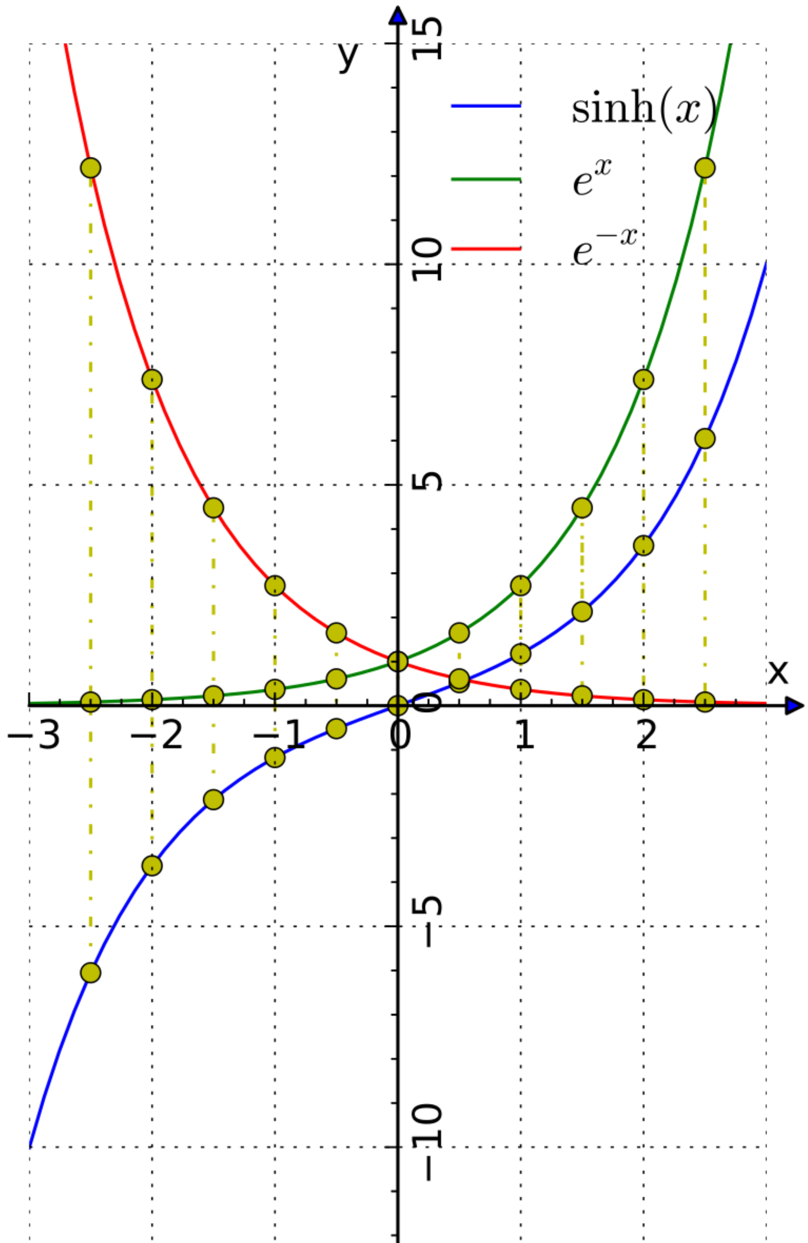

Figure 1.3: $\sinh x$ is half the difference of e^x and e^{-x}

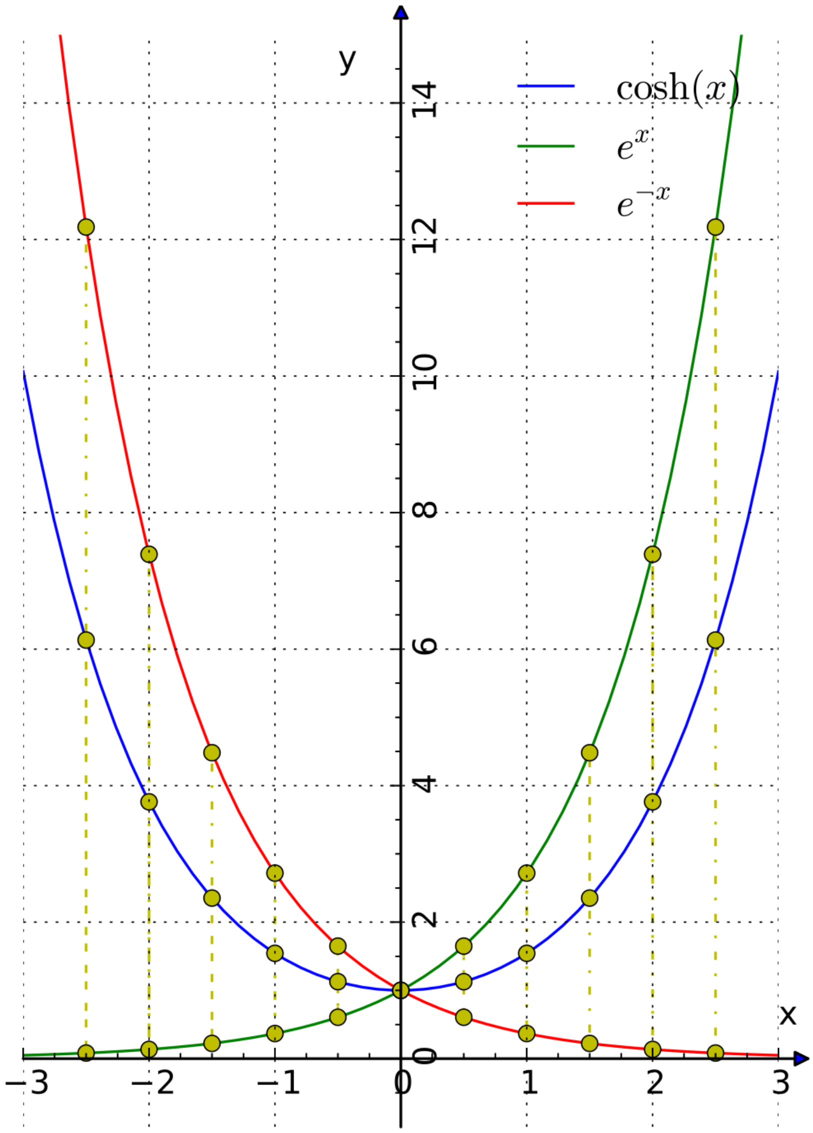

Figure 1.4: $\cosh x$ is the average of e^x and e^{-x}

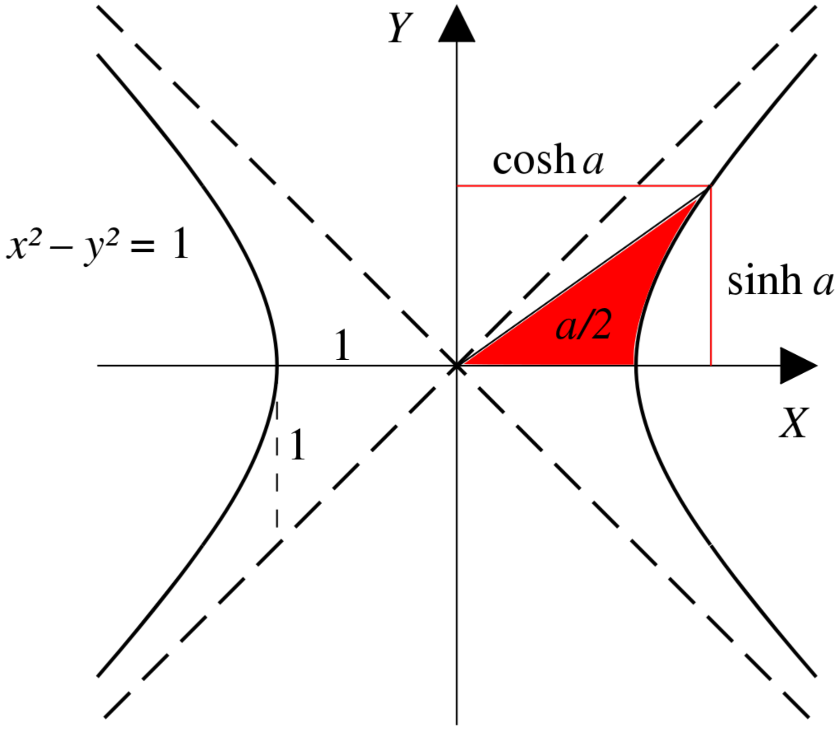

Figure 1.5: A ray through the unit hyperbola $x^2 - y^2 = 1$ at the point $(\cosh a, \sinh a)$, where a is twice the area between the ray, the hyperbola, and the x-axis. For points on the hyperbola below the x-axis, the area is considered negative (see animated version with comparison with the trigonometric (circular) functions).

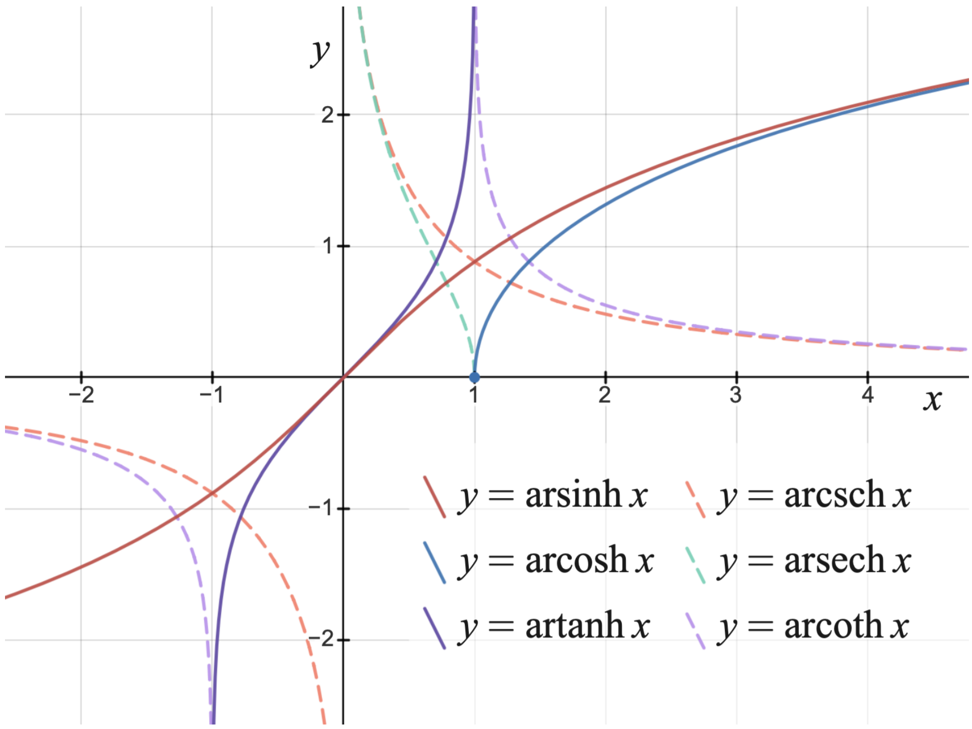

Figure 1.6

1.2.3 Useful relations

The hyperbolic functions satisfy many identities, all of them similar in form to the trigonometric identities.

