Boshernitzan's criterion for unique ergodicity of an interval exchange transformation

WILLIAM A. VEECH

*Department of Mathematics, Wiess School of Natural Sciences, Rice University,
P.O. Box* 1892, *Houston, Texas* 77251, *USA*

Abstract. Confirming a conjecture by Boshernitzan, it is proved that if T is a minimal non-uniquely ergodic interval exchange, the minimum spacing of the partition determined by T^n is $O(1/n)$.

1. *Introduction*

In [1] Boshernitzan introduces a condition, which he calls 'Property P', in the presence of which a minimal interval exchange transformation must be uniquely ergodic. He goes on to conjecture that an apparently significant weakening of this condition remains strong enough to guarantee unique ergodicity. The primary purpose of this note is to verify Boshernitzan's conjecture.

Let $T:[0, 1) \to [0, 1)$ be an exchange on m intervals with division points $0 = \beta_0 < \beta_1 < \cdots < \beta_{m-1} < \beta_m = 1$. Write D_1 for the set $\{\beta_j \mid 0 < j < m\} = D_1$, and for all $n > 0$ define

$$(1.1) \qquad\qquad D_n = \bigcup_{k=0}^{n-1} T^{-k} D_1.$$

D_n defines a partition of $[0, 1)$, and we write $\varepsilon_n = \varepsilon_n(T)$ for the length of the shortest interval of this partition. The following theorem affirms Boshernitzan's conjecture:

(1.2) THEOREM. *Let T be a minimal interval exchange transformation which is not uniquely ergodic. If $\varepsilon_n(T)$ is defined as above, then*:

$$(1.3) \qquad\qquad \lim_{n \to \infty} n\varepsilon_n(T) = 0.$$

Theorem 1.2 has direct application to a restricted class of skew products realizable as interval exchange transformations. To explain, let $X = \mathbb{R}/\mathbb{Z}$ be viewed interchangeably with $[0, 1)$, and let $\theta \in X$ be a fixed irrational. Let F be a finite set, and suppose given permutations γ_j, $1 \le j \le r$, of F. Let X be partitioned by $0 = t_0 < t_1 < \cdots < t_r = 1$, and define a permutation valued function, $\varphi(x) = \gamma_j$, $t_{j-1} \le x < t_j$, $1 \le j \le r$. Using θ and φ, set up on $X \times F$ the skew product

$$(1.4) \qquad\qquad S(x, \gamma) = (x + \theta, \varphi(x)\gamma), \qquad (x, \gamma) \in X \times F.$$

We apply theorem 1.2 to obtain:

(1.5). THEOREM. *With notation as above, suppose the discontinuities of φ lie in $\mathbb{Z}\theta$ (modulo 1). If the skew product (1.4) is minimal, then it is uniquely ergodic.*

A proof of theorem 1.5 (for the case of one discontinuity) was sketched in the correspondence [6]. The theorem has been used to aid the study of a restricted class of rational billiard tables by Boshernitzan [1] and Gutkin [3]. (In fact, the original purpose of this note was to put a proof of theorem 1.5 in the literature. The proof of theorem 1.2 is in spirit the same as our original proof of theorem 1.5, which had its origin in [5].)

(1.6) *Remark.* Let $F = \mathbb{Z}_2 = \{\pm 1\}$, let $0 < t < 1$, and let $\varphi_t = \chi_{[0,t)} - \chi_{[t,1)}$ ($= \varphi$ above for the partition $\{0, t, 1\}$ and $\gamma_1 = +1$, $\gamma_2 = -1$). If $t \notin \mathbb{Z}\theta$, it is automatic that the skew product in (1.4), here denoted S_t, is minimal. If $n > 0$, define $\delta_n = \delta_n(S_t)$ by

$$\delta_n = \text{Min} \left(\min_{0 < k < n} \|k\theta\|, \min_{|k| < n} \|k\theta - t\| \right).$$

In [4] it is proved that if S_t is not uniquely ergodic (t irrational), then $\lim_{\|n\theta\| \to 0} n\delta_n = 0$. This condition readily implies the condition $\lim_{n \to \infty} \delta_n = 0$. If $X \times \mathbb{Z}_2$ is identified with $[0, 2)$ in an obvious way, then S_t becomes an exchange on five intervals, and δ_n is the same as the quantity ε_n defined above. Thus theorem 1.2 contains as a special case the criterion used in [4]. We recall also from [4] that if θ has unbounded partial quotients, there is an uncountable set of t (of Hausdorff dimension 0) such that S_t fails to be uniquely ergodic. In particular, the assumption on the placement of discontinuities in theorem 1.5 cannot be dispensed with entirely.

