A CAT(0)-VALUED POINTWISE ERGODIC THEOREM

TIM AUSTIN

ABSTRACT. In this note we prove the following pointwise ergodic theorem for functions taking values in a separable complete CAT(0)-space:

Theorem. Suppose that $G$ is an amenable locally compact group with left Haar measure $m_G$ and that $(F_n)_{n \geq 1}$ is a tempered Følner sequence of compact subsets of $G$, that $T : G \curvearrowright (\Omega, \mathcal{F}, P)$ is a jointly measurable, probability-preserving action of $G$ on a probability space, that $(X, d)$ a separable complete CAT(0)-space with barycentre map $b$, and that $f : \Omega \to X$ is a measurable function such that for some (and hence any) fixed $x \in X$ we have

$$\int_{\Omega} d(f(\omega), x)^2 P(d\omega) < \infty.$$ 

Then as $n \to \infty$ the functions of empirical barycentres

$$\omega \mapsto b\left(\frac{1}{m_G(F_n)} \int_{F_n} \delta_{f(T_g \omega)} m_G(dy)\right)$$

converge pointwise for almost every $\omega$ to a $T$-invariant function $\bar{f} : \Omega \to X$.

1. INTRODUCTION

Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space and $(X, d)$ a complete separable CAT(0)-space (see, for instance, Bridson and Haefliger [1]). We write $L^2(P; X)$ for the space of all measurable maps $f : \Omega \to X$ such that for some fixed point $x \in X$ we have

$$\int_{\Omega} d(f(\omega), x)^2 P(d\omega) < \infty.$$ 

It is easy to see that in this case this actually holds for every $x \in X$, and that if $f, g \in L^2(P; X)$ then also

$$\int_{\Omega} d(f(\omega), g(\omega))^2 P(d\omega) < \infty.$$ 

If we now define

$$d_2(f, g) := \sqrt{\int_{\Omega} d(f(\omega), g(\omega))^2 P(d\omega)},$$

then this is a metric on $L^2(P; X)$ that is easily seen to be also complete and CAT(0), and separable if $\mathcal{F}$ is countably generated up to $P$-negligible sets.

In addition, let $P_2(X)$ be the collection of probability measures on $X$ with finite second moment, in that sense that $\mu \in P_2(X)$ if for some (and hence every) $x \in X$ we have

$$\int_X d(y, x)^2 \mu(dy) < \infty.$$ 

In these terms the condition that $f \in L^2(P; X)$ is equivalent to $f_\# P \in P_2(X)$, where $f_\# P$ is the pushforward measure of $P$ under $f$. 

1
The important geometric property of complete CAT(0)-spaces that motivates our work is that they support a sensible notion of averaging. More specifically, it has been known essentially since work of Cartan [2] that for any \( \mu \in P_2(X) \) there is a unique point \( x \in X \) for which the above integral is minimized (see Chapter II.2 of Bridson and Haefliger). We will refer to this as the barycentre of \( \mu \) and denoted it by \( b(\mu) \).

In terms of these barycentres we can now define a CAT(0)-notion of ergodic averages. Suppose that \( G \) is an amenable locally compact group with left-invariant Haar measure \( m_G \) that acts on \( (\Omega, F, P) \) through a jointly measurable, \( P \)-preserving action \( g \mapsto T^g \). Given this, for a measurable subset \( E \subseteq G \) with \( m_G(E) < \infty \) and a point \( \omega \in \Omega \) we will write
\[
\nu_{f,E}(\omega) = \frac{1}{m_G(E)} \int_E f(T^g \omega) m_G(dg)
\]
for the empirical measure of \( f \) across the associated orbit patch of \( T \): more explicitly, this is defined by
\[
\nu_{f,E}(\omega)(A) := \frac{1}{m_G(E)} m_G\{g \in E : f(T^g \omega) \in A\}.
\]
It is natural to view the function of barycentres of the empirical measures of \( f \),
\[
\omega \mapsto b(\nu_{f,E}(\omega))
\]
as a CAT(0) analog of the ergodic averages
\[
\frac{1}{m_G(E)} \int_E f(T^g \omega) m_G(dg)
\]
available in case \( f : \Omega \to \mathbb{R} \).