Odd and even functions: $\cosh x$ and $\operatorname{sech} x$ are even functions; the others are odd functions:

$$\sinh(-x) = -\sinh x \tag{1.16}$$
$$\cosh(-x) = \cosh x \tag{1.17}$$
$$\tanh(-x) = -\tanh x \tag{1.18}$$
$$\coth(-x) = -\coth x \tag{1.19}$$
$$\operatorname{sech}(-x) = \operatorname{sech} x \tag{1.20}$$
$$\operatorname{csch}(-x) = -\operatorname{csch} x \tag{1.21}$$

Relations of inverses:

$$\operatorname{arsech} x = \operatorname{arcosh}\left(\frac{1}{x}\right) \tag{1.22}$$
$$\operatorname{arcsch} x = \operatorname{arsinh}\left(\frac{1}{x}\right) \tag{1.23}$$
$$\operatorname{arcoth} x = \operatorname{artanh}\left(\frac{1}{x}\right) \tag{1.24}$$

Hyperbolic sine and cosine satisfy:

$$\cosh x + \sinh x = e^x \tag{1.25}$$
$$\cosh x - \sinh x = e^{-x} \tag{1.26}$$
$$\cosh^2 x - \sinh^2 x = 1 \tag{1.27}$$

the last of which is similar to the Pythagorean trigonometric identity. One also has

$$\operatorname{sech}^2 x = 1 - \tanh^2 x \tag{1.28}$$
$$\operatorname{csch}^2 x = \coth^2 x - 1 \tag{1.29}$$

for the other functions.

1.2.4 Hyperbolic Identities

Sums of arguments:

$$\sinh(x+y) = \sinh x \cosh y + \cosh x \sinh y \tag{1.30}$$
$$\cosh(x+y) = \cosh x \cosh y + \sinh x \sinh y \tag{1.31}$$
$$\tanh(x+y) = \frac{\tanh x + \tanh y}{1 + \tanh x \tanh y} \tag{1.32}$$

Particularly,

$$\cosh(2x) = \sinh^2 x + \cosh^2 x = 2\sinh^2 x + 1 = 2\cosh^2 x - 1 \tag{1.33}$$

$$\sinh(2x) = 2\sinh x \cosh x \tag{1.34}$$

$$\tanh(2x) = \frac{2\tanh x}{1 + \tanh^2 x} \tag{1.35}$$

$$\sinh x + \sinh y = 2\sinh\left(\frac{x+y}{2}\right) \tag{1.36}$$

$$\cosh\left(\frac{x-y}{2}\right) \cosh x + \cosh y = 2\cosh\left(\frac{x+y}{2}\right)\cosh\left(\frac{x-y}{2}\right) \tag{1.37}$$

Square formulas:

$$\sinh^2 x = \frac{1}{2}(\cosh 2x - 1)$$
$$\cosh^2 x = \frac{1}{2}(\cosh 2x + 1)$$

1.2.5 Derivatives

$$\frac{d}{dx}\sinh x = \cosh x \tag{1.38}$$

$$\frac{d}{dx}\cosh x = \sinh x \tag{1.39}$$

$$\frac{d}{dx}\tanh x = 1 - \tanh^2 x = \operatorname{sech}^2 x = \frac{1}{\cosh^2 x} \qquad \text{for } x \neq 0 \tag{1.40}$$

$$\frac{d}{dx}\coth x = 1 - \coth^2 x = -\operatorname{csch}^2 x = -\frac{1}{\sinh^2 x} \qquad \text{for } x \neq 0 \tag{1.41}$$

$$\frac{d}{dx}\operatorname{sech} x = -\tanh x \operatorname{sech} x \tag{1.42}$$

$$\frac{d}{dx}\operatorname{csch} x = -\coth x \operatorname{csch} x \tag{1.43}$$

$$\frac{d}{dx}\operatorname{arsinh} x = \frac{1}{\sqrt{x^2+1}} \tag{1.44}$$

$$\frac{d}{dx}\operatorname{arcosh} x = \frac{1}{\sqrt{x^2-1}} \qquad \text{for } 1 < x \tag{1.45}$$

$$\frac{d}{dx}\operatorname{artanh} x = \frac{1}{1-x^2} \qquad \text{for } |x| < 1 \tag{1.46}$$

$$\frac{d}{dx}\operatorname{arcoth} x = \frac{1}{1-x^2} \qquad \text{for } 1 < |x| \tag{1.47}$$

$$\frac{d}{dx}\operatorname{arsech} x = -\frac{1}{x\sqrt{1-x^2}} \qquad \text{for } 0 < x < 1 \tag{1.48}$$

$$\frac{d}{dx}\operatorname{arcsch} x = -\frac{1}{|x|\sqrt{1+x^2}} \qquad \text{for } x \neq 0 \tag{1.49}$$

Each of the functions sinh and cosh is equal to its second derivative, that is:

$$\frac{d^2}{dx^2}\sinh x = \sinh x \tag{1.50}$$

$$\frac{d^2}{dx^2}\cosh x = \cosh x. \tag{1.51}$$

All functions with this property are linear combinations of sinh and cosh, in particular the exponential functions e^x and e^{-x}.

1.2.6 Taylor series expressions

It is possible to express explicitly the Taylor series at zero (or the Laurent series, if the function is not defined at zero) of the above functions.

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} \tag{1.52}$$

This series is convergent for every complex value of x. Since the function $\sinh x$ is odd, only odd exponents for x occur in its Taylor series.

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} \tag{1.53}$$

This series is convergent for every complex value of x. Since the function $\cosh x$ is even, only even exponents for x occur in its Taylor series. The sum of the sinh and cosh series is the infinite series expression of the exponential function.

1.2.7 Relationship to the exponential function

The decomposition of the exponential function in its even and odd parts gives the identities

$$e^x = \cosh x + \sinh x \tag{1.54}$$

and

$$e^{-x} = \cosh x - \sinh x. \tag{1.55}$$

Combined with Euler's formula

$$e^{ix} = \cos x + i \sin x \tag{1.56}$$

this gives

$$e^{x+iy} = (\cosh x + \sinh x)(\cos y + i \sin y) \tag{1.57}$$

for the general complex exponential function. Additionally,

$$e^x = \sqrt{\frac{1 + \tanh x}{1 - \tanh x}} = \frac{1 + \tanh \frac{x}{2}}{1 - \tanh \frac{x}{2}} \tag{1.58}$$

1.2.8 Composition of hyperbolic and inverse hyperbolic functions

$$\sinh(\operatorname{arcosh} x) = \sqrt{x^2-1} \qquad \text{for } |x|>1 \tag{1.59}$$
$$\sinh(\operatorname{artanh} x) = \frac{x}{\sqrt{1-x^2}} \qquad \text{for } -1<x<1 \tag{1.60}$$
$$\cosh(\operatorname{arsinh} x) = \sqrt{1+x^2} \tag{1.61}$$
$$\cosh(\operatorname{artanh} x) = \frac{1}{\sqrt{1-x^2}} \qquad \text{for } -1<x<1 \tag{1.62}$$
$$\tanh(\operatorname{arsinh} x) = \frac{x}{\sqrt{1+x^2}} \tag{1.63}$$
$$\tanh(\operatorname{arcosh} x) = \frac{\sqrt{x^2-1}}{x} \qquad \text{for } |x|>1 \tag{1.64}$$

1.3 Differential Geometry

Manifold. An d-dimensional manifold $\mathcal{M}$ is a topological space that locally resembles the topological space $\mathbb{R}^d$ near each point. More concretely, for each point x on $\mathcal{M}$, we can find a diffeomorphism (continuous bijection with continuous inverse) between a neighborhood of $\mathbf{x}$ and $\mathbb{R}^d$. The notion of manifold is a generalization of surfaces in high dimensions.

Tangent space. Intuitively, if we think of $\mathcal{M}$ as a d-dimensional manifold embedded in $\mathbb{R}^{d+1}$, the tangent space $\mathcal{T}_\mathbf{x}\mathcal{M}$ at point $\mathbf{x}$ on $\mathcal{M}$ is a d-dimensional hyperplane in $\mathbb{R}^{d+1}$ that best approximates $\mathcal{M}$ around $\mathbf{x}$. Another possible interpretation for $\mathcal{T}_\mathbf{x}\mathcal{M}$ is that it contains all the possible directions of curves on $\mathcal{M}$ passing through $\mathbf{x}$. The elements of $\mathcal{T}_\mathbf{x}\mathcal{M}$ are called tangent vectors and the union of all tangent spaces is called the tangent bundle $\mathcal{T}\mathcal{M} = \cup_{\mathbf{x}\in\mathcal{M}}\mathcal{T}_\mathbf{x}\mathcal{M}$.

Riemannian manifold. A Riemannian manifold is a pair $(\mathcal{M}, \mathbf{g})$, where $\mathcal{M}$ is a smooth manifold and $\mathbf{g} = (g_\mathbf{x})_{\mathbf{x}\in\mathcal{M}}$ is a Riemannian metric, that is a family of smoothly varying inner products on tangent spaces, $g_\mathbf{x} : \mathcal{T}_\mathbf{x}\mathcal{M} \times \mathcal{T}_\mathbf{x}\mathcal{M} \to \mathbb{R}$. Riemannian metrics can be used to measure distances on manifolds.

Distances and geodesics. Let $(\mathcal{M}, \mathbf{g})$ be a Riemannian manifold. For $\mathbf{v} \in \mathcal{T}_\mathbf{x}\mathcal{M}$, define the norm of $\mathbf{v}$ by $\|\mathbf{v}\|_\mathbf{g} := \sqrt{g_\mathbf{x}(\mathbf{v},\mathbf{v})}$. Suppose $\gamma : [a,b] \to \mathcal{M}$ is a smooth curve on $\mathcal{M}$. Define the length of γ by:

$$L(\gamma) := \int_a^b \left\|\gamma'(t)\right\|_\mathbf{g} dt \tag{1.65}$$

Now with this definition of length, every connected Riemannian manifold becomes a metric space and the distance $d : \mathcal{M} \times \mathcal{M} \to [0,\infty)$ is defined as:

$$d(\mathbf{x},\mathbf{y}) := \inf_\gamma\{L(\gamma) : \gamma \text{ is a continuously differentiable curve joining } \mathbf{x} \text{ and } \mathbf{y}\} \tag{1.66}$$

Geodesic distances are a generalization of straight lines (or shortest paths) to non-Euclidean geometry. A curve $\gamma : [a,b] \to \mathcal{M}$ is geodesic if

$$d(\gamma(t),\gamma(s)) = L\left(\gamma|_{[t,s]}\right) \forall (t,s) \in [a,b] (t<s).$$

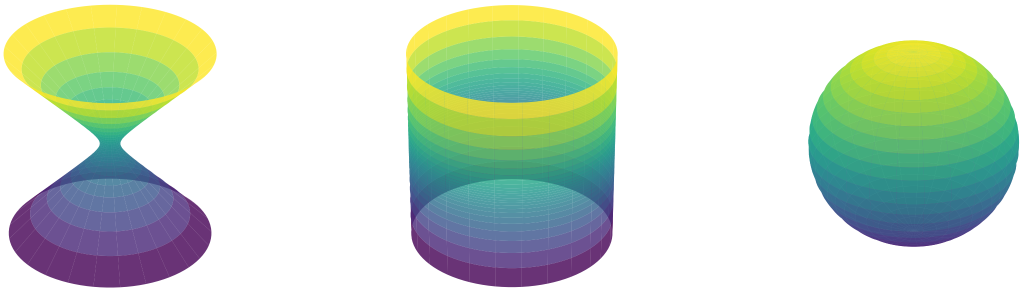

Figure 1.7

It must be emphasized that it is local.

Parallel transport. Parallel transport is a generalization of translation to non-Euclidean geometry. Given a smooth manifold $\mathcal{M}$, parallel transport $P_{\mathbf{x}\to\mathbf{y}}(\cdot)$ maps a vector $\mathbf{v} \in \mathcal{T}_{\mathbf{x}}\mathcal{M}$ to $P_{\mathbf{x}\to\mathbf{y}}(\mathbf{v}) \in \mathcal{T}_{\mathbf{y}}\mathcal{M}$. In Riemannian geometry, parallel transport preserves the Riemannian metric tensor (norm, inner products...).

