The pointwise ergodic Theorem
Also called Birkhoff's ergodic Theorem
or just the ergodic theorem

<u>Theorem</u>: $f : (X, \mathcal{B}, \mu) \circlearrowleft$ is a mpt, $\alpha \in L^1(\mu)$ then

(1) $\exists\, \alpha^* \in L^1(\mu)$ so that

$$\frac{1}{n} \sum_{i=0}^{n-1} \alpha(f^i(x)) \to \alpha^*(x) \quad \text{for a.e. } x$$

(time average converges a.e.)

(2) $\alpha^* \circ f = \alpha^*$ a.e.

(3) $\int_X \alpha^* \, d\mu = \int_X \alpha \, d\mu$

NOTE: The ergodic theorem doesn't mention ergodicity!

probability measure

CORR $f : (X, \mathcal{B}, \mu) \circlearrowleft$ is an *ergodic* mpt and $\alpha \in L^1(\mu)$ then

$$\frac{1}{n} \sum_{i=0}^{n-1} \alpha(f^i(x)) \to \int_X \alpha \, d\mu \quad \text{a.e.}$$

(time average = space average a.e.)

Proof (given Birkhoff). By (2) of Birkhoff and previous theorems α^* is constant a.e. Say with value c.

then $c = \int_X \alpha^* d\mu = \int_X \alpha \, d\mu$ by (3)

and so by (1) $\frac{1}{n} \sum_{i=0}^{n-1} \alpha(f^i(x)) \to \alpha^*(x) = c = \int \alpha \, d\mu$

a.e. ∎

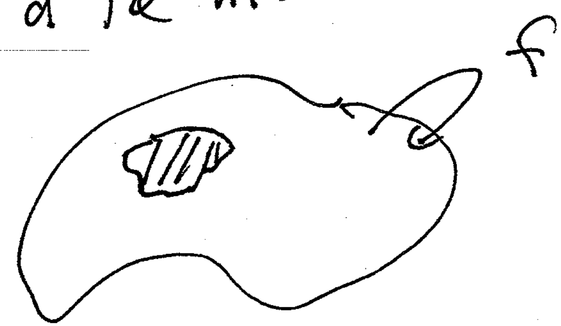
Two applications

(1) Say f is mpt and $A \in \mathcal{B}$, how much time on average does $O^+(x, f)$ spend in A. Assume f is ergodic. Now $\chi_A(f^i(x)) = 1 \Leftrightarrow f^i(x) \in A$.

Using the ergodic theorem

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_A(f^i(x)) = \int \chi_A \, d\mu = \mu(A) \quad \text{a.e. } x$$

So the average number of iterates which land in A tends to $\mu(A)$ which since $\mu(X) = 1$ say, is the relative proportion of the measure of X which lives in A.



(2) (a little informal)

- Recall that $D: S^1 \circlearrowleft$ given by $D(\theta) = 2\theta \mod 1$ is ergodic. This implies that $T: [0,1) \circlearrowleft$ given by $T(x) = 2x \mod 1$ is also ergodic w.r.t. Lebesgue measure
- Now let $Y \subseteq [0,1)$ be all y with a well-defined unique binary expansion so $Y = [0,1)$ minus a countable set, so $\mu(Y) = 1$ (μ is lebesgue)
- Now If $y = \frac{a_1}{2} + \frac{a_2}{2^2} + \dots$ is the unique binary expansion of some $y \in Y$ then

$$T(y) = \frac{a_2}{2} + \frac{a_3}{2^2} + \dots$$

notice $T(Y) = Y$

upper case Y

- Now let $\alpha(x) = \chi_{[1/2,1)}(x)$, so $\alpha(x) = 1$ exactly when $x = \frac{1}{2} + \dots$ i.e. $x = .1\dots$

- Thus $\alpha(T^i(x)) = \alpha\left(\frac{a_{i+1}}{2} + \frac{a_{i+2}}{2^2} + \cdots\right)$

$$= \begin{cases} 1 & \Leftrightarrow a_{i+1} = 1 \\ 0 & \Leftrightarrow a_{i+1} = 0 \end{cases}$$

- Thus $\sum_{i=0}^{n-1} \alpha(T^i(x))$ counts the number of ones in the first n digits of the binary expansion of $x \in Y$

and $\frac{1}{n}\sum_{i=0}^{n-1} \alpha(T^i(x))$ is the average number of ones in the first n digits of $x \in Y$

and $\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \alpha(T^i(x))$ is the asymptotic average of the number of ones in the binary expansion of $x \in Y$

- The ergodic theorem says

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \alpha(T^i(x)) \longrightarrow \int \chi_{[1/2,1)}(x)\, d\mu = 1/2$$

for a.e. $x \in Y$

- So we have Borel's Theorem on normal numbers

<u>Theorem</u>: Almost every (with respect to Lebesgue measure) number in $[0,1)$ has the asymptotic frequency of the occurrence of 1 in its expansion equal to 1/2

[and obviously the same for 0's]

The main estimate for the Ergodic Theorem comes from the MAXIMAL Ergodic Theorem. We first state the concrete version that is actually used. It follows from a more general version about operator averages

Theorem: $f\{(X, \mathcal{B}, \mu)\}$ is an mpt $\alpha \in L^1_{\mathbb{R}}(\mu)$.

