
Before we begin let's fix notation.

For $n \in \mathbb{N}$, define

$$D_n = \{ \underline{v} \in \mathbb{Z}_+^d \mid \underline{v} < (n, n, \ldots, n) \}$$

For $\underline{n} \in \mathbb{Z}_+^d$, define

$$D_{\underline{n}} = \{ \underline{v} \in \mathbb{Z}_+^d \mid \underline{v} < \underline{n} \}$$

Note: We will use an over line to denote a member of $\mathbb{Z}_+^d$.

We will be working over $(X, d, \mathcal{T})$, where $(X, d)$ is a compact metric space, and $\mathcal{T}$ is a continuous $\mathbb{Z}_+^d$-action.

Let $\mathcal{M}(X)$ denote the space of Borel regular probability measures on $X$.

$\mathcal{M}(X, \mathcal{T})$ denote the subspace of $\mathcal{M}(X)$ of $\mathcal{T}$-invariant measures.

If $\alpha = (A_1, \ldots, A_k)$ is a partition of $X$, $\gamma \in \mathbb{Z}_+^d$, $\mu \in \mathcal{M}(X, \mathcal{T})$

$$\alpha^\gamma = \bigvee_{\gamma \in \mathcal{C}} \phi^\gamma \alpha$$

$$H_{\mu} (\alpha) = \sum_{i=1}^{k} \varphi (\mu(A_i)) \quad \varphi (x) = x \log x$$

$$h_{\mu} (\gamma, \alpha) = \lim_{n \to \infty} \frac{1}{n^d} H(\alpha^{D_n})$$

$$h_{\mu} (\gamma) = \sup_{\alpha} h_{\mu} (\gamma, \alpha)$$
Let $S_{\alpha}, E \subseteq X$, is called $(n, \delta)$-separated if

$$\forall x, y \in E, x \neq y, \exists \zeta \in D_n \text{ st } d(T^\zeta(x), T^\zeta(y)) > \delta.$$ 

$E \subseteq X$ is called $(n, \delta)$-spanning if

$$\forall x \in X, \exists y \in E, \forall \zeta \in D_n \text{ st } d(T^\zeta(x), T^\zeta(y)) < \delta.$$ 

Given $f \in C(X), S_{\alpha},$ define:

$$S_{\alpha}f(x) = \sum_{\zeta \in D_n} f(T^\zeta(x))$$

$$P_{\alpha, \delta}(T, f) = \sup \{ \sum_{x \in E} S_{\alpha}f(x) : E \text{ is } (n, \delta) \text{-separated} \}$$

$$P_{\delta}(T, f) = \limsup_{n \to \infty} \log \left( \frac{P_{\alpha, \delta}(T, f)}{n^d} \right)$$

$$P(T, f) = \lim_{\delta \to 0} P_{\delta}(T, f)$$

**Lemma:** Let $a_1, \ldots, a_k \in \mathbb{R}.$ If $\sum_{i=1}^k P_i = 1,$ then

$$\sum_{i=1}^k P_i \log P_i \leq \log \left( \sum_{i=1}^k e^{a_i} \right)$$

with equality iff

$$P_i = \frac{e^{a_i}}{\sum_{j=1}^k e^{a_j}}$$
Theorem (Variational Principle)

\[ P(T, f) = \sup \{ h_\mu(T) + \int_T f \, d\mu \mid \mu \in M(X, T) \} \]

We will prove this 2 steps.

1. \[ P(T, f) \geq \sup \{ h_\mu(T) + \int_T f \, d\mu \mid \mu \in M(X, T) \} \]
2. \[ P(T, f) \leq \sup \{ h_\mu(T) + \int_T f \, d\mu \mid \mu \in M(X, T) \} \]

Proof of 1:

Let \( \mu \in M(X, T) \) be a partition. Let \( \alpha > 0, \beta > 0 \) such that

\[ \forall x, y \in X, \quad |x - y| < \alpha \quad \text{and} \quad |f(x) - f(y)| < \beta \]

Since \( \mu \) is regular, we have \( f \) compact \( B_j \subseteq A_j \) s.t.