Curvature. At a high level, curvature measures how much a geometric object such as surfaces deviate from a flat plane. For instance, the Euclidean space has zero curvature while spheres have positive curvature. We illustrate the concept of curvature in Figure 6. (Figure 6: From left to right: a surface of negative curvature, a surface of zero curvature, and a surface of positive curvature.)

Chapter 2

Hyperbolic Spaces

The *hyperbolic space* in d dimensions is the unique complete, simply connected d-dimensional Riemannian manifold with constant negative sectional curvature.

There exist several models of hyperbolic space such as the *Poincaré model* or the *hyperboloid model* (also known as the *Minkowski model* or the *Lorentz model*). In what follows, we review the Poincaré and the hyperboloid models of hyperbolic space as well as connections between these two models.

2.1 Geometry of Hyperboloid model

An inner product (see [Axl15, 6.3 Definition]) on a real vector space V is a function that associates a real number $\langle u, v\rangle$ with each pair of vectors u and v in V in such a way that the following properties are satisfied for all vectors u, v and z in V and all scalars k.

1. symmetry: $\langle u, v\rangle = \langle u, v\rangle$;
2. additivity in first slot: $\langle u + v, z\rangle = \langle u, z\rangle + \langle v, z\rangle$;
3. homogeneity in first slot: $\langle ku, v\rangle = k\langle u, v\rangle$;
4. positivity: $\langle v, v\rangle \geqslant 0$;
5. definiteness: $\langle v, v\rangle = 0$ if and only if $v = 0$.

A real vector space with an inner product is called a real inner product space.

Definition 1: Minkowski (pseudo) Inner Product

Consider the bilinear map

$$\langle \cdot, \cdot\rangle_{\mathcal{M}} : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \to \mathbb{R}$$

defined by

$$\langle u, v\rangle_{\mathcal{M}} = -u_0 v_0 + \sum_{i=1}^{n} u_i v_i = u^\top J v \tag{2.1}$$

where $J = \mathrm{diag}(-1, 1, \ldots, 1) \in \mathbb{R}^{(n+1)\times(n+1)}$. It is called the Minkowski (pseudo) inner product on $\mathbb{R}^{n+1}$.

This is not an inner product on $\mathbb{R}^{n+1}$ because J has one negative eigenvalue, but it is a pseudo-inner product because all eigenvalues of J are nonzero. Gavin a constant $K > 0$, the equation

$$\langle x, x\rangle_{\mathcal{M}} = -K, \tag{2.2}$$

i.e.,

$$x_0^2 = K + \sum_{i=1}^{n} x_i^2 \geqslant K, \tag{2.3}$$

defines two connected components, determined by the sign of x_0. Note that $x_0 \geqslant K$ or $x_0 \leqslant -K$. The condition $x_0 > 0$ selects one of them.

Definition 2: $\mathcal{H}^{n,K}$

Consider the following subset of $\mathbb{R}^{n+1}$:

$$\begin{aligned} \mathcal{H}^{n,K} &:= \left\{x \in \mathbb{R}^{n+1} : \langle x, x \rangle_{\mathcal{M}} = -K \text{ and } x_0 > 0\right\} && (2.4)\\ &= \left\{x \in \mathbb{R}^{n+1} : x_0^2 = K + x_1^2 + \cdots + x_n^2 \text{ and } x_0 > 0\right\} && (2.5)\\ &= \left\{x \in \mathbb{R}^{n+1} : h(x) := \langle x, x \rangle_{\mathcal{M}} + K = 0\right\} \text{ with } x_0 > 0. && (2.6) \end{aligned}$$

The defining function $h(x) = \langle x, x \rangle_{\mathcal{M}} + K$ has differential

$$\mathrm{D}h(x)[u] = 2\langle x, u \rangle_{\mathcal{M}} = (2Jx)^\top u. \tag{2.7}$$

Notice that $x_0 \neq 0$ for all $x \in \mathcal{H}^{n,K}$; hence, $2Jx \neq 0$ for all $x \in \mathcal{H}^{n,K}$ (J is invertible matrix, thus $Jx = 0$ if and only if $x = 0$). This implies that differential $\mathrm{D}h(x) : \mathbb{R}^{n+1} \to \mathbb{R}$ is surjective (i.e., $\operatorname{rank} \mathrm{D}h(x) = 1$) for all $x \in \mathcal{H}^{n,K}$. By [Bou23, Definition 3.10 & Theorem 3.15], we conclude the following proposition.

Proposition 1: tangent Space $\mathrm{T}_x\mathcal{H}^{n,K}$

Given any constant $K > 0$, the set $\mathcal{H}^{n,K}$ is an embedded submanifold of $\mathbb{R}^{n+1}$ of dimension n with tangent space:

$$\begin{aligned} \mathrm{T}_x\mathcal{H}^{n,K} &= \ker \mathrm{D}h(x) && (2.8)\\ &= \left\{u \in \mathbb{R}^{n+1} : \langle x, u \rangle_{\mathcal{M}} = 0\right\} && (2.9)\\ &= \{u \in \mathbb{R}^{n+1} : x_0 u_0 = \sum_{i=1}^{n} x_i u_i\} && (2.10) \end{aligned}$$

which is an n-dimensional subspace of $\mathbb{R}^{n+1}$. The tangent bundle of $\mathcal{H}^{n,K}$ is given as

$$T\mathcal{H}^{n,K} := \left\{(x, u) \mid x \in M, u \in T_x\mathcal{H}^{n,K}\right\}. \tag{2.11}$$

Example 1. For $n = 1, K = 1$, the manifold $\mathcal{H}^{1,1}$ is one sheet of a hyperbola of two sheets in $\mathbb{R}^2$:

$$\mathcal{H}^{1,1} = \left\{x \in \mathbb{R}^2 : x_0^2 - x_1^2 = 1 \text{ and } x_0 > 0\right\}. \tag{2.12}$$

Proposition 2: $\langle \cdot, \cdot \rangle_{\mathcal{M}}$ is a inner product on $\mathrm{T}_x\mathcal{H}^{n,K}$

$\langle \cdot, \cdot \rangle_{\mathcal{M}}$ is only a pseudo-inner product on $\mathbb{R}^{n+1}$, however, it is an inner product restricted to the tangent spaces of $\mathcal{H}^{n,K}$, i.e.,

$$\langle \cdot, \cdot \rangle_{\mathcal{M}} : \mathrm{T}_x\mathcal{H}^{n,K} \times \mathrm{T}_x\mathcal{H}^{n,K} \to \mathbb{R} \tag{2.13}$$

is a well-defined inner product on $\mathrm{T}_x\mathcal{H}^{n,K}$ for all $x \in \mathcal{H}^{n,K}$. Then, $\|u\|_{\mathcal{M}} = \sqrt{\langle u, u\rangle_{\mathcal{M}}}$ is a well-defined norm on it.

Proof. Symmetry, additivity and homogeneity hold since $\langle u, v\rangle_{\mathcal{M}} = u^\top J v$ with diagonal J. We next show the positivity and definiteness. For all $(x, u) \in \mathrm{T}\mathcal{H}^{n,K}$, we have

$$
\begin{aligned}
\langle u, u\rangle_{\mathcal{M}} &= (\sum_{i=1}^{n} u_i^2) - u_0^2 &(2.14)\\
&= (\sum_{i=1}^{n} u_i^2) - \frac{1}{x_0^2}(\sum_{i=1}^{n} x_i u_i)^2 \text{ (by } u \in T_x\mathcal{H}^{n,K}\text{, i.e., } \langle x, u\rangle_{\mathcal{M}} = 0 \text{ and } x_0 > 0) &(2.15)\\
&\geqslant (\sum_{i=1}^{n} u_i^2) - \frac{1}{x_0^2}(\sum_{i=1}^{n} x_i^2)(\sum_{i=1}^{n} u_i^2) \text{ (by Cauchy-Schwarz inequality)} &(2.16)\\
&= (\sum_{i=1}^{n} u_i^2)\left(1 - \frac{\sum_{i=1}^{n} x_i^2}{x_0^2}\right) &(2.17)\\
&= (\sum_{i=1}^{n} u_i^2)\left(1 - \frac{x_0^2 - K}{x_0^2}\right) \text{ (by } x \in \mathcal{H}^{n,K}\text{, i.e., } \langle x, x\rangle_{\mathcal{M}} = -K) &(2.18)\\
&= \frac{K}{x_0^2}\left(u_1^2 + \cdots + u_n^2\right) \geqslant 0. &(2.19)
\end{aligned}
$$

Note that $0 < \frac{K}{x_0^2} \leqslant 1$ here. If $\langle u, u\rangle_{\mathcal{M}} = 0$, then by (2.19) we have $\sum_{i=1}^{n} u_i^2 = 0$; thus $u_i = 0$ for $i = 1, \ldots, n$. For $i = 0$, $u_0^2 = \sum_{i=1}^{n} u_i^2 = 0$. This completes the proof. $\square$

Remark 1. $\langle u, u\rangle_{\mathcal{M}}$ can be negative if u does not belong to any tangent space of $\mathcal{H}^{n,K}$.

Definition 3: Hyperbolic Space $\mathcal{H}^{n,K}$

The restriction of $\langle \cdot, \cdot\rangle_{\mathcal{M}}$ to each tangent space $\mathrm{T}_x\mathcal{H}^{n,K}$ defines a Riemannian metric on $\mathcal{H}^{n,K}$, turning it into a Riemannian manifold. With this Riemannian structure, we call $\mathcal{H}^{n,K}$ a *hyperbolic space* in the *hyperboloid model*. The main geometric trait of $\mathcal{H}^{n,K}$ with $n \geqslant 2$ is that its *sectional curvatures* are negative constant, equal to

$$-\frac{1}{K} \text{ for some } K > 0. \tag{2.20}$$

Manifolds with that property are called *hyperbolic spaces*. There are several other models that share this trait, namely the *Beltrami-Klein model*, the *Poincaré ball model* and the *Poincaré half-space model*. For more about curvature and these models, see [Lee18, page 62].

Definition 4: North Pole Point

The point $o := (\sqrt{K}, 0, 0, \ldots, 0) \in \mathcal{H}^{n,K}$ is called the north pole point of $\mathcal{H}^{n,K}$.

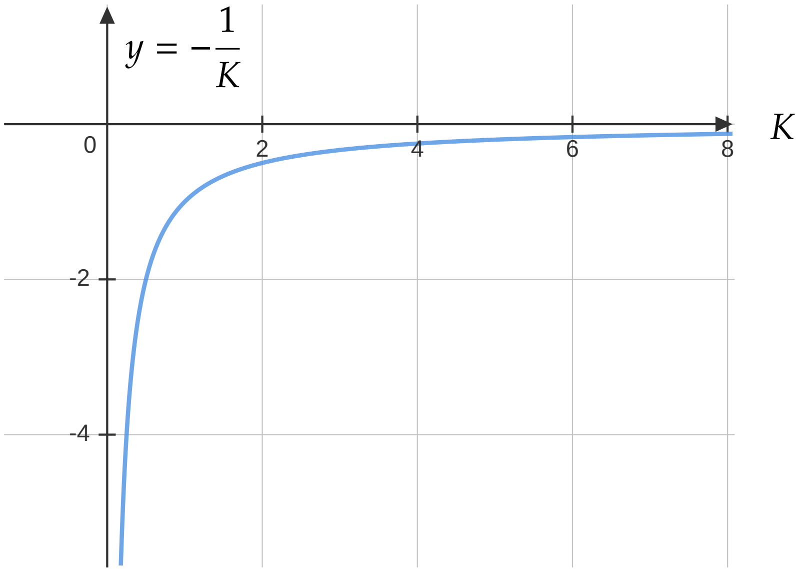

Figure 2.1: Curvatures $-\frac{1}{K}$ determined by $K > 0$.

