Addendum: The Ergodic Hierarchy, Randomness and Hamiltonian Chaos

Roman Frigg and Charlotte Werndl

In this note we correct a few inaccuracies in 'The Ergodic Hierarchy, Randomness and Hamiltonian Chaos' by Berkovitz, Frigg and Kronz (2006), 'BFK' for short. None of the points mentioned here affect the central claims of the paper.

1. Construction of the σ-algebra $\sigma_{n,r}$ (p. 665).

When discussing K-systems, BFK say how to construct the minimal σ-algebra $\sigma_{n,r}$. The problem is that the method of construction as stated does not in general yield a set that has the structure of a sigma algebra. However, BFK do not need this. All that is needed for the analysis in the paper is that the sets specified in the 'construction' are *members of* the minimal sigma algebra $\sigma_{n,r}$, and this is the case.

The problem is that a sigma algebra like $\sigma_{n,r}$ cannot, in general, be constructed at all. The minimal σ-algebra of a set $E \subseteq P(X)$, $\sigma(E)$, is defined as

$$\sigma(E) := \bigcap_{A \subseteq P(X),\ E \subseteq A,\ A \text{ is a } \sigma\text{-algebra}} A, \tag{1}$$

where $P(X)$ is the power set of X. Now per definition $\sigma_{n,r} = \sigma(E')$, where $E' := \{T^k A_i \,|\, k \geq n;\, i = 1 \ldots r\}$. If E' had always a finite number of elements, then $\sigma(E')$ would be constructible. This is so because, first, there is an explicit algorithm for constructing the minimal algebra of a set E, $\alpha(E)$, and second, for finite E it holds that $\sigma(E) = \alpha(E)$ (see standard textbooks on measure theory). However, E' is not generally finite. For instance, for the Bernoulli partition of the baker's transformation E' is countably infinite. And for the minimal σ-algebra of a countably infinite set E the suggested construction method cannot work simply because *no* construction method can. If the suggested method worked, it would also work for the Borel σ-algebra of $[0, 1]$, which is the minimal σ-algebra of the countable set of all open intervals in $[0, 1]$ with rational endpoints, i.e. it is the set $\sigma(E'')$, where $E'' := \{(a, b) \,|\, a, b \in \mathbb{Q} \cap [0, 1]\}$. But the Borel σ-algebra of $[0, 1]$ is provably not constructible. Thus the suggested construction method cannot work.

2. The relation between the KS-entropy and K-systems (p. 667)

There are systems with positive KS-entropy that are not K-systems. Hence it is not the case that a system is a K-system *iff* it has positive KS-entropy. Instead, the relation between being a K-system and having positive KS-entropy is the following. For a dynamical system $[X, \Sigma, \mu, T]$ the entropy of T with respect to the partition $\alpha = \{\alpha_1, \ldots, \alpha_n\}$, $n \in \mathbb{N}$, is defined as

$$H(T, \alpha) := \lim_{k\to\infty} (1/k) H(\alpha \vee T\alpha \vee \ldots \vee T^{k-1}\alpha), \quad (2)$$

where $H(\alpha) := \sum_{i=1}^{n} \mu(\alpha_i) \log(\mu(\alpha_i))$ with $\mu(\alpha_i) \log(\mu(\alpha_i)) := 0$ for $\mu(\alpha_i) = 0$ (Cornfeld et al. 1982, pp. 246-250). A dynamical system $[X, \Sigma, \mu, T]$ is a K-system iff it has completely positive entropy, meaning that $H(T, \alpha)$ is positive for *every* non-trivial partition α, where a nontrivial partition is a partition which does not have an atom of measure one. See *ibid.*, p. 283, which is also cited by BFK.

The KS-entropy is defined as the supremum of $H(T, \alpha)$ over *all* partitions α. From this follows immediately that every K-system has positive KS-entropy. However, the converse is not true. That is, there are systems for which there is a non-trivial partition α' such that $H(T, \alpha') = 0$ and yet the KS-entropy is positive because there is another partition α'' for which $H(T, \alpha'') > 0$. Take, for instance, the dynamical system where $X := [-1, 1]$, Σ is the Lebesgue σ-algebra on $[-1, 1]$, μ is the normalised Lebesgue measure and

$$T(x) := 2x \text{ if } |x| \leq \frac{1}{2} \quad \text{and} \quad 2\,\mathrm{sgn}(x)(1 - |x|) \text{ if } \frac{1}{2} < |x| \leq 1, \quad (3)$$

where $\mathrm{sgn}(x)$ gives the sign of x. In fact, the system consists of the tent map and its mirror image (mirrored at the origin). One can prove that for the partition $\alpha' := \{[-1, 0), [0, 1]\}$ $H(T, \alpha') = 0$ and that for the partition $\alpha'' := \{[-1, 0), [0, 1/2), [1/2, 1]\}$ $H(T, \alpha'') > 0$. Hence this system has positive KS-entropy but is no K-system.

