Ergod. Th. & Dynam. Sys. (2008), **28**, 697–705
doi: 10.1017/S0143385707000387
© *2008 Cambridge University Press*
Printed in the United Kingdom

On the entropy of linear factors

BENJAMIN WEISS
Hebrew University of Jerusalem, Jerusalem 91904, Israel
(e-mail: weiss@math.huji.ac.il)

(*Received* 29 *April* 2007 *and accepted in revised form* 1 *June* 2007)

Abstract. In general, the entropy $h(\mathcal{Y})$ of a stationary process $\mathcal{Y}$, which is a factor of the process $\mathcal{X}$, can take any value in the interval $(0, h(\mathcal{X})]$. For linear factors the situation is completely different. In fact it even matters how the class of linear factors is defined. We will investigate both moving averages of bounded sequences by coefficients in l^1 and the natural notion in L^2.

1. *Introduction*

Bill Parry was one of the first in the West to work extensively in the entropy theory that A. Kolmogorov, V. Rohlin and their schools developed in the Soviet Union. Many learned the theory from his lovely monograph [**P1**] which can serve as a reference for many of the results mentioned below. I too am indebted to Bill for many things and sadly dedicate this small étude to his memory.

As is well known, if a process $\{X_n\}$ has entropy h then among the non-trivial factors of the process, $\{Y_n\}$, can be found factors whose entropy has any value in the interval $(0, h]$. A natural question to ask is what are the possible values of the entropy of *linear* factors, i.e. processes where the Y_n are linear functions of the X_n. There are two rather natural ways to formalize this notion of linear factors. One of them is to assume that the $\mathcal{X}$ process is bounded and that $\mathcal{Y}$ is given by the convolution

$$Y_n = \sum_{i=-\infty}^{\infty} c_i X_{n-i},$$

where c_n denotes a non-zero element of $l^1(\mathbb{Z})$. Such processes are called moving averages of the $\mathcal{X}$ process.

A second formalization is to assume that the X_n are in L^2 and that the Y_n are in the closed linear span of the X_n as elements of L^2. It came as a surprise to me that the answers in these two settings are rather different, even when one restricts to finite-valued processes. It turns out that for an arbitrary finite-valued process X_n the entropy of a non-trivial l^1 moving average is always *equal* to the entropy of the original process while this is definitely not the case for L^2 moving averages.

https://doi.org/10.1017/S0143385707000387 Published online by Cambridge University Press

My original motivation was a question concerning the special case when the X_n are independent with infinite entropy. There it was plausible to conjecture that the entropy of the Y_n process is also infinite. While I was unable to prove this in general I was able to establish this in some special cases including the case when X_0 is uniformly distributed on [0, 1] via the above result about finite-valued processes.

Studying these linear factors led me to another surprise. Let us say that two elements of l^∞, ξ_n and η_n are orthogonal if

$$\lim_{N\to\infty} \frac{1}{2N+1} \sum_{-N}^{+N} \xi_n \eta_n = 0.$$

We will say that two closed shift invariant subsets of $[-1, +1]^{\mathbb{Z}}$, $X$and Y are linearly disjoint if each $\xi \in X$ is orthogonal to each $\eta \in Y$. Simple examples of these are provided by spaces formed by almost periodic functions with disjoint spectra. These examples have zero entropy. The surprise is that there exist linearly disjoint spaces with positive entropy! These can be used to give examples of perfect filtering where both the signal and the noise have positive entropy. The reason for this is that Furstenberg's well-known result on perfect filtering based on his notion of disjointness [**F**] really only requires linear disjointness. As he showed in [**F**], disjointness of two processes entails that at least one of them has zero entropy, but as we will see this is not necessarily the case for linear disjointness.

In §2, I will prove the basic theorem about l^1 averages of finite alphabet processes while in §3, I will give the application to some infinite entropy independent processes. Next, in §4, I will discuss linear disjointness and filtering, while the last section is devoted to the L^2 theory.