In the case of real-valued functions, it is known that if \( f \in L^2(P) \) then these ergodic averages converge for \( P \)-almost every \( \omega \in \Omega \) as \( E \) increases along a suitably chosen Følner sequence of subsets of \( G \). This follows from the classical pointwise ergodic theorem of Birkhoff in case \( G = \mathbb{Z} \), and has also long been known for many other concrete groups such as \( \mathbb{Z}^d \) or \( \mathbb{R}^d \). The general case was only established quite recently by Lindenstrauss, who found that the appropriate condition to place on the Følner sequence \( (F_n)_{n \geq 1} \) is that it be tempered: this holds if for some fixed \( C > 0 \) and all \( n \geq 1 \) we have
\[
m_G\left( \bigcup_{k<n} F_k^{-1} F_n \right) \leq C m_G(F_n)
\]
(this is also referred to as the 'Shulman condition').

By thinning out an initially-given Følner sequence if necessary it follows that any amenable group does admit Følner sequences satisfying this condition. Given such a sequence, Lindenstrauss proves (alongside other results) that for any \( f \in L^1(P) \) there is a \( T \)-invariant function \( \bar{f} : \Omega \to \mathbb{R} \) such that
\[
\frac{1}{m_G(F_n)} \int_{F_n} f(T^g \omega) m_G(dg) \to \bar{f}(\omega)
\]
for \( P \)-almost every \( \omega \in \Omega \).

In the present note we will show that this result can be extended to maps in \( L^2(P; X) \) for a separable CAT(0)-space \( X \), replacing ergodic averages with the empirical measure barycentres introduced above:
Theorem 1.1. If $T: G \curvearrowright (\Omega, \mathcal{F}, P)$, $(X, d)$ are as above, then for any $f \in L^2(P; X)$ there is a $T$-invariant function $\bar{f}: \Omega \to X$ such that for any tempered Følner sequence $(F_n)_{n \geq 1}$ of compact subsets of $G$ we have
\[
b(\nu_{f,E}(\omega)) \to \bar{f}(\omega)
\]
for $P$-a.e. $\omega \in \Omega$.

In particular, if $T$ is ergodic then $\bar{f}$ is just the constant function with value $b(f\#P) \in X$.

Remark  We restrict to separable $X$ in order to avoid discussing the nuances between different notions of ‘measurability’ for $f$, but provided the right notion is chosen this seems to make no real restriction.

We will find that this theorem follows quite quickly from an appeal to the real-valued pointwise ergodic theorem, together with an approximation argument based on a maximal ergodic theorem for such group actions and Følner sequences also obtained by Lindenstrauss in [4]. On the other hand, since any Hilbert space is CAT(0) with barycentre map simply given by averaging, Theorem 1.1 does contain the pointwise ergodic theorem for square-integrable maps into a separable Hilbert space as a special case. It seems to be harder to find a theorem about CAT(0) targets that encompasses the real-valued theorem for arbitrary functions in $L^1(P)$, since the condition that $\mu \in P_2(X)$ already appears in the definition of the barycentre $b(\mu)$.

Theorem 1.1 seems to be new even in the classical case $G = \mathbb{Z}$, but it does bear comparison with various other works relating probability-preserving actions to the geometry of CAT(0)-spaces; we refer the reader in particular to the proof by Karlsson and Margulis in [3] of an analog of Oseledets’ Theorem for cocycles over a probability-preserving action taking values in the semigroup of contractions of a CAT(0)-space.

In the next section we will recall the maximal ergodic theorem we need from Lindenstrauss [4] and derive from it a useful consequence for CAT(0)-valued maps, and then in Section 3 we will use these results to prove Theorem 1.1.

Acknowledgements  My thanks go to Lior Silberman for helpful suggestions.

2. A CAT(0)-valued maximal ergodic theorem

An important innovation behind Lindenstrauss’ proof of his pointwise ergodic theorem was a proof of the weak-(1, 1) maximal ergodic theorem in the setting of general amenable groups and tempered Følner sequences.

Let us introduce the standard notation $Mf$ for the ergodic maximal function associated to a Følner sequence and a function $f: \Omega \to \mathbb{R}$:
\[
Mf(\omega) := \sup_{n \geq 1} \frac{1}{m_G(F_n)} \int_{F_n} f(T^g\omega) \, m_G(\,dg).
\]
(see, for example, Peterson’s book [5] for background on this maximal function in the classical case $G = \mathbb{Z}$, or Section 3 of Lindenstrauss [4]).

