Journal of Inequalities in Pure and Applied Mathematics

http://jipam.vu.edu.au/

Volume 5, Issue 3, Article 54, 2004

# A METRIC INEQUALITY FOR THE THOMPSON AND HILBERT GEOMETRIES

ROGER D. NUSSBAUM AND CORMAC WALSH

MATHEMATICS DEPARTMENT
RUTGERS UNIVERSITY
NEW BRUNSWICK
NJ 08903.
nussbaum@math.rutgers.edu

INRIA ROCQUENCOURT
B.P. 105,
78153 LE CHESNAY CEDEX
FRANCE.
cormac.walsh@inria.fr

*Received 25 September, 2003; accepted 02 April, 2004*
*Communicated by J.M. Borwein*

ABSTRACT. There are two natural metrics defined on an arbitrary convex cone: Thompson's part metric and Hilbert's projective metric. For both, we establish an inequality giving information about how far the metric is from being non-positively curved.

*Key words and phrases:* Hilbert geometry, Thompson's part metric, Cone metric, Non-positive curvature, Finsler space.

2000 *Mathematics Subject Classification.* 53C60.

## 1. INTRODUCTION

Let $C$ be a cone in a vector space $V$. Then $C$ induces a partial ordering on $V$ given by $x \leq y$ if and only if $y - x \in C$. For each $x \in C\backslash\{0\}$, $y \in V$, define $M(y/x) := \inf\{\lambda \in \mathbb{R} : y \leq \lambda x\}$. *Thompson's part metric* on $C$ is defined to be

$$d_T(x, y) := \log \max \left(M\left(x/y\right), M\left(y/x\right)\right)$$

and *Hilbert's projective metric* on $C$ is defined to be

$$d_H(x, y) := \log \left(M\left(x/y\right) M\left(y/x\right)\right).$$

Two points in $C$ are said to be in the same part if the distance between them is finite in the Thompson metric. If $C$ is *almost Archimedean*, then, with respect to this metric, each part of $C$ is a complete metric space. Hilbert's projective metric, however, is only a pseudo-metric: it is possible to find two distinct points which are zero distance apart. Indeed it is not difficult to see

ISSN (electronic): 1443-5756

131-03

that $d_H(x, y) = 0$ if and only if $x = \lambda y$ for some $\lambda > 0$. Thus $d_H$ is a metric on the space of rays of the cone. For further details, see Chapter 1 of the monograph [23].

Suppose $C$ is finite dimensional and let $S$ be a cross section of $C$, that is $S := \{x \in C : l(x) = 1\}$, where $l : V \to \mathbb{R}$ is some positive linear functional with respect to the ordering on $V$. Suppose $x, y \in S$ are distinct. Let $a$ and $b$ be the points in the boundary of $S$ such that $a$, $x$, $y$, and $b$ are collinear and are arranged in this order along the line in which they lie. It can be shown that the Hilbert distance between $x$ and $y$ is then given by the logarithm of the cross ratio of these four points:

$$d_H(x, y) = \log \frac{|bx|\,|ay|}{|by|\,|ax|}.$$

Indeed, this was the original definition of Hilbert. If $S$ is the open unit disk, the Hilbert metric is exactly the Klein model of the hyperbolic plane.

An interesting feature of the two metrics above is that they show many signs of being non-positively curved. For example, when endowed with the Hilbert metric, the Lorentz cone $\{(t, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} : t^2 > x_1^2 + \cdots + x_n^2\}$ is isometric to $n$-dimensional hyperbolic space. At the other extreme, the positive cone $\mathbb{R}^n_+ := \{(x_1, \ldots, x_n) : x_i \geq 0 \text{ for } 1 \leq i \leq n\}$ with either the Thompson or the Hilbert metric is isometric to a normed space [11], which one may think of as being flat. In between, for Hilbert geometries having a strictly-convex $C^2$ boundary with non-vanishing Hessian, the methods of Finsler geometry [28] apply. It is known that such geometries have constant flag curvature $-1$. More general Hilbert geometries were investigated in [17] where a definition was given of a point of positive curvature. It was shown that no Hilbert geometries have such points.

However, there are some notions of non-positive curvature which do not apply. For example, a Hilbert geometry will only be a CAT(0) space (see [6]) if the cone is Lorentzian. Another notion related to negative curvature is that of Gromov hyperbolicity [15]. In [2], a condition is given characterising those Hilbert geometries that are Gromov hyperbolic. This notion has also been investigated in the wider context of uniform Finsler Hadamard manifolds, which includes certain Hilbert geometries [12].

Busemann has defined non-positive curvature for *chord spaces* [7]. These are metric spaces in which there is a distinguished set of geodesics, satisfying certain axioms. In such a space, denote by $m_{xy}$ the midpoint along the distinguished geodesic connecting the pair of points $x$ and $y$. Then the chord space is non-positively curved if, for all points $u$, $x$, and $y$ in the space,

$$d(m_{ux}, m_{uy}) \leq \frac{1}{2} d(x, y), \tag{1.1}$$

where $d$ is the metric.

In the case of the Hilbert and Thompson geometries on a part of a closed cone $C$, there will not necessarily be a unique minimal geodesic connecting each pair of points. However, it is known that, setting $\beta := M(y/x; C)$ and $\alpha := 1/M(x/y; C)$, the curve $\phi : [0, 1] \to C$ :

$$\phi(s; x, y) := \begin{cases} \left(\dfrac{\beta^s - \alpha^s}{\beta - \alpha}\right) y + \left(\dfrac{\beta\alpha^s - \alpha\beta^s}{\beta - \alpha}\right) x, & \text{if } \beta \neq \alpha, \\ \alpha^s x, & \text{if } \beta = \alpha \end{cases} \tag{1.2}$$

is always a minimal geodesic from $x$ to $y$ with respect to both the Thompson and Hilbert metrics. We view these as distinguished geodesics. If the cone $C$ is finite dimensional, then each part of $C$ will be a chord space under both the Thompson and Hilbert metrics. Notice that the geodesics above are projective straight lines. If the cone is strictly convex, these are the only geodesics that are minimal with respect to the Hilbert metric. For Thompson's metric, if two points are in the same part of $C$ and are linearly independent, then there are infinitely many minimal geodesics between them.

In this paper we investigate whether inequalities similar to (1.1) hold for the Hilbert and Thompson geometries with the geodesics given in (1.2). We prove the following two theorems.

**Theorem 1.1.** *Let $C$ be an almost Archimedean cone. Suppose $u, x, y \in C$ are in the same part. Also suppose that $0 < s < 1$ and $R > 0$, and that $d_H(u,x) \leq R$ and $d_H(u,y) \leq R$. If the linear span of $\{u, x, y\}$ is 1- or 2-dimensional, then $d_T(\phi(s;u,x), \phi(s;u,y)) \leq s d_T(x,y)$. In general*

$$d_T\big(\phi(s;u,x), \phi(s;u,y)\big) \leq \left[\frac{2(1-e^{-Rs})}{1-e^{-R}} - s\right] d_T(x,y). \tag{1.3}$$

Note that the bracketed value on the right hand side of this inequality is strictly increasing in $R$. As $R \to 0$, this value goes to $s$, which reflects the fact that in small neighborhoods the Thompson metric looks like a norm. As $R \to \infty$, the bracketed value goes to $2 - s$.

