Counter-examples to an infinitesimal version of the Furstenberg conjecture

BENOÎT R. KLOECKNER

*Université Paris-Est, Laboratoire d'Analyse et de Matématiques Appliquées
(UMR 8050), UPEM, UPEC, CNRS, F-94010, Créteil, France
(e-mail: benoit.kloeckner@u-pec.fr)*

Abstract. In this paper we observe that one of our main results in 'Optimal transport and dynamics of circle expanding maps acting on measures' [*Ergod. Th. & Dynam. Sys.* **33**(2) (2013), 529–548] has an interesting consequence: an infinitesimal version of the Furstenberg conjecture is false in a very strong way. More precisely, we find deformations of the Lebesgue measure on the circle which are first-order invariant simultaneously for all integer multiplications modulo 1. We also correct an error in a lemma of the mentioned article. Both the proof and the statement must be corrected, but the main results of the article are not affected.

1. An infinitesimal Furstenberg conjecture

1.1. *Furstenberg conjecture.* While circle expanding maps have in many respects become toy models in the category of hyperbolic dynamical systems, a prominent question concerning them is still open for over half a century.

CONJECTURE 1.1. (Furstenberg) *If an atomless probability measure μ on the circle $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ is invariant under both*

$$\Phi_2 : x \mapsto 2x \mod 1 \quad and \quad \Phi_3 : x \mapsto 3x \mod 1,$$

then μ is equal to the Lebesgue measure λ.

In the above conjecture, one can replace 2 and 3 by two multiplicatively independent integers. Even the above case is wide open in general, though a theorem of Rudolph asserts that Furstenberg's conjecture holds for measures μ having positive entropy for one of the maps Φ_2 or Φ_3 [**Rud90**] (see also [**Joh92**]). Many other results related to this question can be found in the literature, among which are [**HS12**, **BLMV09**]; the interested reader can for example use the answers to the MathOverflow question [**Math14**] as pointers.

1.2. *A differential calculus perspective.* In [**Klo13**], we started a study of the action of circle expanding maps on the set of probability measure $\mathcal{P}(\mathbb{S}^1)$ in a differentiable setting relying on optimal transport. This setting does not truly give $\mathcal{P}(\mathbb{S}^1)$ the structure of an (infinite-dimensional) manifold, but it makes it possible to define differentiable maps and their derivatives. The main result of [Klo13] is that the push-forward action on measures $\Phi_\#$ of a C^2 expanding circle map Φ is derivable at its absolutely continuous invariant measure, with an explicit computation of its derivative. To see how relevant this result can be in the context of the Furstenberg conjecture, let us consider how one can approach this kind of problem in a differential geometric setting.

Furstenberg's conjecture is a strong rigidity statement; in differential geometry, a common strategy to attack such questions is to aim at weaker rigidity statements. A first weakening would be to ask whether the point known to have a given property of interest (here: λ) is, rather than unique, at least isolated among points with this property? If this stays out of reach, then can we prove that it is not possible to deform this point, i.e. to find a non-constant continuous path starting at this point inside the set defined by the given property? A further weakening is to ask for first-order rigidity, i.e. to ask whether we can use the tangent space and derivatives to prove that no C^1 deformation can exist in the considered set? In the case of the Furstenberg conjecture, we have a space $\mathcal{P}(\mathbb{S}^1)$ and two rather rich subspaces, the sets of atomless invariant measures for Φ_2 and Φ_3. Let us denote these sets of fixed measures by I_2 and I_3; then the conjecture is that $I_2 \cap I_3 = \lambda$. Imagine for a moment that I_2 and I_3 are some sort of differentiable submanifolds of $\mathcal{P}(\mathbb{S}^1)$; then the various above weakenings of Furstenberg's conjecture would take the form of the following questions:

(1) is λ isolated in $I_2 \cap I_3$?;
(2) is λ the sole point in its path-connected component inside $I_2 \cap I_3$?;
(3) must a C^1 curve starting at λ and lying inside $I_2 \cap I_3$ be constant?

Finally, to provide a positive answer to this third weakening, the most common approach would be to prove that the intersection $I_2 \cap I_3$ is 'first-order rigid' at λ, in the sense that the tangent spaces $T_\lambda I_2$ and $T_\lambda I_3$ intersect trivially.

