Journal of Modern Dynamics
Vol. 19, 2023, pp. 835–845
doi:10.3934/jmd.2023024

THE BRIN PRIZE WORKS OF TIM AUSTIN

JEAN-PAUL THOUVENOT
(Communicated by Giovanni Forni)

ABSTRACT. The mathematical activity of Tim Austin has been, since the very beginning, quite abundant and versatile. We will describe and comment on three of his results which were selected as most representative for the Brin Prize Award and which culminated in the proof of the weak Pinsker structure theorem.

1.

The first work of Tim Austin ***On the norm convergence of non-conventional ergodic averages*** has an unusual history. The focus there was on the Furstenberg non-conventional averages: given an invertible measure preserving transformation T of the Lebesgue space $(X,\mathscr{A},m)$ and an integer k together with k L^∞ functions $(f_1, f_2,\dots, f_k)$, the objects of study are the averages:

$$A_N(f_1, f_2,\dots, f_k)(x) = \frac{1}{N}\sum_{n=1}^{N} f_1(T^n(x))f_2(T^{2n}(x))\dots f_k(T^{kn}(x)). \tag{1}$$

Historically, in view of the applications to combinatorial number theory, the important result was that, given $A \in \mathscr{A}$ such that $m(A) > 0$,

$$\liminf_{N\to+\infty}\int A_N(1_A, 1_A,\dots, 1_A)\,dm > 0. \tag{2}$$

The interest in the convergence in L^2 of the averages (1) developed much later [7, 32, 14]. However, it was already shown in Furstenberg's paper [10] (and a first building block in the proof of (2)) that when T is weakly mixing the limit in (1) exists in L^2 and equals

$$\prod_{i=1}^{k}\int f_i\,dm.$$

(Weak mixing implies weak mixing of all orders.) The remarkable answer was given by Host and Kra [15] and independently Ziegler [33] who showed, based

Received August 22, 2023; revised August 27, 2023.
2020 *Mathematics Subject Classification:* Primary: 3702 37A35, 28D20, 37A50, 60F99; Secondary: 37A20, 51F99, 60G10.
Key words and phrases: Non-conventional ergodic averages, weak Pinsker property, amenable groups, compression exponent, scenery entropy.

on a general structure theorem involving the precise description of "characteristic factors", the L^2 convergence of the averages (1). Rapidly after his founding result, Furstenberg, in a joint work with Katznelson [11] extended (2) to the following situation: considering the action of k commuting automorphisms $T_1, T_2, \dots, T_k$ of the Lebesgue space $(X, \mathscr{A}, m)$ together with k L^∞ functions $(f_1, f_2, \dots, f_k)$, the averages

$$A_N(T_1, T_2, \dots, T_k, f_1, f_2, \dots, f_k)(x) = \frac{1}{N}\sum_{n=1}^{N} f_1(T_1^n(x)) f_2(T_2^n(x)) \dots f_k(T_k^n(x)) \tag{3}$$

satisfy, when $m(A) > 0$,

$$\liminf_{N\to+\infty} \int A_N(T_1, T_2, \dots, T_k, 1_A, 1_A, \dots, 1_A)\, dm > 0,$$

opening the gate to a new generation of results in combinatorial number theory. The question of the L^2 limit of the averages in (3) remained open for some time as a result analogous to the Host–Kra structure theorem did not seem to be available. Then, Tao [28] produced a remarkable and quite unusual proof of the L^2 convergence of these averages using, in sharp contrast with the "structural viewpoint", a purely finitary and combinatorial approach. It became some kind of a challenge to the people in the field to produce a more conventional and ergodic theoretical proof of this result. Actually, this was achieved by Tim Austin who obtained a more general statement [2]:

THEOREM 1.1. *Let $r > 0$, $(X, \mathscr{A}, m, T_i)$, $1 \le i \le k$, be k commuting $\mathbb{Z}^r$ actions on the Lebesgue space $(X, \mathscr{A}, m)$, $(a_n)_{n\ge1}$, a family of base points and $(I_n)_{n\ge1}$ a Fölner sequence in $\mathbb{Z}^r$, $f_i, 1 \le i \le d$, L^∞ functions, then*

$$\frac{1}{|\, I_N|} \sum_{n\in a_N + I_N} f_1(T_1^n(x)) f_2(T_2^n(x)) \dots f_k(T_k^n(x)) \tag{4}$$

converges in L^2 when $N \to +\infty$ independently of the choice of the sequences a_n and I_n. Here, each T_i should be understood as a $\mathbb{Z}^r$ action.