**Theorem 1.2.** *Let $C$ be an almost Archimedean cone. Suppose $u, x, y \in C$ are in the same part. Also suppose that $0 < s < 1$ and $R > 0$ and that $d_H(u,x) \leq R$ and $d_H(u,y) \leq R$. If the linear span of $\{u, x, y\}$ is 1- or 2-dimensional, then $d_H(\phi(s;u,x), \phi(s;u,y)) \leq s d_H(x,y)$. In general*

$$d_H\big(\phi(s;u,x), \phi(s;u,y)\big) \leq \left[\frac{1-e^{-Rs}}{1-e^{-R}}\right] d_H(x,y). \tag{1.4}$$

Again, the bracketed value on the right hand side increases strictly with increasing $R$. This time, it ranges between $s$ as $R \to 0$ and 1 as $R \to \infty$.

Our method of proof will be to first establish the results when $C$ is the positive cone $\mathbb{R}^N_+$, with $N \geq 3$. It will be obvious from the proofs that the bounds given are the best possible in this case. A crucial lemma will state that any finite set of $n$ elements of a Thomson or Hilbert geometry can be isometrically embedded in $\mathbb{R}^{n(n-1)}_+$ with, respectively, its Thompson or Hilbert metric. This lemma will allow us to extend the same bounds to more general cones, although in the general case the bounds may no longer be tight.

A special case of Theorem 1.2 was proved in [29] using a simple geometrical argument. It was shown that if two particles start at the same point and travel along distinct straight-line geodesics at unit speed in the Hilbert metric, then the Hilbert distance between them is strictly increasing. This is equivalent to the special case of Theorem 1.2 when $d_H(u,x) = d_H(u,y)$ and $R$ approaches infinity.

A consequence of Theorems 1.1 and 1.2 is that both the Thompson and Hilbert geometries are semihyperbolic in the sense of Alonso and Bridson [1]. Recall that a metric space is semi-hyperbolic if it admits a bounded quasi-geodesic bicombing. A bicombing is a choice of path between each pair of points. We may use the one given by

$$\zeta_{(x,y)}(t) := \begin{cases} \phi\left(\dfrac{t}{d(x,y)}; x, y\right), & \text{if } t \in [0, d(x,y)] \\ y, & \text{otherwise} \end{cases}$$

for each pair of points $x$ and $y$ in the same part of $C$. Here $d$ is either the Thompson or Hilbert metric. This bicombing is geodesic and hence quasi-geodesic. To say it is bounded means that there exist constants $M$ and $\epsilon$ such that

$$d(\zeta_{(x,y)}(t), \zeta_{(w,z)}(t)) \leq M \max(d(x,w), d(y,z)) + \epsilon$$

for each $x, y, w, z \in C$ and $t \in [0, \infty)$.

**Corollary 1.3.** *Each part of $C$ is semihyperbolic when endowed with either Thompson's part metric or Hilbert's projective metric.*

It should be pointed out that for some cones there are other good choices of distinguished geodesics. For example, for the cone of positive definite symmetric matrices $\mathrm{Sym}(n)$, a natural choice would be $\phi(s; X, Y) := X^{1/2}(X^{-1/2}YX^{-1/2})^s X^{1/2}$ for $X, Y \in \mathrm{Sym}(n)$ and $s \in [0, 1]$. It can be shown that, with this choice, $\mathrm{Sym}(n)$ is non-positively curved in the sense of Busemann under both the Thompson and Hilbert metrics. This result has been generalized to both symmetric cones [16] and to the cone of positive elements of a $C^*$-algebra [10].

Although Hilbert's projective metric arose in geometry, it has also been of great interest to analysts. This is because many naturally occurring maps in analysis, both linear and non-linear, are either non-expansive or contractive with respect to it. Perhaps the first example of this is due to G. Birkhoff [3, 4], who noted that matrices with strictly positive entries (or indeed integral operators with strictly positive kernels) are strict contractions with respect to Hilbert's metric. References to the literature connecting this metric to positive linear operators can be found in [14, 13]. It has also been used to study the spectral radii of elements of Coxeter groups [20]. Both metrics have been applied to questions concerning the convergence of iterates of non-linear operators [8, 16, 23, 24, 25]. The two metrics have been used to solve problems involving non-linear integral equations [27, 30], linear operator equations [8, 9], and ordinary differential equations [5, 25, 31, 32]. Thompson's metric has also been usefully applied in [24, 26] to obtain "DAD theorems", which are scaling results concerning kernels of integral operators. Another application of this metric is in Optimal Filtering [19], while Hilbert's metric has been used in Ergodic Theory [18] and Fractal Diffusions [21].

## 2. PROOFS

A cone is a subset of a (real) vector space that is convex, closed under multiplication by positive scalars, and does not contain any vector subspaces of dimension one. We say that a cone is almost Archimedean if the closure of its restriction to any two-dimensional subspace is also a cone.

The proofs of Theorems 1.1 and 1.2 will involve the use of some infinitesimal arguments. We recall that both the Thompson and Hilbert geometries are *Finsler* spaces [22]. If $C$ is a closed cone in $\mathbb{R}^N$ with non-empty interior, then $\operatorname{int} C$ can be considered to be an $N$-dimensional manifold and its tangent space at each point can be identified with $\mathbb{R}^N$. If a norm

$$|v|_x^T := \inf\{\alpha > 0 : -\alpha x \le v \le \alpha x\}$$

is defined on the tangent space at each point $x \in \operatorname{int} C$, then the length of any piecewise $C^1$ curve $\alpha : [a, b] \to \operatorname{int} C$ can be defined to be

$$L^T(\alpha) := \int_a^b |\alpha'(t)|_{\alpha(t)}^T \, dt.$$

The Thompson distance between any two points is recovered by minimizing over all paths connecting the points:

$$d_T(x, y) = \inf\{L^T(\alpha) : \alpha \in PC^1[x, y]\},$$

where $PC^1[x, y]$ denotes the set of all piecewise $C^1$ paths $\alpha : [0, 1] \to \operatorname{int} C$ with $\alpha(0) = x$ and $\alpha(1) = y$. A similar procedure yields the Hilbert metric when the norm above is replaced by the semi-norm

$$|v|_x^H := M(v/x) - m(v/x).$$

Here $M(v/x)$ is as before and $m(v/x) := \sup\{\lambda \in \mathbb{R} : v \geq \lambda x\}$. The Hilbert geometry will be Riemannian only in the case of the Lorentz cone. The Thompson geometry will be Riemannian only in the trivial case of the one-dimensional cone $\mathbb{R}_+$.

Our strategy will be to first prove the theorems for the case of the positive cone $\mathbb{R}^N_+$, and then extend them to the general case. The proof in the case of $\mathbb{R}^N_+$ will involve investigation of the map $g : \operatorname{int}\mathbb{R}^N_+ \to \operatorname{int}\mathbb{R}^N_+$:

$$g(x) := \phi(s; \mathbb{1}, x) = \begin{cases} \left(\dfrac{b^s - a^s}{b-a}\right) x + \left(\dfrac{ba^s - ab^s}{b-a}\right)\mathbb{1}, & \text{if } b \neq a, \\ a^s \mathbb{1}, & \text{if } b = a, \end{cases} \tag{2.1}$$

where $b := b(x) := \max_i x_i$ and $a := a(x) := \min_i x_i$. Here $s \in (0,1)$ is fixed and we are using the notation $\mathbb{1} := (1, \dots, 1)$. The derivative of $g$ at $x \in \operatorname{int}\mathbb{R}^N_+$ is a linear map from $\mathbb{R}^N \to \mathbb{R}^N$. Taking $|\cdot|^T_x$ as the norm on the domain and $|\cdot|^T_{g(x)}$ as the norm on the range, the norm of $g'(x)$ is

$$||g'(x)||_T := \sup\{|g'(x)(v)|^T_{g(x)} : |v|^T_x \leq 1\}.$$

If, instead, we take the appropriate infinitesimal Hilbert semi-norms on the domain and range, then the norm of $g'(x)$ is given by

$$||g'(x)||_H := \sup\{|g'(x)(v)|^H_{g(x)} : |v|^H_x \leq 1\}.$$

For each pair of distinct integers $I$ and $J$ contained in $\{1, \dots, N\}$, let

$$U_{I,J} := \Big\{x \in \operatorname{int}\mathbb{R}^N_+ : 0 < x_I < x_i < x_J \text{ for all } i \in \{1, \dots, N\}\backslash\{I, J\}\Big\}.$$

