NOTE ON ROBUSTNESS OF PERIODIC MEASURES IN ERGODIC OPTIMIZATION

JAIRO BOCHI AND YIWEI ZHANG

Not intended for publication.

Let us recall the basic setting of ergodic optimization, referring the interested reader to [Je] for more information.

Let (X, d) be a compact metric space, and $T\colon X \to X$ be a continuous tranformation. Let $\mathcal{M}(X)$ denote the set of all Borel probability measures on X, and let $\mathcal{M}_T$ denote the subset of T-invariant ones. If $\mu \in \mathcal{M}_T$ is supported on a periodic measure then it is called a *periodic measure*.

Given a continuous function $f\colon X \to \mathbb{R}$, the *ergodic supremum* of f is defined as

$$\operatorname{erg\,sup}(f) \coloneqq \sup_{\mu \in \mathcal{M}_T} \langle f, \mu \rangle \,,$$

where angle brackets denote integration. If the sup is attained at $\mu \in \mathcal{M}_T$ then we say that the measure μ is *maximizing* for f. Such measures always exist.

Let $C^{\mathrm{Lip}}(X)$ denote the space of Lipschitz functions, endowed with the *Lipschitz norm* $\|\cdot\|_{\mathrm{Lip}} \coloneqq \|\cdot\|_\infty + \mathrm{Lip}(\cdot)$ that makes it a Banach space. Notice that the space $C^{\mathrm{Lip}}(X)$ is nonseparable unless X is countable, because the subset $\{d(x, \cdot)\,;\, x \in X\}$ is discrete.

We define subsets

$$C^{\mathrm{Lip}}(X) \supset \mathsf{P} \supset \mathsf{L}$$

as follows: P is the set of $f \in C^{\mathrm{Lip}}(X)$ that have a periodic maximizing measure μ. If in addition μ is the unique maximizing measure for f and for every function sufficiently close to f in the Lipschitz norm, then we write $f \in \mathsf{L}$. The letter L stands for "locking" property[1].

The aim of this note is to prove the following:

Proposition 1 (Yuan and Hunt)**.** *The set* L *equals the interior of* P*, and it is dense in* P*.*

This is basically Remark 4.5 in the paper [YH] by Yuan and Hunt, though these authors impose hyperbolicity hypotheses on the dynamics, and leave for the reader the task of adapting the arguments of their proof of a related fact. Though any specialist in ergodic optimization shouldn't have any difficulty in providing those details himself, we decided to write them for the following reasons:

- it was a nice exercise;
- we wanted to state the fact in our paper [BZ] (though we actually don't use it there);
- we weren't able to find any precise reference.

Date: January 21, 2015.

[1] A free translation of the *verrouillage* term used in [Bo1, §8].

We define the *Wasserstein distance* on $\mathcal{M}(X)$ as follows:

$$d_{\mathrm{W}}(\nu, \mu) := \sup_f \frac{\langle f, \nu - \mu \rangle}{\mathrm{Lip}(f)}, \tag{1}$$

where f runs over all non-constant Lipschitz functions. Wasserstein distances are an extensively studied subject: see e.g. [Vi] and references therein. Let us mention that d_{W} is indeed a distance function, which induces the weak topology on $\mathcal{M}(X)$; we won't need these facts, however. We also mention that $d_{\mathrm{W}}(\nu, \mu)$ equals the minimum "transport cost" between μ and ν when costs are proportional to distances; actually d_{W} is usually defined in this way, and then (1) becomes a consequence.

The Wasserstein distance was used in the context of ergodic optimization in the paper [Bo2].

We will need the following:

Lemma 2. *Let $\mu \in \mathcal{M}_T$ be a periodic measure, and let $\mathcal{O}_\mu$ be its support. Then there exists $C_\mu \ge 1$ such that for all $\nu \in \mathcal{M}_T$ we have*

$$d_{\mathrm{W}}(\nu, \mu) \le C_\mu \langle d(\cdot, \mathcal{O}_\mu), \nu \rangle. \tag{2}$$

Proof. Let p be the period of the orbit $\mathcal{O}_\mu$. If $p = 1$, i.e. $\mathcal{O}_\mu$ contains an unique point x_0, then for every $\nu \in \mathcal{M}(X)$ and $f \in C^{\mathrm{Lip}}(X)$ we have

$$\langle f, \nu \rangle \le \langle f(x_0) + \mathrm{Lip}(f) d(\cdot, x_0), \nu \rangle = \langle f, \mu \rangle + \mathrm{Lip}(f) \langle d(\cdot, x_0), \nu \rangle,$$

so (2) holds with $C_\mu := 1$.

From now on assume that $p > 1$. Let D be the diameter of X and let δ be the minimal distance between distinct points in $\mathcal{O}_\mu$. By uniform continuity, there exists $\varepsilon \in (0, D)$ such that:

$$\left. \begin{array}{l} x, y \in X, \ d(x,y) < \varepsilon \\ i \in \{0, 1, \dots, p-1\} \end{array} \right\} \Rightarrow d(T^i x, T^i y) < \frac{\delta}{2}.$$

Define $C_\mu := D/\varepsilon$. Let us check that inequality (2) is satisfied for every $\nu \in \mathcal{M}_T$. It is sufficient to consider ergodic ν; the general case will follow using ergodic decompositions and the fact that $d_{\mathrm{W}}(\mu, \cdot)$ is convex.