The proof is very innovative in its own way. The L^2 limit theorem gives rise to a joining of the product of k copies of the original $\mathbb{Z}^r$ action (this interpretation was first used by Furstenberg) and, therefore, in particular, to an action of T_1 enlarged to a bigger space. By an inductive use of these extensions (starting from the von Neumann ergodic theorem) and with a very clever control of the successive algebras of invariant sets for the extensions of T_1 and the $T_i T_j^{-1}$, Austin produced a limit object for which the averages (4) converge.

Quite opposite to what happens in the proof of the $\mathbb{Z}$ case, this new object is an extension of the original action, and, furthermore, it is not canonical.

An interest remains in giving some kind of identification of the limit in Theorem 1.1. For this let us stick to the case $r = 1$. One way is to consider the Cesaro averages of the images Δ_n by $(Id \times T_1^n \times T_2^n \times \cdots \times T_k^n)$ of the diagonal measure Δ on the product of $k+1$ copies of the original action. The convergence of these joinings is equivalent to the weak convergence of the averages (4) (this was first used by Ryzhikov to give a joining proof that weak mixing implies weak mixing of all orders). The limit joining gives rise to a new action of $(Id \times T_1 \times T_2 \times \cdots \times T_k)$ on the product of $k+1$ copies of $(X, \mathscr{A}, m)$. The question then is to understand how this new action is built. For instance, is it true that the algebra of invariant sets for this new action is entirely contained in the first coordinate of the preceding product? (This implies that the weak convergence of the averages (4) is in fact L^2.)

2.

The second work of Tim Austin which we are going to look at is ***Amenable groups with very poor compression into Lebesgue spaces***.

Originating from insights of M. Gromov, the compression exponent $\alpha^*_{\mathfrak{X}}(X,\rho)$, of a metric space (X,ρ) is defined, given a Banach space $\mathfrak{X}$, as the supremum of the α's for which there exists an injection $f : X \to \mathfrak{X}$ such that

$$\rho(x,y)^\alpha \lesssim ||f(x) - f(y)|| \lesssim \rho(x,y)$$

with the usual meaning "up to a multiplicative constant" for $\lesssim$. It measures how well (X,ρ) can be "coarsely Lipschitzly embedded" into $\mathfrak{X}$. In what follows, (X,ρ) will be a finitely generated group G equipped with a left invariant word metric d and the Banach spaces $\mathfrak{X}$ will only be L^p's. In this situation, which was first considered and studied in [12], we will denote $c_p(G) = \alpha^*_{L^p}(G,d)$. The compression exponent has been mainly studied for amenable groups (however there is a nice result [1] that $c_2(F) = 1/2$ where F is the Thompson group). It was shown in [12] that when the larger quantity $c_2^\#(G)$ is larger than $1/2$, this forces G to be amenable ($c_2^\#(G)$ is obtained by restricting to the case where the f's in the definition of the compression exponent are equivariant mappings from G to the affine isometries of $\mathfrak{X}$). The $1/2$ value looked for a long time as a threshold for amenable groups. The question first appeared in [1] where it was asked whether $c_p(G) < 1/2$ could be possible for amenable groups. The remarkable and unexpected result of Tim Austin is the following theorem [3]:

THEOREM 2.1. *There exists a finitely generated amenable group G that does not admit any embedding into L^p with a positive compression exponent for any $p \in [1,\infty)$ (i.e., $c_p(G) = 0$).*

As it is playing a great role, let us describe the distortion rate of a mapping $f : X \to Z$ between two metric spaces (X,ρ) and (Z,θ):

$$\text{distortion}(f) = \sup_{u,v\in X, u\neq v} \frac{\theta(f(u), f(v))}{\rho(u,v)} \cdot \sup_{u,v\in X, u\neq v} \frac{\rho(u,v)}{\theta(f(u), f(v))}.$$

