Math. Res. Lett. **15** (2008), *no.* 5, 893–900

A PARTIAL ORDER ON ×2-INVARIANT MEASURES

Oliver Jenkinson

Dedicated to the memory of Bill Parry

Abstract. We introduce a partial order on the set of ×2-invariant probability measures, describing relative dispersion, and show that the minimal elements for this order are the Sturmian measures of Morse & Hedlund.

1. Introduction

Iteration of the "×2 map" $T(x) = 2x \pmod 1$ provides a favourite example of a dynamical system exhibiting a variety of complicated behaviour. The richness of the dynamics of T is reflected by the diversity of its invariant sets, and more generally of its invariant (Borel) probability measures. Much is known about the set of ×2-invariant measures, though open problems remain, such as the determination of those members which are also ×3-invariant, see e.g. [13, 18, 32].

Here we shall be interested in the *relative dispersion* of ×2-invariant measures, declaring one measure to be more diffuse than another if, roughly speaking, its mass lies closer to the boundary of $[0, 1)$. With this in mind, it is convenient to extend T to the closed interval $X = [0, 1]$ by defining $T(1) = 1$, thus effecting a (weak-$*$) compactification of the set of invariant measures. The resulting simplex $\mathcal{M}$ of T-invariant Borel probability measures on X is naturally identified with the set of shift-invariant Borel probability measures on $\{0, 1\}^{\mathbb{N}}$.

An established notion of the relative dispersion of two probability measures μ and ν is the following: we say that μ is *majorized*[1] by ν, and write $\mu \prec \nu$, if $\mu(f) \leq \nu(f)$ for every convex function $f : X \to \mathbb{R}$. Among all (not necessarily invariant) probability measures on X, the least diffuse are the Dirac masses: indeed Jensen's inequality implies that the Dirac mass δ_ϱ is majorized by every probability measure μ whose barycentre $b(\mu) := \int x \, d\mu(x)$ equals ϱ. The most diffuse probability measures are those concentrated on the boundary of X: $\nu_\varrho := (1 - \varrho)\delta_0 + \varrho\delta_1$ majorizes every μ with $b(\mu) = \varrho$.

If, however, we restrict attention to probability measures invariant under some given transformation of X, it is usually a non-trivial problem to determine which measures in this subset are the most, or the least, diffuse. In the case of the ×2 map, the *most* diffuse measures are in fact easily identified: both endpoints 0 and

Received by the editors September 18, 2007.

[1] While its roots lie in economics [12, 25, 29], the mathematical foundations of majorization were laid by Hardy, Littlewood & Pólya [15], and later developed notably in [3, 9, 10]. Majorization is the basis for numerous inequalities in pure mathematics (see [27]), while in various applications it is regarded as the proper indicator of relative dispersion (see e.g. [27, 31]).

1 are fixed, so each measure ν_ϱ is invariant, and majorizes all other members of $\mathcal{M}_\varrho := \{\mu \in \mathcal{M} : b(\mu) = \varrho\}$.[2] In other words, each ordered set $(\mathcal{M}_\varrho, \prec)$ has a greatest element, namely ν_ϱ, and these measures are precisely the maximal elements of $(\mathcal{M}, \prec)$. Determining the *minimal* elements of $(\mathcal{M}, \prec)$ is more difficult, however, since the Dirac measures δ_ϱ are not invariant (unless $\varrho = 0$ or 1). In particular it is unclear, a priori, whether or not each $(\mathcal{M}_\varrho, \prec)$ has a least element.