(1.7) *Remark.* Fix $m \geq 1$, and let π be a fixed irreducible permutation on $\{1, 2, \ldots, m\}$ (i.e. $\pi(\{1, \ldots, k\}) = \{1, \ldots, k\}$ implies $k = m$). Let

$$\Delta_{m-1} = \{\lambda \in \mathbb{R}^m \mid \lambda_j > 0, 1 \leq j \leq m, \sum \lambda_j = 1\},$$

and for each $\lambda \in \Delta_{m-1}$ let T_λ be the (λ, π) interval exchange. Fix $\varepsilon > 0$, and for each $n > 0$ let $u(n, \varepsilon)$ be the $(m-1)$-Lebesgue measure of the set of λ such that $n\varepsilon_n(T_\lambda) \geq \varepsilon$. In [1] Boshernitzan proves, by an ingenious argument, the assertion $\lim_{\varepsilon \to 0} u(n, \varepsilon) = 0$ uniformly in n. Coupled with theorem 1.2 one obtains an even more direct proof that a.e. λ the (λ, π) interval exchange is uniquely ergodic.

(1.8) *Remark.* In [1] Boshernitzan announces a generalization of theorem 1.5. If T is an exchange on m intervals, and if the lengths of these intervals span a two dimensional vector space over $\mathbb{Q}$, then T is uniquely ergodic as soon as T is minimal. Boshernitzan has pointed out to us that his theorem, which appears in [2], is also a consequence of theorem 1.2.

2. *Ergodicity*

The proof of theorem 1.2 is based upon an elementary lemma. The set D_n of (1.1) will be identified with the partition of $X = [0, 1)$ it determines. μ denotes Lebesgue measure on $\mathbb{R}$.

(2.1) LEMMA. *Let T be a minimal interval exchange, and let I be an interval of the partition D_n. If J is a subinterval of I, of length at least ε_{2n}, then*:

$$(2.2) \qquad \mu\left(\bigcup_{j=0}^{n-1} T^j J \right) \geq n\varepsilon_{2n}.$$

The lemma will be proved, in a slightly stronger formulation, in § 3. In this section it will be used as a vehicle for establishing ergodicity. Define, in terms of T and $\varepsilon > 0$, a set $S(\varepsilon) = \{n > 0 \mid n\varepsilon_{2n} > \varepsilon\}$. (The subscript $2n$ is intentional.)

(2.3) LEMMA. *Let T be a minimal interval exchange, and suppose $A \subseteq X$ is a measurable set such that $TA = A$. Given $\varepsilon > 0$ and $\delta > 0$, there exists $q = q(A, \varepsilon, \delta)$ with the following property: If $q \leq n \in S(\varepsilon)$, there exists for each interval I of D_n a choice, $B = B(I) = A$ or A^c, such that if J is any subinterval of I, of length at least ε_{2n}, then*:

$$(2.4) \qquad \mu(J \cap B) \geq (1 - \delta)\mu(J).$$

Proof. We may suppose $\delta < \tfrac{1}{2}$. Both sides of (2.4) are continuous functions of the endpoints of J. Therefore, if a choice of $B = A$ or A^c can be made for each interval $J \subseteq I$, of length at least ε_{2n}, so that (2.4) is true, then the choice *must* be independent of J, i.e. a function of I.

We may also suppose $\delta < \varepsilon$. There exists $\gamma > 0$ and a measurable set $K \subseteq X$ with the following properties: $\mu(K) > 1 - \delta$, and if $x \in K$, there is a choice, $B = B(x) = A$ or A^c, such that (2.4) holds for every interval J satisfying $x \in J$ and $\mu(J) < \gamma$.