1.3. *The infinitesimal Furstenberg conjecture.* Since I_2 and I_3 are defined (if we forget momentarily the atomless condition) as sets of fixed points for $\Phi_{2\#}$ and $\Phi_{3\#}$, the first-order rigidity question would reduce to ask whether the spaces $E_2, E_3 \subset T_\lambda \mathcal{P}(\mathbb{S}^1)$ of invariant vectors for the derivatives $D_\lambda \Phi_{2\#}$ and $D_\lambda \Phi_{3\#}$ intersect trivially. Even if all the above speculation turns out to be wrong (e.g. I_2 and I_3 could not be anything close to submanifolds), this last question is perfectly defined in the differential setting alluded to above, and can be considered an infinitesimal version of Furstenberg's conjecture. The main purpose of this paper is to prove that this question has a negative answer.

THEOREM 1.2. *The vector space $E_2 \cap E_3 \subset T_\lambda \mathcal{P}(\mathbb{S}^1)$ of tangent vectors at λ that are simultaneously invariant under both $D_\lambda \Phi_{2\#}$ and $D_\lambda \Phi_{3\#}$ is infinite-dimensional.*

The vector space $\bigcap_{d=2}^\infty E_d$ of tangent vectors at λ that are simultaneously invariant under all the $D_\lambda \Phi_{d\#}$ is two-dimensional.

In order to keep this paper short, and to avoid repetition, we refer to [**Klo13**] for the definition of derivatives and tangent spaces in the set of measures; but such a result could feel very abstract and potentially artificial, so let us give a direct corollary that contains no reference to optimal transport or abstract differential geometric setting. The idea behind this corollary goes back to an insight of Otto [**Ott01**] related to the point of view of Benamou and Brenier [**BB00**] and developed in [**AGS08**]: by integration, smooth test functions can serve as a kind of (weak) coordinates on $\mathcal{P}(\mathbb{S}^1)$; for simplicity, this corollary is phrased in a restricted way, using only that $\bigcap_{d=2}^{\infty} E_d$ is not reduced to 0.

COROLLARY 1.3. *There exists a path of probability measures $(\mu_t)_{t \in (-\varepsilon, \varepsilon)}$ with $\mu_0 = \lambda$, continuous in the weak topology, with μ_t atomless for almost all t, such that*

$$\frac{d}{dt} \int_{\mathbb{S}^1} \psi_0 \, d\mu_t \bigg|_{t=0} \neq 0$$

for some smooth function $\psi_0 : \mathbb{S}^1 \to \mathbb{R}$, and

$$\frac{d}{dt} \int_{\mathbb{S}^1} \psi \, d\mu_t \bigg|_{t=0} = \frac{d}{dt} \int_{\mathbb{S}^1} \psi \, d(\Phi_{d\#}\mu_t) \bigg|_{t=0}$$

for all smooth functions $\psi : \mathbb{S}^1 \to \mathbb{R}$ and all integers $d \geq 2$.

Remark 1.4.

(1) The first condition ensures that μ_t depends significantly on t (in particular, it avoids the degenerate and obvious choice $\mu_t \equiv \lambda$), while the second condition expresses that for small t, μ_t is 'almost invariant' under all the push-forward maps $\Phi_{d\#}$. Of course, this condition can be rewritten

$$\frac{d}{dt} \int_{\mathbb{S}^1} \psi \, d\mu_t \bigg|_{t=0} = \frac{d}{dt} \int_{\mathbb{S}^1} \psi \circ \Phi_d \, d\mu_t \bigg|_{t=0}.$$

(2) This corollary is intrinsically much weaker than the theorem, as differentiability in the sense of Wasserstein distance implies differentiability of the integrals of test functions, but the converse implication does not hold. For example, a curve of the form $(t\mu + (1-t)\nu)_t$ is usually not differentiable (or even rectifiable) in the differential structure induced by W_2, while the integral of any test function depends affinely on t. Nevertheless, we do not know a simpler way to get Corollary 1.3 even when restricting d to $\{2, 3\}$. Even if the Furstenberg conjecture were false and there were an atomless probability measure $\mu \neq \lambda$ invariant by Φ_2 and Φ_3, the curve $(t\lambda + (1-t)\mu)_t$ would not work as these measures are not positive for negative t.