2. *Finite alphabets*

In this section we will deal with the case of linear maps of processes with a finite number of values. Let c_i be a fixed non-trivial summable sequence which will serve as the coefficients of the moving averages. Let X_n be a finite-valued process and let Y_n be defined by the convolution

$$Y_n = \sum_{i=-\infty}^{\infty} c_i X_{n-i}.$$

In order to show that the entropy of Y is the same as that of the entropy of X we need to study the relative entropy of the factor mapping π defined by this convolution. To this end, we observe that the images of points under π^{-1} are translates of a fixed linear space which is $\pi^{-1}(0)$, where 0 denotes the null sequence. Denote by $\mathcal{N}_c$ the space of all bounded sequences u_i such that $c * u = 0$. This is a linear space invariant under the shift and we want to show that it cannot have positive entropy. For this we will see that it is deterministic, in the sense that for $u \in \mathcal{N}_c$ the values u_n for $n < 0$ determine all of the values of u. This is the content of the following lemma which is the discrete analogue of a theorem of Newman [**N**]. I include a proof for the convenience of the reader.

LEMMA 1. *If c is a non-trivial element of l^1 and u is a bounded sequence such that $c * u = 0$ and $u_n = 0$ for all $n < 0$ then u is identically zero.*

https://doi.org/10.1017/S0143385707000387 Published online by Cambridge University Press

Proof. Let us denote by $c(t)$ the Fourier series with coefficients given by c. Clearly c is a continuous function that is not identically zero, and therefore there is some interval where c does not vanish, say $J = [a, b]$. Let $f(t)$ be a 'triangle' function with support in the middle half of J. As is well known, all of the Fourier coefficients of f are different from zero and they are summable. By Wiener's lemma [**W**, p. 90], $f(t)/c(t)$ has an absolutely convergent Fourier series; we denote the coefficients of this series by b. If ϕ denotes the Fourier coefficients of $f(t)$ then we see that $\phi = c * b$ and hence for any $u \in \mathcal{N}_c$ we also have that $\phi * u = 0$. In particular, we have that

$$\sum_{n=0}^{+\infty} \phi_{-n} u_n = 0.$$

Now the same holds for f translated by any number s less than $\frac{1}{4}|b - a|$. It follows that we also have

$$\sum_{n=0}^{+\infty} e^{isn} \phi_{-n} u_n = 0$$

for all such s. But

$$A(z) = \sum_{n=0}^{+\infty} z^n \phi_{-n} u_n$$

represents an analytic function in the unit disc, and multiplying together a finite number of rotates of A will yield an analytic function which vanishes identically on the boundary. Such a function is clearly zero identically and thus we see (recall that all ϕ_n are different from zero) that all $u_n = 0$. □

Returning to our processes X_n and Y_n (which is a factor of the $\mathcal{X}$ process) we will now make use of one of the basic formulas for the relative entropy. We will use the notation X_i^j for $i < j$ to denote the random variables $X_i, X_{i+1}, \ldots, X_j$, $h(\mathcal{X})$ to denote the entropy of the process $\mathcal{X}$ and $H(X_0|\mathcal{F})$ the conditional entropy of X_0 given the σ-field $\mathcal{F}$. When a collection of random variables appears instead of a σ-field, the intention is the smallest σ-field generated by these random variables. The formula in question is

$$h(\mathcal{X}) = h(\mathcal{Y}) + H(X_0|X_{-\infty}^{-1} \vee Y_{-\infty}^{+\infty}).$$

In the usual proofs of this formula one assumes that for both processes the entropies of the individual random variables X_i and Y_i are finite. However, it is not hard to see that one really needs this assumption only for the X_i. For them, this hypothesis is crucial since it is not very hard to construct examples of countably valued processes $\mathcal{X}$ with finite and positive process entropy which are nonetheless predictable (i.e. X_0 is a function of the $\{X_i : i < 0\}$), which would contradict the above formula with $\mathcal{Y}$ being the trivial factor. ¿From the lemma it follows that the second expression vanishes since, given the full sequence of the Y_n, $X_{-\infty}^{-1}$ determines all of the X_n and, in particular, X_0.

This establishes the following theorem.

THEOREM 1. *If $\mathcal{X}$ is a finite-valued process, or more generally a stationary process with $H(X_0) < \infty$, and $\mathcal{Y}$ is a non-trivial moving average of $\mathcal{X}$ then the entropy of $\mathcal{Y}$ equals the entropy of $\mathcal{X}$.*

https://doi.org/10.1017/S0143385707000387 Published online by Cambridge University Press

Example I. Even though the factor process must have full entropy it may be non-trivial. A simple example of this can be seen with the 2-shift on the symbols $\{-1, +1\}$ and $Y_0 = X_0 X_1$, where the factor mapping is everywhere two to one.