We observe that

$$T_o\mathcal{H}^{n,K} = \left\{u \in \mathbb{R}^{n+1} : \langle o, u\rangle_{\mathcal{M}} = -\sqrt{K}u_0 = 0\right\} \tag{2.21}$$
$$= \left\{u \in \mathbb{R}^{n+1} : u_0 = 0\right\} \tag{2.22}$$
$$= \left\{\left(0, u'\right) \in \mathbb{R}^{n+1} : u' \in \mathbb{R}^n\right\} \tag{2.23}$$
$$\simeq \mathbb{R}^n. \tag{2.24}$$

Thus, if we fix the dimension n, then for different curvatures $-\frac{1}{K}$, the manifolds $H^{n,K}$ have different north pole points but share the same tangent space at their north pole points. See Figure 2.2.

Proposition 3: Riemannian distance on $\mathcal{H}^{n,K}$

The distance function induced by Riemannian metric $\langle\cdot,\cdot\rangle_{\mathcal{M}}$ is

$$d^K_{\mathcal{M}}(x, y) = \sqrt{K}\,\mathrm{arcosh}\left(-\langle x, y\rangle_{\mathcal{M}}/K\right) \tag{2.25}$$

for all $x, y \in \mathcal{H}^{n,K}$.

We now check the well-definedness of (2.25). Recall that the natural domain of the inverse hyperbolic function

$$\mathrm{arcosh}(z) = \ln\left(z + \sqrt{z^2 - 1}\right) \tag{2.26}$$

is $[1, \infty)$. We need to show that $-\langle x, y\rangle_{\mathcal{M}}/K \geqq 1$ for all $x, y \in \mathcal{H}^{n,K}$. Define two vectors in $\mathbb{R}^{n+1}$:

$$x' := \left(\sqrt{K}, x_1, \ldots, x_n\right), y' := \left(\sqrt{K}, y_1, \ldots, y_n\right). \tag{2.27}$$

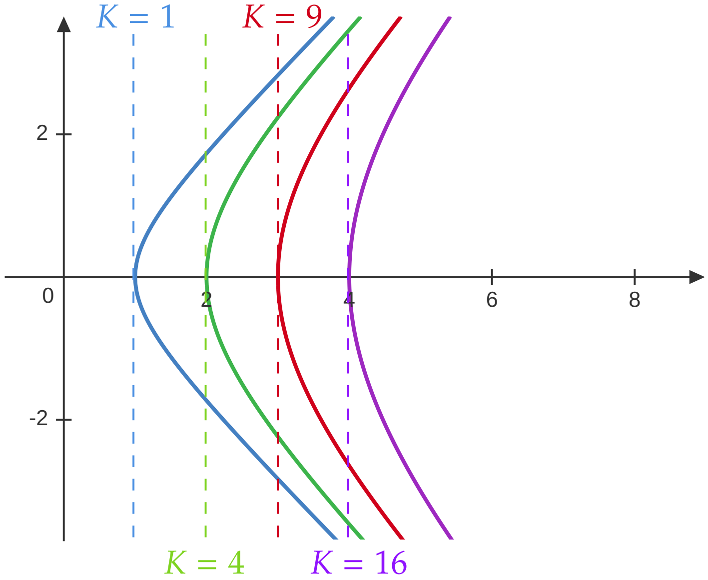

Figure 2.2: Graphs of $\mathcal{H}^{1,K}$ with $K = 1, 4, 9, 16$. Dashed lines denote the tangent spaces at north pole points.

Since $x, y \in \mathcal{H}^{n,K}$ and (2.3), $\|x'\|_2 = x_0 > 0$ and $\|y'\|_2 = y_0 > 0$. Then

$$-\langle x, y\rangle_{\mathcal{M}} = x_0 y_0 - \sum_{i=1}^{n} x_i y_i \tag{2.28}$$

$$= \|x'\|_2 \|y'\|_2 - (\sum_{i=1}^{n} x_i y_i + \sqrt{K}\sqrt{K}) + K \tag{2.29}$$

$$= \|x'\|_2 \|y'\|_2 - \langle x', y'\rangle + K \text{ (here } \langle\cdot,\cdot\rangle \text{ is usual inner product on } \mathbb{R}^{n+1}) \tag{2.30}$$

$$\geqq K \text{ (by Cauchy-Schwarz inequality)} \tag{2.31}$$

Therefore, $d_{\mathcal{M}}^{K}(x, y)$ is well-defined for any pair of x, y in $\mathcal{H}^{n,K}$.

Recall that the tangent space $\mathrm{T}_x\mathcal{H}^{n,K}$ is an n-dimensional subspace of $\mathbb{R}^{n+1}$. We consider its orthogonal complement in $\mathbb{R}^{n+1}$ as below.

Proposition 4: Normal Space $\mathrm{N}_x\mathcal{H}^{n,K}$

The orthogonal complement of $\mathrm{T}_x\mathcal{H}^{n,K}$ with respect to $\langle\cdot,\cdot\rangle_{\mathcal{M}}$ is the one-dimensional normal space:

$$\mathrm{N}_x\mathcal{H}^{n,K} := \left\{v \in \mathbb{R}^{n+1} : \langle u, v\rangle_{\mathcal{M}} = 0 \text{ for all } u \in \mathrm{T}_x\mathcal{H}^{n,K}\right\} \tag{2.32}$$

$$= \operatorname{span}(x). \tag{2.33}$$

Proof. From $\dim \mathrm{T}_x\mathcal{H}^{n,K} = n$, we know that $\dim \mathrm{N}_x\mathcal{H}^{n,K} = 1$, which means the normal space $\mathrm{N}_x\mathcal{H}^{n,K}$ at point $x \in \mathcal{H}^{n,K}$ is a one-dimensional subspace in $\mathbb{R}^{n+1}$. By definition, we have $x \in \mathrm{N}_x\mathcal{H}^{n,K}$. Since x always is nonzero, we clam that $\mathrm{T}_x\mathcal{H}^{n,K} = \operatorname{span}(x)$. □

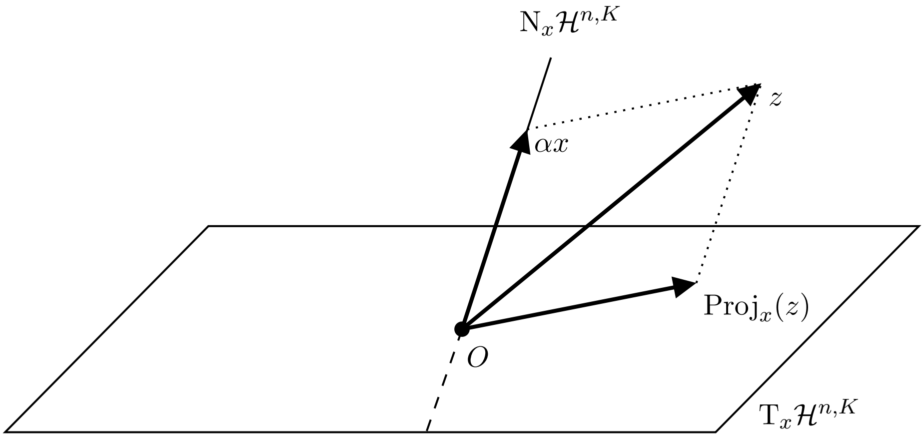

Figure 2.3: The unique decomposition of z under the direct sum of tangent space and normal space.

Proposition 5: Orthogonal Projector Proj_x

The orthogonal projector $\mathrm{Proj}_x : \mathbb{R}^{n+1} \to \mathrm{T}_x\mathcal{H}^{n,K}$ is given by

$$\mathrm{Proj}_x(z) = z + \frac{1}{K}\langle x, z\rangle_{\mathcal{M}} \cdot x. \tag{2.34}$$

for $z \in \mathbb{R}^{n+1}$.

Proof. Let $z \in \mathbb{R}^{n+1}$. Since tangent space and normal space consist a direct sum of $\mathbb{R}^{n+1}$, then we have the unique decomposition of z in the form (See Figure 2.3):

$$z = \alpha x + \mathrm{Proj}_x(z) \tag{2.35}$$

for some $\alpha \in \mathbb{R}$. From (2.32), the normal part of z is αx. By definition of orthogonal projection, $\mathrm{Proj}_x(z)$ is exactly the tangent part of z. By rearranging terms and redefining the symbol α, we have

$$\mathrm{Proj}_x(z) = z + \alpha x \tag{2.36}$$

for some $\alpha \in \mathbb{R}$. Since $\mathrm{Proj}_x(z) \in \mathrm{T}_x\mathcal{H}^{n,K}$, equation $\langle \mathrm{Proj}_x(z), x\rangle_{\mathcal{M}} = 0$ implies that

$$\langle z + \alpha x, x\rangle_{\mathcal{M}} = \langle z, x\rangle_{\mathcal{M}} + \alpha\langle x, x\rangle_{\mathcal{M}} = \langle z, x\rangle_{\mathcal{M}} - \alpha K = 0. \tag{2.37}$$

Thus, $\alpha = \frac{1}{K}\langle z, x\rangle_{\mathcal{M}}$. □

With this tool in hand, we can construct a useful formula to compute gradients of functions on $\mathcal{H}^{n,K}$.

Proposition 6: Compute Riemannian Gradients

Let $\bar{f} : \mathbb{R}^{n+1} \to \mathbb{R}$ be a smooth function on the Euclidean space $\mathbb{R}^{n+1}$ with the usual inner product $\langle u, v\rangle = u^\top v$. Let $f = \bar{f}|_{\mathcal{H}^{n,K}}$ be the restriction of $\bar{f}$ to $\mathcal{H}^{n,K}$ with the Riemannian structure as described above. The gradient of f is related to that of $\bar{f}$ as follows:

$$\mathrm{grad}\, f(x) = \mathrm{Proj}_x(J\,\mathrm{egrad}\,\bar{f}(x)), \tag{2.38}$$

where $J = \mathrm{diag}(-1, 1, \ldots, 1)$ and Proj_x is defined by (2.34).