Remarkably, it turns out that each $(\mathcal{M}_\varrho, \prec)$ does have a least element, which moreover can be identified explicitly:

Theorem 1. *For every $\varrho \in [0,1]$, the ordered set $(\mathcal{M}_\varrho, \prec)$ has a least element. This least element is the Sturmian measure S_ϱ of rotation number ϱ.*

The Sturmian measure of rotation number ϱ is defined in terms of a standard symbolic coding procedure induced by the circle rotation of angle ϱ. Specifically, for $\varrho \in (0,1)$ it is the push forward of Lebesgue measure on X under the map $x \mapsto \sum_{n\geq 0} \chi_{[1-\varrho,1)}(\{x+n\varrho\})/2^{n+1}$, where $\{\,\cdot\,\}$ denotes reduction modulo 1, and for $\varrho = 0$ or 1 it is the Dirac measure supported on the corresponding fixed point. For rational ϱ the Sturmian measure is supported on a single periodic orbit, while for irrational ϱ its support is a uniquely ergodic Cantor set. For example the Sturmian measures of rotation numbers 1/2, 1/3, 2/5, 3/8 and 5/13 are, respectively, the periodic orbit measures corresponding to the strings

$$01\,,\ 001\,,\ 00101\,,\ 00100101\,,\ 0010010100101\,,$$

while the Sturmian measure of rotation number $(3-\sqrt{5})/2$ is the unique invariant probability measure supported by the shift orbit closure of

$$0010010100100101001010010010100101\ldots$$

The terminology *Sturmian* follows Morse & Hedlund [28], who coined the term to describe (symbolically) the points in the support of Sturmian measures, so-called *Sturmian sequences*. A Sturmian sequence of rotation number ϱ is a sequence on the alphabet $\{0,1\}$ in which the symbol 1 occurs with frequency ϱ, and as "regularly" as possible[3]. For various other characterizations see e.g. [1, 4, 8, 20, 26, 28, 30].

A result of Cartier [9] asserts that $\mu \prec \nu$ if and only if ν is a *dilation* of μ, i.e. there exists a family of probability measures $(D_x)_{x\in X}$, with each $b(D_x) = x$, such that if $f : X \to \mathbb{R}$ is bounded and Borel then so is $x \mapsto D_x(f)$, and $\nu(f) = \int D_x(f)\,d\mu(x)$. Combining this characterization with Theorem 1 gives:

Corollary 1. *If $f : X \to \mathbb{R}$ is strictly convex then for every $\varrho \in [0,1]$, the Sturmian measure S_ϱ has strictly smaller f-integral than any other measure in $\mathcal{M}_\varrho$.*

The proof follows from the fact that Jensen's inequality is strict, i.e. $D_x(f) > f(x)$, whenever f is strictly convex and D_x is not the Dirac measure at x. In particular the *variance* $\mathrm{var}(\mu) = \int (x - b(\mu))^2\,d\mu(x)$ is minimized precisely when μ is Sturmian:

[2] An equivalent definition of $\mathcal{M}_\varrho$ is as the set of invariant measures giving weight ϱ to the symbol "1" (i.e. to the half-interval $[1/2,1]$), since the identity on X differs from the indicator function $\chi_{[1/2,1]}$ by $T - \mathrm{id}$, which has zero mean for each invariant measure.

[3] Morse & Hedlund used the term *mechanical* to reflect the regularity of such sequences. Many authors prefer the term *balanced*, reserving the terminology Sturmian for the case of irrational ϱ (see e.g. [1, 26, 30]).

Corollary 2. *For every $\varrho \in [0,1]$, the Sturmian measure of rotation number ϱ is the unique measure with smallest variance in $\mathcal{M}_\varrho$.*

Another consequence is that, among all periodic orbits with a given arithmetic mean, the Sturmian orbit has largest geometric mean:

Corollary 3. *Let $x_1, \ldots, x_Q$ be any non-Sturmian periodic orbit. If $s_1, \ldots, s_q$ denotes the Sturmian orbit of rotation number $p/q = Q^{-1}\sum_{i=1}^{Q} x_i$, then $(\prod_{i=1}^{q} s_i)^{1/q} > (\prod_{i=1}^{Q} x_i)^{1/Q}$.*

The original motivation for Theorem 1 was its application to ergodic optimization (see e.g. [5, 6, 11, 21]): for a bounded Borel function $f : X \to \mathbb{R}$, a measure $\mu \in \mathcal{M}$ is said to be *minimizing* if $\mu(f) = \inf_{m \in \mathcal{M}} m(f)$ and *maximizing* if $\mu(f) = \sup_{m \in \mathcal{M}} m(f)$. Since $\varrho \mapsto S_\varrho(f) = \inf_{\mu \in \mathcal{M}_\varrho} \mu(f)$ is convex, and strictly so when f is strictly convex (a simple consequence of Corollary 1), we deduce[4]:

Corollary 4. *If $f : X \to \mathbb{R}$ is convex then it has a Sturmian minimizing measure. If it is strictly convex then it has a unique minimizing measure, and this measure is Sturmian.*

The appearance of Sturmian measures as minimizing and maximizing measures for natural classes of functions is not altogether new: Bousch [4] has shown that every degree-one trigonometric polynomial has unique minimizing and maximizing measures, both of which are Sturmian (this corresponds to the fact that Sturmian measures describe the boundary of the "poisson" $\{\int \exp(2\pi i t)\, d\mu(t) : \mu \in \mathcal{M}\} \subset \mathbb{C}$). A significant difference is that, for convex functions, the Sturmian nature of the minimizing measure is a consequence of solving a family of restricted variational problems. For degree-one trigonometric polynomials f there is no analogue of Theorem 1: it is not the case that every Sturmian measure has smaller f-integral than all other invariant measures of same barycentre.

Investigation of the ordered set $(\mathcal{M}, \prec)$ is facilitated by the fact that certain ×2-invariant measures are known "explicitly" (Lebesgue measure, periodic orbit measures), while others can be well-approximated, either by periodic orbit measures or via explicit symbolic coding. For example, part of $(\mathcal{M}_{1/2}, \prec)$ is depicted in the Hasse diagram[5] of Figure 1; here each binary string represents the invariant probability measure supported by the corresponding periodic orbit, so for example Morse measure (concentrated on the shift orbit closure of the Morse sequence 0110100110010110…) majorizes $01 = \{1/3, 2/3\}$, is majorized by $000111 = \{1/9, \ldots, 8/9\}$, and is incomparable with $0011 = \{1/5, 2/5, 3/5, 4/5\}$.

To derive the above relations, and more generally to decide if and how two measures in $\mathcal{M}$ are related, one may exploit the characterisation (see [23, 24, 31]) that $\mu \prec \nu$ if and only if $b(\mu) = b(\nu)$ and $\int_0^t \mu[0,x]\, dx \le \int_0^t \nu[0,x]\, dx$ for all $t \in [0,1]$. In the case where μ and ν are purely atomic and each atom has equal weight (e.g. if they are periodic orbit measures of the same period), this recovers the famous Hardy-Littlewood-Pólya criterion for majorization (see [15, 16]): if $b(\mu) = b(\nu)$ then $\mu \prec \nu$

[4]Note that determining the *maximizing* measure for a convex function is trivial, because the endpoints of X are both fixed by T.

[5]See [22] for a Hasse diagram of a larger portion of $(\mathcal{M}_{1/2}, \prec)$.

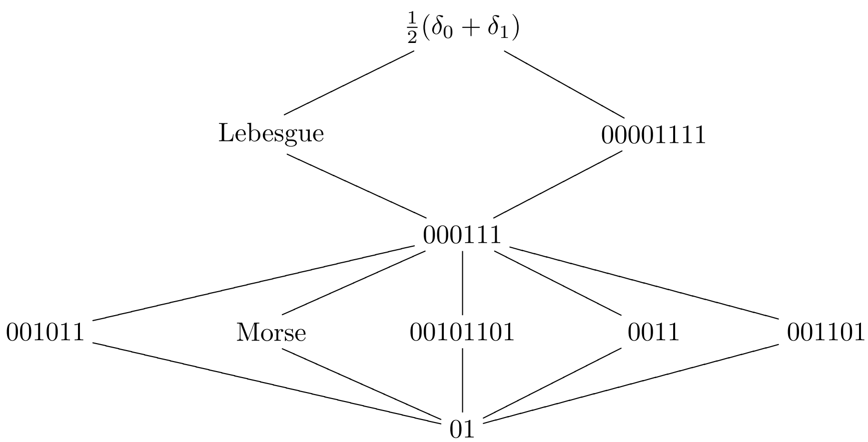

FIGURE 1. Hasse diagram of part of $\mathcal{M}_{1/2}$

if and only if $\sum_{i=1}^n \mu_i \geq \sum_{i=1}^n \nu_i$ for all $1 \leq n \leq Q-1$, where $\mu := Q^{-1}\sum_{i=1}^Q \delta_{\mu_i}$, $\nu := Q^{-1}\sum_{i=1}^Q \delta_{\nu_i}$, with $\mu_1 \leq \ldots \leq \mu_Q$ and $\nu_1 \leq \ldots \leq \nu_Q$. For example, combining this criterion with Theorem 1 yields:

Corollary 5. *Let $\mu_1 < \ldots < \mu_Q$ be any periodic orbit, and $p/q = Q^{-1}\sum_{i=1}^Q \mu_i$ its arithmetic mean, where $1 \leq p < q$ are coprime integers. If $s_1 < \ldots < s_q$ are the points in the Sturmian orbit of rotation number p/q, and $s'_{i+(j-1)q} := s_j$ for $1 \leq i \leq q$, $1 \leq j \leq Q/q$, then*

$$\sum_{i=1}^n s'_i \geq \sum_{i=1}^n \mu_i \quad \text{for all} \quad 1 \leq n \leq Q-1 \,. \tag{1}$$

If the orbit $\mu_1 < \ldots < \mu_Q$ is non-Sturmian then the extreme cases $n = 1$ and $n = Q-1$ of (1) clearly become strict inequalities, $s_1 > \mu_1$ and $s_q < \mu_Q$; these two inequalities were first proved, using different methods, by Bernhardt [2] (see also [14]). In fact a modification of the proof below yields a stronger result, namely that s_1 is strictly greater than $\min(\mathrm{supp}(\mu))$, and s_q is strictly smaller than $\max(\mathrm{supp}(\mu))$, for any non-Sturmian measure $\mu \in \mathcal{M}_{p/q}$.

Notation. The Lebesgue measure of a set E is denoted by $|E|$, and the Lebesgue integral of a function g by $\int g$. For other measures μ we write either $\mu(g)$ or $\int g\, d\mu$ for the integral of g.

2. Proof of Theorem 1

If ϱ equals 0 or 1 then $\mathcal{M}_\varrho$ is a singleton, and Theorem 1 is trivially true, so suppose $\varrho \in (0,1)$. Since C^2 convex functions are weakly dense among all convex functions, it suffices to show that if f is C^2 convex then $S_\varrho(f) \leq \mu(f)$ for all $\mu \in \mathcal{M}_\varrho$. If $f_\theta : X \to \mathbb{R}$ is defined, for $\theta \in \mathbb{R}$, by $f_\theta(x) = f(x) + \theta x$, this is equivalent to

$$S_\varrho(f_\theta) \leq \mu(f_\theta) \quad \text{for all } \mu \in \mathcal{M}_\varrho \,. \tag{2}$$

The strategy, therefore, will be to find, for each $\varrho \in (0,1)$, a value $\theta \in \mathbb{R}$ for which (2) holds. In fact θ can be chosen with the stronger property that S_ϱ is a minimizing measure for f_θ, a result described by the following Theorem 2, which in particular implies Theorem 1.

Theorem 2. *Let $f : X \to \mathbb{R}$ be a C^2 convex function. For every $\varrho \in (0,1)$ there exists $\theta \in \mathbb{R}$ such that the Sturmian measure S_ϱ is a minimizing measure for the function f_θ defined by $f_\theta(x) = f(x) + \theta x$.*

Proof. There exists $\gamma \in (0,1/2)$ such that the support of S_ϱ is contained in $H_\gamma = [\gamma, \gamma + 1/2]$ (see e.g. [7, 8, 14, 34]).[6] Define $\tau = \tau_\gamma : X \to X$ by

$$\tau_\gamma(x) = \begin{cases} (x+1)/2 & \text{if } x \in [0, 2\gamma) \\ x/2 & \text{if } x \in [2\gamma, 1]\,. \end{cases}$$