Fix a point $x \in X$ such that the Birkhoff averages of every continuous function f along the orbit of x converge to $\langle f, \nu \rangle$. We will inductively define a *transport sequence* $(y_i)_{i \ge -1}$ in $\mathcal{O}_\mu$. As an auxiliary device for the definition of the sequence, each integer $i \ge -1$ will be labelled as *good* or *bad*. The definition is as follows: The point $y_{-1} \in \mathcal{O}_\mu$ is chosen arbitrarily. The time -1 is labelled bad. Assume by induction that $y_{-1}, \dots, y_i$ are already defined (but y_{i+1} is not) and that the times $-1, \dots, i$ are already labelled (but $i + 1$ is not); then:

- If $d(T^{i+1}x, \mathcal{O}_\mu) < \varepsilon$ then each time $j \in \{i+1, i+2, \dots, i+p\}$ is labelled good, and y_j is defined as the unique point in $\mathcal{O}_\mu$ that is closest to $T^j x$. Notice that $y_j = T^{j-i} y_i$, and in particular each point of $\mathcal{O}_\mu$ appears exactly once in the list $y_{i+1}, y_{i+2}, \dots, y_{i+p}$.
- Else if $d(T^{i+1}x, \mathcal{O}_\mu) \ge \varepsilon$ then the time $i + 1$ is labelled bad, and we define y_{i+1} as $T(y_k)$, where k is the biggest bad time less than or equal to i.

This completes the definition of the transport sequence. Notice that it is equidistributed in the sense that:

$$\forall y \in \mathcal{O}_\mu, \quad \lim_{n \to \infty} \frac{1}{n} \# \big\{ i \in \{0, 1, \dots, n-1\} \, ; \, y_i = y \big\} = \frac{1}{p}.$$

Also notice that for all $i \geq 0$, the distance $d(T^i x, \mathcal{O}_\mu)$ equals $d(T^i x, y_i)$ if i is a good time, and is at least ε otherwise. In either case we have

$$d(T^i x, y_i) \leq C_\mu d(T^i x, \mathcal{O}_\mu) .$$

Using these properties we obtain, for every $f \in C^{\mathrm{Lip}}(X)$,

$$\begin{aligned}
\langle f, \nu \rangle &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) \\
&\leq \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \big[f(y_i) + \mathrm{Lip}(f)\, d(T^i x, y_i) \big] \\
&\leq \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \big[f(y_i) + C_\mu \,\mathrm{Lip}(f)\, d(T^i x, \mathcal{O}_\mu) \big] \\
&= \langle f, \mu \rangle + C_\mu \,\mathrm{Lip}(f) \, \langle d(\cdot, \mathcal{O}_\mu), \nu \rangle ,
\end{aligned}$$

which yields inequality (2). ☐

Remark 3. The lemma wouldn't be true replacing $\nu \in \mathcal{M}_T$ with $\nu \in \mathcal{M}(X)$. (Exercise.)

Proof of Proposition 1. By definition, the set L is open and is contained in P; once we show that it is dense in P it will follow that it is also the interior of P. So we are left to prove denseness.

Let $f \in \mathsf{P}$, let μ be a periodic maximizing measure for f, and let $\mathcal{O}_\mu$ be its support (which is finite). For $t > 0$, consider $f_t \coloneqq f - t d(\cdot, \mathcal{O}_\mu)$. These functions belong to the Banach space $C^{\mathrm{Lip}}(X)$ and converge to f as $t \to 0$. Moreover, for any $g \in C^{\mathrm{Lip}}(X)$ and $\nu \in \mathcal{M}_T$, using Lemma 2 and definition (1) we obtain

$$\begin{aligned}
\langle f_t + g, \nu \rangle &= \langle f, \nu \rangle + \langle g, \nu \rangle - t \langle d(\cdot, \mathcal{O}_\mu), \nu \rangle \\
&\leq \langle f, \mu \rangle + \langle g, \mu \rangle + (C_\mu \,\mathrm{Lip}(g) - t) \langle d(\cdot, \mathcal{O}_\mu), \nu \rangle \\
&= \langle f_t + g, \mu \rangle + (C_\mu \,\mathrm{Lip}(g) - t) \langle d(\cdot, \mathcal{O}_\mu), \nu \rangle .
\end{aligned}$$

Therefore if $\mathrm{Lip}(g) < t/C_\mu$ then μ is the unique maximizing measure for $f_t + g$. This shows that $f_t \in \mathsf{L}$ for any $t > 0$. So f belongs to the closure of L, as we wanted to show. ☐

References

[BZ] Bochi, J.; Zhang, Y. Ergodic optimization of prevalent super-continuous functions. Preprint arXiv:1501.00961

[Bo1] Bousch, T. La condition de Walters. *Ann. Sci. École Norm. Sup.* 34 (2001), no. 2, 287–311.

[Bo2] ———. Nouvelle preuve d'un théorème de Yuan et Hunt. *Bull. Soc. Math. France* 136 (2008), no. 2, 227–242.

[Je] Jenkinson, O. Ergodic optimization. *Discrete Contin. Dyn. Syst.* 15 (2006), no. 1, 197–224.

[Vi] Villani, C. *Optimal transport. Old and new.* Grundlehren der Mathematischen Wissenschaften, 338. Springer-Verlag, Berlin, 2009.

[YH] Yuan, G.; Hunt, B.R. Optimal orbits of hyperbolic systems. *Nonlinearity* 12 (1999), no. 4, 1207–1224.

Facultad de Matemáticas, Pontificia Universidad Católica de Chile