Proof. By definition, $\mathrm{grad}\, f(x)$ is the unique vector in $\mathrm{T}_x\mathcal{H}^{n,K}$ such that $\mathrm{D} f(x)[u] = \langle \mathrm{grad}\, f(x), u\rangle_{\mathcal{M}}$ for all $u \in \mathrm{T}_x\mathcal{H}^{n,K}$. Since $\bar{f}$ is a smooth extension of f, we can compute

$$
\begin{aligned}
\mathrm{D}\, f(x)[u] &= \mathrm{D}\bar{f}(x)[u] && (2.39)\\
&= \langle \mathrm{egrad}\, \bar{f}(x), u\rangle \text{ (by definition of egrad } \bar{f}(x)) && (2.40)\\
&= \langle J\, \mathrm{egrad}\, \bar{f}(x), u\rangle_{\mathcal{M}} \text{ (by definition of } \langle\cdot,\cdot\rangle_{\mathcal{M}}) && (2.41)\\
&= \left\langle J\, \mathrm{gegrad}\, \bar{f}(x), \mathrm{Proj}_x(u)\right\rangle_{\mathcal{M}} \text{ (because } u \text{ is tangent at } x) && (2.42)\\
&= \langle \mathrm{Proj}(J\, \mathrm{egrad}\, \bar{f}(x)), u\rangle_{\mathcal{M}}. \text{ (}\mathrm{Proj}_x \text{ is self-adjoint with respect to } \langle\cdot,\cdot\rangle_{\mathcal{M}}) && (2.43)
\end{aligned}
$$

The claim follows by uniqueness. □

Some notes from `hyperbolicfactory.m` in Manopt:

- $\mathcal{H}^{n,K}$ is an embedded submanifold of $\mathbb{R}^{n+1}$ equipped with the usual inner product. Thus, when defining the Euclidean gradient for example (`problem.egrad`), it should be specified as if the function were defined in Euclidean space directly. The tool `M.egrad2rgrad` will automatically convert that gradient to the correct Riemannian gradient, as needed to satisfy the metric. The same is true for the Euclidean Hessian and other tools that manipulate elements in the embedding space.
- Importantly, $\mathcal{H}^{n,K}$ *is not a Riemannian submanifold of* $\mathbb{R}^{n+1}$, because its metric is not obtained simply by restricting the Euclidean metric to the tangent spaces.
- However, $\mathcal{H}^{n,K}$ is a *semi*-Riemannian submanifold of *Minkowski space*, that is, $\mathbb{R}^{n+1}$ equipped with the Minkowski inner product.
- Minkowski space itself can be seen as a (linear) semi-Riemannian manifold embedded in Euclidean space.

Note that $J\, \mathrm{egrad}\, \bar{f}(x)$ in (2.38) is the gradient of $\bar{f}$ in the Minkowski space $\mathbb{R}^{n+1}$ with pseudo-inner product $\langle\cdot,\cdot\rangle_M$. See [O'n83] for a general treatment of submanifolds of spaces equipped with pseudo-inner products.

Proposition 7: Geodesic

For arbitrary $(x, u) \in \mathrm{T}\mathcal{H}^{n,K}$ with $u \neq 0$,

$$
\begin{aligned}
c(t) &:= \begin{cases} \cosh\left(\frac{\|tu\|_{\mathcal{M}}}{\sqrt{K}}\right)\cdot x + \sqrt{K}\sinh\left(\frac{\|tu\|_{\mathcal{M}}}{\sqrt{K}}\right)\cdot \frac{tu}{\|tu\|_{\mathcal{M}}} & \text{if } t \neq 0 \\ x & \text{if } t = 0 \end{cases} && (2.44)\\
&= \cosh\left(\frac{t\|u\|_{\mathcal{M}}}{\sqrt{K}}\right)\cdot x + \sqrt{K}\sinh\left(\frac{t\|u\|_{\mathcal{M}}}{\sqrt{K}}\right)\cdot\frac{u}{\|u\|_{\mathcal{M}}} \quad \text{for all } t \in \mathbb{R} && (2.45)
\end{aligned}
$$

defines the unique geodesic on $\mathcal{H}^{n,K}$ such that $c(0) = x$ and $c'(0) = u$.

Let $K = 1$, then

$$c(t) = \cosh\left(t\|u\|_{\mathcal{M}}\right)\cdot x + \frac{\sinh\left(t\|u\|_{\mathcal{M}}\right)}{\|u\|_{\mathcal{M}}}\cdot u. \quad (2.46)$$

Let $K = 1$ and $u \in \mathrm{T}_x\mathcal{H}^{n,K}$ be unit-speed, then

$$c(t) = \cosh(t)\, x + \sinh(t)\, u. \quad (2.47)$$

Proof. (1) *We first show that (2.44) is equal to (2.45).* When $t \neq 0$, we have

$$
\begin{aligned}
c(t) &:= \cosh\left(\frac{\|tu\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot x + \sqrt{K} \sinh\left(\frac{\|tu\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot \frac{tu}{\|tu\|_{\mathcal{M}}} && (2.48)\\
&= \cosh\left(\frac{|t| \cdot \|u\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot x + \sqrt{K} \sinh\left(\frac{|t| \cdot \|u\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot \frac{tu}{|t| \cdot \|u\|_{\mathcal{M}}} && (2.49)\\
&= \cosh\left(\frac{\pm t \cdot \|u\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot x + \sqrt{K} \sinh\left(\frac{\pm t \cdot \|u\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot \frac{tu}{\pm t \cdot \|u\|_{\mathcal{M}}} && (2.50)\\
&= \cosh\left(\frac{t \cdot \|u\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot x \pm \sqrt{K} \sinh\left(\frac{\pm t \cdot \|u\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot \frac{u}{\|u\|_{\mathcal{M}}} && (2.51)\\
&= \cosh\left(\frac{t\|u\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot x + \sqrt{K} \sinh\left(\frac{t\|u\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot \frac{u}{\|u\|_{\mathcal{M}}}. && (2.52)
\end{aligned}
$$

Furthermore, substituting $t = 0$ into (2.52) yields $\cosh(0)x + \sinh(0)u = x$.

(2) *Show that $c(t)$ is a curve on $\mathcal{H}^{n,K}$, i.e., $c(t) \in \mathcal{H}^{n,K}$ for all $t \in \mathbb{R}$.* Let $p := \frac{t\|u\|_{\mathcal{M}}}{\sqrt{K}}$. Then,

$$
\langle c(t), c(t)\rangle_{\mathcal{M}} = \left\langle \cosh(p) \cdot x + \sqrt{K} \sinh(p) \cdot \frac{u}{\|u\|_{\mathcal{M}}},\ \cosh(p) \cdot x + \sqrt{K} \sinh(p) \cdot \frac{u}{\|u\|_{\mathcal{M}}} \right\rangle_{\mathcal{M}} \tag{2.53}
$$

$$
\begin{aligned}
&= \cosh^2(p)\langle x, x\rangle_{\mathcal{M}} + \frac{K \sinh^2(p)}{\|u\|^2_{\mathcal{M}}}\langle u, u\rangle_{\mathcal{M}} + 2\sqrt{K}\frac{\cosh(p)\sinh(p)}{\|u\|_{\mathcal{M}}}\langle x, v\rangle_{\mathcal{M}} && (2.54)\\
&= \cosh^2(p)\langle x, x\rangle_{\mathcal{M}} + \frac{K \sinh^2(p)}{\|u\|^2_{\mathcal{M}}}\langle u, u\rangle_{\mathcal{M}} && (2.55)\\
&= -K\cosh^2(p) + K\sinh^2(p) && (2.56)\\
&= -K(\cosh^2(p) - \sinh^2(p)) && (2.57)\\
&= -K. && (2.58)
\end{aligned}
$$

$$
\begin{aligned}
\langle c(t), c(t)\rangle_{\mathcal{M}} &= \left\langle \cosh(p) \cdot x + \frac{\sinh(p)}{p} t \cdot u,\ \cosh(p) \cdot x + \frac{\sinh(p)}{p} t \cdot u \right\rangle_{\mathcal{M}} && (2.59)\\
&= \cosh^2(p)\langle x, x\rangle_{\mathcal{M}} + \frac{t^2 \sinh^2(p)}{p^2}\langle u, u\rangle_{\mathcal{M}} + 2\frac{\cosh(p)\sinh(p)}{p} t\langle x, v\rangle_{\mathcal{M}} && (2.60)\\
&= \cosh^2(p)\langle x, x\rangle_{\mathcal{M}} + \frac{t^2 \sinh^2(p)}{p^2}\langle u, u\rangle_{\mathcal{M}} && (2.61)\\
&= -K\cosh^2(p) + K\sinh^2(p) && (2.62)\\
&= -K(\cosh^2(p) - \sinh^2(p)) && (2.63)\\
&= -K. && (2.64)
\end{aligned}
$$

(2) *Show that $c(t)$ is the geodesic on $\mathcal{H}^{n,K}$ such that $c(0) = x$ and $c'(0) = u$.*

The (extrinsic) velocity and acceleration of c in $\mathbb{R}^d$ are easily derived:

$$\dot{c}(t) = \frac{\|u\|_{\mathcal{M}}}{\sqrt{K}} \cdot \sinh\left(\frac{t\|u\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot x + \cosh\left(\frac{t\|u\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot u, \tag{2.65}$$

$$\ddot{c}(t) = \frac{\|u\|_{\mathcal{M}}^2}{K} \cdot \cosh\left(\frac{t\|u\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot x + \frac{\|u\|_{\mathcal{M}}}{\sqrt{K}} \cdot \sinh\left(\frac{t\|u\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot u \tag{2.66}$$

$$= \|u\|_{\mathcal{M}}^2 \cdot c(t). \tag{2.67}$$

The velocity $c'(t)$ matches $\dot{c}(t)$. Owing to (5.23), to get the (intrinsic) acceleration of c on S^{d-1}, we project:

$$c''(t) = \mathrm{Proj}_{c(t)}\, \ddot{c}(t) = \left(I_d - c(t)c(t)^\top\right) \ddot{c}(t) = 0. \tag{2.68}$$

$$c''(t) = \mathrm{Proj}_{c(t)}(\ddot{c}(t)) \tag{2.69}$$

$$= \ddot{c}(t) + \frac{1}{K}\langle c(t), \ddot{c}(t)\rangle_{\mathcal{M}} \cdot c(t) \tag{2.70}$$

$$= \|u\|_{\mathcal{M}}^2 \cdot c(t) + \frac{1}{K}\langle c(t), \|u\|_{\mathcal{M}}^2 \cdot c(t)\rangle_{\mathcal{M}} \cdot c(t) \tag{2.71}$$

$$= \|u\|_{\mathcal{M}}^2 \cdot c(t) + \frac{\|u\|_{\mathcal{M}}^2}{K}\langle c(t), \cdot c(t)\rangle_{\mathcal{M}} \cdot c(t) \tag{2.72}$$

$$= 0. \tag{2.73}$$

Thus, c is a curve with zero acceleration on $\mathcal{H}^{n,K}$. □

Remark 2 (Compare with the geodesics on the sphere)**.** [Bou23, Example 5.37.] Consider the sphere $\mathrm{S}^{d-1} = \left\{x \in \mathbb{R}^d : x^\top x = 1\right\}$ equipped with the Riemannian submanifold geometry of $\mathbb{R}^d$ with the canonical metric. For a given $x \in \mathrm{S}^{d-1}$ and $v \in \mathrm{T}_x\, \mathrm{S}^{d-1}$ (nonzero), consider the curve

$$c(t) = \cos(t\|v\|) \cdot x + \frac{\sin(t\|v\|)}{\|v\|} \cdot v, \tag{2.74}$$

which traces a so-called great circle on the sphere. Then c is a geodesic on S^{d-1} such that $c(0) = x$ and $c'(0) = v$. Compare (2.46) with above equation.

Notice that this geodesic $c(t)$ is defined for all t. Thus, we have the next result.

Lemma 1: Completeness of $\mathcal{H}^{n,K}$

Hyperbolic Space $\mathcal{H}^{n,K}$ is a complete Riemannian manifold.

Mapping between tangent space and hyperbolic space is done by exponential and logarithmic maps. Given $x \in \mathcal{H}^{n,K}$ and a tangent vector $v \in \mathrm{T}_x\mathcal{H}^{n,K}$, the exponential map

$$\exp_x^K : \mathrm{T}_x\mathcal{H}^{n,K} \to \mathcal{H}^{n,K} \tag{2.75}$$

assigns to v the point $\exp_x^K(v) := \gamma(1)$, where γ is the unique geodesic satisfying $\gamma(0) = x$ and $\dot{\gamma}(0) = v$. The logarithmic map

$$\log_x^K : \mathcal{H}^{n,K} \to \mathrm{T}_x\mathcal{H}^{n,K} \tag{2.76}$$

is the reverse map that maps back to the tangent space at x such that

$$\log_x^K \circ \exp_x^K = \text{identity map on } \mathrm{T}_x\mathcal{H}^{n,K}, \tag{2.77}$$
$$\exp_x^K \circ \log_x^K = \text{identity map on } \mathcal{H}^{n,K}. \tag{2.78}$$

Remark 3. In general Riemannian manifolds these operations are only locally defined, but in hyperbolic space they form a bijection between $\mathcal{H}^{n,K}$ and $\mathrm{T}_x\mathcal{H}^{n,K}$ at any point x.