The construction of the group G by Tim Austin is quite remarkable. It is exhausted by a succession of finite groups with growing diameters which embed with bounded distortions in G but which have high distortions in L^p. These high distortions are the main tool which makes good compression embeddings impossible. These finite groups, the building blocks of the construction, were first described by Khot and Naor [18]. They have produced for all integers d

subgroups $V_d \subset \mathbb{Z}_2^d$ such that $\mathbb{Z}_2^d / V_d$ have distortion rates in L^p growing like d. The group G, starting with an amenable countable group K with exponential growth, is produced by Tim Austin as a semidirect product of $(\mathbb{Z}_6 \wr K)$ (the wreath product of K by $\mathbb{Z}_6$) with $(\mathbb{Z}_2^{\oplus(\mathbb{Z}_6 \wr K)}/V)$ where V is a $\mathbb{Z}_2$-subspace in $\mathbb{Z}_2^{\oplus(\mathbb{Z}_6/K)}$ in such a way that a family $\mathbb{Z}_2^{d_n}/V_{d_n}$ of the type described above is embedded in this semi direct product with uniformly bounded distortions and controlled expansion ratios. The construction is quite intricate and subtle as antagonistic constraints have to be taken care of (it is also difficult to describe).

We see that the final group G

$$(\mathbb{Z}_2^{\oplus(\mathbb{Z}_6/K)}/V) \rtimes (\mathbb{Z}_6 \wr K)$$

is a two fold abelian extension of an amenable group K with exponential growth. It can actually be made four step solvable.

3.

I come now to the third article of this selection, ***Measure concentration and the weak Pinsker property***.

It deals with the fundamental of abstract ergodic theory. Following the early developments which started with Kolmogorov and Sinai, a first conjecture was made by Pinsker [25] that every ergodic automorphism of a Lebesgue space would be isomorphic to the direct product of a zero entropy system with a K-automorphism. This followed the proof by the same Pinsker that, expressed in modern language, 0-entropy and K-automorphisms are disjoint. This conjecture was refuted by Ornstein [23] who constructed a mixing transformation for which it failed (this followed a previous example of his [22] using a K-automorphism with no square root, leading to a counterexample with an atomic Pinsker algebra).

The Pinsker algebra is a canonical object associated to every transformation: it is its greatest 0 entropy factor. The Bernoulli (and non Bernoulli) theory as developed by Ornstein [21], and then by Feldman, Rudolph, Shields, Smorodinsky, Weiss and many others had evolved in various directions, one of them being its relativization which is closely linked to the possibility of expressing a transformation as a direct product, one of the factors of this product being isomorphic to a Bernoulli shift [29]. It was then first noticed that the Ornstein–Shields K non Bernoulli examples [24] could be expressed as such products, and furthermore, that the entropy of the non Bernoulli factor could be made arbitrarily small. It was quickly realized that all known examples (which were not numerous at that time) satisfied this property, naturally called the weak Pinsker property. It is clearly weaker than the product described by the Pinsker conjecture as it does not single out a precise deterministic factor, but is at the same time stronger as the large entropy factors are just Bernoulli shifts which are the simplest possible positive entropy transformations, entirely described by their entropies (as a consequence of the Ornstein isomorphism theorem). The real meaning of this property is that there is no specific complexity attached to positive entropy and that the structure of every transformation is entirely described by factors of arbitrarily small entropy.

The objects which we are going to deal with are $(X,\mathscr{A},m,T)$, where T is an invertible measure preserving transformation of the Lebesgue space $(X,\mathscr{A},m)$. We shall frequently condense, when there is no ambiguity, the term $(X,\mathscr{A},m,T)$ into, simply, T. In this category of measure preserving actions on Lebesgue spaces, whose theory was developed by Rokhlin [26], we are going to use the following facts concerning factors and the decomposition into direct products:

(A) Given $(X,\mathscr{A},m,T)$ and a factor map f to $(Y,\mathscr{B},\mu,S)$ (i.e., $S.f = f.T$), calling $\tilde{\mathscr{B}} = f^{-1}(\mathscr{B})$ which is a T invariant sub σ-algebra of $\mathscr{B}$, the restriction of T to $\tilde{\mathscr{B}}$ which we denote $T_{\tilde{\mathscr{B}}}$ is isomorphic to $(Y,\mathscr{B},\mu,S)$. We shall therefore identify the restrictions of T to invariant sub-σ algebras to factor actions onto a Lebesgue space;

(B) In the same way, it is equivalent, for an invertible transformation T of a Lebesgue space $(X,\mathscr{A},m)$ to be isomorphic to a direct product $(Y_1,\mathscr{B}_1,\mu_1,T_1)\times(Y_2,\mathscr{B}_2,\mu_2,T_2)$ or to possess two T-invariant sub-σalgebras of $\mathscr{A}$, $\mathscr{A}_1$ and $\mathscr{A}_2$ such that $\mathscr{A}_1$ and $\mathscr{A}_2$ are independent and $\mathscr{A}_1\bigvee\mathscr{A}_2=\mathscr{A}$.