(3) One could try to extend this infinitesimal argument to the construction of families of counter-examples to the Furstenberg conjecture: if one of the invariant vectors we found could be extended to a vector field preserved by both Φ_2 and Φ_3, then the integral curve issued from λ would be entirely made of invariant measures for both Φ_2 and Φ_3. However, it would be incredibly bold to conjecture this extension to be possible: we do not even know whether $\Phi_{2\#}$ is differentiable at any non-absolutely continuous measure. Note also that this extension cannot be expected at all for the full semigroup $\mathbb{N}$, as it is known that the Lebesgue measure is the only atomless measure invariant under all Φ_d (this holds more generally for large enough sub-semigroups

of $\mathbb{N}$; see [**EF10**]). One can still dream of making this approach work for finitely generated multiplicative sub-semigroups, as this case is very different from larger sub-semigroups: in the former case, the remainder in the first-order Taylor formula for $\Phi_{d\#}$ at λ can be made uniform over the generators (for a fixed simultaneously invariant tangent vector).

2. *Proofs*

Given the results of [**Klo13**], the only merit of Theorem 1.2 is its statement: its proof is completely straightforward.

Proof of Theorem 1.2. We only need to look closely at the expressions of $\mathcal{L}_d := D_\lambda \Phi_{d\#}$. As indicated in the proof of [**Klo13**, Proposition 4.4], setting $c_k(x) = \cos(2\pi k x)$ and $s_k(x) = \sin(2\pi k x)$, we have

$$\mathcal{L}_d(c_k) = \mathcal{L}_d(s_k) = 0$$

when k is not a multiple of d, and

$$\mathcal{L}_d(c_{md}) = d c_m, \quad \mathcal{L}_d(s_{md}) = d s_m$$

when m is a positive integer. Note that the union of $(c_k)_{k \geq 1}$ and $(s_k)_{k \geq 1}$ is a Hilbert basis of $T_\lambda \mathcal{P}(\mathbb{S}^1)$.

It follows that the 1-eigenspace E_d of $\mathcal{L}_d$ is generated by the functions

$$\sum_{m \geq 0} d^{-m} c_{d^m j} \quad \text{and} \quad \sum_{m \geq 0} d^{-m} s_{d^m j},$$

where j runs over positive integers not multiples of d.

Taking intersections, we first see that $E_2 \cap E_3$ is generated by the functions

$$\sum_{n,m \geq 1} 2^{-n} 3^{-m} c_{2^n 3^m j} \quad \text{and} \quad \sum_{n,m \geq 1} 2^{-n} 3^{-m} s_{2^n 3^m j},$$

where j is any positive integer prime with 6. In particular, $E_2 \cap E_3$ is infinite-dimensional.

We also get that $\bigcap_{d \geq 2} E_d$ is generated by

$$\sum_{p \geq 1} p^{-1} c_p \quad \text{and} \quad \sum_{p \geq 1} p^{-1} s_p$$

and is thus two-dimensional. Note that these functions are indeed in $L^2(\lambda)$, and therefore represent tangent vectors at λ. $\qquad \square$

The proof of the corollary is also easy; it relies mostly on the following pointwise version of the continuity equation. We do not claim any novelty in the lemma below, but still provide a simple proof of the simple case we need.

LEMMA 2.1. *Assume that $(\mu_t)_t$ is a curve of probability measures on $\mathbb{S}^1$ which is differentiable at 0 with tangent vector $v \in T_{\mu_0} \mathcal{P}(\mathbb{S}^1)$, in the sense that*

$$W_2(\mu_t, \mu_0 + tv) = o(t)$$

(be reminded that $\mu_0 + tv := (\mathrm{Id} + tv)_{\#} \mu_0$ is the image of tv by the exponential map at μ_0).

Then, for all smooth functions ψ, we have

$$\frac{d}{dt}\int \psi\, d\mu_t\bigg|_{t=0} = \int \psi' v\, d\mu_0.$$

Note that this lemma is mostly relevant in the case when μ_0 is regular in the sense of Gigli (i.e. in dimension one, atomless) since then all 'tangent vectors' at μ_0 are indeed represented by a vector field $v \in L^2(\mu_0)$. In a more general manifold, the same result would hold with ψ a compactly supported function, and $\nabla \psi$ instead of ψ'.

Proof. First, we observe that, denoting by Π_t an optimal transport plan from μ_t to $\mu_0 + tv$, we have

$$\left|\int \psi\, d\mu_t - \int \psi\, d(\mu_0 + tv)\right| = \left|\int \psi(x)\, d\Pi_t(x, y) - \int \psi(y)\, d\Pi_t(x, y)\right|$$

$$\leq \int |\psi(x) - \psi(y)|\, d\Pi(x, y)$$

$$\leq \sup(\psi')\int d(x, y)\, d\Pi(x, y)$$

$$\leq \sup(\psi')\left(\int d(x, y)^2\, d\Pi(x, y)\right)^{1/2}\left(\int 1\, d\Pi(x, y)\right)^{1/2}$$

$$= \sup(\psi')\, W_2(\mu_t, \mu_0 + tv)$$

$$= o(t),$$

so that we can use $\mu_0 + tv$ to estimate the derivative of the integral of ψ.