\[ \mu(A_j \setminus B_j) < \epsilon \]

We define a new partition \( B \) by

\[ B = \{ B_0, B_1, \ldots, B_k \} \quad \text{with} \quad B_0 = X \setminus \bigcup_{i=1}^k B_i \]

Since \( B_i \) are compact and disjoint,

\[ b = \min_{1 \leq i \neq j \leq k} d(B_i, B_j) > 0 \]

Since \( X \) is compact, \( f \) is uniformly cts. So \( \forall \delta > 0, \delta < \frac{b}{2} \), it

\[ \forall x, y \in X, \quad d(x, y) < \delta \Rightarrow |f(x) - f(y)| < \beta. \]
Let \( n \in \mathbb{Z}_+ \) and let \( E \) be \((n,s)\)-separated wrt \( T \), i.e.

\[
\forall x \neq y \in E, \ \forall t \in \mathcal{D}_n, \ d(T^t(x), T^t(y)) \geq s
\]

Why suppose \( E \) fails to be \((n,s)\)-separated if any point is added. So \( E \) is \((n,s)\)-spanning, i.e.

\[
\forall x \in X, \ \exists y \in E, \ \forall t \in \mathcal{D}_n \ s.t. \ d(T^t(x), T^t(y)) \leq s
\]

For \( C \in \beta^{\mathcal{D}_n} \), define

\[
\delta_C := \sup \{(S_n f)(x) | x \in C\}
\]

Now let's estimate!

\[
\begin{align*}
H_n(\beta^{\mathcal{D}_n}) &+ \int S_n f \ dm \\
\leq & \sum_{C \in \beta^{\mathcal{D}_n}} -\psi(\mu(C)) + \int_C S_n f \ dm \\
\leq & \sum_{C \in \beta^{\mathcal{D}_n}} -\psi(\mu(C)) + \delta_C \mu(C) \\
= & \sum_{C \in \beta^{\mathcal{D}_n}} \mu(C) (\delta_C - \log \mu(C)) \\
\leq & \log \left( \sum_{C \in \beta^{\mathcal{D}_n}} e^{\delta_C} \right), \ \text{by lemma}
\end{align*}
\]

By extreme value theorem, \( \delta_C \) is obtained for some \( x \in C \). Since \( E \) is \((n,s)\)-spanning, \( \exists y \in E \) such that

\[
d(T^t(x), T^t(y)) \leq s, \ \forall t \in \mathcal{D}_n
\]

4
\[ |x - S_n f(y_c)| = |S_n f(x) - S_n f(y_c)| \]
\[ \leq \sum_{c \in \mathcal{C}_n} |f(T^c(x)) - f(T^c(y_c))| \quad \text{triangle inequality} \]
\[ \leq \sum_{c \in \mathcal{C}_n} \varepsilon \quad \text{since} \quad d(T^c(x), T^c(y_c)) < \varepsilon \]
\[ = n^d \varepsilon \quad \text{E is (n, \varepsilon) - spanning}. \]

So we have
\[ |x - S_n f(y_c)| + n^d \varepsilon \]

Since \( \frac{a}{2} < b \), we have each ball of radius \( a \) meets the closures of at most 2 members of \( \beta \). For each \( y \in E \), define
\[ A_y = \{ c \in \beta^{D_n} \mid y_c = y \} \]

Since each \( c \in \beta^{D_n} \) is of the form
\[ c = \bigcap_{i \in \mathcal{I}} T^{c_i} (B_{j_i}) \quad \text{for some} \quad j_i \in \mathcal{E}_0, \ldots, k_3. \]

For each \( c \), s.t. \( y_c \neq y \), there are at most 2 choices of \( j_i \). Thus
\[ |A_y| \leq 2^{10n_1} = 2^{n^4} \]