We have the following direct expressions of the exponential and the logarithmic maps.

Proposition 8: Exponential and Logarithmic Maps

For $x \in \mathcal{H}^{n,K}, v \in \mathrm{T}_x\mathcal{H}^{n,K}$ and $y \in \mathcal{H}^{n,K}$ such that $v \neq 0$ and $y \neq x$, the exponential and logarithmic maps are given by:

$$\exp_x^K(v) = \cosh\left(\frac{\|v\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot x + \sqrt{K}\sinh\left(\frac{\|v\|_{\mathcal{M}}}{\sqrt{K}}\right) \cdot \frac{v}{\|v\|_{\mathcal{M}}}, \tag{2.79}$$

and

$$\log_x^K(y) = d_{\mathcal{M}}^K(x,y) \cdot \frac{y + \frac{1}{K}\langle x,y\rangle_{\mathcal{M}} \cdot x}{\left\|y + \frac{1}{K}\langle x,y\rangle_{\mathcal{M}} \cdot x\right\|_{\mathcal{M}}} \tag{2.80}$$
$$= d_{\mathcal{M}}^K(x,y) \cdot \frac{\mathrm{Proj}_x(y)}{\|\mathrm{Proj}_x(y)\|_{\mathcal{M}}}. \tag{2.81}$$

Proof. We only verify that exp, log are well-defined.

(1) *Show that* $\exp_x^K(v) \in \mathcal{H}^{n,K}$ *for all* $(x,v) \in \mathrm{T}\mathcal{H}^{n,K}$. Define two numbers a, b as below:

$$\exp_x^K(v) = \underbrace{\cosh\left(\frac{\|v\|_{\mathcal{M}}}{\sqrt{K}}\right)}_{a} \cdot x + \underbrace{\sqrt{K}\sinh\left(\frac{\|v\|_{\mathcal{M}}}{\sqrt{K}}\right)\frac{1}{\|v\|_{\mathcal{M}}}}_{b} \cdot v, \tag{2.82}$$

By introducing $t := \frac{\|v\|_{\mathcal{M}}}{\sqrt{K}}$, we have

$$\langle \exp_x^K(v), \exp_x^K(v)\rangle_{\mathcal{M}} = \langle ax + bv, ax + bv\rangle_{\mathcal{M}} \tag{2.83}$$
$$= \cosh^2(t)\langle x,x\rangle_{\mathcal{M}} + K\sinh^2(t)\frac{1}{\|v\|_{\mathcal{M}}^2}\langle v,v\rangle_{\mathcal{M}} + 2ab\langle x,v\rangle_{\mathcal{M}} \tag{2.84}$$
$$= \cosh^2(t)\langle x,x\rangle_{\mathcal{M}} + K\sinh^2(t)\frac{1}{\|v\|_{\mathcal{M}}^2}\langle v,v\rangle_{\mathcal{M}} \tag{2.85}$$
$$= \cosh^2(t)\langle x,x\rangle_{\mathcal{M}} + K\sinh^2(t) \tag{2.86}$$
$$= -K\left(\cosh^2(t) - \sinh^2(t)\right) \tag{2.87}$$
$$= -K. \tag{2.88}$$

Thus, $\exp_x^K(v) \in \mathcal{H}^{n,K}$.

(2) *Show that* $\log_x^K\left[\exp_x^K(v)\right] = v$ *for all* $v \in \mathrm{T}_x\mathcal{H}^{n,K}$. Define $y := \exp_x^K(v)$, namely, $y := ax + bv$. Then we have the following results in turn:

$$\langle x,y\rangle_{\mathcal{M}} = \langle x, ax+bv\rangle_{\mathcal{M}} = a\langle x,x\rangle_{\mathcal{M}} = -aK, \tag{2.89}$$

and

$$d_{\mathcal{M}}^{K}(x,y) = \sqrt{K}\,\operatorname{arcosh}\left(-\langle x,y\rangle_{\mathcal{M}}/K\right) \tag{2.90}$$
$$= \sqrt{K}\,\operatorname{arcosh}\left(a\right) \tag{2.91}$$
$$= \sqrt{K}\,\operatorname{arcosh}\left[\cosh\left(\frac{\|v\|_{\mathcal{M}}}{\sqrt{K}}\right)\right] = \|v\|_{\mathcal{M}}, \tag{2.92}$$

and

$$\operatorname{Proj}_x(y) = \operatorname{Proj}_x(ax) + \operatorname{Proj}_x(bv) = \operatorname{Proj}_x(bv) = bv. \tag{2.93}$$

Hence,

$$\log_x^K[y] = d_{\mathcal{M}}^{K}(x,y)\cdot\frac{\operatorname{Proj}_x(y)}{\|\operatorname{Proj}_x(y)\|_{\mathcal{M}}} \tag{2.94}$$
$$= \|v\|_{\mathcal{M}}\frac{bv}{\|bv\|_{\mathcal{M}}} \tag{2.95}$$
$$= \|v\|_{\mathcal{M}}\frac{v}{\|v\|_{\mathcal{M}}}\ \text{(since } b > 0) \tag{2.96}$$
$$= v. \tag{2.97}$$

This shows that $\log_x^K \circ \exp_x^K$ is the identity map on $\mathrm{T}_x\mathcal{H}^{n,K}$.

(3) *Show that* $\exp_x^K\left[\log_x^K(y)\right] = y$ *for all* $y \in \mathcal{H}^{n,K}$. Define $v := \log_x^K(y)$. Then we have the following results in turn:

$$\|v\|_{\mathcal{M}} = \left\| d_{\mathcal{M}}^{K}(x,y)\cdot\frac{\operatorname{Proj}_x(y)}{\|\operatorname{Proj}_x(y)\|_{\mathcal{M}}}\right\|_{\mathcal{M}} \tag{2.98}$$
$$= d_{\mathcal{M}}^{K}(x,y) \tag{2.99}$$
$$= \sqrt{K}\,\operatorname{arcosh}\left(-\langle x,y\rangle_{\mathcal{M}}/K\right), \tag{2.100}$$

and then

$$\frac{\|v\|_{\mathcal{M}}}{\sqrt{K}} = \operatorname{arcosh}\left(-\langle x,y\rangle_{\mathcal{M}}/K\right), \tag{2.101}$$
$$\cosh\left(\frac{\|v\|_{\mathcal{M}}}{\sqrt{K}}\right) = -\langle x,y\rangle_{\mathcal{M}}/K, \tag{2.102}$$

and

$$\|\operatorname{Proj}_x(y)\|_{\mathcal{M}}^2 = \left\langle y + \frac{1}{K}\langle x,y\rangle_{\mathcal{M}}\cdot x,\ y + \frac{1}{K}\langle x,y\rangle_{\mathcal{M}}\cdot x\right\rangle_{\mathcal{M}} \tag{2.103}$$
$$= \langle y,y\rangle_{\mathcal{M}} + \frac{1}{K^2}\langle x,y\rangle_{\mathcal{M}}^2\langle x,x\rangle_{\mathcal{M}} + \frac{2}{K}\langle x,y\rangle_{\mathcal{M}}^2 \tag{2.104}$$
$$= -K - \frac{1}{K}\langle x,y\rangle_{\mathcal{M}}^2 + \frac{2}{K}\langle x,y\rangle_{\mathcal{M}}^2 \tag{2.105}$$
$$= \frac{1}{K}\langle x,y\rangle_{\mathcal{M}}^2 - K, \tag{2.106}$$

and

$$\sqrt{K}\sinh\left(\frac{\|v\|_{\mathcal{M}}}{\sqrt{K}}\right)\cdot\frac{v}{\|v\|_{\mathcal{M}}} = \sqrt{K}\sinh\left(\frac{\|v\|_{\mathcal{M}}}{\sqrt{K}}\right)\cdot\frac{\operatorname{Proj}_x(y)}{\|\operatorname{Proj}_x(y)\|_{\mathcal{M}}} \tag{2.107}$$

$$= \sqrt{K}\sinh\left(\operatorname{arcosh}\left(-\langle x,y\rangle_{\mathcal{M}}/K\right)\right)\cdot\frac{\operatorname{Proj}_x(y)}{\|\operatorname{Proj}_x(y)\|_{\mathcal{M}}} \tag{2.108}$$

$$= \sqrt{K}\sqrt{(\langle x,y\rangle_{\mathcal{M}}/K)^2-1}\cdot\frac{\operatorname{Proj}_x(y)}{\|\operatorname{Proj}_x(y)\|_{\mathcal{M}}} \tag{2.109}$$

$$= \operatorname{Proj}_x(y). \tag{2.110}$$

Finally, we have

$$\exp_x^K[v] = \cosh\left(\frac{\|v\|_{\mathcal{M}}}{\sqrt{K}}\right)\cdot x + \sqrt{K}\sinh\left(\frac{\|v\|_{\mathcal{M}}}{\sqrt{K}}\right)\cdot\frac{v}{\|v\|_{\mathcal{M}}} \tag{2.111}$$

$$= -\langle x,y\rangle_{\mathcal{M}}/K\cdot x + \operatorname{Proj}_x(y) \tag{2.112}$$

$$= y. \tag{2.113}$$

This shows that $\exp_x^K \circ \log_x^K =$ is the identity map on $\mathcal{H}^{n,K}$. □

Example 2 (Mapping from Euclidean to hyperbolic spaces [CYRL19])**.** Let $\mathrm{x}^E \in \mathbb{R}^n$ denote input Euclidean features. Let $o := (\sqrt{K},0,0,\ldots,0)$ denote the north pole in $\mathcal{H}^{n,K}$, which we use as a reference point to perform tangent space operations. We interpret $\left(0,x^E\right)$ as a point in $\mathrm{T}_o\mathcal{H}^{n,K}$ and have

$$x^H := \exp_o^K\left(\left(0,x^E\right)\right) \tag{2.114}$$

$$= \cosh\left(\frac{\|\left(0,x^E\right)\|_{\mathcal{M}}}{\sqrt{K}}\right)\cdot o + \sqrt{K}\sinh\left(\frac{\|\left(0,x^E\right)\|_{\mathcal{M}}}{\sqrt{K}}\right)\cdot\frac{\left(0,x^E\right)}{\|\left(0,x^E\right)\|_{\mathcal{M}}} \tag{2.115}$$

$$= \cosh\left(\frac{\|x^E\|_2}{\sqrt{K}}\right)\cdot o + \sqrt{K}\sinh\left(\frac{\|x^E\|_2}{\sqrt{K}}\right)\cdot\frac{\left(0,x^E\right)}{\|x^E\|_2} \tag{2.116}$$

$$= \left(\sqrt{K}\cosh\left(\frac{\|x^E\|_2}{\sqrt{K}}\right), \sqrt{K}\sinh\left(\frac{\|x^E\|_2}{\sqrt{K}}\right)\cdot\frac{x^E}{\|x^E\|_2}\right). \tag{2.117}$$

Foe the last equality, notice the position of the zero elements in o and $\left(0,x^E\right)$ as vectors of $\mathbb{R}^n$.