On each set $U_{I,J}$, the map $g$ is $C^1$ and is given by the formula

$$g(x) = \left(\frac{x_J^s - x_I^s}{x_J - x_I}\right) x + \left(\frac{x_J x_I^s - x_I x_J^s}{x_J - x_I}\right)\mathbb{1}.$$

Let $U$ denote the union of the sets $U_{I,J}$; $I, J \in \{1, \dots, N\}$, $I \neq J$. If $x \in \mathbb{R}^N_+\backslash U$, then there must exist distinct integers $m, n \in \{1, \dots, N\}$ with either $x_n = x_m = \max_i x_i$ or $x_n = x_m = \min_i x_i$. The set $x \in \mathbb{R}^N_+$ with $x_n = x_m$ has ($N$-dimensional) Lebesgue measure zero, so the complement of $U$ in $\mathbb{R}^N_+$ has Lebesgue measure zero.

We recall the following results from [22]. The first is a combination of Corollaries 1.3 and 1.5 from that paper.

**Proposition 2.1.** *Let $C$ be a closed cone with non-empty interior in a finite dimensional normed space $V$. Suppose $G$ is an open subset of $\operatorname{int} C$ such that $\phi(s; x, y) \in G$ for all $x, y \in G$ and $s \in [0,1]$. Suppose also that $f : G \to \operatorname{int} C$ is a locally Lipschitzian map with respect to the norm on $V$. Then*

$$\inf\{k \geq 0 : d_T(f(x), f(y)) \leq k d_T(x,y) \text{ for all } x, y \in G\} = \operatorname*{ess\,sup}_{x \in G} ||f'(x)||_T.$$

It is useful in this context to recall that every locally Lipschitzian map is Fréchet differentiable Lebesgue almost everywhere. The next proposition is a special case of Theorem 2.5 in [22].

**Proposition 2.2.** *Let $C$ be a closed cone with non-empty interior in a normed space $V$ of finite dimension $N$. Let $l$ be a linear functional on $V$ such that $l(x) > 0$ for all $x \in \operatorname{int} C$, and define $S := \{x \in C : l(x) = 1\}$. Let $G$ be a relatively-open convex subset of $S$. Suppose that $f : G \to \operatorname{int} C$ is a locally Lipschitzian map with respect to the norm on $V$. Then*

$$\inf\{k \geq 0 : d_H(f(x), f(y)) \leq k d_H(x,y) \text{ for all } x, y \in G\} = \operatorname*{ess\,sup}_{x \in G} ||f'(x)||_{\tilde{H}},$$

*where* $||f'(x)||_{\tilde{H}} := \sup\{|f'(x)(v)|^H_{f(x)} : |v|^H_x \leq 1, l(v) = 0\}$. *Here the essential supremum is taken with respect to the* $N-1$*-dimensional Lebesgue measure on* $S$.

Since we wish to apply Propositions 2.1 and 2.2 to the map $g$, we must prove that it is locally Lipschitzian.

**Lemma 2.3.** *The map* $g : \operatorname{int}(\mathbb{R}^N_+) \to \operatorname{int}(\mathbb{R}^N_+)$ *defined by (2.1) is locally Lipschitzian.*

*Proof.* We use the supremum norm $||x||_\infty := \max_i |x_i|$ on $\mathbb{R}^N$. Clearly, $|b(x) - b(y)| \leq ||x - y||_\infty$ and $|a(x) - a(y)| \leq ||x - y||_\infty$ for all $x, y \in \operatorname{int}(\mathbb{R}^N_+)$. Therefore both $a$ and $b$ are Lipschitzian with Lipschitz constant 1.

Let $\gamma : [0, \infty) \to [0, \infty)$ be defined by

$$\gamma(t) := \begin{cases} \dfrac{t^s - 1}{t - 1}, & \text{for } t \neq 1, \\ s, & \text{for } t = 1. \end{cases}$$

Then $g$ may be expressed as

$$g(x) = a^{s-1}\gamma(b/a)\, x + a^s\Big(1 - \gamma(b/a)\Big)\mathbb{1}.$$

The Binomial Theorem gives that

$$\gamma(t) = \sum_{k=1}^{\infty} \binom{s}{k} (t-1)^k \qquad \text{for } |t-1| < 1$$

and so $\gamma$ is $C^\infty$ on a neighborhood of 1. Hence it is $C^\infty$ on $[0, \infty)$, and thus locally Lipschitzian. It follows that $g$ is also locally Lipschitzian. □

2.1. **Thompson's Metric.** We have the following bound on the norm of $g'(x)$ with respect to the Thompson metric.

**Lemma 2.4.** *Consider the Thompson metric on* $\operatorname{int}\mathbb{R}^N_+$. *Let* $x \in U_{1,N}$. *If* $N = 1$ *or* $N = 2$ *then the norm of* $g'$ *at* $x$ *is given by* $||g'(x)||_T = s$. *If* $N \geq 3$, *then*

$$||g'(x)||_T = \frac{x_N - x_{N-1}}{x_N - x_1}\theta\left(\frac{x_N}{x_1}\right)\frac{x_1^{s+1}}{E_{N-1}} + \frac{(x_N^s - x_1^s)x_{N-1}}{E_{N-1}} + \frac{x_{N-1} - x_1}{x_N - x_1}\theta\left(\frac{x_1}{x_N}\right)\frac{x_N^{s+1}}{E_{N-1}} \tag{2.2}$$

*where* $\theta(t) := (1-s) - t^s + st$ *and* $E_i(x) := E_i := x_i(x_N^s - x_1^s) + x_N x_1^s - x_1 x_N^s$.

*Proof.* If $N = 1$ and $x > 0$, then $g(x) = x^s$. We leave the proof in this case to the reader and assume that $N \geq 2$.

For $x \in U_{1,N}$,

$$g(x) = \left(\frac{x_N^s - x_1^s}{x_N - x_1}\right) x + \left(\frac{x_N x_1^s - x_1 x_N^s}{x_N - x_1}\right)\mathbb{1}.$$

Let

$$h_{ij}(x) := \frac{x_j}{g_i(x)}\frac{\partial g_i}{\partial x_j}(x).$$

Straightforward calculation gives, for each $j \in \{1, \dots, N\}$,

$$h_{1j}(x) = s\delta_{1j}$$

$$\text{and} \qquad h_{Nj}(x) = s\delta_{Nj}.$$

Here $\delta_{ij}$ is the Kronecker delta function which takes the value $1$ if $i = j$ and the value $0$ if $i \neq j$. Clearly, $h_{ij}(x) = 0$ for $1 < i < N$ and $j \notin \{1, i, N\}$. For $1 < i < N$,

$$h_{i1}(x) = \tfrac{x_N - x_i}{x_N - x_1}\, \theta\left(\tfrac{x_N}{x_1}\right) \tfrac{x_1^{s+1}}{E_i} \geq 0, \tag{2.3}$$

$$h_{ii}(x) = \tfrac{x_N^s - x_1^s}{E_i} x_i \geq 0, \tag{2.4}$$

$$h_{iN}(x) = -\tfrac{x_i - x_1}{x_N - x_1}\, \theta\left(\tfrac{x_1}{x_N}\right) \tfrac{x_N^{s+1}}{E_i} \leq 0. \tag{2.5}$$

Inequalities (2.3) – (2.5) rely on the fact that $\theta(t) \geq 0$ for $t \geq 0$. This may be established by observing that $\theta(1) = \theta'(1) = 0$ and $\theta''(t) > 0$ for $t \geq 0$.