If $\theta := -\int \sum_{n=1}^\infty 2^{-n} f' \circ \tau^n$ then the L^∞ function $\sum_{n=1}^\infty 2^{-n} f_\theta' \circ \tau^n$ has Lebesgue integral zero, so is the almost everywhere derivative of a Lipschitz function $\varphi_\theta : X \to \mathbb{R}$ satisfying $\varphi_\theta(0) = \varphi_\theta(1)$. Now $(f_\theta + \varphi_\theta - \varphi_\theta \circ T)' = 0$ Lebesgue almost everywhere on H_γ, because $\tau \circ T$ is, Lebesgue almost everywhere, the identity function on H_γ. Therefore, since $f_\theta + \varphi_\theta - \varphi_\theta \circ T$ is absolutely continuous on X, its restriction to H_γ is constant. We claim that the constant value taken by $f_\theta + \varphi_\theta - \varphi_\theta \circ T$ on H_γ is in fact its minimum, i.e. that

$$(f_\theta + \varphi_\theta)(s) \le \begin{cases} (f_\theta + \varphi_\theta)(s + 1/2) & \text{for } s \in (\gamma, 1/2]\,, \\ (f_\theta + \varphi_\theta)(s - 1/2) & \text{for } s \in (1/2, \gamma + 1/2)\,. \end{cases} \tag{3}$$

From this it follows that S_ϱ is a minimizing measure for $f_\theta + \varphi_\theta - \varphi_\theta \circ T$, and hence for f_θ.

To prove the first inequality in (3) let $s \in (\gamma, 1/2]$, so that

$$\begin{aligned} (f_\theta + \varphi_\theta)(s) - (f_\theta + \varphi_\theta)(s + 1/2) &= \int_\gamma^s (f_\theta + \varphi_\theta)' - \int_{\gamma+1/2}^{s+1/2} (f_\theta + \varphi_\theta)' \\ &= \sum_{n=0}^\infty \left[\int_E 2^{-n} f_\theta' \circ \tau^n - \int_{E'} 2^{-n} f_\theta' \circ \tau^n \right] \\ &= \sum_{n=0}^\infty \left[\int_{\tau^n(E)} f_\theta' - \int_{\tau^n(E')} f_\theta' \right] \\ &= \int C_s \cdot f_\theta'\,, \end{aligned}$$

where

$$C_s := \sum_{n=0}^\infty \left[\chi(\tau^n E) - \chi(\tau^n E') \right],\ E = E_s := (\gamma, s],\ E' = E_s' := (\gamma + 1/2, s + 1/2]\,.$$

Now $C_s(0) = 0$, and C_s is Lebesgue-integrable, with $\int C_s = 0$, because $|\tau^n E| = 2^{-n}|E| = |\tau^n E'|$ for all $n \ge 0$. If $B_s(t) := \int_0^t C_s$ then $B_s(0) = 0 = B_s(1)$, so

[6] The value γ is unique if and only if ϱ is irrational. S_ϱ is the unique invariant measure for the continuous degree-one map of the circle ($[0,1]$ with endpoints identified) whose restriction to H_γ is T and which is constant on the complement of H_γ.

integration by parts yields $(f_\theta+\varphi_\theta)(s)-(f_\theta+\varphi_\theta)(s+1/2) = -\int B_s\cdot f_\theta''$. Now f is convex, therefore $f_\theta''=f''\ge 0$, so the required inequality will follow if it can be shown that B_s is non-negative on X.

For this, first note that C_s is identically zero on $[0,\gamma]$, hence so is B_s. Since E' and H_γ are disjoint, $\tau^m(E')\cap\tau^n(E')=\emptyset$ for $m\neq n$, so $\sum_{n=0}^\infty \chi(\tau^n(E'))\le 1$. In particular, $C_s\ge 1-\sum_{n=0}^\infty\chi(\tau^n(E'))\ge 0$ on $E=(\gamma,s]$, hence $B_s\ge 0$ on $(\gamma,s]$ as well. So $B_s(t)\ge 0$ for $t\in[0,s]$. Now C_s is identically zero on $(s+1/2,1]$, and equal to -1 on $E'=(\gamma+1/2,s+1/2]$, so $B_s(t)=\int_0^t C_s = -\int_t^1 C_s = s+1/2-t>0$ for $t\in[\gamma+1/2,s+1/2)$, and $B_s(t)=0$ for $t\in[s+1/2,1]$. If $t\in(s,\gamma+1/2)$ then

