Proper Gromov Transforms of Metrics are Metrics

Andreas Dress*

August 6, 2001

Abstract

In phylogenetic analysis, a standard problem is to approximate a given metric by an *additive* metric. Here it is shown that, given a metric D defined on some finite set X and a non-expansive map $f : X \to \mathbb{R}$, the one-parameter family of the Gromov transforms $D^{\Delta,f}$ of D relative to f and Δ that starts with D for large values of Δ and ends with an additive metric for $\Delta = 0$ consists exclusively of metrics. It is expected that this result will help to better understand some standard tree reconstruction procedures considered in phylogenetic analyis.

Keywords and Phrases: Metrics, additive metrics, Δ additive metrics, ultra metrics Farris transforms, Gromov transforms, phylogenetic analysis, phylogenetic combinatorics

1 Introduction

Given a finite set X of cardinality $n \geq 2$, a symmetric map

$$D : X \times X \to \mathbb{R} : (x, y) \mapsto xy$$

from $X \times X$ into $\mathbb{R}$, a map $f : X \to \mathbb{R}$, and a non-negative real number Δ, the *Gromov transform* $D^{\Delta,f}$ of D relative to f and Δ has been defined in [2] (see also [4, 1]) where also the biological motivation for this construction within the context of phylogenetic analysis has been discussed:

First, one considers the *Farris transform* (cf.[3]) D_f of D relative to f defined by

$$D_f : X \times X \to \mathbb{R} : (x, y) \mapsto xy_f := xy - f(x) - f(y).$$

Then, one forms the (unique!) largest symmetric map D' from $X \times X$ into $\mathbb{R}$ below D_f — also denoted by $(D_f)^\Delta$ — that satisfies the inequalities

$$D'(x, z) \leq \max(D'(x, y), D'(y, z)) + \Delta$$

*GK Strukturbildungsprozesse, FSP Mathematisierung, Universität Bielefeld, D-33501 Bielefeld, Germany, email: dress@mathematik.uni-bielefeld.de. This work was supported by the *Graduiertenkolleg Strukturbildungsprozesse*.

for all $x, y \in X$. And then one defines $D^{\Delta,f}$ as the Farris transform $D'_g = ((D_f)^\Delta)_g$ of $(D_f)^\Delta$ relative to the map $g := -f$.

It has been observed in [2] that $D^{\Delta,f}$ coincides with the largest (Δ, f) additive map below D in $\mathcal{S}_2(X)$, the space of all symmetric maps from $X \times X$ into $\mathbb{R}$, i.e. it coincides with the (necessarily unique) largest map D'' in $\mathcal{S}_2(X)$ satisfying the inequalities

$$D''(x,y) \le D(x,y)$$

and

$$D''(x,y)_f \le \max(D''(x,z)_f, D''(y,z)_f) + \Delta$$

for all $x, y, z \in X$, and that $D^f := D^{0,f}$ is an additive[1] metric for every metric $D \in \mathcal{S}_2(X)$ and for every f in the tight span

$$T(D) := \{f \in \mathbb{R}^X : f(x) = \max(xy - f(y) : y \in X) \text{ for all } x \in X\}$$

of D. More generally, it was shown there that, given some $D \in \mathcal{S}_2(X)$ and some $f \in \mathbb{R}^X$, the inequality $D^{\Delta,f}(x,y) \ge 0$ holds for all $\Delta \ge 0$ and all $x, y \in X$ if and only if $D^f(x,y) \ge 0$ holds for all $x, y \in X$ if and only if $D^f(x,x) \ge 0$ holds for all $x \in X$ if and only if the inequality $f(y) \le xy + f(x)$ or, equivalently,

$$|f(y) - f(x)| = \max(f(y) - f(x), f(x) - f(y)) \le xy$$

holds for all $x, y \in X$, i.e. if and only if the map

$$X \times X \to \mathbb{R} : (x,y) \mapsto |f(x) - f(y)|$$

is a map below D or, still equivalently, if and only if the map $f : X \to \mathbb{R}$ is a *non-expansive* map considered as a map from the metric space (X, D) into $\mathbb{R}$, endowed with its standard metric.