Proposition 2.1 (Theorem 3.2 of [4]). Let $T: G \curvearrowright (\Omega, \mathcal{F}, P)$ be an action as above and \((F_n)_{n \geq 1}\) a tempered Følner sequence. Then there is a $c > 0$, depending on the sequence...
(\(F_n\))\(_{n \geq 1}\) but not on \(X\), such that for any \(f \in L^1(P)\) we have
\[
P\{\omega \in \Omega : Mf(\omega) > \alpha\} \leq \frac{c}{\alpha}\|f\|_1
\]
for all \(\alpha > 0\).

We will make use of this result to prove the following analog for CAT(0)-valued maps.

**Theorem 2.2 (CAT(0)-valued maximal ergodic theorem).** If \(f, h : \Omega \rightarrow X\) are two members of \(L^2(P; X)\) then
\[
P\{\omega \in \Omega : \sup_{n \geq 1} d(b(\nu_{f,F_n}(\omega)), b(\nu_{h,F_n}(\omega))) > \alpha\} \leq \frac{c}{\alpha^2}d_2(f, h)^2
\]
for every \(\alpha \in (0, \infty)\), where \(c\) is the same constant as in Proposition 2.1.

In order to prove this, we need an elementary result controlling the behaviour of barycentres in terms of the Wasserstein metric on \(P_2(X)\). Recall that this is defined for \(\mu, \nu \in P_2(X)\) by
\[
W_2(\mu, \nu) := \inf_{\lambda} \sqrt{\int X^2 d(x, y)^2 \lambda(dx, dy)}.
\]
As is standard, this defines a metric on \(P_2(X)\) (see, for instance, Villani [6]), and in conjunction with the CAT(0) condition it controls barycentres as follows.

**Lemma 2.3.** If \(\mu, \nu \in P_2(X)\) then
\[
d(b(\mu), b(\nu)) \leq W_2(\mu, \nu).
\]

**Proof** Since \(b(\nu)\) is the barycentre of \(\nu\), for any other fixed point \(x \in X\) we have
\[
d(x, b(\nu))^2 \leq \int_X d(x, y)^2 \nu(dy),
\]
and so integrating with respect to \(\mu\) we obtain
\[
\int_X d(x, b(\nu))^2 \mu(dx) \leq \int_X d(x, y)^2 \nu(dy)\mu(dx).
\]
On the other hand, by the definition of \(b(\mu)\) we have
\[
d(b(\mu), y)^2 \leq \int_X d(x, y)^2 \mu(dx)
\]
for any fixed \(y \in X\); taking \(y := b(\nu)\) and concatenating the above inequalities gives the result.

**Proof of Theorem 2.2** For any \(\alpha > 0\) the above lemma implies that
\[
\{\omega \in \Omega : \sup_{n \geq 1} d(b(\nu_{f,F_n}(\omega)), b(\nu_{h,F_n}(\omega))) > \alpha\} \subseteq \{\omega \in \Omega : \sup_{n \geq 1} W_2(\nu_{f,F_n}(\omega), \nu_{h,F_n}(\omega)) > \alpha\}.
\]
On the other hand, for each \(n\) the measure
\[
\lambda_n := \frac{1}{m_G(F_n)} \int_{F_n} \delta_{(f(T^s \omega), h(T^s \omega))} m_G(dy)
\]
on $X^2$ is clearly a joining of $\nu_{f,F_n}(\omega)$ and $\nu_{h,F_n}(\omega)$, and so

$$W_2(\nu_{f,F_n}(\omega),\nu_{h,F_n}(\omega))^2 \leq \frac{1}{m_G(F_n)} \int_{F_n} d(f(T^2\omega),h(T^2\omega))^2 m_G(dg).$$

If we now write $F(\omega) := d(f(T^2\omega),h(T^2\omega))^2$, then combining the above observations gives

$$\{\omega \in \Omega : \sup_{n \geq 1} d(b(\nu_{f,F_n}(\omega)),b(\nu_{h,F_n}(\omega))) > \alpha^2\} \subseteq \{\omega \in \Omega : MF(\omega) > \alpha\},$$

and so finally observing that $\|F\|_1 = d_2(f,h)^2$ completes the proof. \hfill \Box