$$\begin{aligned} B_s(t) &= \int_0^t C_s = \int_\gamma^t C_s \\ &= \sum_{n=0}^\infty \int_\gamma^t \chi(\tau^n(E))-\chi(\tau^n(E')) \\ &\ge |(\gamma,t]\cap E| - |(\gamma,t]\cap E'| - \sum_{n=1}^\infty |(\gamma,t]\cap\tau^n E'| \\ &\ge |E| - \sum_{n=1}^\infty |\tau^n E'| = |E|(1-\sum_{n=1}^\infty 2^{-n}) = 0\,. \end{aligned}$$

The second inequality in (3) is proved similarly: if $s\in(1/2,\gamma+1/2)$ then

$$(f_\theta+\varphi_\theta)(s)-(f_\theta+\varphi_\theta)(s-1/2) = \int \tilde{C}_s\cdot f_\theta' = -\int \tilde{B}_s\cdot f''\,,$$

where $\tilde{C}_s := \sum_{n=0}^\infty \chi(\tau^n(D'))-\chi(\tau^n(D))$, $D'=[s-1/2,\gamma)$, $D=[s,\gamma+1/2)$, $\tilde{B}_s(t):=\int_0^t\tilde{C}_s$, and an argument analogous to the one above can be used to show that $\tilde{B}_s$ is non-negative on X. □

Remark 1.
(a) The choice of θ in Theorem 2 is inspired by Bousch's *précondition de Sturm* [4], cf. [17].
(b) If $f_\theta+\varphi_\theta$ happens to itself be convex (in general it is not, despite its second derivative being Lebesgue almost everywhere positive), the key inequality (3) is immediate.

Acknowledgments

This project began in 1997, from work [19, 20] towards the conjecture that Sturmian measures describe the boundary of the "poisson", a result discovered independently and subsequently proved by Thierry Bousch [4]. It gained impetus when Peter Gibson and Mike Taylor informed me of the work of Hardy, Littlewood & Pólya relating doubly stochastic matrices to majorization. I am grateful to Thierry Bousch for encouraging me to pursue the project, and for sharing his ideas on how to resolve it; in particular, the proof given here follows a suggestion of Thierry. Since 2003 I have benefited from EPSRC support in the form of an Advanced Research Fellowship.

References

[1] J.-P. Allouche & J. Shallit, *Automatic sequences. Theory, applications, generalizations*, Cambridge University Press, Cambridge, 2003.

[2] C. Bernhardt, *A class of endomorphisms of the circle*, Proc. London Math. Soc., **45** (1982), 258–280.

[3] D. Blackwell, *Equivalent comparisons of experiments*, Ann. Math. Statist. **24** (1953), 265–272.

[4] T. Bousch, *Le poisson n'a pas d'arêtes*, Ann. Inst. Henri Poincaré (Proba. et Stat.) **36**, (2000), 489–508.

[5] ———, *La condition de Walters*, Ann. Sci. ENS, **34**, (2001), 287–311.

[6] T. Bousch & O. Jenkinson, *Cohomology classes of dynamically non-negative C^k functions*, Invent. Math. **148** (2002), 207–217.

[7] T. Bousch & J. Mairesse, *Asymptotic height optimization for topical IFS, tetris heaps, and the finiteness conjecture*, Jour. Amer. Math. Soc. **15** (2002), 77–111.

[8] S. Bullett & P. Sentenac, *Ordered orbits of the shift, square roots, and the devil's staircase*, Math. Proc. Camb. Phil. Soc. **115** (1994), 451–481.

[9] P. Cartier, J. M. G. Fell, & P.-A. Meyer, *Comparaison des mesures portées par un ensemble convexe compact*, Bull. Soc. Math. France, **92** (1964), 435–445.