Consequently, a Gromov transform $D^{\Delta,f}$ of a map D in $\mathcal{S}_2(X)$ relative to some $\Delta \ge 0$ and some map $f : X \to \mathbb{R}$ will be called a *proper* Gromov transform of D if $|f(x) - f(y)| \le xy$ holds for all $x, y \in X$. Here, we now want to show that proper Gromov transforms of metrics always are metrics, too. More precisely, the following result will be established:

Theorem 1.1 *Given a symmetric map $D : X \times X \to \mathbb{R} : (x,y) \mapsto xy$ and a map $f : X \to \mathbb{R}$, the map $D^{\Delta,f}$ is a Δ-additive metric for every $\Delta \ge 0$ if and only if the following two (obviously necessary) conditions are satisfied:*

(i) *D is a metric and*

(ii) *$|f(y) - f(x)| \le xy$ holds for all $x, y \in X$.*

[1] A map D in $\mathcal{S}_2(X)$ is called additive if $xy + uv \le \max(xu + yv, xv + yu)$ holds for all $x, y, u, v \in X$ or, equivalently, if it is $(0, f)$ additive for every map f of the form $f = h_a : X \to \mathbb{R} : x \mapsto xa$ for some $a \in \mathbb{R}$.

2 Basic Definitions and Results

In view of the fact that $D = D^{\Delta,f}$ holds for every sufficiently large number Δ, it is obvious that Condition *(i)* in Theorem 1.1 must hold if $D^{\Delta,f}$ is a metric for all $\Delta \geq 0$. Further, Condition *(ii)* must hold, too, in view of the results from [2] quoted above.

To establish the converse, we'll need the following definitions and results:

Definition 2.1 *Given a symmetric map $D : X \times X \to \mathbb{R}$ and a map $f : X \to \mathbb{R}$, put*

$$\Delta_f(D) := \max\{xy_f - \max(xa_f, ay_f) : x, y, a \in X\}$$

and

$$X(x,y) = X_{D,f}(x,y) := \{a \in X : xy_f - \max(xa_f, ay_f) = \Delta_f(D)\}$$

The next three results are crucial for the proof of our theorem in the next section.

Lemma 2.2 *If $X(x,y) \neq \emptyset$ holds for some $x, y \in X$, we have necessarily*

$$|f(x) - f(y)| \leq xy - \Delta_f(D).$$

Proof: Choose some $a \in X(x,y)$ and note that $xa - f(a) \geq -f(x)$ and $ya - f(a) \geq -f(y)$ implies

$$xy - \Delta_f(D) = f(x) + f(y) + xy_f - \Delta_f(D) = f(x) + f(y) + \max(xa_f, ya_f) =$$

$$= \max(f(y) + xa - f(a), f(x) + ya - f(a))$$
$$\geq \max(f(y) - f(x), f(x) - f(y)) = |f(x) - f(y)|. \qquad \blacksquare$$

Lemma 2.3 *Given some $D \in \mathcal{S}_2(X)$ and some $f \in \mathbb{R}^X$ with $\Delta_f(D) > 0$, there exists some positive real number $\epsilon = \epsilon(D,f) \leq \Delta_f(D)$ and some symmetric map $k_f : X \times X \to \mathbb{N}_0$ such that*

$$D^{\Delta,f}(x,y) = xy - k_f(x,y)(\Delta_f(D) - \Delta)$$

holds for all $x, y \in X$ and all Δ in the interval $[\Delta_f(D) - \epsilon, \Delta_f(D)]$. The map k_f satisfies the condition

$$k_f(x,y) \geq 1 + k_f(x,y|a)$$

with

$$k_f(x,y|a) := \min\{k_f(t,a) : t \in \{x,y\}, xy_f = ta_f + \Delta_f(D)\}$$

for all $a \in X(x,y)$, and it is the smallest such map, i.e. we have

$$k_f(x,y) = \max(1 + k_f(x,y|a) : a \in X(x,y))$$

in case $X(x,y) \neq \emptyset$ and

$$k_f(x,y) = 0$$

in case $X(x,y)=\emptyset$.