3. Bernoulli systems (pp. 670-671 and p. 677)

A and B in (C−B) do not range over all sets in Σ, but only over the atoms of the 'Bernoulli partition' α or one of its iterates $T^k\alpha$. Remember that for the Bernoulli partition α the independence condition holds: $\mu(T^n\alpha_i \cap T^m\alpha_j) = \mu(T^n\alpha_i)\mu(T^m\alpha_j)$ for all $\alpha_i, \alpha_j \in \alpha$ and $m \neq n$, $m, n \in \mathbb{Z}$. Moreover, (C−B) holds not only for positive integers, but for positive and negative integers. Hence a correct statement of (C−B) is:

$$C(T^n B, A) = 0 \text{ for all } n \in \mathbb{Z} \setminus \{0\} \text{ and for all } A, B \in \beta, \quad (4)$$

where $\beta := T^k\alpha$ for the *some* $k \in \mathbb{Z}$ (i.e. β is either the Bernoulli partition α or one of its iterates).

That the condition yields wrong results when applied to sets that are *not* atoms of the Bernoulli partition or one of its iterates can be seen in the following example. Consider the dynamical system of the baker's transformation, where X is the unit square, Σ is the Lebesgue σ-algebra, μ is the Lebesgue-measure and T is given by

$$T(x,y) := (2x, \frac{y}{2}) \text{ if } 0 \leq x < \frac{1}{2},\ (2x-1, \frac{y+1}{2}) \text{ if } \frac{1}{2} \leq x \leq 1. \tag{5}$$

Let $A := \{(x,y)\,|\,0 \leq x \leq 1,\ 1/2 \leq y \leq 3/4\}$ and $B := \{(x,y)\,|\,1/2 \leq x \leq 1,\ 0 \leq y \leq 1/2\}$. Then one easily calculates that $C(TB, A) = 3/16 \neq 0$.

The condition (C$-$B*) should be amended accordingly. The correct version is: since *all* $T^i\alpha$ are independent, for finite or infinite intersections of the form $T^nA_1 \cap T^{n+1}A_2 \cap \ldots$ the following condition holds:

$$C(T^nA_1 \cap T^{n+1}A_2 \cap \ldots, A_0) = 0, \tag{6}$$

for all $n \in \mathbb{N}$ and all $A_0, A_1, \ldots \in \beta$, where β is defined as above.

The claims of Section 3.5.4 (p. 677) need to be changed accordingly.

At this point it is also worth explaining in some detail why this condition holds (this point is only briefly mentioned in Footnote 8 of the paper, in which BFK refer the reader to Mañé's (1983, p. 87) formulation of the independence condition, which is basically the independence condition Petersen (1983) gives). The argument is as follows. Consider Petersen's (1983, p. 275) definition of independence of a partition: α is independent iff for any distinct powers $i_1, \ldots, i_r \in \mathbb{Z}$ and not necessarily distinct $A_j \in \alpha$, $j = 1, \ldots, r$,

$$\mu(T^{i_1}A_1 \cap \ldots \cap T^{i_r}A_r) = \mu(A_1)\ldots\mu(A_r). \tag{7}$$

For finite intersections (6) follows immediately. For infinite intersections consider the infinite intersection $B := A_0 \cap T^nA_1 \cap T^{n+1}A_2 \ldots$ with $A_i \in \beta$ arbitrary, where β is defined as above, and set $B_l := A_0 \cap T^nA_1 \cap \ldots T^{n+l}A_{l+1}$. Then $B_l \searrow B$ for $l \to \infty$, where $A_n \searrow A$ is defined as $A_0 \supseteq A_1 \supseteq \ldots$ and $\cap_{n\geq 0}A_n = A$. Since per assumption $\mu(X) < \infty$, μ is continuous from above, i.e. for $A_n \searrow A$ it holds that $\mu(A) = \lim_{n\to\infty}\mu(A_n)$. Therefore, we get from (7):

$$\mu(B) = \lim_{l\to\infty}\mu(B_l) = \lim_{l\to\infty}\mu(A_0)\prod_{m=0}^{l}\mu(A_{m+1}) = \mu(A_0)\prod_{m=0}^{\infty}\mu(A_{m+1}). \tag{8}$$