Now we have
\[ H_m(\beta^{dn}) + \int S_n f d\mu \]
\[ \leq \log \sum_{C \in \beta^{dn}} e^{dC} \]
\[ \leq \log \sum_{C \in \beta^{dn}} e^{S_n f(g_C)} + n^d e \]
\[ = \log \left( e^{n^d e} \sum_{g \in G} |A_{g,C}| e^{S_n f(g_C)} \right) \]
\[ \leq \log \left( e^{n^d e} \sum_{g \in G} e^{S_n f(g)} \right) \]
\[ = n^d e + n^d \log 2 + \log \left( \sum_{g \in G} e^{S_n f(g)} \right) \]
\[ \leq n^d a + n^d \log 2 + \log (P_{n,s}(C,f)) \quad \text{since } e \leq a \]

Dividing both sides by \( n^d \),

\[ \frac{1}{n^d} H_m(\beta^{dn}) + \int f d\mu \leq a + \log 2 + \frac{1}{n^d} \log (P_{n,s}(C,f)) \]
\[ \quad \text{since } \mu \in M(C,T) \]

Letting \( n \to \infty \), we get

\[ h_m(C,\beta) + \int f d\mu \leq a + \log 2 + P_s(C,f) \leq a + \log 2 + P(C,f). \]

We need to turn that \( \beta \) into an \( \alpha \). Recall:

\[ h_m(C,\alpha) \leq h_m(T,\beta) + H_m(\alpha|\beta) \]

So we have

\[ h_m(C,\alpha) + \int f d\mu \leq a + \log 2 + H(\alpha|\beta) + P(C,f). \]
Intuitively, the elements in $\beta$ are very close approximations of $\alpha$, so we should have $H(\alpha | \beta)$ be very small.

\[
H_{\mu}(\alpha | \beta) = -\sum_{i=0}^{|\kappa|} \mu(\beta_i) \sum_{j=1}^{i-1} \varphi(\mu(A_j | \beta_i))
\]

\[
= -\mu(\beta_0) \sum_{j=1}^{i-1} \varphi(\mu(A_j | \beta_0)) \quad \text{since for } 1 \leq i, j \leq k
\]
\[
\quad \mu(A_j | \beta_i) \in \{0, 1\}
\]
\[
\leq \mu(\beta_0) \log |\kappa|
\]
\[
\leq \varepsilon \cdot |\kappa| \log |\kappa| , \quad \text{since } \beta_0 = \bigcup_{i=1}^{|\kappa|} \mu(A_i | \beta_i), \mu(A_i | \beta_i) \leq \varepsilon
\]
\[
< \vartheta
\]

Thus
\[
h(\mathcal{Z}) + \int f d\mu \leq 2\vartheta + \log 2 + P(\mathcal{Z}, f).
\]

Since
\[
h_{\mu}(\mathcal{Z}^n) = \sum_{i=0}^{n-1} h_{\mu}(\mathcal{Z}) , \quad \mathcal{Z}^n = (\mathcal{Z}, \ldots, \mathcal{Z})
\]
\[
\int S_f d\mu = \sum_{i=0}^{n-1} \int f d\mu , \quad \mu \in M(C, \mathcal{Z}).
\]
\[
P(\mathcal{Z}^n, S_n f) = \sum_{i=0}^{n-1} P(\mathcal{Z}, f)
\]

We get by replacing $\mathcal{Z}$ with $\mathcal{Z}^n$ and $f$ with $S_n f$ that, for all $n \in \mathbb{N}$,

\[
h(\mathcal{Z}) + \int f d\mu \leq \frac{2\vartheta + \log 2 + P(\mathcal{Z}, f)}{h_{\mu}}
\]

\[
= \frac{1}{h_{\mu}}
\]

\[
h(\mathcal{Z}) + \int f d\mu \leq P(\mathcal{Z}, f)
\]
Proof of (2):

Let $\theta_0$, we will find a $\mu \in M(X, \mathcal{U})$ such that

$$h_\mu (\mathcal{U}) + \int \log \frac{P_\mu (\mathcal{U}, f)}{P_0 (\mathcal{U}, f)} \, d\mu \geq 0.$$ 

Let $E_n$ be an $(n, s)$- separated set with

$$\log \sum_{y \in E_n} e^{S_n f(y)} \geq \log P_{n, s} (\mathcal{U}, f) - 1$$

Let us define $\nu_n \in M(X)$ that is atomic on $E_n$ by

$$\nu_n = \sum_{y \in E_n} e^{S_n f(y)} \delta_y \frac{\sum_{z \in E_n} e^{S_n f(z)}}{\sum_{z \in E_n} e^{S_n f(z)}}$$