DEFINITION 3.1. *If a transformation T is isomorphic to a direct product $T_1\times T_2$, where T_2 is isomorphic to a Bernoulli shift, we say that the factor σ-algebra of T which corresponds to T_1 splits.*

The theorem of Tim Austin which solves a question that goes back to 1976 is the following [4]:

THEOREM 3.1. *Let $(X,\mathscr{A},m,T)$ be an ergodic invertible transformation with entropy $h(T)>0$. Let $\epsilon>0$. Then there exists an invariant factor σ-algebra $\mathscr{A}_1$ which splits such that $T_{\mathscr{A}_1}$ (the restriction of T to $\mathscr{A}_1$) has entropy smaller than ϵ.*

There are two extensions of this theorem (which Tim concentrates in only one statement):

THEOREM 3.2. *Let $(X,\mathscr{A},m,T)$ be an ergodic invertible transformation and $\mathscr{B}$ an invariant factor σ-algebra of $\mathscr{A}$ such that the relative entropy of the action of T conditioned on $\mathscr{B}$ is positive. Then, given $\epsilon>0$, there exists another factor σ-algebra $\mathscr{C}$ such that $\mathscr{B}\subset\mathscr{C}$, the entropy of the restriction of the action of T to $\mathscr{C}$ conditioned on $\mathscr{B}$ is $<\epsilon$ and $\mathscr{C}$ splits.*

Furthermore:

THEOREM 3.3. *All the preceding results extend to the action of countable amenable groups.*

It is the consequence of a result by Fieldsteel [8] that these theorems for $\mathbb{Z}$ actions extend to flows.

The main Theorem 3.1 is extremely striking. It has a considerable value as it is universal and gives very strong informations concerning the structure of measure preserving transformations; it says that one cannot hope for rigidity properties of positive entropy measure preserving transformations. They have, coming from the Bernoulli factor in the decomposition into a direct product, a big centralizer, they cannot be coalescent (a coalescent transformation is one which cannot have strict factors which are isomorphic to itself), etc. It was not even known, before Tim's result, whether every ergodic positive entropy transformation could be written as a direct product.

As one might expect, the proof is extremely remarkable. and quite intricate. Tim Austin proceeds with tools coming from information theory as created by Shannon (which is some kind of return to the origin), an important part does not contain any dynamics and only deals with measures on finite partitions in (finite) product spaces. In this framework, I am just going to give a list of the most important statements and information theoretic concepts which are used.

First of all the Kantorovich–Rubinstein duality linking two apparently different distances on measures on metric spaces. Let (K,d) be a compact metric space and μ and ν two probability measures on K. A coupling of μ and ν is a probability measure on $K\times K$ with respective marginals μ and ν. Let

$$\bar{d}(\mu,\nu)=\inf\left\{\int d(x,y)\,d\lambda : \lambda \text{ coupling of } \mu \text{ and } \nu\right\}$$

the Wasserstein transportation metric and

$$d^*(\mu,\nu)=\sup_{f\in \mathrm{Lip}_1(K)}\int f\,d\lambda - f\,d\mu$$

the d^* metric which was first devised by Fortet and Mourier [9]. The duality theorem of Kantorovich and Rubinstein is that $d^*=\bar{d}$.

The next important object is the Kullback–Leibler (K.L.) divergence between two probability measure μ and ν, it is $+\infty$ if ν is not absolutely continuous with respect to μ. otherwise,

$$D(\nu||\mu)=\int \frac{d\nu}{d\mu}\log\frac{d\nu}{d\mu}\,d\mu.$$

We recall the Hamming metric: let A be a finite set, n an integer, the Hamming distance between two elements of A^n , $x=(x_i, i\leq N)$ and $x'=(x'_i, i\leq N)$ is

$$d_n(x,x')=\frac{1}{n}\sum_{i=1}^{n}\delta(x_i,y_i).$$

Here $\delta(x_i,y_i)=1$ if $x_i\neq y_i$ and $\delta(x_i,y_i)=0$ if $x_i=y_i$.

As a first nice example of a connection between the K.L. divergence and the Wasserstein distance of measures associated to the Hamming metric on product spaces, let us mention the following result of Marton [19]:

THEOREM 3.4. *If A is a finite set and μ a product measure on A^n, then*

$$\bar{d}_n(\nu,\mu)\leq\sqrt{\frac{1}{n}D(\nu||\mu)}$$

for all probability measures ν on A^n.