Next, we have

$$\int \psi\, d(\mu_0 + tv) - \int \psi\, d\mu_0 = \int (\psi \circ (\mathrm{Id} + tv) - \psi)\, d\mu_0$$

$$= \int (\psi(x + tv(x)) - \psi(x))\, d\mu_0$$

$$= \int \left(t\psi'(x)v(x) + \frac{1}{2}\psi''(c_x)t^2 v(x)^2\right) d\mu_0$$

$$= t\int \psi' v\, d\mu_0 + O(t^2).$$

Note that the finiteness of $\|v\|_{L^2(\mu_0)}$ is part of the definition of the tangent space at μ_0.

Now the claimed equality follows readily from these two estimates. $\qquad \square$

Proof of Corollary 1.3. Let $v \in \bigcap_{d \geq 2} E_d$ be a non-zero tangent vector at λ invariant under all the $D_\lambda \Phi_{d\#}$, and define $\mu_t = \lambda + tv$.

By definition of the tangent space at λ, $\int v\, d\lambda = 0$, so that v has a well-defined antiderivative. Let ψ_0 be a smooth approximation of one of its antiderivatives, so that $\int \psi_0' v\, d\lambda \simeq \int v^2\, d\lambda$ is non-zero.

Then the pointwise continuity equation implies that

$$\frac{d}{dt}\int \psi_0\, d\mu_t\bigg|_{t=0} = \int \psi_0' v\, d\lambda \neq 0.$$

Moreover, the invariance of v means that the curve $(\Phi_{d\#}\mu_t)$ is also differentiable at $t = 0$, with tangent vector v. In consequence, we get for all smooth functions ψ:

$$\frac{d}{dt}\int \psi \, d(\Phi_{d\#}\mu_t)\Big|_{t=0} = \int \psi' v \, d\lambda = \frac{d}{dt}\int \psi \, d\mu_t\Big|_{t=0}.$$

The weak continuity of (μ_t) is obvious, and the fact that μ_t is atomless for almost all t was proved in [**Klo13**]. $\qquad\qquad\square$

3. *Corrigendum*

The proof of [**Klo13**, Lemma 4.2] is flawed: the Wasserstein distance between $\rho\lambda$ and $\bar{\rho}\lambda$ is independent of t, so that the right-hand side of the first inline equation should be $\varepsilon t + 2^{-3/2}\varepsilon$ rather than $(1 + 2^{-3/2})\varepsilon t$. This mistake can be corrected by estimating how well a piecewise-constant density with k pieces of equal length can approximate the given density. Then the issue is moved to the main argument: in order to ultimately get a $o(t)$ remainder, we need to take advantage of the presence of many overlaps (as in [**Klo13**, Figure 2]), which only exist if k increases not too fast with respect to t. This can be ensured by adding a regularity hypothesis.

LEMMA 3.1. (Corrected version of [**Klo13**, Lemma 4.2]) *Given a* Hölder continuous *and* positive *density ρ, vector fields $v_1, \ldots, v_n \in L^2(\rho\lambda)$ and positive numbers $\alpha_1, \ldots, \alpha_n$ summing up to* 1*, one has*

$$\mathrm{W}\bigg(\rho\lambda + t\sum_i \alpha_i v_i, \sum_i \alpha_i(\rho\lambda + tv_i)\bigg) = o(t).$$

The positivity assumption may not be necessary, but at the very least simplifies the proof. Note that in [**Klo13**], we only used Lemma 4.2 for positive C^1 densities, so all main results stand without modification as soon as we prove the corrected lemma.

Proof. We prove the case $n = 2$, since the general case can then be deduced by induction.

Let ε be any positive number, and consider vector fields $\bar{v}_i$ ($i = 1, 2$) that are constant on the intervals $[j/k_1, (j+1)/k_1]$ for some k_1 and all $j < k_1$ and such that $\|\bar{v}_i - v_i\|_{L^2(\rho\lambda)} \le \varepsilon$. Note that k_1 and the $\bar{v}_i$ are chosen to depend only on ε, not on t; in particular, $\|\bar{v}_1 - \bar{v}_2\|_\infty$ is finite and independent of t.