Let

$$\tilde{B}^T := \left\{ v \in \mathbb{R}^N : \max_j |v_j| \leq 1 \right\}.$$

We wish to calculate

$$||g'(x)||_T = \sup \left\{ \left| \sum_j h_{ij} v_j \right| : 1 \leq i \leq N, v \in \tilde{B}^T \right\}. \tag{2.6}$$

For $i = 1$ or $i = N$, we have $|\sum_j h_{ij} v_j| \leq s$ for any choice of $v \in \tilde{B}^T$. If $N = 2$, then it follows that $||g'(x)||_T = s$ for all $x \in U_{1,N}$.

For the rest of the proof we shall therefore assume that $N \geq 3$. For $1 < i < N$, it is clear from inequalities (2.3) – (2.5) that $|\sum_j h_{ij} v_j|$ is maximized when $v_1 = v_i = 1$ and $v_N = -1$. In this case

$$\left| \sum_j h_{ij} v_j \right| = \frac{1}{E_i} \left[ \frac{x_N - x_i}{x_N - x_1} \theta\left(\frac{x_N}{x_1}\right) x_1^{s+1} + (x_N^s - x_1^s) x_i + \frac{x_i - x_1}{x_N - x_1} \theta\left(\frac{x_1}{x_N}\right) x_N^{s+1} \right] \tag{2.7}$$

$$= \frac{c_1 x_i + c_2}{c_3 x_i + c_4}, \tag{2.8}$$

where $c_1$, $c_2$, $c_3$, and $c_4$ depend on $x_1$ and $x_N$ but not on $x_i$. Observe that $c_3 x_i + c_4 \neq 0$ for $x_1 \leq x_i \leq x_N$. Given this fact, the general form of expression (2.8) leads us to conclude that it is either non-increasing or non-decreasing when regarded as a function of $x_i$. When we substitute $x_i = x_1$, we get $|\sum_j h_{ij} v_j| = s$. When we substitute $x_i = x_N$, we get

$$\left| \sum_j h_{ij} v_j \right| = \frac{2\Big(1 - (x_1/x_N)^s\Big)}{1 - (x_1/x_N)} - s. \tag{2.9}$$

Now, writing $\Gamma(t) := 2(1 - t^s)/(1 - t) - s$, we have $\Gamma'(t) = -2t^s \theta(t^{-1})/(1 - t)^2 < 0$, in other words $\Gamma$ is decreasing on $(0, 1)$. In particular, $\Gamma(x_1/x_N) \geq \lim_{t \to 1} \Gamma(t) = s$. Therefore expression (2.7) is non-decreasing in $x_i$. So, the supremum in (2.6) is attained when $v$ is as above and $i = N - 1$. Recall that $x_{N-1}$ is the second largest component of $x$. The conclusion follows. □

**Corollary 2.5.** *Let $R > 0$. If $N = 1$ or $N = 2$, then* ess sup$\{||g'(x)||_T : x \in \operatorname{int} \mathbb{R}_+^N\} = s$. *If $N \geq 3$, then*

$$\text{ess sup}\{||g'(x)||_T : d_H(x, \mathbb{1}) \leq R\} = \frac{2(1 - e^{-Rs})}{1 - e^{-R}} - s.$$

*Proof.* Note that if $\sigma : \mathbb{R}_+^N \to \mathbb{R}_+^N$ is some permutation of the components, then $g \circ \sigma(x) = \sigma \circ g(x)$ for all $x \in \mathbb{R}_+^N$. Furthermore, $\sigma$ will be an isometry of both the Thompson and Hilbert metrics. It follows that, given any $x \in U_{I,J}$ with $I, J \in \{1, \ldots, N\}$, $I \neq J$, we may reorder the components of $x$ to find a point $y$ in $U_{1,N}$ such that $||g'(y)||_T = ||g'(x)||_T$. Recall, also, that the complement of $U$ in $\operatorname{int}\mathbb{R}^N_+$ has $N$-dimensional Lebesgue measure zero. From these two facts, it follows that the essential supremum of $||g'(x)||_T$ over $\overline{B}_R(\mathbb{1}) := \{x \in \operatorname{int}\mathbb{R}^N_+ : d_H(x,\mathbb{1}) \leq R\}$ is the same as its supremum over $\overline{B}_R(\mathbb{1}) \cap U_{1,N}$.

In the case when $N = 1$ or $N = 2$, the conclusion follows immediately.

For $N = 3$, we must maximize expression (2.2) under the constraints $x_1 < x_{N-1} < x_N$ and $x_1/x_N \geq \exp(-R)$. First, we maximize over $x_{N-1}$, keeping $x_1$ and $x_N$ fixed. In the proof of the previous lemma, we showed that expression (2.2) is non-decreasing in $x_{N-1}$, and so it will be maximized when $x_{N-1}$ approaches $x_N$. Here it will attain the value

$$\frac{2\Big(1-(x_1/x_N)^s\Big)}{1-(x_1/x_N)} - s = \Gamma(x_1/x_N). \tag{2.10}$$

We also showed that $\Gamma$ is decreasing on $(0,1)$. Therefore (2.10) will be maximized when $x_1/x_N = \exp(-R)$, where it takes the value

$$\frac{2(1-e^{-Rs})}{1-e^{-R}} - s.$$

□

**Lemma 2.6.** *Let $C$ be an almost Archimedean cone and let $\{x_i : i \in I\}$ be a finite collection of elements of $C$ of cardinality $n$, all lying in the same part. Denote by $W$ the linear span of $\{x_i : i \in I\}$ and write $C_W := C \cap W$. Denote by $\operatorname{int} C_W$ the interior of $C_W$ as a subset of $W$, using on $W$ the unique Hausdorff linear topology. Then each of the points $x_i; i \in I$ is contained in $\operatorname{int} C_W$. Furthermore, there exists a linear map $F : W \to \mathbb{R}^{n(n-1)}$ such that $F(\operatorname{int} C_W) \subset \operatorname{int}\mathbb{R}^{n(n-1)}_+$ and*

$$M(x_i/x_j;C) = M(F(x_i)/F(x_j);\mathbb{R}^{n(n-1)}_+) \tag{2.11}$$

*for each $i,j \in I$.*

*Proof.* Since the points $\{x_i : i \in I\}$ all lie in the same part of $C$, they also all lie in the same part of $C_W$. Therefore there exist positive constants $a_{ij}$ such that $x_j - a_{ij}x_i \in C_W$ for all $i,j \in I$. If we define $a := \min\{a_{ij} : i,j \in I\}$ it follows that $x_j + \delta x_i \in C_W$ whenever $|\delta| \leq a$ and $i,j \in I$. Now select $i_1,\dots,i_m \in I$ such that $\{x_{i_k} : 1 \leq k \leq m\}$ form a linear basis for $W$. For each $y \in W$, we define $||y|| := \max\{|b_k| : 1 \leq k \leq m\}$, where $y = \sum_{k=1}^m b_k x_{i_k}$ is the unique representation of $y$ in terms of this basis. The topology on $W$ generated by this norm is the same as the one we have been using. If $||y|| \leq a/m$ and $j \in I$, then $x_j + m b_k x_{i_k} \in C_W$ for $1 \leq k \leq m$. It follows that

$$x_j + y = \frac{1}{m}\sum_{k=1}^m (x_j + m b_k x_{i_k}) \in C_W$$

whenever $||y|| \leq a/m$. This proves that $x_j \in \operatorname{int} C_W$ for all $j \in I$.