This implies that if ν is the restriction of μ to a set of μ measure $> 2^{-n\epsilon}$, then $\bar{d}_n(\nu,\mu) < \sqrt{\epsilon}$, which is a nice presentation of the "concentration of measure" phenomenon.

Let us introduce the total correlation (T.C.) defined for a measure μ on a product space A^n. Let $\mu_i, 1 \le i \le n$, be the marginals of μ in A^n. Then

$$TC(\mu) = D(\mu||\mu_1 \times \mu_2 .. \times \mu_n).$$

The other very important quantity is the dual total correlation (D.T.C.) of a finite family of partitions $P_i, 1 \le i \le n$, of a measure space $(X, \mathscr{A}, m)$. Recall that for a k sets partition P of X whose atoms $p_1, p_2, \dots, p_k$ have measures $m(p_1), m(p_2), \dots, m(p_k)$, its Shannon entropy is

$$H(P) = -\sum_{i=1}^{k} m(p_i) \log m(p_i).$$

The conditional entropy of P given Q is then $H(P|Q) = H(P \bigvee Q) - H(Q)$. Coming back to our n partitions P_i, define

$$\hat{P}_i = \bigvee_{j \ne i} P_j.$$

Then [13]:

$$DTC(P_1, P_2, \dots, P_n) = H\Big(\bigvee_{i=1}^{i=n} P_i\Big) - \sum_{i=1}^{i=n} H(P_i|\hat{P}_i).$$

Note that $DTC(P_1, P_2, \dots, P_n) \ge 0$ and that if $DTC(P_1, P_2, \dots, P_n) = 0$, the P_i's are independent.

Tim Austin also needs a substitute to finite partitions, which he calls fuzzy partitions (commonly called partitions of unity): A fuzzy partition of a space (X, μ) is a finite sequence $\rho = (f_1, f_2, \dots, f_n)$ of positive functions which satisfy

$$\sum_{i=1}^{n} f_i = 1.$$

A related important number is, calling μ_i the measure $\frac{1}{\int f_i d\mu} f_i \mu$,

$$I_\mu(f_1, f_2, \dots, f_n) = \sum_{i=1}^{n} \int f_i D(\mu_i||\mu).$$

If the fuzzy partition is an actual partition P, i.e., $f_i = 1_{p_i}$,

$$I_\mu(1_{p_1}, 1_{p_2}, \dots, 1_{p_n}) = H(P).$$

He then introduces a main quantity which will be essentially put to work:

Let (K, d, μ) be a metric probability space, κ and $r > 0$. (K, d, μ) is said to have the property $T(\kappa, r)$ if every probability measure ν on K satisfies

$$\bar{d}(\nu, \mu) \le \frac{1}{\kappa} D(\nu||\mu) + r.$$

Let us note that, in case $X = A^n$ and $d = d_n$, when κ is large compared to n and r small, this means that under these conditions, the measure μ satisfies that, for every set E in A^n whose measure is not exponentially small, $\bar{d}(\mu_{|E},\mu)$ is small, which is an instance of the important extremality condition.

We quote now a proposition which presents the DTC decrement argument which constitutes a fundamental ingredient in the proof:

LEMMA 3.1. *Let $n > 0$, A be a finite set and μ be a probability measure on A^n which does not satisfy $T(rn/200, r)$. Then there is a fuzzy partition (ρ_1, ρ_2) such that:*

$$I_\mu(\rho_1,\rho_2) \geq r^2 n^{-1} e^{-n}$$

and

$$\int \rho_1 d\mu . DTC(\mu_{|\rho_1}) + \int \rho_2 d\mu . DTC(\mu_{|\rho_2}) \leq DTC(\mu) - \frac{1}{2} I_\mu(\rho_1,\rho_2).$$

This says that, in a quantified way, one can split μ into two parts, its restriction to every one being more concentrated. Let us note that, in case $DTC(\mu) < r^2 n^{-1} e^{-n}$, then μ satisfies $T(rn/200, r)$.