Proposition 9: Riemannian Connection ∇ of $\mathcal{H}^{n,K}$

For all smooth vector fields V on $\mathcal{H}^{n,K}$ and all $(x,u)\in \mathrm{T}\mathcal{H}^{n,K}$, define the operator ∇ as

$$\nabla_u V := \operatorname{Proj}_x(\mathrm{D}\bar{V}(x)[u]) \tag{2.118}$$

where $\bar{V}$ is any smooth extension of V to a neighborhood of $\mathcal{H}^{n,K}$ in $\mathbb{R}^{n+1}$ and $\mathrm{D}\bar{V}(x)[u]$ is the usual directional derivative. It is an exercise to check that ∇ is the Riemannian connection for $\mathcal{M}$.

It is instructive to compare this with [Bou23, Theorem 5.9] where we make the same claim under the assumption that the embedding space is Euclidean. Here, the embedding space is not Euclidean, but the result stands. Again, see [O'n83] for a general treatment.

Definition 5: Covariant Derivative $\frac{\mathrm{D}}{\mathrm{d}t}$ of $\mathcal{H}^{n,K}$ (induced by ∇)

The covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$ (induced by ∇) for a smooth vector field Z along a smooth curve $c : I \to \mathcal{H}^{n,K}$ is given by

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t) = \mathrm{Proj}_{c(t)} \left(\frac{\mathrm{d}}{\mathrm{d}t} Z(t) \right) \tag{2.119}$$

where $\frac{\mathrm{d}}{\mathrm{d}t} Z(t)$ is the usual derivative of Z understood as a map from I to $\mathbb{R}^{n+1}$ this makes use of the fact that $Z(t) \in \mathrm{T}_{c(t)}\mathcal{H}^{n,K} \subset \mathbb{R}^{n+1}$. Compare this with [Bou23, Proposition 5.31].

We proceed to construct a formula for the Hessian of a function on $\mathcal{H}^{n,K}$ based on the gradient and Hessian of a smooth extension.

Proposition 10

The Hessian of f is related to that of $\bar{f}$ as follows:

$$\mathrm{Hess}\, f(x)[u] = \mathrm{Proj}_x(J\, \mathrm{ehess}\, \bar{f}(x)[u]) + \langle x, J\, \mathrm{egrad}\, \bar{f}(x) \rangle_{\mathcal{M}} \cdot u, \tag{2.120}$$

where $J = \mathrm{diag}(-1, 1, \ldots, 1)$ and Proj_x is defined by (2.34).

Proof. Let $\bar{G}$ be any smooth extension of $\mathrm{grad}\, f$ to a neighborhood of $\mathcal{H}^{n,K}$ in $\mathbb{R}^{n+1}$. Then,

$$\bar{G}(x) = \mathrm{Proj}_x(J\, \mathrm{egrad}\, \bar{f}(x)) = J\, \mathrm{egrad}\, \bar{f}(x) + \frac{1}{K} \langle J\, \mathrm{egrad}\, \bar{f}(x), x \rangle_{\mathcal{M}} \cdot x. \tag{2.121}$$

Thus, for all $(x, u) \in \mathrm{T}\mathcal{M}$ we have

$$\begin{aligned} \mathrm{Hess}\, f(x)[u] &= \nabla_u\, \mathrm{grad}\, f && (2.122) \\ &= \mathrm{Proj}_x(\mathrm{D}\bar{G}(x)[u]) && (2.123) \\ &= \mathrm{Proj}_x \left(J\, \mathrm{ehess}\, \bar{f}(x)[u] + qx + \langle J\, \mathrm{egrad}\, \bar{f}(x), x \rangle_{\mathcal{M}} \cdot u \right) && (2.124) \\ &= \mathrm{Proj}_x(J\, \mathrm{ehess}\, \bar{f}(x)[u]) + \langle J\, \mathrm{egrad}\, \bar{f}(x), x \rangle_{\mathcal{M}} \cdot u, && (2.125) \end{aligned}$$

where q is the derivative of $\langle J\, \mathrm{egrad}\, \bar{f}(x), x \rangle_{\mathcal{M}}$ at x along u — and we do not need to compute it since qx is in the normal space, hence it vanishes through the projector. □

Proposition 11: Parallel Transport

If two points x and y on the hyperboloid $\mathcal{H}^{d,1}$ are connected by a geodesic, then the parallel transport of a tangent vector $v \in \mathrm{T}_x\mathcal{H}^{d,1}$ to the tangent space $\mathrm{T}_y\mathcal{H}^{d,1}$ is:

$$P_{x \to y}(v) = v - \frac{\langle \log_x(y), v \rangle_{\mathcal{L}}}{d_{\mathcal{L}}^1(x, y)^2} \left(\log_x(y) + \log_y(x) \right) \tag{2.126}$$

Definition 6: Projections

Finally, we recall projections to the hyperboloid manifold and its corresponding tangent spaces. A

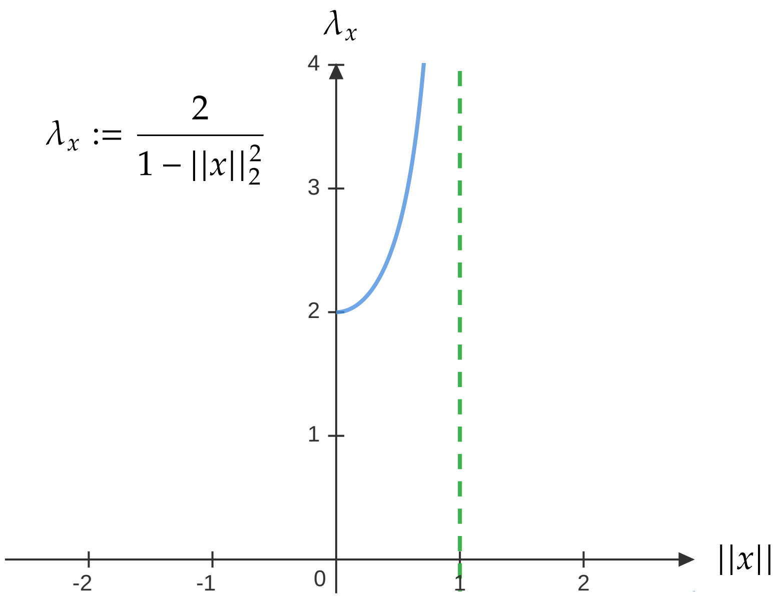

Figure 2.4

point $x = (x_0, x_{1:d}) \in \mathbb{R}^{d+1}$ can be projected on the hyperboloid manifold $\mathcal{H}^{d,1}$ with:

$$\Pi_{\mathbb{R}^{d+1} \to \mathcal{H}^{d,1}}(x) := \left(\sqrt{1 + \|x_{1:d}\|_2^2}, x_{1:d} \right). \tag{2.127}$$

2.2 Geometry of Poincaré ball model

Let $\|.\|_2$ be the Euclidean norm.

Definition 7: Poincaré ball model

The Poincaré ball model with unit radius and constant negative curvature -1 in d dimensions is the Riemannian manifold $\left(\mathbb{D}^{d,1}, (g_x)_x\right)$ where

$$\mathbb{D}^{d,1} := \left\{ x \in \mathbb{R}^d : \|x\|^2 < 1 \right\}, \tag{2.128}$$

and

$$g_x = \lambda_x^2 I_d \tag{2.129}$$

where $\lambda_x := \frac{2}{1-\|x\|_2^2}$ and I_d is the identity matrix. This means that Riemannian metric $\langle u, v \rangle := u^T g_x v$.

Proposition 12

The induced distance between two points (x, y) in $\mathbb{D}^{d,1}$ can be computed as:

$$d_{\mathbb{D}}^1(x, y) = \operatorname{arcosh}\left(1 + 2\frac{\|x - y\|_2^2}{\left(1 - \|x\|_2^2\right)\left(1 - \|y\|_2^2\right)} \right). \tag{2.130}$$

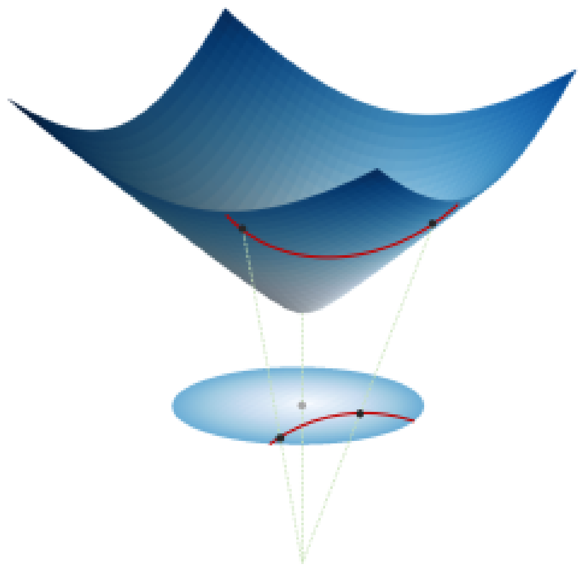

Figure 2.5

2.3 Connection between the Poincaré ball model and the hyperboloid model

While the hyperboloid model tends to be more stable for optimization than the Poincaré model [30], the Poincaré model is very interpretable and embeddings can be directly visualized on the Poincaré disk. Fortunately, these two models are isomorphic (cf. Figure 7) and there exist a diffeomorphism $\Pi_{\mathcal{H}^{d,1}\to \mathrm{D}^{d,1}}(\cdot)$ mapping one space onto the other:

$$\begin{aligned} \Pi_{\mathcal{H}^{d,1}\to\mathbb{D}^{d,1}}(x_0,\ldots,x_d) &= \frac{(x_1,\ldots,x_d)}{x_0+1} \\ \text{and } \Pi_{\mathbb{D}^{d,1}\to\mathcal{H}^{d,1}}(x_1,\ldots,x_d) &= \frac{\left(1+\|x\|_2^2, 2x_1,\ldots,2x_d\right)}{1-\|x\|_2^2}. \end{aligned} \tag{2.131}$$

Chapter 3

Hadamard manifold

See `https://en.wikipedia.org/wiki/Hadamard_manifold#cite_note-Li2102-1`

Definition 8: (Cartan-)Hadamard manifold

In mathematics, a Hadamard manifold, named after Jacques Hadamard - more often called a Cartan-Hadamard manifold, after Élie Cartan — is a Riemannian manifold (M, g) that is complete and simply connected and has everywhere non-positive sectional curvature.

By Cartan–Hadamard theorem, all Cartan-Hadamard manifolds are diffeomorphic to the Euclidean space $\mathbb{R}^n$.

Furthermore it follows from the Hopf-Rinow theorem that every pairs of points in a Cartan-Hadamard manifold may be connected by a unique geodesic segment.

Thus Cartan-Hadamard manifolds are some of the closest relatives of $\mathbb{R}^n$.

Example 3. The Euclidean space $\mathbb{R}^n$ with its usual metric is a Cartan-Hadamard manifold with constant sectional curvature equal to 0.

Standard n-dimensional hyperbolic space $\mathbb{H}^n$ is a Cartan-Hadamard manifold with constant sectional curvature equal to -1.

Proposition 13

In Cartan-Hadamard manifolds, the map $\exp_p : \mathrm{T}M_p \to M$ is a diffeomorphism for all $p \in M$.