It is easy to see that $\beta_{ij} := M(x_i/x_j;C) = M(x_i/x_j;C_W)$ for all $i,j \in I$, $i \neq j$. Observe that $\beta_{ij}x_j - x_i \in \partial C_W$. Since $\operatorname{int} C_W$ is a non-empty open convex set which does not contain $\beta_{ij}x_j - x_i$, the geometric version of the Hahn-Banach Theorem implies that there exists a linear functional $f_{ij} : W \to \mathbb{R}$ and a real number $r_{ij}$ such that $f_{ij}(\beta_{ij}x_j - x_i) \leq r_{ij} < f_{ij}(z)$ for all $z \in \operatorname{int} C_W$. Because $0$ is in the closure of $\operatorname{int} C_W$ and $f_{ij}(0) = 0$, we have $r_{ij} \leq 0$. On the other hand, if $f_{ij}(z) < 0$ for some $z \in \operatorname{int} C_W$, then considering $f_{ij}(tz)$ we see that $f_{ij}$ would not be bounded below on $\operatorname{int} C_W$. It follows that $r_{ij} = 0$. Since $\beta_{ij}x_j - x_i$ is in the closure of $\operatorname{int} C_W$, we must have $f_{ij}(\beta_{ij}x_j - x_i) = 0$.

Now, define

$$F : W \to \mathbb{R}^{n(n-1)} : z \mapsto (f_{ij}(z))_{i,j\in I, i\neq j},$$

so that $f_{ij}(z); i,j \in I, i \neq j$ are the components of $F(z)$. Clearly, $F$ is linear and maps $\operatorname{int} C_W$ into $\operatorname{int} \mathbb{R}_+^{n(n-1)}$. Also, for all $i,j \in I$, $i \neq j$,

$$M(F(x_i)/F(x_j); \mathbb{R}_+^{n(n-1)}) = \inf\{\lambda > 0 : f_{kl}(\lambda x_j - x_i) \geq 0 \text{ for all } k,l \in I,\, k \neq l\}.$$

For $\lambda \geq \beta_{ij}$, we have $\lambda x_j - x_i \in \operatorname{cl} C_W$ and so $f_{kl}(\lambda x_j - x_i) \geq 0$ for all $k,l \in I$, $k \neq l$. On the other hand, for $\lambda < \beta_{ij}$, we have $f_{ij}(\lambda x_j - x_i) < 0$ since $f_{ij}(x_j) > 0$. We conclude that $M(F(x_i)/F(x_j); \mathbb{R}_+^{n(n-1)}) = \beta_{ij}$. □

**Lemma 2.7.** *Theorem 1.1 holds in the special case when $C = \mathbb{R}_+^N$ with $N \geq 3$.*

*Proof.* Each part of $\mathbb{R}_+^N$ consists of elements of $\mathbb{R}_+^N$ all having the same components equal to zero. Thus each part can be naturally identified with $\operatorname{int} \mathbb{R}_+^n$, where $n$ is the number of strictly positive components of its elements. We may therefore assume initially that $\{x,y,u\} \subset \operatorname{int} \mathbb{R}_+^N$.

Define $L : \mathbb{R}^N \to \mathbb{R}^N$ by $L(z) := (u_1 z_1, \dots, u_N z_N)$. Its inverse is given by $L^{-1}(z) := (u_1^{-1} z_1, \dots, u_N^{-1} z_N)$. Both $L$ and $L^{-1}$ are linear maps which leave $\mathbb{R}_+^N$ invariant. It follows that $L$ and $L^{-1}$ are isometries of $\mathbb{R}_+^N$ with respect to both the Thompson and Hilbert metrics. Therefore, for $u, z \in \operatorname{int} \mathbb{R}_+^N$,

$$L^{-1}(\phi(s; u, z)) = \phi(s; L^{-1}(u), L^{-1}(z)).$$

Thus, we may as well assume that $u = \mathbb{1}$.

We now wish to apply Proposition 2.1 with $f := g$ and $G := B_{R+\epsilon}(\mathbb{1}) = \{z \in \mathbb{R}_+^N : d_H(z, \mathbb{1}) < R + \epsilon\}$. It was shown in [23] that $G$ is a convex cone, in other words that it is closed under multiplication by positive scalars and under addition of its elements. Since $\phi(s; w, z)$ is a positive combination of $w$ and $z$, it follows that $\phi(s; w, z)$ is in $G$ if $w$ and $z$ are. If we now apply Lemma 2.3, Proposition 2.1, and Corollary 2.5, and let $\epsilon$ approach zero, we obtain the desired result. □

**Lemma 2.8.** *Theorem 1.1 holds in the special case when the linear span of $\{x,y,u\}$ is one- or two-dimensional.*

*Proof.* Let $W$ denote the linear span of $\{x,y,u\}$, in other words the smallest linear subspace containing these points. By Lemma 2.6, $x$, $y$, and $u$ are in the interior of $C \cap W$ in $W$. It is easy to see that $M(z/w; C) = M(z/w; C \cap W)$ for all $w, z \in \operatorname{int}(C \cap W)$. Therefore, we can work in the cone $C \cap W$.

It is not difficult to show [14] that if $m := \dim W$ is either one or two, then there is a linear isomorphism $F$ from $W$ to $\mathbb{R}^m$ taking $\operatorname{int}(C \cap W)$ to $\operatorname{int} \mathbb{R}_+^m$. It follows that $F$ is an isometry of both the Thompson and Hilbert metrics and $F(\phi(s; z, w)) = \phi(s; F(z), F(w))$ for all $z, w \in \operatorname{int}(C \cap W)$. We may thus assume that $C = \mathbb{R}_+^m$ and $u, x, y \in \operatorname{int} C$.

As in the proof of Lemma 2.7, we may assume that $u = \mathbb{1}$.

To obtain the required result, we apply Lemma 2.3, Corollary 2.5, and Proposition 2.1 with $f := g$ and $G := \operatorname{int} \mathbb{R}_+^m$. □

*of Theorem 1.1.* Let $W$ denote the linear span of $\{x,y,u\}$. Lemma 2.8 handles the case when these three points are not linearly independent; we will therefore assume that they are. Thus the five points $x$, $y$, $u$, $\phi(s; u, x)$, and $\phi(s; u, y)$ are distinct. We apply Lemma 2.6 and obtain a linear map $F : W \to \mathbb{R}_+^{20}$ with the specified properties. From (2.11), it is clear that $d_T(z,w) = {d_T}'(F(z), F(w))$ for each $z, w \in \{x, y, u, \phi(s; u, x), \phi(s; u, y)\}$. Here we are using ${d_T}'$ to denote the Thompson metric on $\mathbb{R}_+^{20}$. Note that $\phi(s; u, x)$ is a positive combination of $u$ and $x$ and that the coefficients of $u$ and $x$ depend only on $s$, $M(u/x; C)$, and $M(x/u; C)$. The latter two quantities are equal to $M(F(u)/F(x); \mathbb{R}_+^{20})$ and $M(F(x)/F(u); \mathbb{R}_+^{20})$ respectively. We conclude that $F(\phi(s; u, x)) = \phi(s; F(u), F(x))$. A similar argument gives $F(\phi(s; u, y)) =$