The proof is extremely interesting, and uses many different ideas and tools, starting with the use, if the suitable inequalities for DTC are not satisfied, of the duality theorem and of an analogue of some "logarithmic Sobolev inequalities". It is also the beginning of a long fractional distillation out of which the "structure" will eventually emerge. It takes the following form (note that part of the work is also to go from DTC to TC):

THEOREM 3.5. *Given $\epsilon, r > 0$, there exist $c > 0$ and $\kappa > 0$ such that for any finite set A the following will hold for sufficiently large n: Let μ be a probability measure on A^n and $E = TC(\mu)$, then there exists a family of disjoint sets in A^n, $U_1, U_2, \ldots, U_m$ such that:*

(1) $m \leq ce^{cE}$,

(2) $\mu(\bigcup_{i=1}^{i=m} U_i) > 1 - \epsilon$,

(3) $\mu_{|U_i}$ *satisfies* $T(\kappa n, r)$ *for all* $1 \leq i \leq m$.

Closely related to the preceding we say that a measure μ on A^n is ϵ,δ extremal if for every partition into less than $e^{n\delta}$ sets, $U_1, U_2, \ldots, U_k$, there is a subcollection S of them whose union has measure $> 1 - \epsilon$ such that for every $U_i \in S$, $\bar{d}_n(\mu_{|U_i},\mu) < \epsilon$. (Actually Tim Austin uses a continuous version of the extremality condition. For this presentation we shall not need it.) The Marton result already quoted says that if μ is a product measure on A^n with all its marginals equal, then it is $\sqrt{\epsilon}, \epsilon$ extremal. We shall now describe in a very sketchy way the dynamical version of this theorem which leads to the final result through the more classical use of the relative Bernoulli theory [29]. We shall only deal with finite entropy transformations therefore equipped, using Krieger's theorem, of a finite generator.

Given an ergodic invertible measure preserving transformation T of the probability space $(X, \mathscr{A}, \mu)$ together with a finite generating partition ξ, and a set $A \in \mathscr{A}$ we will look at $\mu_{|A}$ restricted to $\xi_{[0,n]}$ which we call $\mu^{\xi,n}_{|A}$ as a measure on ξ^n.

THEOREM 3.6. *Given $(X, \mathscr{A}, m, T)$ ergodic with positive entropy and ξ as above, for every ϵ, there exists n_0, there exists δ, such that for all $n > n_0$, one can produce a family of disjoint sets in ξ^n, $U_1, U_2, \ldots, U_m$, $m \leq e^{n\epsilon}$ such that*

$$\mu(U_1 \cup U_2 \cup \cdots \cup U_m) > 1 - \epsilon$$

and all the $\mu^{\xi,n}_{|U_j}$ are ϵ, δ extremal.

In a way, there is meaning to say that what has been achieved by Tim Austin is some kind of extremal decomposition. (This is written by a former student who attended in 1966 a Choquet course in Paris.)

Extremality is closely related to the Bernoulli property. Actually, for a system $(X, \mathscr{A}, m, T)$ equipped with a finite generator ξ, the Ornstein isomorphism theorem can be stated in the following way: it is equivalent for this system to be isomorphic to a Bernoulli shift with entropy $h(T)$ or to satisfy that for every ϵ there exists n_0 and $\delta > 0$ such that for all $n > n_0$, m restricted to ξ^n is ϵ, δ extremal. Using then a relative version of extremality, and a relative version of the above isomorphism, an almost direct corollary of the previous Theorem 3.6 is that there exists for every ϵ a factor algebra $\mathscr{B}$ of $(X, \mathscr{A}, m, T)$ restricted to which the entropy of T is smaller than ϵ and which ϵ-splits. This means that there exists a joining (λ) of $(X, \mathscr{A}, m, T)$ with the direct product $(T_{\mathscr{B}} \times S)$ where S is a Bernoulli shift in such a way that λ when restricted to $\mathscr{B}$ is the identity and ξ is ϵ-embedded in this direct product, that is $\xi \overset{\epsilon}{\subset} \mathscr{B} \times S \quad (\lambda)$. This is then promoted to exact splitting, as in [30].

Let us point out one fact about the organization of the splitting factors inside a transformation: there is an example, due to Kalikow [16], of a transformation T which possesses two splitting factors $\mathscr{B}_1$ and $\mathscr{B}_2$ such that $\mathscr{B}_1 \wedge \mathscr{B}_2 = \nu$. However, it is always true that, given T, if $\mathscr{B}_1$ and $\mathscr{B}_2$ are two splitting factors of equal entropy, $T_{\mathscr{B}_1}$ and $T_{\mathscr{B}_2}$ are isomorphic. There are many questions concerning the structure of the splitting factors. For instance, does there exist a K-automorphism T such that any factor $\mathscr{B}$ such that $T_{\mathscr{B}}$ is not isomorphic to a Bernoulli shift is isomorphic to a splitting factor?