Then define $\lambda_n \in M(X)$ by

$$\lambda_n = \frac{1}{h} \sum_{E_n} \nu_n \otimes T^{-1}$$

Now by Reisz-Representation theorem the space of Borel-regular measures is isometrically isomorphic to $C(X)^\ast$, so we can impose the weak $^\ast$ topology on $M(X)$. Since the space of probability measures are closed subset of the unit ball, with is compact by Banach-Alaoğlu, we have $M(X)$ is compact.
So there is a subsequence $n_j$ such that

$$M_{n_j} \to M, \ j \to \infty$$

Weakly for some $\mu \in M(X)$.

I claim that $\mu \in M(X, C)$. Let's show this. Fix $\bar{K} \in \mathbb{R}^d$.

$$\left| \int f \circ T^{\bar{K}} \, d\mu - \int f \, d\mu \right|
\leq \lim_{j \to \infty} \left| \int f \circ T^{\bar{K}} \, dM_{n_j} - \int f \, dM_{n_j} \right|
\leq \lim_{j \to \infty} \frac{1}{n^d} \int \sum_{\bar{z} \in D_{n_j} + \bar{K}} \sum_{\bar{z} \in D_{n_j} \cap \Delta D_{n_j}} f \circ T^{\bar{z}} \, d\sigma^n
\leq \lim_{j \to \infty} \frac{1}{n^d} \left| D_{n_j} + \bar{K} \Delta D_{n_j} \right| \|f\|_{\infty}
= \lim_{j \to \infty} \frac{1}{n^d} \left( n^d - \text{vol}((\bar{n} - \bar{K})) \right) \|f\|_{\infty}
= 0$$

Thus $\mu \in M(X, C)$. 
**Lemma:** \( \exists \) a \( s \)-fine, \( \mu \)-good partition \( \mathcal{A} = \{ A_1, \ldots, A_n \} \).

\[ \text{ie} \quad \text{diam} (A_i) < s \quad \forall i \quad \mu (\partial A_i) = 0. \]

**pf:** First let's show why we can find a \( s \)-fine, \( \mu \)-good ball.

Let \( r < \frac{s}{2} \) \( \forall x \in X \),

\[ B(x, r) = \bigcup_{r' < r} \partial B(x, r') \]

Since the union is disjoint and uncountable, \( \exists r' < r \) st

\[ \mu (\partial B(x, r')) = 0 \]

By compactness we can cover \( X \) by balls \( B_1, \ldots, B_k \) st

\[ \text{diam} (B_i) < s \]

Let \( A_i = B_i \), \( A_n = B_n \setminus \bigcup_{i=1}^{n-1} B_i \). The result follows since

\[ \partial A_n \subseteq \bigcup_{i=1}^{n-1} \partial B_i \]

So let \( \mathcal{A} = \{ A_1, \ldots, A_n \} \) be a \( \mu \)-good, \( s \)-fine partition.

Since \( \mathcal{A} \) is \( s \)-fine we have each element of \( \mathcal{A} \) has

at most one element from \( E_n \).

Now lets do some algebra!
\[ H_{\sigma_n}(x^{D_n}) + \int S_n f \, d\sigma_n \]
\[ = \sum_{y \in E_n} -\sigma_n(y) \log \sigma_n(y) + S_n f(y) \sigma_n(y) \]
\[ = \sum_{y \in E_n} \sigma_n(y) \left( S_n f(y) - \log \sigma_n(y) \right) \]
\[ = \log \left( \sum_{y \in E_n} S_n f(y) \right) \]
\[ \geq \log P_{n,8}(\tau, f) - 1 \]

So \[ \log P_{n,8}(\tau, f) \leq H_{\sigma_n}(x^{D_n}) + \int S_n f \, d\sigma_n + 1 \]

We need to wrestle the \( M_n \) on the right hand side. We will get clever.