Example II. If c has finite support then clearly the $\mathcal{Y}$ process is also finite valued. Less obvious is the next example of a $c \in l^1$ with infinite support and a finite-valued positive entropy process $\mathcal{X}$ such that $c * X$ is still finite valued. Take any polynomial $P(z)$ of degree d with real coefficients p_j with no roots on the unit circle that is symmetric, i.e. $P(z) = z^d P(1/z)$. Then if c denotes the coefficients of the Laurent series for $1/P$ we have that the c_k are real valued and, by Wiener's theorem [**K**, p. 239], $c \in l^1$. Start with the full shift on a finite number of real values $V = \{v_1, v_2, \ldots, v_M\}$ and a product measure and convolve with the coefficients of P to get a process $\mathcal{X}$ which clearly is finite valued. When the $\mathcal{X}$ process realizations are convolved with c we get back the V shift, since $c * p = \delta_0$.

3. *Moving averages of independent processes*

In this section we shall assume that the X_i are independent and identically distributed (i.i.d.) and in some special cases shall show that any non-trivial moving average has the same entropy. In finite entropy this is covered by the result in the preceding section and so we will concentrate now on the case of infinite entropy. Perhaps the simplest case to analyze is the one in which the distribution of X_0 is uniform on [0, 1]. Using the dyadic expansion of real numbers in the unit interval we can write

$$X_n = \sum_{k=1}^{\infty} X_{n,k}/2^k \tag{1}$$

where the $X_{n,k}$ are independent random variables taking the values $\{0, 1\}$ with equal probability. As before we fix c_i, a non-trivial summable sequence which will serve as the coefficients of the moving averages, and define the process $\mathcal{Y}$ by

$$Y_n = \sum_{i=-\infty}^{\infty} c_i X_{n-i}. \tag{2}$$

Our goal is to show that the entropy of the process $\mathcal{Y}$ is infinite. To this end, notice that the sum of the first N terms in equation (1) defines a finite-valued random variable with entropy N. Fix an N and write $X_n = X_n^0 + X_n^1$ for the decomposition of the right-hand side of (1) into the sum of the first N terms and the remainder. The main thing to observe here is that the distribution of $\mathcal{X}^1$ is the same as that of $2^{-N}\mathcal{X}$. Since the definition of $\mathcal{Y}$ is linear, applying the moving averages to this decomposition gives a corresponding decomposition $Y_n = Y_n^0 + Y_n^1$. According to Theorem 1 we know that the entropy of the $\mathcal{Y}^0$ process also equals N. Furthermore, by our earlier observation the entropy of $\mathcal{Y}^1$ equals that of $\mathcal{Y}$.

Now assume that $h(\mathcal{Y}) = h < \infty$, and choose N to be greater than $2h$. Clearly $\mathcal{Y}^0$ is a function of $\mathcal{Y}$ and $\mathcal{Y}^1$ and this gives a contradiction. We have established the following proposition.

PROPOSITION 1. *A non-trivial moving average of i.i.d. random variables uniformly distributed on* [0, 1] *has infinite entropy.*

https://doi.org/10.1017/S0143385707000387 Published online by Cambridge University Press

This idea of decomposing the $\mathcal{X}$ process into the sum of mutually independent processes, one of which is finite valued, can work even if there is no homogeneity. For example, we can establish the following proposition.

PROPOSITION 2. *Suppose that the distribution of the bounded i.i.d. random variables X_n can be expressed as an infinite convolution of independent finite-valued random variables and that the entropy of $\mathcal{X}$ is infinite. If $\mathcal{Y}$ is any non-trivial moving average of the X_n it has infinite entropy.*

Proof. We will reuse equation (2) to define the $\mathcal{Y}$ process. Since the $\mathcal{X}$ process has infinite entropy, it is clear that, when decomposing each X_n into an initial finite sum of finite-valued random variables X_n^0 and the rest as before, the entropy of $\mathcal{X}^0$ tends to infinity. As before, if we suppose that the entropy of $\mathcal{Y}$ is finite and equal to h, we can choose the decomposition in such a way that the entropy h_0 of the corresponding $\mathcal{Y}^0$ exceeds $3h$. Observe that since $\mathcal{Y}^0$ is a function of the pair $(\mathcal{Y}, \mathcal{Y}^1)$, we have $h(\mathcal{Y}^1) = h_1 \geq h_0 - h > 2h$. Then consider the pair of processes $(\mathcal{Y}^0, \mathcal{Y}^1)$. They are independent and therefore the entropy of the pair equals $h_0 + h_1$. On the other hand, this pair of processes is a function of the pair $(\mathcal{Y}, \mathcal{Y}^0)$ and has entropy at most $h + h_0$, which gives a contradiction. □