4.

Although it is not part of the Brin Prize selection, I want to mention briefly a fourth work of Tim Austin which very closely belongs to the circle of ideas which have appeared in this presentation. It is ***Scenery entropy as an invariant of RWRS processes***. This (unfortunately unpublished) work deals with a generalization of the famous T, T^{-1} transformation which we now describe.

Given a probability space (P, π), the Bernoulli action $B(\pi)$ is defined on the product space $X = P^{\mathbb{Z}}$, equipped with the product measure $m = \pi^{\oplus \mathbb{Z}}$ as the shift $T : (x_n), n \in \mathbb{Z} \to (x_{n-1}), n \in \mathbb{Z}$. When P is the two sets space $(+1 -1)$ equipped with the measure $\pi = (1/2, 1/2)$, the T, T^{-1} skew product is the map $\hat{T}$ from $X \times X$ to $X \times X$ defined by $(x, y) \to (Tx, T^{x_0} y)$. It is clearly a "random walk on a random scenery" (R.W.R.S). It was shown by Meilijson [20] that $\hat{T}$ is a K-automorphism and by Kalikow [17] that it is not Bernoulli.

It is easy to give a more general form to the preceding skew-product, namely, given $a > 0$ and T_a the Bernoulli shift with entropy a acting on Y to consider $\hat{T}_a$ acting on $(X \times Y)$ by $(x, y) \to (Tx, T_a^{x_0} y)$. Clearly, for every $a > 0$, these transformations remain K with entropy $\log 2$. The theorem established by T. Austin is, in this simplest description, that when $a \neq b$, T_a is not isomorphic to T_b. This is quite remarkable and a new source of examples to the fact, already mentioned in this note [24], that there exist uncountably many K-automorphisms, with the same entropy, pairwise not isomorphic. Actually, the real statement of the theorem, resting on a continuous version of the T, T^{-1} transformation [27], is the following [5]:

THEOREM 4.1. *Let $(Y, \mathscr{B}, \mu, S)$ be a subshift of finite type equipped with an S-invariant measure μ which is a Gibbs measure for an Hölder continuous potential on Y, then if σ, $Y \to R, y \to \sigma(y)$ is Hölder continuous, is not a coboundary, and satisfies $\int \sigma(y) d\mu = 0$ together with an "enhanced invariance principle" and if (X_1, T_t^1, m_1) and (X_2, T_t^2, m_2) are two ergodic flows such that $h(T_1^1) \neq h(T_1^2)$, the skew products $(Y \times X_i, \hat{S}_i, m \times \mu_i)(i = 1, 2)$ defined by $\hat{S}_i(y, x) = (Sy, T^i_{\sigma(y)} x)$ are not isomorphic.*

The enhanced invariance principle was shown by Bromberg [6] (also, unfortunately, unpublished) entailing the following quite remarkable consequence:

THEOREM 4.2. *For any $h > 0$, there is a compact manifold with a smooth volume form that admits continuously many smooth volume preserving K-automorphisms of entropy h which are pairwise non isomorphic.*

As Tim Austin explains himself, the idea of building an invariant analogous to the “slow entropy” by counting the number of distributions that exhibit a high concentration does not seem to be relevant and the proof is closer to the “secondary entropy" as defined by Vershik [31]. However it is beautiful, quite elaborate, and the reader enjoys its meanders.

REFERENCES

[1] G. N. Arzhantseva, V. S. Guba and M. V. Sapir, Metrics on diagram groups and uniform embeddings in a Hilbert space, *Comment. Math. Helv.*, **81** (2006), 911–929.

[2] T. Austin, On the norm convergence of non-conventional ergodic averages, *Ergodic Theory Dynam. Systems*, **30** (2010), 321–338.

[3] T. Austin, Amenable groups with very poor compression into Lebesgue spaces, *Duke Math. J.*, **159** (2011), 187–222.

[4] T. Austin, Measure concentration and the weak Pinsker property, *Publ. Math. Inst. Hautes Études Sci.*, **128** (2018), 1–119.

[5] T. Austin, Scenery entropy as an invariant of RWRS processes, preprint, arXiv:1405.1468, 2014.

[6] M. Bromberg, Invariance principle for local time by quasi-compactness, preprint, arXiv:1511.01746, 2015.

[7] J.-P. Conze and E. Lesigne, Sur un théorème ergodique pour des mesures diagonales, *C. R. Acad. Sci. Paris Sér. I Math.*, **306** (1988), 491–493.

