Basics about entropy

The summary below is superceded by some later lectures, but at this point can serve as an introduction.

Let $(X, \mathcal{F}, \mu, T)$ be an ergodic measure preserving system.

Partitions

A partition is a collection $\alpha = \{A_1, \ldots, A_k\}$ of pairwise disjoint measurable sets whose union is X (up to measure 0). We assume always that our partitions are finite. A partiiton $\beta = \{B_1, \ldots, B_m\}$ refines α if every B_i is a subset of some A_j (i.e. every A_j is a union of some of the B_i). The join of α, β is the partition

$$\alpha \vee \beta = \{A \cap B \,:\, A \in \alpha\,,\, B \in \beta\}$$

We write

$$T^{-k}\alpha = \{T^{-k}A_1, \ldots, T^{-k}A_k\}$$

This is also a partition (the fact that it covers X up to measure 0 uses the fact that T prserves μ). Note that

$$x \in T^{-k}A_i \iff T^k x \in A_i$$

Write

$$\alpha^n = \alpha \vee T^{-1}\alpha \vee \ldots \vee T^{-(n-1)}\alpha = \bigvee_{i=0}^{n-1} T^{-i}\alpha$$

Let $\alpha(x) = A_i$ if $x \in A_i$; this defines $\alpha(x)$ uniquely since the elements of α are disjoint and cover X.

Entropy

For every partition α, there is a (unique) number $h = h_\mu(T, \alpha) \geq 0$ such that

$$\mu(\alpha^n(x)) \approx 2^{-hn} \qquad \text{for } \mu\text{-a.e. } x$$

i.e.

$$\lim_{n\to\infty} -\frac{1}{n} \log \mu\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha(x)\Big) = h \qquad \text{for } \mu\text{-a.e. } x$$

This is not the usual definition of $h_\mu(T, \alpha)$, rather it is the Shannon-McMillan-Breiman theorem, but it is equivalent. Note that we have not defined entropy for non-ergodic systems (it is possible to do so but we do not need it).

The entropy of T is

$$h_\mu(T) = \sup\{h_\mu(T, \alpha) \,:\, \alpha \text{ a partition of } X\}$$

Generators and approximate generators

A partition α is called a generator for T if there is a set $X_0 \subseteq X$ of full measure such that for every $x, y \in X_0$ if $x \neq y$ then $(T^{-k}\alpha)(x) \neq (T^{-k}\alpha)(y)$ for some k. This is equivalent (under mild technical assumptions on the measure space) to $\sigma(T^{-k}\alpha \,:\, k \in \mathbb{N}) = \mathcal{F} \bmod \mu$.

If $\alpha = \{A_1, \ldots, A_k\}$ is a generator then $h_\mu(T, \alpha) = h_\mu(T)$ (i.e. the supremum in the definition of entropy is attained for α) and $h_\mu(T) \leq \log k$. It is a nontrivial fact that if $h_\mu(T) < \infty$ then there is a generator with $\lfloor h_\mu(T) + 1 \rfloor$ elements.

If β refines α then $h_\mu(T, \beta) \geq h_\mu(T, \alpha)$. If $\alpha_1, \alpha_2, \alpha_3, \ldots$ are a refining sequence of partitions which together generate the σ-algebra $\mathcal{F}$ up to measure 0, then $h_\mu(T) = \lim h_\mu(T, \alpha_n)$.

Entropy-typical sequences, and the "AEP"

For every $\varepsilon > 0$, let

$$\alpha_\varepsilon^n = \{A \in \alpha^n \,:\, 2^{-n(h+\varepsilon)} < \mu(A) < 2^{-n(h-\varepsilon)}\}$$

Then

$$|\alpha_\varepsilon^n| < 2^{n(h+\varepsilon)}$$

and given $\varepsilon > 0$, for μ-a.e. x we have $\alpha^n(x) \in \alpha_\varepsilon^n$ for all large enough n. Furthermore if $\beta_n \subseteq \alpha^n$ and $|\beta_n| \leq 2^{n(h-\varepsilon)}$ then for μ-a.e. x we have $\alpha^n(x) \notin \beta_n$ for all large enough n (Proof:

$$\mu(\bigcup\{A \in \beta_n \cap \alpha_\varepsilon^n\}) \leq |\beta_N| \cdot 2^{-n(h+\varepsilon)} < 2^{-2\varepsilon n}$$

and this is summable. By Borel-Cantelli, μ-a.e. x satisfies $\alpha^n(x) \notin \beta_n \cap \alpha_\varepsilon^n$ for all large enough n. But also $\alpha^n(x) \in \alpha_\varepsilon^n$ for all large enough n; together, this means $\alpha^n(x) \notin \beta_n$ for all large enough n).