This idea can also be pursued even when the decomposition of the $\mathcal{X}$ process is not into independent summands. As an example of this consider the case of a bounded but countably valued process with infinite entropy. Denote by $V = \{v_m : m \in \mathbb{N}\}$ the set of values that X_i takes with positive probability. Without loss of generality, we may assume that zero is not a member of V. For a fixed M denote by V_0 the set $\{v_m : m < M\}$ and set $V_1 = V \backslash V_0$. Now let ${}_0X_i$ be defined as X_i in the case when $X_i \in V_0$ and as zero otherwise, and put ${}_1X_i = X_i - {}_0X_i$. Once again the $\mathcal{X}^0$ process is finite valued and its entropy will tend to infinity if the $\mathcal{X}$ process has infinite entropy. The difference between this decomposition and the preceding ones is that now the processes $\mathcal{X}^0$ and $\mathcal{X}^1$ are not independent. However, if we introduce a third process, $\mathcal{X}^2$, which takes just two values $\{0, 1\}$, according to whether X_i belongs to V_0 or V_1, then the $\mathcal{X}^2$ process is a common factor of both $\mathcal{X}^0$ and $\mathcal{X}^1$ and upon conditioning on this common factor the processes become independent.

As before, corresponding to any such decomposition, we can decompose the moving average $\mathcal{Y}$ given by equation (2) into $Y_i = {}_0Y_i + {}_1Y_i$. Even though these do not have a natural common factor we can still consider the processes of pairs $({}_0Y_i, {}_2X_i)$, $({}_1Y_i, {}_2X_i)$ and these are relatively independent over their common factor $\mathcal{X}^2$. Now we can repeat the argument that proved the preceding proposition as follows. We suppose that the entropy of the $\mathcal{Y}$ process is finite and equal to h say, where we can clearly assume that $h > 1$. Choose now the M in the above decomposition so that the entropy h_0 of the corresponding $\mathcal{Y}^0$ satisfies $h_0 > 5h$. As before we can deduce that $h(\mathcal{Y}^1) = h_1 \geq h_0 - h$. As $\mathcal{X}^2$ takes only two values its entropy is at most equal to one, and so the relative entropy of the pair $({}_1Y_i, {}_2X_i)$ over the factor process $\mathcal{X}^2$ is at least $h_1 - 1$. The relative independence yields that the entropy of the combined process $({}_0Y_i, {}_1Y_i, {}_2X_i)$ is at least $h_0 + h_1 - 1$ but, however, as it is also equal to $(Y_i, {}_0Y_i, {}_2X_i)$ its entropy is at most $h + h_0 + 1$ which gives a contradiction. We have established the following proposition.

https://doi.org/10.1017/S0143385707000387 Published online by Cambridge University Press

PROPOSITION 3. *Suppose that the distribution of the bounded i.i.d. random variables X_n has countable support and that the entropy of $\mathcal{X}$ is infinite. If $\mathcal{Y}$ is any non-trivial moving average of the X_n it has infinite entropy.*

The two methods that were used can be combined to establish further special cases. However, I haven't been able to extend them to prove the general result that any non-trivial moving average of an infinite entropy i.i.d. process has infinite entropy and I leave this as an open problem. For infinite entropy non i.i.d. processes it can happen that a non-trivial moving average has finite entropy. We will give an example of this in the next section.

4. *Orthogonality and linear disjointness*

It will be convenient to denote the moving average of a process $\mathcal{X}$ as a convolution,

$$Y_n = (c * X)_n = \sum_{i=-\infty}^{\infty} c_{n-i} X_i, \tag{3}$$

where as usual the sequence c is in l^1 and the process $\mathcal{X}$ is bounded. If a moving average of the $\mathcal{Y}$ process is formed with another l^1 sequence b,

$$Z_n = (b * Y)_n = \sum_{i=-\infty}^{\infty} b_{n-i} Y_i, \tag{4}$$

then it is easy to see that we have $Z = a * X$, where $a = b * c$ is the usual convolution of two summable sequences. Now it is very easy to give examples of two non-trivial summable sequences b, c such that $b * c = 0$. All that one needs to do is to take two even, infinitely differentiable periodic functions on the circle $B(t), C(t)$ with disjoint support so that their product BC vanishes and let b_n be the Fourier coefficients of $B(t)$ and c_n the Fourier coefficients of $C(t)$. Let $\mathcal{X}$ be any finite-valued process with positive entropy and let $\mathcal{Y}$ be defined by applying the c_n to the $\mathcal{X}$ process. By Theorem 1 this process also has positive entropy and thus, in particular, is not trivial. The process defined by the $Z_n = (b * Y)_n$ vanishes identically but b is not identically zero. Note that since the functions B, C were even the sequences of coefficients are real and symmetric $b_{-n} = b_n$, $c_{-n} = c_n$.