Now consider any value of t, to be taken small enough a few times below. Let $k = k(t)$ be a multiple of k_1 having the magnitude of $(1/t)^{1/(1+\beta/2)}$, where β is the Hölder exponent of ρ, say $kt^{1/(1+\beta/2)} \in [1, 2]$.

We define $\bar{\rho}$ as the density that is constant on each $I_j = [j/k, (j+1)/k)$, of value $\bar{\rho}_j := k\int_{I_j} \rho \, d\lambda$. Denoting by C the Hölder constant of ρ, we get

$$\|\rho - \bar{\rho}\|_\infty \le Ck^{-\beta}.$$

We denote by $\bar{v}_i(j)$ the value of $\bar{v}_i$ on I_j; observe that when t is small, these values are the same on many successive intervals since k is much larger than k_1.

Let us first bound above $\mathrm{W}(\rho\lambda, \bar{\rho}\lambda)$. We consider the monotone rearrangement fixing 0 as transport plan; by definition of $\bar{\rho}$, it preserves each I_i. To simplify notation, let us bound

the cost due to the mass located in I_0, the other intervals behaving in exactly the same way. The cumulative distribution functions of $\rho\lambda$ and $\bar{\rho}\lambda$ are given by

$$F(x) = \int_0^x \rho \, d\lambda \quad \text{and} \quad G(x) = x\bar{\rho}_0.$$

The monotone rearrangement is given on I_0 by the map $T = G^{-1} \circ F$, so that the contribution of I_0 to its cost is

$$
\begin{aligned}
\int_0^{1/k} |T(x) - x|^2 \rho(x) \, dx &= \int_0^{1/k} \left| \frac{1}{\bar{\rho}_0} \int_0^x \rho \, d\lambda - \frac{1}{\bar{\rho}_0} \int_0^x \bar{\rho}_0 \, d\lambda \right|^2 \rho(x) \, dx \\
&\leq \int_0^{1/k} \left| \frac{1}{\bar{\rho}_0} \int_0^x |\rho - \bar{\rho}_0| \, d\lambda \right|^2 \rho(x) \, dx \\
&\leq \int_0^{1/k} \frac{C^2 x^2}{\bar{\rho}_0^2 k^{2\beta}} \rho(x) \, dx \\
&\leq \frac{C^2}{\bar{\rho}_0 k^{3+2\beta}}.
\end{aligned}
$$

Since the mass lying in I_0 is $\bar{\rho}_0/k$ (for both densities), the ratio cost per mass is bounded above by

$$\frac{C^2}{\bar{\rho}_0^2 k^{2+2\beta}} \leq \frac{C^2}{(\min \rho)^2 \, k^{2+2\beta}}.$$

Since this holds in all intervals I_i, the overall cost is bounded by the same value, so that

$$W(\rho\lambda, \bar{\rho}\lambda) \leq \frac{C}{\min \rho} \frac{1}{k^{1+\beta}}.$$

The same argument also yields

$$W(\rho\lambda + v, \bar{\rho}\lambda + v) \leq \frac{C}{\min \rho} \frac{1}{k^{1+\beta}}$$

for any vector field v which is constant on each I_j: indeed, if Π is a transport plan from $\rho\lambda$ to $\bar{\rho}\lambda$, then $(\text{Id} + v, \text{Id} + v)_\# \Pi$ is a transport plan from $\rho\lambda + v$ to $\bar{\rho}\lambda + v$ whose cost is not greater than the cost of Π (for each bit of mass moved from x to $T(x)$ by Π, this new plan moves the same amount of mass from $x + v(x)$ to $T(x) + v(T(x))$; the hypothesis that v is constant on each I_j then ensures that $v(T(x)) = v(x)$). Applying this to $v = t \sum \alpha_i \bar{v}_i$, we get

$$W\left(\rho\lambda + t \sum \alpha_i \bar{v}_i, \bar{\rho}\lambda + t \sum \alpha_i \bar{v}_i\right) \leq \frac{C}{\min \rho} \frac{1}{k^{1+\beta}}.$$

Applying the same reasoning to each $v = t\bar{v}_i$ separately and concatenating the corresponding transport plan also yields

$$W\left(\sum \alpha_i (\rho\lambda + t\bar{v}_i), \sum \alpha_i (\bar{\rho}\lambda + t\bar{v}_i)\right) \leq \frac{C}{\min \rho} \frac{1}{k^{1+\beta}}.$$