$\phi(s; F(u), F(y))$. Inequality (1.3) follows by applying Lemma 2.7 to the points $F(x)$, $F(y)$, and $F(u)$ in the cone $\mathbb{R}^{20}_+$. □

2.2. **Hilbert's Metric.** We shall continue to use the same notation. Thus, for a given $N \in \mathbb{N}$ and $s \in (0,1)$, we use $g$ to denote the function in (2.1) and $U$ to denote the union of sets $U_{I,J}$ with $I, J \in \{1, \dots, N\}$, $I \neq J$. We also use the functions $\theta(t) := (1-s) - t^s + st$ and $E_i(x) := E_i := x_i(x_N^s - x_1^s) + x_N x_1^s - x_1 x_N^s$, and write $h_{ij}(x) := (x_j/g_i(x))\partial g_i/\partial x_j(x)$. As was noted earlier, $\theta(t) > 0$ if $t > 0$ and $t \neq 1$. Also, $\gamma(t) := (1-t^s)/(1-t)$, $\gamma(1) := s$ is strictly decreasing on $[0, \infty)$. We shall also use the simple but useful observation that if $c_1$, $c_2$, $c_3$, and $c_4$ are constants such that $c_3 t + c_4 \neq 0$ for $a \le t \le b$, then the function $t \mapsto (c_1 t + c_2)/(c_3 t + c_4)$ is either increasing on $[a,b]$ (if $c_1 c_4 - c_2 c_3 \ge 0$) or decreasing on $[a,b]$ (if $c_1 c_4 - c_2 c_3 \le 0$). Either way, the function attains is maximum over $[a,b]$ at $a$ or $b$.

Recall that if $g$ is Fréchet differentiable at $x \in \operatorname{int} \mathbb{R}^N_+$ then $||g'(x)||_H$ denotes the norm of $g'(x)$ as a linear map from $(\mathbb{R}^N, ||\cdot||_x^H)$ to $(\mathbb{R}^N, ||\cdot||_{g(x)}^H)$, although, of course, $||\cdot||_x^H$ and $||\cdot||_{g(x)}^H$ are semi-norms rather than norms.

**Lemma 2.9.** *Consider the Hilbert metric on* $\operatorname{int} \mathbb{R}^N_+$ *with* $N \ge 2$. *Let* $x \in U_{1,N}$. *If* $N = 2$ *then the norm of* $g'$ *at* $x$ *is given by* $||g'(x)||_H = s$. *If* $N \ge 3$, *then*

$$||g'(x)||_H = \frac{x_N - x_{N-1}}{x_N - x_1} \theta\left(\frac{x_N}{x_1}\right) \frac{x_1^{s+1}}{E_{N-1}} + \frac{(x_N^s - x_1^s) x_{N-1}}{E_{N-1}}. \tag{2.12}$$

*Proof.* The norm of $g'(x)$ as a map from $(\mathbb{R}^N, ||\cdot||_x^H)$ to $(\mathbb{R}^N, ||\cdot||_{g(x)}^H)$ is given by

$$||g'(x)||_H = \sup_{v \in \tilde{B}^H} \max_{i,k} \sum_j (h_{ij} - h_{kj}) v_j,$$

where

$$\tilde{B}^H := \left\{ v \in \mathbb{R}^N : \max\nolimits_j v_j - \min\nolimits_j v_j \le 1 \right\}.$$

To calculate $||g'(x)||_H$ we will need to determine the sign of $h_{ij} - h_{kj}$ for each $i, j, k \in \{1, \dots, N\}$. We introduce the notation

$$L_{ik} := \sup_{v \in \tilde{B}^H} \sum_j (h_{ij} - h_{kj}) v_j. \tag{2.13}$$

Note that $g$ is homogeneous of degree $s$, in other words $g(\lambda x) = \lambda^s g(x)$ for all $x \in \mathbb{R}^N_+$ and $\lambda > 0$. Therefore,

$$\sum_j x_j \frac{\partial g_i}{\partial x_j}(x) = s g_i(x)$$

for each $i \in \{1, \dots, N\}$. Thus $\sum_j h_{ij} = s$ for each $i \in \{1, \dots, N\}$, a fact that could also have been obtained by straightforward calculation. It follows that

$$\sum_j (h_{ij} - h_{kj}) v_j = \sum_j (h_{ij} - h_{kj})(v_j + c) \tag{2.14}$$

for any constant $c \in \mathbb{R}$.

It is clear that an optimal choice of $v$ in (2.13) would be to take $v_j := 1$ for each component $j$ such that $h_{ij} - h_{kj} > 0$ and $v_j := 0$ for each component such that $h_{ij} - h_{kj} < 0$. Alternatively, we may choose $v_j := 0$ when $h_{ij} - h_{kj} > 0$ and $v_j := -1$ when $h_{ij} - h_{kj} < 0$. That the optimal value is the same in both cases follows from (2.14). Also, it is easy to see that $L_{ik} = L_{ki}$.

Fix $i, k \in \{1, \dots, N\}$ so that $i < k$. There are four cases to consider.

- **Case 1.** $1 < i < k < N$. Recall that $h_{1j}(x) = s\delta_{1j}$ and $h_{Nj}(x) = s\delta_{Nj}$. A calculation using equations (2.3) – (2.5) gives

$$E_i(x)E_k(x)(h_{i1}(x) - h_{k1}(x)) = x_N^s x_1^{s+1}(x_k - x_i)\theta\Big(\frac{x_N}{x_1}\Big) \geq 0$$

  and

$$E_i(x)E_k(x)(h_{iN}(x) - h_{kN}(x)) = x_1^s x_N^{s+1}(x_k - x_i)\theta\Big(\frac{x_1}{x_N}\Big) \geq 0. \tag{2.15}$$

  We also have that $h_{ii}(x) - h_{ki}(x) = h_{ii}(x) > 0$ and $h_{ik}(x) - h_{kk}(x) = -h_{kk}(x) < 0$. So an optimal choice of $v \in \tilde{B}^H$ in equation (2.13) is given by $v_j := -\delta_{jk}$. We conclude that $L_{ik} = h_{kk}$ in this case.
- **Case 2.** $1 = i < k < N$. We will show that $h_{k1}(x) \leq h_{11}(x) = s$. Consider $x_1$ and $x_N$ as fixed and $x_k$ as varying in the range $x_1 \leq x_k \leq x_N$. From equation (2.3), $h_{k1}(x) = (c_1x_k+c_2)/(c_3x_k+c_4)$, where $c_1$, $c_2$, $c_3$, and $c_4$ depend on $x_1$ and $x_N$, and both $c_3$ and $c_4$ are positive. A simple calculation shows that $c_1c_4-c_2c_3 = -\theta(x_N/x_1)x_1^{s+1}x_N^s$, which is negative. Hence $h_{k1}$ is decreasing in $x_k$ and takes its maximum value when $x_k = x_1$. Here it achieves the value

$$\frac{x_1}{x_N - x_1}\theta\Big(\frac{x_N}{x_1}\Big) = s - \frac{x_1^{1-s}(x_N^s - x_1^s)}{x_N - x_1} < s.$$

  Thus we conclude that $h_{11}(x) - h_{k1}(x) > 0$. We also have that $h_{1k}(x) - h_{kk}(x) = -h_{kk}(x) \leq 0$ and $h_{1N}(x) - h_{kN}(x) = -h_{kN}(x) \geq 0$. Thus the optimal choice of $v \in \tilde{B}^H$ is given by $v_j := -\delta_{jk}$. We conclude that in this case $L_{1k}(x) = h_{kk}(x)$.
- **Case 3.** $1 < i < k = N$. Here $h_{i1} \geq h_{N1} = 0$, $h_{ii} \geq h_{Ni} = 0$, and $h_{iN} \leq h_{NN} = s$. So the optimal $v \in \tilde{B}^H$ is given by $v_j := \delta_{j1} + \delta_{ji}$. We conclude that $L_{iN} = h_{i1} + h_{ii}$.
- **Case 4.** $i = 1$ and $k = N$. Here $s = h_{11} \geq h_{N1} = 0$ and $0 = h_{1N} \leq h_{NN} = s$. Thus the optimal $v \in \tilde{B}^H$ is given by $v_j := \delta_{1j}$. We conclude that $L_{1N} = s$.