The connection between this phenomenon and orthogonality is provided by the following simple analytical lemma.

LEMMA 2. *Let x_i and y_i be bounded real valued sequences, and c a symmetric, summable sequence such that $y * c = 0$. Then $x * c$ and y are orthogonal.*

Proof. We have to examine the expression

$$\lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{+N} (x * c)_n y_n$$

and show that it tends to zero as $N \to \infty$. Upon substitution we obtain

$$\lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{+N} \left(\sum_{k=-\infty}^{+\infty} x_k c_{n-k} \right) y_n$$

https://doi.org/10.1017/S0143385707000387 Published online by Cambridge University Press

or, taking into account the symmetry,

$$\lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{+N} \left(\sum_{k=-\infty}^{+\infty} x_k c_{-n+k} \right) y_n.$$

Now we replace $-n+k$ by l and interchange the order of summation to obtain

$$\lim_{N\to\infty} \sum_{k=-\infty}^{+\infty} x_k \left(\frac{1}{2N+1} \sum_{l=k-N}^{k+N} c_l y_{k-l} \right).$$

To complete the proof we will show that, uniformly in k,

$$\sum_{k=-\infty}^{+\infty} \left| \frac{1}{2N+1} \sum_{l=k-N}^{k+N} c_l y_{k-l} \right|$$

tends to zero as $N \to \infty$. To this end fix some $\varepsilon > 0$. When the interval over which the inner sum ranges, $[k-N, k+N]$, includes the interval $[-\varepsilon N, +\varepsilon N]$ we replace the inner sum by minus the sum of $c_l y_{k-l}$ over the complement. This is possible since by our hypothesis the inner sum over all l vanishes. The grand sum can then be estimated by

$$2\|y\| \left(\sum_{|l|>\varepsilon N} |c_l| + \varepsilon \sum_{|l|<\varepsilon N} |c_l| \right)$$

where $\|y\|$ is the sup norm of y. Clearly this completes the proof. □

To get a particularly simple example, start with the 2-shift $\Omega = \{-1, +1\}^{\mathbb{Z}}$, and let B and C be two functions on the circle as above. If X is the set of all convolutions of elements of Ω with the sequence of Fourier coefficients of B then X is a closed shift invariant set of bounded sequences, as is Y, the analogous set defined by the Fourier coefficients of C. Each of these subshifts supports invariant measures of positive entropy, namely the image of the independent measure, but, by the above, the subshifts are linearly orthogonal and both have positive entropy.

An application of the ergodic theorem shows that the processes are orthogonal and, as I have already remarked, Furstenberg's proof in [**F**] applies to such processes. The result of the discussion gives an example of two bounded real valued processes each of them with positive entropy such that whoever observes the sum process $\{X_n + Y_n\}$ can recover full information about both summands. This phenomenon by itself can of course be easily arranged in trivial ways by separating the values that each of the summands can assume but here we have this in a much less obvious way.

5. *The L^2 theory*

In this section we will change the setting to L^2 and ask about factors that are generated by functions in the closed linear span of $\mathcal{X}$. We start with an example which illustrates the new phenomenon. In the direct product of an irrational rotation with the Bernoulli 2-shift take a partition into two sets defined by partitioning the circle into two equal parts. Perturb that partition by a small amount to get X_0, a two-valued function such that the process generated by X_0 has positive entropy but X_0 has a non-zero projection onto the rotation

https://doi.org/10.1017/S0143385707000387 Published online by Cambridge University Press

factor so that the spectral type of the linear span of the $\{X_n\}$ includes a discrete component. This means that in the linear span of the $\{X_n\}$ there is a function with discrete spectrum and clearly this linear factor is non-trivial and has zero entropy. In the case when the process $\mathcal{X}$ is a K-process this cannot happen and indeed in that case the entropy cannot decrease at all! In fact, we have the following general theorem.

