Saifuddin Syed

Math 601D: Variation Principle.

Before we begin let's fix notation.

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

$$D_n = \{ \bar{\imath} \in \mathbb{Z}_+^d \mid \bar{\imath} < (n, n, \dots, n) \}$$

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

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

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

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

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

$M(X, \tau)$ denote the subspace of $M(X)$ of τ-invariant measures.

If $\alpha = (A_1, \dots, A_k)$ is a partition of X, $G \subseteq \mathbb{Z}_+^d$, $\mu \in M(X, \tau)$

$$\alpha^G = \bigvee_{\bar{g} \in G} T^{-\bar{g}} \alpha$$

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

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

$$h_\mu(\tau) = \sup_\alpha h_\mu(\tau, \alpha)$$

Let $\delta > 0$, $E \subseteq X$, is called (n,δ) - separated if

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

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

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

Given $f \in C(X)$, $\delta > 0$, define:

$$S_n f(x) := \sum_{\bar{i} \in D_n} f(T^{\bar{i}}(x))$$

$$P_{n,\delta}(T,f) = \sup \left\{ \sum_{x \in E} e^{S_n f(x)} : E \text{ is } (n,\delta)\text{-separated} \right\}$$

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

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

<u>Lemma:</u> Let $a_1, \ldots, a_k \in \mathbb{R}$. If $p_i \geq 0$, $\sum_{i=1}^{k} p_i = 1$, then

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

with equality iff

$$p_i = \frac{e^{a_i}}{\sum_{j=1}^{k} e^{a_j}}$$

Theorem: (Variational Principle)

$$P(T,f) = \sup\{ h_\mu(T) + \int f\,d\mu \mid \mu \in M(X,T)\}$$

We will prove this 2 steps.

① $P(T,f) \geq \sup\{ h_\mu(T) + \int f\,d\mu \mid \mu \in M(X,T)\}$

② $P(T,f) \leq \sup\{ h_\mu(T) + \int f\,d\mu \mid \mu \in M(X,T)\}$

Proof of ①:

Let $\mu \in M(X,T)$ and $\alpha = \{A_1, \dots, A_k\}$ be a partition.
Let $a>0$, $a \geq \varepsilon > 0$ such that

$$\varepsilon |\alpha| \log |\alpha| < a$$

Since μ is regular, we have $\exists$ compact $B_j \subseteq A_j$ st

$$\mu(A_j \setminus B_j) < \varepsilon$$

We define a new partition β by

$$\beta = \{B_0, B_1, \dots, B_k\}, \quad B_0 = X \setminus \bigcup_{i=1}^{k} B_i$$

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 $\exists\, \delta > 0$, $\delta < \frac{b}{2}$ st

$\forall x, y \in X$, $d(x,y) < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.

Let $n \in \mathbb{Z}_+$ and let E be (n,δ) separated wrt σ. ie.

$$\forall x \neq y \in E, \ \exists \bar{\imath} \in D_n \ d(T^{\bar{\imath}}(x), T^{\bar{\imath}}(y)) \geq \delta$$

Wlog suppose E fails to be (n,δ)-separated if any point is added.

So E is (n,δ)-spanning, ie.

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

For $C \in \beta^{D_n}$, define

$$\alpha_C := \sup \{ (S_n f)(x) \mid x \in C \}$$

Now let's estimate!

$$H_\mu(\beta^{D_n}) + \int S_n f \, d\mu$$

$$= \sum_{C \in \beta^{D_n}} -\varphi(\mu(C)) + \int_C S_n f \, d\mu$$

$$\leq \sum_{C \in \beta^{D_n}} -\varphi(\mu(C)) + \alpha_C \mu(C)$$

$$= \sum_{C \in \beta^{D_n}} \mu(C) (\alpha_C - \log \mu(C))$$

$$\leq \log \Big(\sum_{C \in \beta^{D_n}} e^{\alpha_C} \Big) \quad \text{, by lemma}$$

By extreme value theorem, α_C is obtained for some $x \in \bar{C}$. Since E is (n,δ)-spanning, $\exists y_C \in E$ such that

$$d(T^{\bar{\imath}}(x), T^{\bar{\imath}}(y_C)) < \delta, \ \forall \bar{\imath} \in D_n$$

δ was chosen to utilize the uniform continuity of f. So lets use it!

$$
\begin{aligned}
& |\alpha_c - S_n f(y_c)| \\
&= |S_n f(x) - S_n f(y_c)| \\
&\le \sum_{\bar{i}\in D_n} |f(T^{\bar{i}}(x)) - f(T^{\bar{i}}(y_c))| && \text{, triangle inequality} \\
&\le \sum_{\bar{i}\in D_n} \varepsilon && \text{, since } d(T^{\bar{i}}(x), T^{\bar{i}}(y_c)) < \delta \\
& && E \text{ is } (n,\delta)\text{-spanning.} \\
&= n^d \varepsilon
\end{aligned}
$$