If $N = 2$ then Case 4 is the only one possible, and so $||g'(x)||_H = s$. So, for the rest of the proof, we will assume that $N \geq 3$.

We know that $h_{i1}(x) + h_{ii}(x) = s - h_{iN}(x) \geq s$ so Case 3 dominates Case 4, that is to say $L_{iN}(x) \geq L_{1N}(x)$ for $i > 1$. Since $h_{i1}(x) \geq 0$ for $i \in \{1, \ldots, N\}$, Case 3 also dominates Cases 1 and 2, meaning that $L_{iN}(x) \geq L_{ik}(x)$ for $k < N$, $i < k$.

The final step is to maximize $L_{iN}(x) = h_{i1}(x)+h_{ii}(x) = s-h_{iN}(x)$ over $i \in \{2, \ldots, N-1\}$. From (2.15), $h_{mN}(x) \geq h_{nN}(x)$ for $m < n$. Thus the maximum occurs when $i = N - 1$. Recall that we have ordered the components of $x$ in such a way that $x_{N-1}$ is the second largest component of $x$. We conclude that

$$||g'(x)||_H = \max_{i,k:i<k} L_{ik} = h_{N-1,1} + h_{N-1,N-1}$$

By substituting the expressions in (2.3) and (2.4), we obtain the required formula. □

**Corollary 2.10.** *Let $R > 0$ and $N \geq 2$. Let $l$ be a linear functional on $\mathbb{R}^N$ such that $l(x) > 0$ for all $x \in \operatorname{int}\mathbb{R}_+^N$ and define $S := \{x \in \mathbb{R}_+^N : l(x) = 1\}$. If $N = 2$, then* ess sup$\{||g'(x)||_H : x \in S\} = s$. *If $N \geq 3$, then*

$$\text{ess sup}\{||g'(x)||_H : d_H(x, \mathbb{1}) \leq R, x \in S\} = \frac{1 - e^{-Rs}}{1 - e^{-R}}.$$

*In both cases, the essential supremum is taken with respect to the $N-1$-dimensional Lebesgue measure on $S$.*

*Proof.* Note that the complement of $U \cap S$ in $S$ has $N-1$-dimensional Lebesgue measure zero. Using the reordering argument in the proof of Corollary 2.5, we deduce the result in the case when $N = 2$.

The case when $N \geq 3$ reduces to maximizing the right hand side of (2.12) subject to the constraints $x_1 < x_{N-1} < x_N$ and $x_1/x_N \geq \exp(-R)$. We can write the expression in (2.12) in the form $s + (c_1 x_{N-1} + c_2)/(c_3 x_{N-1} + c_4)$, where $c_1$, $c_2$, $c_3$, and $c_4$ depend only on $x_1$ and $x_N$ and $c_1 \geq 0$, $c_2 \leq 0$, $c_3 \geq 0$, $c_4 \geq 0$. It follows that, if we view $x_1$ and $x_N$ as fixed and $x_{N-1}$ as variable, the expression is maximized when $x_{N-1} = x_N$. The value obtained there will be

$$\frac{1-(x_1/x_N)^s}{1-(x_1/x_N)} = \gamma(x_1/x_N).$$

If we recall that $\gamma$ is decreasing on $[0,1)$ and $x_1/x_N \geq \exp(-R)$, we see that

$$||g'(x)||_H \leq \frac{1-e^{-Rs}}{1-e^{-R}}.$$

If $x_1/x_N = \exp(-R)$, then, by choosing $x \in U_{1,N}$ with $x_{N-1}$ close to $x_N$, we can arrange that $||g'(x)||_H$ is as close as desired to this value. ☐

**Lemma 2.11.** *Theorem 1.2 holds in the special case when $C = \mathbb{R}^N_+$ with $N \geq 3$.*

*Proof.* As in the proof of Lemma 2.7, we may assume that $x, y \in \operatorname{int} \mathbb{R}^N_+$ and $u = \mathbb{1}$. Define $l : \mathbb{R}^N \to \mathbb{R}$ by $l(z) := \sum_{i=1}^N z_i/N$ and let $S := \{x \in \mathbb{R}^N_+ : l(x) = 1\}$. Then $l$ is a linear functional and $l(z) > 0$ for all $z \in \operatorname{int} \mathbb{R}^N_+$. It is easy to check that $\phi(s; \lambda z, \mu w) = \lambda^{1-s}\mu^s \phi(s; z, w)$ for all $\lambda, \mu > 0$ and $z, w \in \operatorname{int} \mathbb{R}^N_+$. Thus

$$d_H\left(\phi\left(s; \frac{u}{l(u)}, \frac{x}{l(x)}\right), \phi\left(s; \frac{u}{l(u)}, \frac{y}{l(y)}\right)\right) = d_H(\phi(s; u, x), \phi(s; u, y)).$$

We also have that $d_H(x/l(x), y/l(y)) = d_H(x, y)$. Therefore we may assume that $x, y \in S$. Let $\epsilon > 0$ and define $G := \{z \in S : d_H(z, \mathbb{1}) < R + \epsilon\}$. It was shown in [23] that $G$ is convex. Also, Lemma 2.3 states that $g$ is locally Lipschitzian. We may therefore apply Proposition 2.2 with $f := g$. Since $g$ is homogeneous of degree $s$, we have that $g'(x)(x) = sg(x)$ for all $x \in G$. This, combined with the fact that $|g(x)|^H_{g(x)} = 0$, implies that $||g'(x)||_{\tilde{H}} = ||g'(x)||_H$. Using Corollary 2.10, and letting $\epsilon$ approach zero, we deduce the required result. ☐

**Lemma 2.12.** *Theorem 1.2 holds in the special case when the linear span of $\{u, x, y\}$ is 1- or 2-dimensional.*

*Proof.* If the linear span of $\{u, x, y\}$ is one-dimensional, then all Hilbert metric distances are zero, so assume that it is two-dimensional. The same argument as was used in Lemma 2.8 shows that it suffices to prove the result for $C = \mathbb{R}^2_+$, $u = \mathbb{1}$, and $x, y \in \operatorname{int} \mathbb{R}^2_+$. As shown in the proof of Lemma 2.11, we may assume that $l(x) = l(y) = 1$ where $l((z_1, z_2)) := (z_1 + z_2)/2$. We now apply Proposition 2.2 with $f := g$ and $G := S := \{z \in \operatorname{int} \mathbb{R}^2_+ : l(z) = 1\}$. Again, $||g'(x)||_{\tilde{H}} = ||g'(x)||_H$ for all $x \in G$. The result follows from the first part of Corollary 2.10. ☐