In particular, $xy-\Delta < f(x)-f(y)$ *implies* $k_f(x,y)=0$ *in view of the previous lemma,*

$$xy_f = ya_f + \Delta_f(D) > xa + \Delta_f(D)$$

implies

$$k_f(x,y) \geq 1 + k_f(y,a),$$

and

$$xy_f = xa_f + \Delta_f(D) = ya_f + \Delta_f(D)$$

implies

$$k_f(x,y) \geq 1 + \min(k_f(x,a), k_f(y,a))$$

Proof: The simple and straight forward proof of this lemma is left to the reader.

Lemma 2.4 *Assume we are given some metric* $D \in \mathcal{S}_2(X)$*, and some* $f \in \mathbb{R}^X$ *with* $|f(u)-f(v)| \geq uv$ *for all* $u,v \in X$*. Assume furthermore that* $xz = xy + yz$ *and* $a \in X(x,y)$ *holds for some fixed elements* $x,y,z,a \in X$ *and put* $\Delta := \Delta_f(D)$*. Then, we have either*

(1) $$az_f < xa_f = xz_f - \Delta,\ xy_f - \Delta = xa_f, \text{ and } yz_f + 2f(y) = 0,$$

or

(2) $$xa_f < az_f = xz_f - \Delta,\ xy_f - \Delta = ay_f, \text{ and } ay + yz = az,$$

or

(3) $$az_f = xa_f = xz_f - \Delta,\ yz_f + 2f(y) = 0,\ xy_f - \Delta = xa_f = ay_f, \text{ and } ay + yz = az.$$

Proof: Clearly, our assumptions $\Delta = \Delta_f(D)$, $xy_f - \Delta = \max(xa_f, ay_f)$, $yz \geq f(z) - f(y)$, and $ay + yz \geq az$ imply

$$\max(xa_f, az_f) \geq xz_f - \Delta, \quad xy_f - \Delta \geq xa_f, ay_f, \quad yz_f + 2f(y) \geq 0,$$

and

$$ay_f + yz_f + 2f(y) \geq az_f,$$

respectively. Thus, we have

$$\max(xa_f, az_f) \geq xz_f - \Delta = xy_f + yz_f + 2f(y) - \Delta \geq xa_f + yz_f + 2f(y) \geq xa_f$$

as well as

$$\max(xa_f, az_f) \geq xz_f - \Delta = xy_f + yz_f + 2f(y) - \Delta \geq ay_f + yz_f + 2f(y) \geq az_f.$$

However, we have either $\max(xa_f, az_f) = xa_f$ or $\max(xa_f, az_f) = az_f$, and the former implies

$$xa_f = xz_f - \Delta,\ xy_f - \Delta = xa_f, \text{ and } yz_f + 2f(y) = 0,$$

while $\max(xa_f, az_f) = az_f$ implies

$$az_f = xz_f - \Delta,\ xy_f - \Delta = ay_f,\ and\ ay_f + yz_f + 2f(y) = az_f,$$

or, equivalently,

$$az_f = xz_f - \Delta,\ xy_f - \Delta = ay_f,\ and\ ay + yz = az.$$

So, we have either

$$az_f < xa_f = xz_f - \Delta,\ xy_f - \Delta = xa_f,\ \text{and}\ \ yz_f + 2f(y) = 0,$$

or

$$xa_f < az_f = xz_f - \Delta,\ xy_f - \Delta = ay_f,\ \text{and}\ \ ay + yz = az,$$