So we have

$$\alpha_c \le S_n f(y_c) + n^d \varepsilon$$

Since $\frac{\delta}{2} < b$, we have each ball of radius δ meets the closures of at most 2 members of β. 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

$$C = \bigcap_{\bar{i}\in D_n} T^{\bar{i}}(B_{j_{\bar{i}}}) \quad , \text{ for some } j_{\bar{i}} \in \{0, \dots, k\}.$$

For each C, s.t. $y_c = y$, there are atmost 2 choices of $j_{\bar{i}}$. thus

$$|A_y| \le 2^{|D_n|} = 2^{n^d}$$

Now we have

$$
\begin{aligned}
& H_\mu(\beta^{D_n}) + \int S_n f \, d\mu \\
& \leq \log \sum_{C \in \beta^{D_n}} e^{\alpha_C} \\
& \leq \log \sum_{C \in \beta^{D_n}} e^{S_n f(y_C) + n^d \varepsilon} \\
& = \log \left(e^{n^d \varepsilon} \sum_{y \in E} |A_y| e^{S_n f(y)} \right) \\
& \leq \log \left(e^{n^d \varepsilon} 2^{n^d} \sum_{y \in E} e^{S_n f(y)} \right) \\
& = n^d \varepsilon + n^d \log 2 + \log \left(\sum_{y \in E} e^{S_n f(y)} \right) \\
& \leq n^d a + n^d \log 2 + \log (P_{n,\delta}(T, f)) \qquad \text{since } \varepsilon < a
\end{aligned}
$$

Dividing both sides by n^d.

$$
\frac{1}{n^d} H_\mu(\beta^{D_n}) + \underset{\substack{\uparrow \\ \text{Since } \mu \in M(X,T)}}{\int f \, d\mu} \leq a + \log 2 + \frac{1}{n^d} \log (P_{n,\delta}(T, f))
$$

Letting $n \to \infty$, we get

$$
\begin{aligned}
h_\mu(T, \beta) + \int f \, d\mu & \leq a + \log 2 + P_\delta(T, f) \\
& \leq a + \log 2 + P(T, f).
\end{aligned}
$$

We need to turn that β into an α. Recall:

$$
h_\mu(T, \alpha) \leq h_\mu(T, \beta) + H_\mu(\alpha | \beta)
$$

So we have

$$
h_\mu(T, \alpha) + \int f \, d\mu \leq a + \log 2 + H(\alpha | \beta) + P(T, f).
$$

· Intuitively the elements in β are very close approximations of α, so we should have $H(\alpha|\beta)$ be very small.

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

$$= -\mu(\beta_0) \sum_{j=1}^{|\alpha|} \varphi(\mu(A_j|B_0)) \quad \text{since for } 1 \le i, j \le k \;\; \mu(A_j|B_i) \in \{0,1\}$$

$$\le \mu(B_0) \log |\alpha|$$

$$\le \varepsilon |\alpha| \log |\alpha|, \quad \text{since } B_0 = \bigcup_{i=1}^{k} (A_i \setminus B_i), \; \mu(A_i \setminus B_i) < \varepsilon$$

$$< a$$

Thus $\quad h(T) + \int f \, d\mu \le 2a + \log 2 + P(T,f)$.

Since
$$h_\mu(T^{\bar{n}}) = n^d h_\mu(T), \quad \bar{n} = (n, \dots, n)$$
$$\int S_n f \, d\mu = n^d \int f \, d\mu, \quad \mu \in \mathcal{M}(X,T)$$
$$P(T^{\bar{n}}, S_n f) = n^d P(T,f)$$

We get by replacing T with $T^{\bar{n}}$ and f with S_n that, $\forall n \in \mathbb{N}$

$$h(T) + \int f \, d\mu \le \frac{2a + \log 2}{n^d} + P(T,f)$$

$$\Rightarrow \quad h(T) + \int f \, d\mu \le P(T,f)$$

Proof of ②:

Let $\delta > 0$, we will find a $\mu \in M(X, T)$ such that

$$h_\mu(T) + \int f \, d\mu \geq P_\delta(T, f)$$

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

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

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

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

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

$$\mu_n = \frac{1}{n^d} \sum_{\bar{i} \in D_n} \sigma_n \circ T^{-\bar{i}}$$

Now by Reisz-Representation theorem the space of Borel-regular measures is isometrically isomorphic to $C(X)^*$, so we can impose the weak* topology on $M(X)$. Since the space of probability measures are closed subset of the unit ball, with is compact by Banach-Alaoglu, we have $M(X)$ is compact.