Because T is minimal, we may select q so that $\varepsilon_n < \gamma$, $n \geq q$. Let such an n also belong to $S(\varepsilon)$, and let I be an interval of D_n. If J is a subinterval of I of measure at least ε_{2n}, then (2.2), the fact $\varepsilon > \delta$, and the choice of K imply there is a j, $0 \leq j < n$, such that $T^j J \cap K \neq \varnothing$. Select any point x in the intersection, and observe that because A and A^c are invariant, $B(x)$ satisfies

$$\mu(B \cap J) = \mu(B \cap T^j J) \geq (1 - \delta)\mu(T^j J) = (1 - \delta)\mu(J).$$

(Note that the definition of D_n implies T^j is linear on I for $0 \leq j < n$.) The lemma is now proved.

We now consider $D_1 = \{\beta_j \mid 0 < j < m\}$. For each j, $0 < j < m$, there exists $s = s_j \geq 0$ such that the sequence $T^{-i}\beta_j$, $i > s_j$, does not intersect D_1. (T can have no periodic orbit.) Now let n be large, and let $L(j, n)$ be the interval, $L = (T^{-n}\beta_j - \varepsilon_n/4, T^{-n}\beta_j + \varepsilon_n/4)$. The images of L under T^i, $0 \leq i < n - s_j$, are pairwise disjoint intervals, by definition of D_n and ε_n. Their total measure is $(n - s_j)\varepsilon_n/2$, which is at least $\varepsilon/4$ if n is large and belongs to $S(\varepsilon)$. If we assume $\delta < \varepsilon/4$, it will follow there exists $B = B_j = A$ or A^c such that (2.4) holds with $T^i L$ in place of J, $0 \leq i < n - s_j$. Fix such an i, and let I' and I'' be the successive intervals of D_n which share $T^i T^{-n}\beta_j$ as an endpoint. If δ is sufficiently small, lemma 2.3 implies readily that $B(I') = B_j = B(I'')$. Letting $\delta \to 0$, we conclude each set, A or A^c, is, modulo μ, a union of intervals with endpoints from the set D_r, where $r = \mathrm{Max}\,(s_1, \ldots, s_{m-1})$, together with 0 and 1. As each set is invariant, and as T is assumed to be minimal, it must be that one of A or A^c has measure 0. A is an arbitrary measurable invariant set, and therefore T

is ergodic. We have proved (assuming lemma 2.1):

(2.5) LEMMA. *If T satisfies the hypotheses of theorem* 1.2, *then T is ergodic relative to Lebesgue measure.*

3. *Unique ergodicity*

In this section we assume D_1 is augmented by 0, and we write $D_n = D_n(T)$ for the set (1.1). Observe the relations

(3.1)
$$D_n(T^{-1}) = T^n D_n(T),$$
$$D_{2n}(T) = T^{-n}(D_n(T) \cup D_n(T^{-1})).$$

Write $D_n^*(T) = T^n D_{2n}(T)$. We observe T^j is linear on each interval of $D_n(T)$ (resp. $D_n(T^{-1})$) for $0 \le j < n$ (resp. $-n \le j \le -1$). In particular, T^j is linear on each interval of $D_n^*(T)$ for $-n \le j < n$. It follows from (3.1) that $D_n^*(T)$ has shortest interval of length ε_{2n}. Finally, we note that if I is an interval of $D_n^*(T)$, the intervals $I, TI, \dots, T^{n-1}I$ are pairwise disjoint.

(3.2) LEMMA. *Let T be a minimal interval exchange, and suppose there exists $\varepsilon > 0$ such that the set $S(\varepsilon)$, defined following lemma* 2.1, *is infinite. If μ_0 is an ergodic T-invariant Borel probability measure on $[0, 1)$, then μ_0 is absolutely continuous relative to μ (Lebesgue measure).*

Proof. Since μ_0 is assumed to be ergodic, and since the minimality of T implies μ_0 is non-atomic, there exists a generic point x for μ_0, that is, a point for which the relation

(3.3)
$$\lim_{N \to \infty} N^{-1} \sum_{j=0}^{N-1} \chi_J(T^j x) = \mu_0(J)$$

holds for each interval J. Let N increase to ∞ through a sequence in $S(\varepsilon)$. The discussion preceding this lemma combines with (3.3) to imply $\mu_0(J) \le \varepsilon^{-1}\mu(J)$. The lemma clearly follows.