or

$$az_f = xa_f = xz_f - \Delta,\ yz_f + 2f(y) = 0,\ xy_f - \Delta = xa_f = ay_f,\ \text{and}$$

$$ay + yz = az,$$

as claimed. ∎

3 Proof of theTheorem

We can now return to the proof of Theorem 1.1: It is easy to see that all we need to show is that the assumptions (i), (ii), and $\Delta := \Delta_f > 0$ imply that $k_f(x,z) \geq k_f(x,y) + k_f(y,z)$ holds for all $x, y, z \in X$ with $xz = xy + yz$. We will do this by "induction" relative to xz_f, i.e. we assmue that our claim holds for all $x', y', z' \in X$ with $x'z' = x'y' + y'z'$ and $x'z'_f < xz_f$.

If $k_f(x,y) + +k_f(y,z) = 0$ holds, there is nothing to prove. Otherwise, we may assume that, say, $k_f(x,y)$ is positive. Thus, we can choose some $a \in X(x,y)$ with

$$k_f(x,y) = 1 + k_f(xy, y|a).$$

Clearly, the four elements x, y, z, a satisfy the assumptions in Lemma 2.4 implying that either one of the three cases considered there must hold.

In case (1), we get

$$k_f(x,z) \geq 1 + k_f(x,a),\ \ k_f(x,y) \leq 1 + k_f(x,a)$$

(in view of $x \in \{t \in \{x,y\} : ta_f = xy_f - \Delta\}$), and $k_f(y,z) = 0$ which together implies our claim

$$k_f(x,z) \geq 1 + k_f(x,a) \geq k_f(x,y) = k_f(x,y) + k_f(y,z).$$

In case (2), we get

$$k_f(x,z) \geq 1 + k_f(z,a),\ \ k_f(x,y) \leq 1 + k_f(y,a)$$

(in view of $y \in \{t \in \{x, y\} : ta_f = xy_f - \Delta\}$), and $k_f(z, a) \geq k_f(z, y) + k_f(y, a)$ (in view of $za_f \leq xz_f - \Delta < xz_f$ and our induction hypothesis). Together, this implies also our claim

$$k_f(x, z) \geq 1 + k_f(z, a) \geq 1 + k_f(z, y) + k_f(y, a) \geq k_f(x, y) + k_f(z, y).$$

In case (3) finally, we get

$$k_f(y, z) = 0 \text{ and } ay + yz = az,$$

as before as well as

$$k_f(x, z) \geq 1 + \min(k_f(x, a), k_f(a, z)) \text{ and } k_f(x, y) = 1 + \min(k_f(x, a), k_f(a, y)).$$

Thus, our induction hypothesis implies

$$k_f(z, a) \geq k_f(z, y) + k_f(y, a) = k_f(y, a)$$

which implies in turn that

$$k_f(x, z) \geq 1 + \min(k_f(x, a), k_f(a, z)) \geq$$

$$1 + \min(k_f(x, a), k_f(a, y)) = k_f(x, y) = k_f(x, y) + k_f(z, y)$$

must hold in this case, too. Together, this establishes our theorem. ∎

References

[1] B. Bowditch, Notes on Gromov's hyperbolicity criterion for path metric spaces, in:E. Ghys et al. Group theory from a geometric viewpoint, (World Scientific, 1991) 64-167.

[2] A. Dress, K.T. Huber, J.H. Koolen, V. Moulton, and J. Weyer-Menkhoff, Δ additive and Δ ultra-additive maps, Gromov's trees, and the Farris transform, submitted

[3] J.S. Farris, On the phenetic approach to vertebrate classification, in Major patterns in vertebrate evolution, (Plenum, New York, 1977) 823-950.

[4] M. Gromov, Hyperbolic Groups, in:Essays in Group Theory, MSRI series vol. 8, (Springer-Verlag, 1988).