Likewise, for $B' := T^n A_1 \cap T^{n+1} A_2 \ldots$ with $A_i \in \beta$ arbitrary, we get that $\mu(B') = \prod_{m=0}^{\infty} \mu(A_{m+1})$. Hence

$$C(T^n A_1 \cap T^{n+1} A_2 \cap \ldots, A_0) = \mu(B) - \mu(A_0)\mu(B') = 0, \tag{9}$$

and this is all we need to show.

4. Probabilistic relevance and the relevance (p. 672)

The definition of the probabilistic relevance and the relevance is incomplete in that it makes no statement for the case $p(B^{t_1}) = 0$. Intuitively, a set that has probability zero cannot have any probabilistic relevance, and so it is natural to augment the definition as follows:

$$(\mathrm{R}_p)\ \ R_p(B^{t_1}, A^t) := p(A^t|B^{t_1}) - p(A^t) \text{ if } p(B^{t_1}) > 0 \text{ and } 0 \text{ otherwise.} \tag{10}$$

The definition of the relevance R should be amended accordingly:

$$(\mathrm{R})\ \ R(B^{t_1}, A^t) := R_p(B^{t_1}, A^t)p(B^{t_1}) = p(A^{t_1}\&B^t) - p(A^t)p(B^{t_1}) \quad \text{if } p(B^{t_1}) \neq 0 \text{ and } 0 \text{ otherwise.} \tag{11}$$

5. Condition (μ-DynT) for discrete systems (p. 674)

On p. 674 BFK state that the discrete instants are separated by unit time intervals. This can but need not be the case. For instance, if the discrete transformation is a Poincaré map of a continuous system, it is not usually the case that the instants of time are separated by unit time intervals. For this reason it is not generally true that $T_{t_i \to t} B_i = T^{t-t_i} B_i$ for all $i \in \{1, \ldots, r\}$.

This problem can be fixed by choosing the t_i such that they mark the times at which T is applied to X; that is, the mapping T is applied to the system at t_1, at t_2, at t_3, etc., and at no other time. Thereby we choose the labelling of the instants of time such that the instances are ordered 'backwards': $t_r < t_{r-1} < \ldots < t_1 < t$, where t is 'now'. It then follows that the map has been iterated i times between t_i and t; hence $T_{t_i \to t} B_i = T^i B_i$ for all $i \in \{1, \ldots, r\}$. Revising (μ-DynT) accordingly yields:

$$\begin{aligned} &\text{For all instants of time } t \text{ and for all } t_1, t_2, \ldots, t_r, \\ &\text{where } t_r < t_{r-1} < \ldots < t_1 < t, \text{ and for all } A, B_1, \ldots, B_r \in \Sigma \\ &p(A^t \& B_1^{t_1} \& \ldots \& B_r^{t_r}) = \mu(A \cap TB_1 \cap \ldots \cap T^r B_r). \end{aligned} \tag{12}$$

When discussing dynamical conditions (like K-mixing) that disregard the first n steps of $t_r < t_{r-1} < ... < t_n < \ldots < t_1 < t$, this condition becomes $p(A^t \& B_n^{t_n} \& \ldots \& B_r^{t_r}) = \mu(A \cap T^n B_n \cap T^{n+1} B_{n+1} \cap \ldots \cap T^r B_r)$.

The discussion that follows in the paper has to be adapted as just done for (μ-DynT), but none of Berkovitz et al.'s claims is affected by this.

Bibliography

Berkovitz, J., Frigg, R. and F. Kronz (2006). The Ergodic Hierarchy, Randomness and Hamiltonian Chaos. *Studies in the History and Philosophy of Modern Physics*, 37, pp. 661-691.

Cornfeld, I.P., Fomin, S.V. and Ya.G. Sinai (1982). *Ergodic Theory.* Berlin and New York: Springer (=Grundlehren der mathematischen Wissenschaften 245).

Mañé, R. (1983). *Ergodic Theory and Differentiable Dynamics.* Berlin and New York: Springer.

Petersen, K. (1983). *Ergodic Theory.* Cambridge: Cambridge University Press (=Cambridge Studies in Advanced Mathematics 2).