THEOREM 2. *Let $\mathcal{X}$ be a finite-valued process, and $\mathcal{Y}$ any process defined by a function Y_0 in the L^2 span of $\mathcal{X}$. Either the entropy of the $\mathcal{Y}$ process is zero or it equals that of $\mathcal{X}$.*

Proof. Denote by L the closed linear span of the $\{X_n\}$ in L^2. Let $\mathcal{G}$ denote the invariant σ-algebra generated by the $\{Y_n\}$, let M denote the subspace of L^2 that consists of all functions measurable with respect to $\mathcal{G}$, and let $M' \subset L \cap M$ be the closed linear span of the $\{Y_n\}$ and let $M'' \subset L$ be its orthogonal complement in L. We now write $X_0 = Z_0' + Z_0''$ where $Z_0' \in M'$, $Z_0'' \in M''$ and consider two possibilities.

(1) One possibility is that Z_0'' is already contained in the L^2 span of the $\{Z_n'' : n < 0\}$. In this case, since X_0 is finite valued, the relative entropy $h(\mathcal{X}|\mathcal{Y})$ which is given by the conditional entropy $H(X_0|X_{-\infty}^{-1} \vee \mathcal{G})$ will vanish, and thus the factor $\mathcal{Y}$ has full entropy.

(2) If Z_0'' is not already contained in the L^2 span of the $\{Z_n'' : n < 0\}$ then by Szegö's theorem (see below) the spectral measure associated to M'' must include the full Lebesgue measure. It follows that the spectral measure associated to M' is singular. In this case $\mathcal{Y}$ must be measurable with respect to the Pinsker algebra and thus has entropy zero (see Remark I below) . This concludes the proof of the theorem. □

Remark I. The result at the end of the proof which we need is stated as follows in [**P2**, p. 72]:

THEOREM 3. *If T is a stationary real-valued process with finite second moments and singular covariance then $h(T) = 0$.*

Parry refers to the papers of Pinsker (1960) and Newton (1971). Both of these authors state the result only for Gaussian processes and the important observation that the result holds for arbitrary processes is apparently due to Parry himself.

Remark II. In fact we need only a particularly simple case of Szegö's theorem, which can be easily formulated as follows. Suppose that U is a unitary operator on a Hilbert space H, and $h \in H$ is an element which is not in the closed linear span of $\{U^{-n}h : n < 0\} = H'$. Set $h = h' + h''$ where $h' \in H'$ and h'' is orthogonal to H' and observe that the $\{U^n h'' : n \in \mathbb{Z}\}$ are orthonormal, and hence the Lebesgue measure is part of the spectral measure associated to the cyclic space spanned by $U^n h$.

Remark III. The same argument will apply if we have a $\mathbb{Z}^d$ action with $d > 1$ if we replace the 'past' $\{n < 0\}$ by the past defined by a standard lexicographic ordering of $\mathbb{Z}^d$. Since the entropy theory with respect to this kind of past is exactly analogous to the entropy theory for $\mathbb{Z}$ actions we can easily generalize the above theorem to random fields, or processes parametrized by $\mathbb{Z}^d$. The l^1 theory of §2, on the other hand, does not easily generalize to higher dimensions since the simple result from harmonic analysis that we relied upon

https://doi.org/10.1017/S0143385707000387 Published online by Cambridge University Press

there fails to hold in higher dimensions. I do not know if the entropy preserving result of Theorem 1 is valid in higher dimensions.

Acknowledgement. I would like to to thank Hillel Furstenberg and Izzy Katznelson for their many helpful insights and Jean-Paul Thouvenot who pointed out to me that the L^2 story differs from the l^1 story.

References

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

[K] Y. Katznelson. *An Introduction to Harmonic Analysis*, 3rd edn. Cambridge University Press, Cambridge, 2004.

[N] D. J. Newman. Translates are always dense on the half line. *Proc. Amer. Math. Soc.* **27** (1969), 511–512.

[P1] W. Parry. *Entropy and Generators in Ergodic Theory*. W. A. Benjamin, Inc., New York, 1969.

[P2] W. Parry. *Topics in Ergodic Theory (Cambridge Tracts in Mathematics, 75)*. Cambridge University Press, Cambridge, 1981.

[W] N. Wiener. *The Fourier Integral and Certain of its Applications*. Cambridge University Press, Cambridge, 1933.

https://doi.org/10.1017/S0143385707000387 Published online by Cambridge University Press