We shall now recast the basic lemma:

(3.4) LEMMA. *Let the notation and assumptions be those of lemma* 2.1, *and (for convenience) assume the set D_1 does not lie in a single orbit. The interval J contains a subinterval L, of length at least ε_{2n}, such that the intervals $TL, T^2L, \dots, T^{n-1}L$ are disjoint from each other and from J.*

Proof. We may suppose J is left closed and right open. Let $Ux = T^{n(x)}x$ be the induced (first return) map of T on J. If $n(x) \ge n$ for all $x \in J$, there is nothing to prove. We next observe that because T^j is linear on J for $0 \le j < n$, and because U is one-to-one, the set where $n(x) < n$, if non-empty, is comprised of one or two intervals, each sharing an endpoint with J, and on each of which $n(x)$ is constant. Suppose first there is only one interval, J_1. If $n(x) = j$ on J_1, and if $T^j = \mathrm{Id} + c_j$ on (I and) J, the complement of J_1 in J has length $|c_j| \ge \varepsilon_{2n}$ (see the first paragraph of this section), and this complement can be taken for L. If there are two intervals, J_1 and J_2 say, the assumed minimality of T and the assumption made about D_1 imply

their union is not all of J. (Indeed, for each element β of D_1 there exists an element x_0 of J which visits β before its return to J. If β is not in the orbit of the left endpoint of J, then $n(x_0) \geq n$.) If $n(x) = j_i$ on J_i, $i = 1, 2$, the interval $J \cap [T^{j_1}J_1 \cup T^{j_2}J_2]^c$ contains the interval $I \cap [T^{j_1}I \cup T^{j_2}I]^c$. The case $I = J$ of the argument (in parentheses) just given implies the latter interval is non-trivial, and the remark in the first paragraph of the section implies it has length at least ε_{2n}. The lemma follows.

(3.5). *Remark.* If T is a minimal interval exchange, and if D_1 lies in a single orbit, one finds that T arises by a stacking construction from a minimal exchange on two intervals. T is then automatically uniquely ergodic, and for this reason we have excluded it from consideration in lemma 3.4.

Proof of theorem 1.2. Lemmas 2.1 and 2.5 imply Lebesgue measure is ergodic for T. Lemma 3.2 implies every ergodic T-invariant Borel probability is absolutely continuous relative to Lebesgue measure. Therefore T admits but one ergodic invariant Borel probability, and the ergodic decomposition of an arbitrary invariant Borel probability has but one component. That is, T is uniquely ergodic.

Proof of theorem 1.5. Evidently S is an interval exchange on the interval $[0, \text{Card } F)$, where F is the given finite set. Let q_k be the kth denominator in the sequence of convergents to θ. If $0 < \alpha < \frac{1}{2}$, and if $n = [q_k \alpha]$, where $[\cdot]$ denotes the greatest integer function, the assumption $D_1 \subset \mathbb{Z}\theta$ implies, for large k and this associated n, $\varepsilon_{2n} \geq \|q_{k-1}\theta\|$. It follows that $n\varepsilon_{2n} \geq \alpha/4$.

REFERENCES

[1] M. Boshernitzan, A condition for minimal interval exchange maps to be uniquely ergodic. *Duke Math. J.* **52** (1985), 723–752.
[2] M. Boshernitzan. In preparation.
[3] E. Gutkin. Billiards on almost integrable polyhedral surfaces. *Ergod. Th. & Dynam. Sys.*, **4** (1984), 569–584.
[5] W. A. Veech. Finite group extensions of irrational rotations. *Israel J. Math.* **21** (1975), 240–259.
[4] W. A. Veech. Strict ergodicity in zero dimensional dynamical systems and the Kronecker-Weyl theorem mod 2. *Trans. Amer. Math. Soc.*, **140** (1969), 1–33.
[6] W. A. Veech. Letter to E. Gutkin, August 10, 1983.