*of Theorem 1.2.* The proof uses Lemmas 2.11 and 2.12 and is exactly analogous to the proof of Theorem 1.1. ☐

*of Corollary 1.3.* We first prove the result for the case of Thompson's metric. We will use the alternative characterization of semihyperbolicity given in Lemma 1.2 of [1]. Suppose $x, y, x', y' \in C$ are all in the same part and are such that neither $d_T(x, x')$ nor $d_T(y, y')$ is greater than 1. Let $t \in [0, \infty)$ and write $z := \zeta_{(x,y)}(t)$ and $w := \phi(d_T(x, z)/d_T(x, y); x, y')$. Observe that

$d_T(y, y') \leq 1$ implies $|d_T(x, y) - d_T(x, y')| \leq 1$. Since $d_T(x, w) = d_T(x, y')d_T(x, z)/d_T(x, y)$, we have

$$|d_T(x, w) - d_T(x, z)| \leq d_T(x, z)/d_T(x, y) \leq 1$$

Similar reasoning allows us to conclude that

$$|d_T(x, w') - d_T(x', z')| \leq 1,$$

where $z' := \zeta_{(x',y')}(t)$ and $w' := \phi(d_T(x', z')/d_T(x', y'); x, y')$. From $d_T(x, z) = \min(t, d_T(x, y))$ and $d_T(x', z') = \min(t, d_T(x', y'))$, we have that

$$|d_T(x, z) - d_T(x', z')| \leq |d_T(x, y) - d_T(x', y')| \leq 2.$$

So

$$d_T(w, w') = |d_T(x, w) - d_T(x, w')| \leq 4.$$

By Theorem 1.1, $d_T(z, w) \leq 2d_T(y, y') \leq 2$ and $d_T(z', w') \leq 2d_T(x, x') \leq 2$. The triangle inequality gives $d_T(z, z') \leq d_T(z, w) + d_T(w, w') + d_T(w', z') \leq 8$. This is the uniform bound required by the characterization of semihyperbolicity we are using.

The proof that $C$ is semihyperbolic when endowed with Hilbert's metric is similar. □

## REFERENCES

[1] J.M. ALONSO AND M.R. BRIDSON, Semihyperbolic groups, *Proc. London Math. Soc.*, **70**(1) (1995), 56–114.

[2] Y. BENOIST, Convexes hyperboliques et fonctions quasisymétriques, *Publ. Math. Inst. Hautes Études Sci.*, **97** (2003),181–237.

[3] G. BIRKHOFF, Extensions of Jentzsch's theorem, *Trans. Amer. Math. Soc.*, **85** (1957), 219–227.

[4] G. BIRKHOFF, Uniformly semi-primitive multiplicative processes, *Trans. Amer. Math. Soc.*, **104** (1962), 37–51.

[5] G. BIRKHOFF AND L. KOTIN, Integro-differential delay equations of positive type, *J. Differential Equations*, **2**(1966), 320–327.

[6] M. BRIDSON AND A. HAEFLIGER, *Metric Spaces of Non-Positive Curvature*, Springer-Verlag, 1999.

[7] H. BUSEMANN AND B.B. PHADKE, *Spaces with distinguished geodesics*, volume 108 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker Inc., New York, 1987.

[8] P.J. BUSHELL, Hilbert's metric and positive contraction mappings in a Banach space, *Arch. Rational Mech. Anal.*, **52** (1973), 330–338.

[9] P.J. BUSHELL, The Cayley-Hilbert metric and positive operators, In *Proceedings of the symposium on operator theory (Athens, 1985)*, Volume 84, pages 271–280, 1986.

[10] G. CORACH, H. PORTA, AND L. RECHT, Convexity of the geodesic distance on spaces of positive operators, *Illinois J. Math.*, **38**(1) (1994), 87–94.

[11] P. de la HARPE, On Hilbert's metric for simplices, *Geometric Group Theory (London Math. Soc. Lecture Notes)*, **181** (1993), 97–119.

[12] D. EGLOFF, Uniform Finsler Hadamard manifolds, *Ann. Inst. H. Poincaré Phys. Théor.*, **66**(3) (1997), 323–357.

[13] S.P. EVESON AND R.D. NUSSBAUM, Applications of the Birkhoff-Hopf theorem to the spectral theory of positive linear operators, *Math. Proc. Cambridge Philos. Soc.*, **117**(3) (1995), 491–512.

[14] S.P. EVESON AND R.D. NUSSBAUM, An elementary proof of the Birkhoff-Hopf theorem, *Math. Proc. Cambridge Philos. Soc.*, **117**(1) (1995), 31–55.

[15] M. GROMOV, Hyperbolic groups, In *Essays in group theory*, volume 8 of *Math. Sci. Res. Inst. Publ.*, pages 75–263. Springer, 1987.

[16] J. GUNAWARDENA AND C. WALSH, Iterates of maps which are non-expansive in Hilberts's metric, *Kybernetika*, **39(**2) (2003), 193–204.

[17] P. KELLY AND E.G. STRAUS, Curvature in Hilbert geometries, *Pacific J. Math.*, **25**(3) (1968), 549–552.

[18] C. LIVERANI, Decay of correlations, *Ann. of Math.(2)*, **142**(2) (1995), 239–301.

[19] C. LIVERANI AND M.P. WOJTKOWSKI, Generalization of the Hilbert metric to the space of positive definite matrices, *Pacific J. Math.*, **166**(2) (1994), 339–355.

[20] C. MCMULLEN, Coxeter groups, Salem numbers and the Hilbert metric, *Publ. Math. IHES.*, **95** (2002), 151–183.

[21] V. METZ, Hilbert's projective metric on cones of Dirichlet forms, *J. Funct. Anal.*, **127**(2) (1995), 438–455.

[22] R.D. NUSSBAUM, Finsler structures for the part metric and hilbert's projective metric and applications to ordinary differential equations, *Diff. and Int. Eqns.*, 7(6) (1994), 1649–1707.

[23] R.D. NUSSBAUM, Hilbert's projective metric and iterated nonlinear maps, *Mem. Amer. Math. Soc.*, **75**(391), 1988.

[24] R.D. NUSSBAUM, Iterated nonlinear maps and Hilbert's projective metric, II, *Mem. Amer. Math. Soc.*, **79**(401), 1989.

[25] R.D. NUSSBAUM, Omega limit sets of nonexpansive maps() finiteness and cardinality estimates, *Differential Integral Equations*, **3**(3) (1990), 523–540.

[26] R.D. NUSSBAUM, Entropy minimization, Hilbert's projective metric, and scaling integral kernels, *J. Funct. Anal.*, **115**(1) (1993), 45–99.

[27] A.J.B. POTTER, Applications of Hilbert's projective metric to certain classes of non-homogeneous operators, *Quart. J. Math. Oxford Ser. (2)*, **28**(109)(1977), 93–99.

[28] Z. SHEN, *Lectures on Finsler Geometry*, World Scientific, 2001.

[29] E. SOCIÉ-MÉTHOU, Behaviour of distance functions in Hilbert-Finsler geometry, *Differential Geom. Appl.*, **20**(1) (2004), 1–10.

[30] A.C. THOMPSON, On certain contraction mappings in a partially ordered vector space, *Proc. Amer. Math. Soc.*, **14** (1963), 438–443.

[31] K. WYSOCKI, Behavior of directions of solutions of differential equations, *Differential Integral Equations*, **5**(2) (1992)281–305.

[32] K. WYSOCKI, Some ergodic theorems for solutions of homogeneous differential equations, *SIAM J. Math. Anal.*, **24**(3) (1993), 681–702.