[8] A. Fieldsteel, Stability of the weak Pinsker property for flows, *Ergodic Theory Dynam. Systems*, **4** (1984), 381–390.

[9] R. Fortet and E. Mourier, Convergence de la répartition empirique vers la répartition théorique, *Ann. Sci. École Norm. Sup.*, **70** (1953), 267–285.
[10] H. Furstenberg, Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. Analyse Math.*, **31** (1977), 204–256.
[11] H. Furstenberg and Y. Katznelson, An ergodic Szemerédi theorem for commuting transformations, *J. Analyse Math.*, **34** (1978), 275–291.
[12] E. Guentner and J. Kaminker, Exactness and uniform embeddability of discrete groups, *J. London Math. Soc.*, **70** (2004), 703–718.
[13] T. S. Han, Linear dependence structure of the entropy space, *Information and Control*, **29** (1975), 337–368.
[14] B. Host and B. Kra, Convergence of Conze-Lesigne averages, *Ergodic Theory Dynam. Systems*, **21** (2001), 493–509.
[15] B. Host and B. Kra, Nonconventional ergodic averages and nilmanifolds, *Ann. of Math.*, **161** (2005), 397–488.
[16] S. Kalikow, Non-intersecting splitting σ-algebras in a non-Bernoulli transformation, *Ergodic Theory Dynam. Systems*, **32** (2012), 691–705.
[17] S. A. Kalikow, T, T^{-1} transformation is not loosely Bernoulli, *Ann. of Math.*, **115** (1982), 393–409.
[18] S. Khot and A. Naor, Nonembeddability theorems via Fourier analysis, *Math. Ann.*, **334** (2006), 821–852.
[19] K. Marton, A simple proof of the blowing-up lemma, *IEEE Trans. Inform. Theory*, **32** (1986), 445–446.
[20] I. Meilijson, Mixing properties of a class of skew-products, *Israel J. Math.*, **19** (1974), 266–270.
[21] D. Ornstein, Bernoulli shifts with the same entropy are isomorphic, *Advances in Math.*, **4** (1970), 337–352.
[22] D. S. Ornstein, A K automorphism with no square root and Pinsker's conjecture, *Advances in Math.*, **10** (1973), 89–102.
[23] D. S. Ornstein, A mixing transformation for which Pinsker's conjecture fails, *Advances in Math.*, **10** (1973), 103–123.
[24] D. S. Ornstein and P. C. Shields, An uncountable family of K-automorphisms, *Advances in Math.*, **10** (1973), 63–88.
[25] M. S. Pinsker, Dynamical systems with completely positive or zero entropy, *Dokl. Akad. Nauk SSSR*, **133** (1960), 1025–1026 (Russian); translated as *Soviet Math. Dokl.*, **1** (1960), 937–938.
[26] V. A. Rohlin, Lectures on the entropy theory of transformations with invariant measure, *(Russian) Uspehi Mat. Nauk*, **22** (1967), 3–56.
[27] D. J. Rudolph, Asymptotically Brownian skew products give non-loosely Bernoulli K-automorphisms, *Invent. Math.*, **91** (1988), 105–128.
[28] T. Tao, Norm convergence of multiple ergodic averages for commuting transformations, *Ergodic Theory Dynam. Systems*, **28** (2008), 657–688.
[29] J.-P. Thouvenot, Quelques propriétés des systèmes dynamiques qui se décomposent en un produit de deux systèmes dont l'un est un schéma de Bernoulli, *Israel J. Math.*, **21** (1975), 177–207.
[30] J.-P. Thouvenot, On the stability of the weak Pinsker property, *Israel J. Math.*, **27** (1977), 150–162.
[31] A. M. Vershik, Dynamic theory of growth in groups: Entropy, boundaries, examples, *Uspekhi Mat. Nauk*, **55** (2000), 59–128.
[32] Q. Zhang, On convergence of the averages, $.\frac{1}{N}\sum_{n=1}^{N} f_1(R^n x) f_2(S^n x) f_3(T^n x)$, *Monatsh. Math.*, **122** (1996), 275–300.
[33] T. Ziegler, Universal characteristic factors and Furstenberg averages, *J. Amer. Math. Soc.*, **20** (2007), 53–97.

JEAN-PAUL THOUVENOT <jean-paul.thouvenot@upmc.fr>: LPSM, Sorbonne Université, 4 Place Jussieu, 75005 Paris, France