If $B_a = \{x \in X : \sup_{n \geq 1} \frac{1}{n} \sum_{i=0}^{n-1} \alpha(f^i(x)) > a\}$

and $f^{-1}(A) = A$ then

$$\int_{B_a \cap A} \alpha \, d\mu \geq a \, \mu(B_a \cap A)$$

- Remark: Looks like Chebychev inequality but is about averages apply to $-\alpha, -a$ yields upper bound!

- The proof and formulation of the abstract MAXIMAL ergodic theorem requires a bit of Linear operator theory. We restrict to Banach spaces

- $L : B_1 \to B_2$, is a linear transformation in the usual sense

- L is bounded if

$$\sup \frac{\| Lx \|_{B_2}}{\| x \|_{B_1}} = M < \infty$$

- In this case $M = \| L \|_{B_1, B_2}$, the usual case is $B_1 = B_2$ and this is just written $\| L \|$

- By linearity $\displaystyle \| L \| = \sup_{\| x \|_{B_1} = 1} \frac{\| Lx \|_{B_2}}{\| x \|_{B_1}}$

- $\| L \|$ is called the operator norm.

- The most common case is $L : L^p(\mathbb{X}, \mathcal{B}, \mu) \circlearrowleft$ and one just writes $\|L\|_p$ or $\|L\|$

- Now let $L^0(\mathbb{X}, \mathcal{B}, \mu)$ be all measurable functions $\boxed{\mathbb{X} \xrightarrow{f} \mathcal{Y} \xrightarrow{\alpha} \mathbb{C}}$

  $\alpha : \mathcal{Y} \to \mathbb{C}$ define

  $U(\alpha) = \alpha \circ f$ which is a linear transformation when $f : \mathbb{X} \to \mathcal{Y}$ is measurable from $(\mathbb{X}, \mathcal{B}, \mu)$ to $(\mathcal{Y}, \mathcal{B}', \mu')$ and $U : L^0(\mathcal{Y}, \mathcal{B}', \mu') \to L^0(\mathbb{X}, \mathcal{B}, \mu)$

  (contravariant in the jargon.)

- We need to understand how U acts on L^p-spaces.

Change of variables

$f : (X, \mathcal{B}, \mu) \to (X', \mathcal{B}', \mu')$ is measure preserving and $U(\alpha) = \alpha \circ f$

then $\int U(\alpha)\, d\mu = \int \alpha\, d\mu'$

$$X \xrightarrow{f} X' \xrightarrow{\alpha} \mathbb{C}$$

$\mu \qquad \mu'$

$$= \int \alpha \circ f\, d\mu$$

where if one side doesn't exist or is infinite the other side has the same ideas.

Proof idea. Let $\alpha = \chi_A$ for $A \in \mathcal{B}'$ and then

$$\int_{X'} \alpha\, d\mu' = \int_{X'} \chi_A\, d\mu' = \mu'(A) = \mu(f^{-1}(A))$$

$$= \int_X \chi_{f^{-1}(A)}\, d\mu = \int_X \chi \circ f\, d\mu = \int U(\alpha)\, d\mu$$

Assume $\alpha \geq 0$, approximate by simple functions, etc.

CORR If $f:(U,\mathcal{B},\mu)\,\circlearrowleft$ is measure preserving then $U(\alpha)=\alpha\circ f$ is a linear operator $U: L^p(U,\mathcal{B},\mu)\,\circlearrowleft$ with $\|U\|_p = 1$

Proof Assume $\alpha \in L^p(U,\mathcal{B},\mu)$, by change of variables

$$\int |\alpha|^p d\mu = \int |\alpha\circ f|^p d\mu$$

so
$$\frac{\|U\alpha\|}{\|\alpha\|} = \left(\frac{\int|\alpha\circ f|^p d\mu}{\int|\alpha|^p d\mu}\right)^{1/p} = 1$$

A comment on change of variables

- Say $(X, \mathcal{B}, \mu)$ is a measure space and $f : X \to Y$ is onto. Define a σ-algebra on Y as $A \in \mathcal{B}' \Leftrightarrow f^{-1}(A) \in \mathcal{B}$ and a measure μ' on $\mathcal{B}'$ by $\mu'(A) = \mu(f^{-1}(A))$, then it is usually written $\mu' = f_* \mu$, the push forward of μ

$$X \xrightarrow{f} Y \xrightarrow{\alpha} \mathbb{C}$$

Change of variables is then

$$\int_X \alpha \circ f \, d\mu = \int_Y \alpha \, d f_* \mu$$

- Sometimes $\alpha \circ f$ is written $f^* \alpha$ the pull back of α. Then $\int f^* \alpha \, d\mu = \int \alpha \, d f_* \mu$
- When $f : (X, \mathcal{B}, \mu) \circlearrowleft$ preserves measure, $f_* \mu = \mu$ so $\int_X \alpha \circ f \, d\mu = \int_X \alpha \, d\mu$

More algebraically, Integration gives a bilinear pairing

$$\text{functions} \times \text{measures} \longrightarrow \mathbb{C}$$
$$\alpha \quad\quad \mu \quad \mapsto \int \alpha\, d\mu$$

writing $[\alpha, \mu] = \int \alpha\, d\mu$

$\Rightarrow$ $[f^*\alpha, \mu] = [\alpha, f_*\mu]$ is the change of variables formula expressed as adjoints.