Let \( m < \frac{n}{2}, \quad \bar{k} \in D_m \). We tile \( D_m \) by copies of \( D_m \) starting at \( \bar{k} \).

\[ q_f = \left( \frac{\bar{k} - \bar{k}}{m} \right) \] number of copies of \( m \) on each axis.
So we have

\[ D_n = \{ k + \bar{m} + j | \bar{c} \in D_{\bar{q}}, j \in D_m \} \]

starting point where within tile outside hilir.

\( F \) is a residual set with

\[ |F| \leq |D_n| - |D_{n-2m}| = n^d - (n-2m)^d = f_{n,m} \]

So we can write \( d_{F,n} \) as, for each \( m < \frac{n}{2}, \bar{c} \in D_{\bar{q}} \)

\[ d_{F,n} = \bigvee_{\bar{c} \in D_{\bar{q}}, j \in D_m} T^{-((\bar{c} + \bar{m} + j))} \land \bigvee_{\bar{c} \in F} T^{-\bar{c}} \]

\[ = \bigvee_{\bar{c} \in D_{\bar{q}}} T^{-(\bar{c} + \bar{m})} \bigvee_{j \in D_m} T^{-j} \land \bigvee_{\bar{c} \in F} T^{-\bar{c}} \]

So now it's a matter of putting the pieces together.

\[ \log P_{\text{lev}}(11) \leq H_{\sigma_n}(d_{F,n}) + \int s \text{f} f_{\sigma_n} + 1 \]

\[ = \sum_{\bar{c} \in D_{\bar{q}}} H_{\sigma_n}(T^{-(\bar{c} + \bar{m})} d_{F,n}) + \sum_{\bar{c} \in F} H_{\sigma_n}(T^{-\bar{c}} d_{F,n}) + \int s \text{f} f_{\sigma_n} + 1 \]

\leq \sum_{\bar{c} \in D_{\bar{q}}} H_{\sigma_n}(T^{-(\bar{c} + \bar{m})} d_{F,n}) + f_{n,m} \log|\alpha| + \int s \text{f} f_{\sigma_n} + 1 \]

We now sum over all \( j \in D_m \), and dividing by \( n^d \).
\[
\frac{m^d \log P_n, s(T, f)}{n^d} \leq \frac{1}{n^d} \sum_{z \in \mathbb{D}_n} H_{\mathbb{D}_n, z} \left(\varphi_{m^n, \alpha} \right) + \frac{m^d}{n^d} \left( f_{m^n, \log} \log 1 + 1 \right) + \frac{m^d}{n^d} \int f \, d\mu
\]

\[
= H_{\mu} (\varphi_{m^n, \alpha}) + \frac{m^d}{n^d} \left( f_{m^n, \log} \log 1 + 1 \right) + m^d \int f \, d\mu
\]

Since \( \rho H_{\mu} + (1 - \rho) H_{\nu} \leq H_{\mu + \rho \nu - \rho \mu} \), \( \rho \in (0, 1) \), \( \mu, \nu \in M(X) \).

Now, by taking limits along \( n_j \) we get

\[
H_{\mu_j} (\varphi_{m^n, \alpha}) \to H_{\mu} (\varphi_{m^n, \alpha}), \quad \text{since } \alpha \text{ is } \mu \text{-good}
\]

\[
\int f \, d\mu_j \to \int f \, d\mu, \quad \text{since } f \in C_b (X)
\]

and \( \mu_j \to \mu \).

Also

\[
\lim_{n \to \infty} \frac{f_{m^n, \log}}{n^d}, \quad \lim_{n \to \infty} \frac{n^d - (n - 2n)^d}{n^d} = 0
\]

So we have

\[
m^d \frac{P_{m^n} (T, f)}{n^d} \leq H_{\mu} (\varphi_{m^n, \alpha}) + m^d \int f \, d\mu
\]

Dividing by \( m^d \) and taking limits as \( m \to \infty \) we get

\[
P_{m^n} (T, f) \leq h_{\mu} (T, \alpha) + \int f \, d\mu
\]

\[
\leq h_{\mu} (T) + \int f \, d\mu.
\]