[10] G. Choquet, *Le théorème de représentation intégrale dans les ensembles convexes compacts*, Ann. Inst. Fourier (Grenoble), **10** (1960), 333–344.

[11] G. Contreras, A. Lopes, & Ph. Thieullen, *Lyapunov minimizing measures for expanding maps of the circle*, Ergod. Th. & Dyn. Sys. **21** (2001), 1379–1409.

[12] H. Dalton, *The measurement of the inequality of incomes*, Econom. J. **30** (1920), 348–361.

[13] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory, **1** (1967), 1–49.

[14] J-M. Gambaudo, O. Lanford III, & C. Tresser, *Dynamique symbolique des rotations*, C. R. Acad. Sc. Paris, **299** (1984), 823–826.

[15] G. H. Hardy, J. E. Littlewood, & G. Pólya, *Some simple inequalities satisfied by convex functions*, Messenger Math. **58** (1929), 145–152.

[16] ———, *Inequalities*, Cambridge University Press, Cambridge, (1st ed., 1934; 2nd ed., 1952).

[17] E. Harriss & O. Jenkinson, *Flattening functions on flowers*, Ergod. Th. & Dyn. Sys. **27** (2007), 1865–1886.

[18] B. Host, *Nombres normaux, entropie, translations*, Israel J. Math. **91** (1995), 419–428.

[19] O. Jenkinson, *Conjugacy rigidity, cohomological triviality, and barycentres of invariant measures*, Ph. D. thesis, University of Warwick, 1996.

[20] ———, *Frequency locking on the boundary of the barycentre set*, Experimental Mathematics, **9** (2000), 309–317.

[21] ———, *Ergodic optimization*, Discrete & Cont. Dyn. Sys. **15** (2006), 197–224.

[22] ———, *Optimization and majorization of invariant measures*, Electron. Res. Announc. Amer. Math. Soc. **13** (2007), 1–12.

[23] J. Karamata, *Sur une inégalité rélative aux fonctions convexes*, Publ. Math. Univ. Belgrade, **1** (1932), 145–148.

[24] S. Karlin & A. Novikoff, *Generalized convex inequalities*, Pacific J. Math. **13** (1963), 1251–1279.

[25] M. O. Lorenz, *Methods of measuring concentration of wealth*, J. Amer. Statist. Assoc. **9** (1905), 209–219.

[26] M. Lothaire, *Algebraic combinatorics on words*, Encyclopaedia of mathematics and its applications, vol. 90, Cambridge University Press, 2002.

[27] A. W. Marshall & I. Olkin, *Inequalities: theory of majorization and its applications*, Mathematics in science and engineering vol. 143, Academic Press, 1979.

[28] M. Morse & G. A. Hedlund, *Symbolic Dynamics II. Sturmian Trajectories*, Amer. J. Math. **62** (1940), 1–42.

[29] A. C. Pigou, *Wealth and welfare*, Macmillan, New York, 1912.

[30] N. Pytheas Fogg, *Substitutions in dynamics, arithmetics and combinatorics*, Springer Lecture Notes in Mathematics vol. 1794, 2002.

[31] M. Rothschild & J. Stiglitz, *Increasing risk I: A definition*, J. of Economic Theory, **2** (1970), 225–243.

[32] D. Rudolph, $\times 2$ *and* $\times 3$ *invariant measures and entropy*, Ergod. Th. & Dyn. Sys. **10** (1990), 395–406.
[33] I. Schur, *Über eine Klasse von Mittelbildungen mit Anwendungen die Determinanten*, Theorie Sitzungsber. Berlin. Math. Gesellschaft, **22** (1923), 9–20 (*Issai Schur colleted works*, (A. Brauer & H. Rohrbach, eds.) Vol. II, pp. 416–427, Springer-Verlag, Berlin, 1973).
[34] P. Veerman, *Symbolic dynamics of order-preserving orbits*, Physica D, **29** (1987), 191–201.

Oliver Jenkinson; School of Mathematical Sciences, Queen Mary, University of London, Mile End Road, London, E1 4NS, UK.
E-mail address: omj@maths.qmul.ac.uk
URL: www.maths.qmul.ac.uk/∼omj