3. PROOF OF THE MAIN THEOREM

Proof of the Theorem 1.1 Let us first prove the assertion of Theorem 1.1 for a finite-valued function $h : \Omega \to X$. To this is associated some finite measurable partition $\Omega = A_1 \cup A_2 \cup \ldots \cup A_m$ and collection of points $x_1, x_2, \ldots, x_k \in X$ such that $h(\omega) = x_i$ when $\omega \in A_i$. This case now follows easily from Lemma 2.3 and the real-valued pointwise ergodic theorem: given $\varepsilon > 0$, for almost every $\omega \in \Omega$ and every $i \leq k$, that theorem gives some $n(\omega, k, \varepsilon) \geq 1$ such that

$$n \geq n(\omega, k, \varepsilon) \implies \left| \frac{1}{m_G(F_n)} m_G\{g \in F_n : T^g\omega \in A_i\} - P(A_i) \right| < \varepsilon,$$

and so choosing $n$ sufficiently this we can make this hold for this $\varepsilon$ and $\omega$ and all $i \leq k$. It follows that the empirical measure

$$\nu_{h,F_n}(\omega) = \frac{1}{m_G(F_n)} \int_{F_n} \delta_{h(T^g\omega)} m_G(dg)$$

satisfies the total variation inequality

$$\|\nu_{h,F_n}(\omega) - P(A_1)\delta_{x_1} - P(A_2)\delta_{x_2} - \cdots - P(A_k)\delta_{x_k}\|_{\text{var}} < \varepsilon,$$

and hence that

$$W_2(\nu_{h,F_n}(\omega), P(A_1)\delta_{x_1} + P(A_2)\delta_{x_2} + \cdots + P(A_k)\delta_{x_k}) < \sqrt{\varepsilon} \text{ diam}\{x_1, x_2, \ldots, x_k\}.$$

Combining this with Lemma 2.3 and noting that

$$P(A_1)\delta_{x_1} + P(A_2)\delta_{x_2} + \cdots + P(A_k)\delta_{x_k}$$

is just the pushforward measure $h_{\#}P$, this shows that for almost every $\omega$, for every $\varepsilon > 0$ we have

$$d(b(\nu_{h,F_n}(\omega)), b(h_{\#}P)) < \sqrt{\varepsilon} \text{ diam}\{x_1, x_2, \ldots, x_k\}$$

for all sufficiently large $n$, and so since $\text{diam}\{x_1, x_2, \ldots, x_k\}$ is a fixed quantity for a given $h$ it follows that

$$b(\nu_{h,F_n}(\omega)) \to b(h_{\#}P)$$

as $n \to \infty$ for almost every $\omega$. 

Finally, the same assertion for an arbitrary \( f \in L^2(P; X) \) follows from a routine approximation argument and appeal to Theorem 2.2. Letting \( c > 0 \) be the constant of Theorem 2.1 and given \( \alpha > 0 \), since \( X \) is separable we can always find a finite-valued function \( \psi : \Omega \rightarrow X \) such that \( d_2(f, h) < \alpha^2 \), and hence

\[
P\left\{ \omega \in \Omega : \sup_{n \geq 1} d(b(\nu_{f_n}(\omega)) \rightarrow b(\eta) \} \leq \frac{c}{\alpha^2} d_2(f, h)^2 < c\alpha^2,
\]

so since we have seen that \( b(\nu_{f_n}(\omega)) \rightarrow b(\eta) \) for a conegligible set of \( \omega \in \Omega \) it follows that outside the above subset of \( \Omega \) of measure at most \( c\alpha^2 \), the sequence \( \{b(\nu_{f_n}(\omega))\}_{n=1}^\infty \) asymptotically oscillates by at most \( \alpha \) in \( X \), and so since \( \alpha \) was arbitrary this sequence must actually converge for almost every \( \omega \). This completes the proof. \( \square \)

**Remark** From another simple approximation by a finite-valued function as in the above proof and an appeal to the Dominated Convergence Theorem, it follows directly from Theorem 1.1 that the empirical barycentres of \( f \) also converge to \( \bar{f} \) in the metric space \((L^2(P; X), d_2)\).

\[ \triangleright \]

**References**

[1] M. R. Bridson and A. Haefliger. *Metric spaces of non-positive curvature*, volume 319 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[2] É. Cartan. *Leçons sur la géométrie des espaces de Riemann*. Gauthier-Villars, Paris, 1951. 2d ed.

[3] A. Karlsson and G. A. Margulis. A multiplicative ergodic theorem and nonpositively curved spaces. *Comm. Math. Phys.*, 208(1):107–123, 1999.

[4] E. Lindenstrauss. Pointwise theorems for amenable groups. *Invent. Math.*, 146(2):259–295, 2001.

[5] K. E. Petersen. *Ergodic Theory*. Cambridge University Press, Cambridge, 1983.

[6] C. Villani. *Topics in Optimal Transportation*. American Mathematical Society, Providence, 2003.