We can now use the main idea of the original lemma [**Klo13**, Figure 2] to get the bound

$$W\left(\bar{\rho}\lambda + t \sum \alpha_i \bar{v}_i, \sum \alpha_i (\bar{\rho}\lambda + t\bar{v}_i)\right) \leq t^{3/2} k^{1/2} \|\bar{v}_1 - \bar{v}_2\|_\infty^{3/2}.$$

Indeed, for each I_j the construction pictured in Figure 2 in [**Klo13**] gives a transport plan from $(\mathrm{Id} + t(\alpha_1 v_1(j) + \alpha_2 v_2(j)))_\# \bar{\rho}_j \lambda_{|I_j}$ to $\alpha_1(\mathrm{Id} + t v_1(j))_\# \bar{\rho}_j \lambda_{|I_j} + \alpha_2(\mathrm{Id} + t v_2(j))_\# \bar{\rho}_j \lambda_{|I_j}$ whose cost is at most $t^3 \bar{\rho}_j |\bar{v}_1(j) - \bar{v}_2(j)|^3$ for a contribution to the mass of $\bar{\rho}_j / k$.

The last estimates we need are

$$\mathrm{W}\Big(\rho\lambda + t \sum \alpha_i v_i, \ \rho\lambda + t \sum \alpha_i \bar{v}_i\Big) \le \varepsilon t$$

and

$$\mathrm{W}\Big(\sum \alpha_i (\rho\lambda + t\bar{v}_i), \ \sum \alpha_i (\rho\lambda + t v_i)\Big) \le \varepsilon t.$$

They are both obtained, as in the original proof, by observing that for any measure μ and any vector fields v, $\bar{v}$ in $L^2(\mu)$, the transport plan $(\mathrm{Id} + v, \mathrm{Id} + \bar{v})_\# \mu$ has cost exactly $\|v - \bar{v}\|_{L^2(\mu)}$.

Using the triangle inequality to combine all these estimates, we get

$$\mathrm{W}\Big(\rho\lambda + t \sum \alpha_i v_i, \ \sum \alpha_i (\rho\lambda + t v_i)\Big)$$
$$\le 2\varepsilon t + t^{3/2} k^{1/2} \|\bar{v}_1 - \bar{v}_2\|_\infty^{3/2} + \frac{2C}{\min \rho} \frac{1}{k^{1+\beta}}$$
$$= 2\varepsilon t + O(t^{(4+3\beta)/(4+\beta)}) + O(t^{(2+2\beta)/(2+\beta)}) = o(t). \qquad \square$$

REFERENCES

[**AGS08**]   L. Ambrosio, N. Gigli and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures (Lectures in Mathematics, ETH Zürich)*, 2nd edn. Birkhäuser, Basel, 2008.

[**BB00**]   J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Math.* **84**(3) (2000), 375–393.

[**BLMV09**]   J. Bourgain, E. Lindenstrauss, P. Michel and A. Venkatesh. Some effective results for $\times a \times b$. *Ergod. Th. & Dynam. Sys.* **29**(6) (2009), 1705–1722.

[**EF10**]   M. Einsiedler and A. Fish. Rigidity of measures invariant under the action of a multiplicative semigroup of polynomial growth on $\mathbb{T}$. *Ergod. Th. & Dynam. Sys.* **30**(1) (2010), 151–157.

[**HS12**]   M. Hochman and P. Shmerkin. Local entropy averages and projections of fractal measures. *Ann. of Math. (2)* **175**(3) (2012), 1001–1059.

[**Joh92**]   A. S. A. Johnson. Measures on the circle invariant under multiplication by a nonlacunary subsemigroup of the integers. *Israel J. Math.* **77**(1–2) (1992), 211–240.

[**Klo13**]   B. Kloeckner. Optimal transport and dynamics of expanding circle maps acting on measures. *Ergod. Th. & Dynam. Sys.* **33**(2) (2013), 529–548.

[**Math14**]   math*overflow*, user39115 *et al* Furstenberg $\times 2 \times 3$ conjecture, bibliography, http://mathoverflow.net/questions/161517/furstenberg-times-2-times-3-conjecture-bibliography, posted 26 March 2014.

[**Ott01**]   F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations* **26**(1–2) (2001), 101–174.

[**Rud90**]   D. J. Rudolph. $\times 2$ and $\times 3$ invariant measures and entropy. *Ergod. Th. & Dynam. Sys.* **10**(2) (1990), 395–406.