Chapter 4

hyperbolicfactory

Listing 4.1: Matlab example

```
function M = hyperbolicfactory(n, m, transposed)
% Factory for matrices whose columns live on the hyperbolic manifold
%
% function M = hyperbolicfactory(n)
% function M = hyperbolicfactory(n, m)
% function M = hyperbolicfactory(n, m, transposed)
%
% Returns a structure M which describes the hyperbolic manifold in
    ↪ Manopt.
% A point on the manifold is a matrix X of size (n+1)-by-m whose
    ↪ columns
% live on the hyperbolic manifold, that is, for each column x of X,
    ↪ we have
%
% -x(1)^2 + x(2)^2 + x(3)^2 + ... + x(n+1)^2 = -1.
%
% The positive branch is selected by M.rand(), that is, x(1) > 0, but
    ↪  all
% tools work on the negative branch as well.
%
% Equivalently, defining the Minkowski (semi) inner product
%
% <x, y> = -x(1)y(1) + x(2)y(2) + x(3)y(3) + ... + x(n+1)y(n+1)
%
% and the induced Minkowski (semi) norm ||x||^2 = <x, x>, we can
    ↪ write
% compactly that each column of X has squared Minkowski norm equal to
    ↪  -1.
%
% The set of matrices X that satisfy this constraint is a smooth
    ↪ manifold.
% Tangent vectors at X are matrices U of the same size as X. If x and
    ↪  u are
```

```
% the kth columns of X and U respectively, then <x, u> = 0.
%
% This manifold is turned into a Riemannian manifold by restricting
    ↪ the
% Minkowski inner product to each tangent space (a simple calculation
% confirms that this metric is indeed Riemannian and not just semi
% Riemannian, that is, it is positive definite when restricted to
    ↪ each
% tangent space). This is the hyperbolic manifold: for m = 1, all of
    ↪ its
% sectional curvatures are equal to -1. This is called the
    ↪ hyperboloid or
% the Lorentz geometry.
%
% This manifold is an embedded submanifold of Euclidean space (the
    ↪ set of
% matrices of size (n+1)-by-m equipped with the usual trace inner
    ↪ product).
% Thus, when defining the Euclidean gradient for example (problem.
    ↪ egrad),
% it should be specified as if the function were defined in Euclidean
    ↪  space
% directly. The tool M.egrad2rgrad will automatically convert that
    ↪ gradient
% to the correct Riemannian gradient, as needed to satisfy the metric
    ↪ . The
% same is true for the Euclidean Hessian and other tools that
    ↪ manipulate
% elements in the embedding space.
%
% Importantly, the resulting manifold is /not/ a Riemannian
    ↪ submanifold of
% Euclidean space, because its metric is not obtained simply by
    ↪ restricting
% the Euclidean metric to the tangent spaces. However, it is a
% semi-Riemannian submanifold of Minkowski space, that is, the set of
% matrices of size (n+1)-by-m equipped with the Minkowski inner
    ↪ product.
% Minkowski space itself can be seen as a (linear) semi-Riemannian
    ↪ manifold
% embedded in Euclidean space. This view is entirely equivalent to
    ↪ the one
% described above (the Riemannian structure of the resulting manifold
    ↪  is
% exactly the same), and it is useful to derive some of the tools
    ↪ this
% factory provides.
%
```

```
% If transposed is set to true (it is false by default), then the
    ↪ matrices
% are transposed: a point X on the manifold is a matrix of size m-by
    ↪ -(n+1)
% and each row is an element in hyperbolic space. It is the same
    ↪ geometry,
% just a different representation.
%
%
% Resources:
%
% 1. Nickel and Kiela, "Learning Continuous Hierarchies in the
    ↪ Lorentz
% Model of Hyperbolic Geometry", ICML, 2018.
%
% 2. Wilson and Leimeister, "Gradient descent in hyperbolic space",
% arXiv preprint arXiv:1805.08207 (2018).
%
% 3. Pennec, "Hessian of the Riemannian squared distance", HAL INRIA,
    ↪  2017.
%
% Ported primarily from the McTorch toolbox at
% https://github.com/mctorch/mctorch.
%
% See also: poincareballfactory spherefactory obliquefactory
    ↪ obliquecomplexfactory


% This file is part of Manopt: www.manopt.org.
% Original authors: Bamdev Mishra <bamdevm@gmail.com>, Mayank
    ↪ Meghwanshi,
% Pratik Jawanpuria, Anoop Kunchukuttan, and Hiroyuki Kasai Oct 28,
    ↪ 2018.
% Contributors: Nicolas Boumal
% Change log:
% May 14, 2020 (NB):
% Clarified comments about distance computation.
% July 13, 2020 (NB):
% Added pairmean function.
% Sep. 24, 2023 (NB):
% Edited out bsxfun() for improved speed.

% Design note: all functions that are defined here but not exposed
% outside work for non-transposed representations. Only the wrappers
% that eventually expose functionalities handle transposition. This
% makes it easier to compose functions internally.

if ~exist('m', 'var') || isempty(m)
```

```
m = 1;
end

if ~exist('transposed', 'var') || isempty(transposed)
transposed = false;
end

if transposed
trnsp = @(X) X';
trnspstr = ', transposed';
else
trnsp = @(X) X;
trnspstr = '';
end

M.name = @() sprintf('Hyperbolic manifold H(%d, %d)%s', n, m,
    ↪ trnspstr);

M.dim = @() n*m;

M.typicaldist = @() sqrt(n*m);

% Returns a row vector q such that q(k) is the Minkowski inner
    ↪ product
% of columns U(:, k) and V(:, k). This is defined in all of Minkowski
% space, not only on tangent spaces. In particular, if X is a point
    ↪ on
% the manifold, then inner_\mathcal{M}inkowski_columns(X, X) should
    ↪ return a
% vector of all -1's.
function q = inner_\mathcal{M}inkowski_columns(U, V)
q = -U(1, :).*V(1, :) + sum(U(2:end, :).*V(2:end, :), 1);
end

% Riemannian metric: we sum over the m copies of the hyperbolic
% manifold, each equipped with a restriction of the Minkowski metric.
M.inner = @(X, U, V) sum(inner_\mathcal{M}inkowski_columns(trnsp(U),
    ↪ trnsp(V)));

% Mathematically, the Riemannian metric is positive definite, hence
% M.inner always returns a nonnegative number when U is tangent at X.
% Numerically, because the inner product involves a difference of
% positive numbers, round-off may result in a small negative number.
% Taking the max against 0 avoids imaginary results.
M.norm = @(X, U) sqrt(max(M.inner(X, U, U), 0));

M.dist = @(X, Y) norm(dists(trnsp(X), trnsp(Y)));
% This function returns a row vector of length m such that d(k) is
```

```
    ↪ the
% geodesic distance between X(:, k) and Y(:, k).
function d = dists(X, Y)
% Mathematically, each column of U = X-Y has nonnegative squared
% Minkowski norm. To avoid potentially imaginary results due to
% round-off errors, we take the max against 0.
U = X-Y;
mink_sqnorms = max(0, inner_\mathcal{M}inkowski_columns(U, U));
mink_norms = sqrt(mink_sqnorms);
d = 2*asinh(.5*mink_norms);
% The formula above is equivalent to
% d = max(0, real(acosh(-inner_\mathcal{M}inkowski_columns(X, Y))));
% but is numerically more accurate when distances are small.
% When distances are large, it is better to use the acosh formula.
end

M.proj = @(X, U) trnsp(projection(trnsp(X), trnsp(U)));
function PU = projection(X, U)
inners = inner_\mathcal{M}inkowski_columns(X, U);
PU = U + X .* inners;
end

M.tangent = M.proj;

% For Riemannian submanifolds, converting the Euclidean gradient into
% the Riemannian gradient amounts to an orthogonal projection. Here
% however, the manifold is not a Riemannian submanifold of Euclidean
% space, hence extra corrections are required to account for the
    ↪ change
% of metric.
M.egrad2rgrad = @(X, egrad) trnsp(egrad2rgrad(trnsp(X), trnsp(egrad))
    ↪ );
function rgrad = egrad2rgrad(X, egrad)
egrad(1, :) = -egrad(1, :);
rgrad = projection(X, egrad);
end

M.ehess2rhess = @(X, egrad, ehess, U) ...
trnsp(ehess2rhess(trnsp(X), trnsp(egrad), trnsp(ehess), trnsp(U)));
function rhess = ehess2rhess(X, egrad, ehess, U)
egrad(1, :) = -egrad(1, :);
ehess(1, :) = -ehess(1, :);
inners = inner_\mathcal{M}inkowski_columns(X, egrad);
rhess = projection(X, ehess + U .* inners);
end

% For the exponential, we cannot separate trnsp() nicely from the
    ↪ main
```

```
% function because the third input, t, is optional.
M.exp = @exponential;
function Y = exponential(X, U, t)
X = trnsp(X);
U = trnsp(U);

if nargin < 3
tU = U; % corresponds to t = 1
else
tU = t*U;
end

% Compute the individual Minkowski norms of the columns of U.
mink_inners = inner_\mathcal{M}inkowski_columns(tU, tU);
mink_norms = sqrt(max(0, mink_inners));

% Coefficients for the exponential. For b, note that NaN's appear
% when an element of mink_norms is zero, in which case the correct
% convention is to define sinh(0)/0 = 1.
a = cosh(mink_norms);
b = sinh(mink_norms)./mink_norms;
b(isnan(b)) = 1;

Y = X .* a + tU .* b;

Y = trnsp(Y);
end

M.retr = M.exp;

M.log = @(X, Y) trnsp(logarithm(trnsp(X), trnsp(Y)));
function U = logarithm(X, Y)
d = dists(X, Y);
a = d./sinh(d);
a(isnan(a)) = 1;
U = projection(X, Y .* a);
end

M.hash = @(X) ['z' hashmd5(X(:))];

M.rand = @() trnsp(myrand());
function X = myrand()
X1 = randn(n, m);
x0 = sqrt(1 + sum(X1.^2, 1)); % selects positive branch
X = [x0; X1];
end

M.normalize = @(X, U) U / M.norm(X, U);
```

```
M.randvec = @(X) M.normalize(X, M.proj(X, randn(size(X))));

M.lincomb = @matrixlincomb;

M.zerovec = @(X) zeros(size(X));

M.transp = @(X1, X2, U) M.proj(X2, U);

M.isotransp = @(X1, X2, U) ...
trnsp(parallel_transport(trnsp(X1), trnsp(X2), trnsp(U)));
function V = parallel_transport(X1, X2, U)
V = inner_\mathcal{M}inkowski_columns(X2, U);
V = V ./ (1 - inner_\mathcal{M}inkowski_columns(X1, X2)) .* (X1 + X2)
   ↪ ;
V = U + V;
end

M.pairmean = @(x1, x2) M.exp(x1, M.log(x1, x2), .5);

% vec returns a vector representation of an input tangent vector
   ↪ which
% is represented as a matrix; mat returns the original matrix
% representation of the input vector representation of a tangent
% vector; vec and mat are thus inverse of each other.
vect = @(X) X(:);
M.vec = @(X, U_\mathcal{M}at) vect(trnsp(U_\mathcal{M}at));
M.mat = @(X, U_vec) trnsp(reshape(U_vec, [n+1, m]));
M.vecmatareisometries = @() false;

end
```

Bibliography

[Axl15] Sheldon Axler. *Linear Algebra Done Right*. Undergraduate Texts in Mathematics. Springer International Publishing, Cham, 2015.

[Bou23] Nicolas Boumal. *An introduction to optimization on smooth manifolds*. Cambridge University Press, 1 edition, March 2023.

[CYRL19] Ines Chami, Zhitao Ying, Christopher Ré, and Jure Leskovec. Hyperbolic graph convolutional neural networks. *Advances in neural information processing systems*, 32, 2019.

[Lee18] John M Lee. *Introduction to Riemannian manifolds*, volume 2. Springer, 2018.

[O'n83] Barrett O'neill. *Semi-Riemannian geometry with applications to relativity*. Academic press, 1983.