So there is a subsequence n_j such that

$$\mu_{n_j} \to \mu \quad , \quad j \to \infty$$

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

I claim that $\mu \in M(X, \tau)$. Let's show this. Fix $\bar{k} \in \mathbb{Z}_+^d$.

$$\left| \int f \circ T^{\bar{k}} d\mu - \int f d\mu \right|$$

$$= \lim_{j \to \infty} \left| \int f \circ T^{\bar{k}} d\mu_j - \int f \mu_{n_j} \right|$$

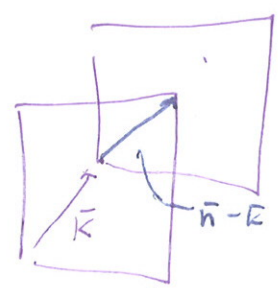
$$= \lim_{j \to \infty} \left| \frac{1}{n^d} \int \left(\sum_{\bar{i} \in D_{n_j} + \bar{k}} f \circ T^{\bar{i}} - \sum_{\bar{i} \in D_{n_j}} f \circ T^{\bar{i}} \right) d\sigma_n \right|$$

$$= \lim_{j \to \infty} \left| \frac{1}{n^d} \int \sum_{\bar{i} \in D_{n_j} + \bar{k} \, \Delta \, D_{n_j}} f \circ T^{\bar{i}} \, d\sigma_n \right|$$

$$\leq \lim_{j \to \infty} \frac{1}{n^d} \left| D_{n_j} + \bar{k} \, \Delta \, D_{n_j} \right| \, \|f\|_\infty$$

$$= \lim_{j \to \infty} \frac{1}{n^d} \left(n^d - \mathrm{vol}(\bar{n} - \bar{k}) \right) \|f\|_\infty$$

$$= 0$$



$\bar{n} = (n, \ldots, n)$

Thus $\mu \in M(X, \tau)$.

<u>Lemma</u>: $\exists$ a δ-fine, μ-good partition $\alpha = \{A_1, \dots, A_k\}$.

ie $\quad \text{diam}(A_i) < \delta$ & $\mu(\partial A_i) = 0$.

<u>pf</u>: First let's show wlog we can find a δ-fine, μ-good ball.

Let $r < \frac{\delta}{2}$ $x \in X$,

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

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

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

By compactness we can cover X by ball $B_1, \dots, B_k$ st

$$\text{diam}(B_i) < \delta$$

Let $A_1 = \bar{B}_1$, $A_n = \bar{B}_n \setminus \bigcup_{i=1}^{n-1} \bar{B}_i$. The result follows since

$$\partial A_n \subseteq \bigcup_{i=1}^{n} \partial B_i$$

$\blacksquare$

So let $\alpha = \{A_1, \dots, A_k\}$ be a μ-good, δ-fine partition.
Since α is δ-fine we have each element of α^{D_n} has at most one element from E_n.

Now lets do some algebra!

$$H_{\sigma_n}(\alpha^{D_n}) + \int S_n f \, d\sigma_n$$

$$= \sum_{y \in E_n} -\sigma_n(\{y\}) \log \sigma_n(\{y\}) + S_n f(y) \sigma_n(\{y\})$$

$$= \sum_{y \in E_n} \sigma_n(\{y\}) \left(S_n f(y) - \log \sigma_n(\{y\}) \right)$$

$$= \log \Big(\sum_{y \in E_n} S_n f(y) \Big)$$, by lemma since $\sigma_n(\{y\}) = \dfrac{e^{S_n f(y)}}{\sum_{z \in E_n} e^{S_n f(z)}}$

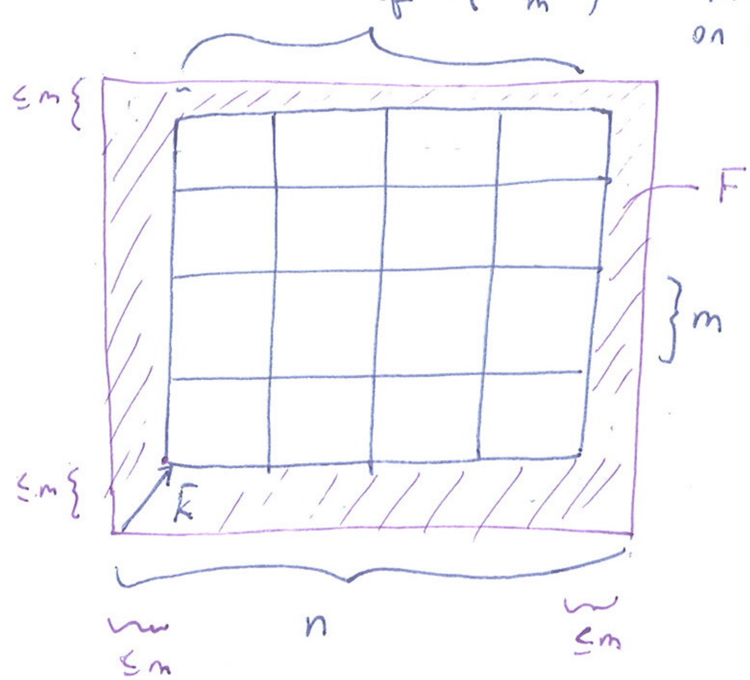
$$\geq \log P_{n,\delta}(\tau, f) - 1$$

So $\log P_{n,\delta}(\tau, f) \leq H_{\sigma_n}(\alpha^{D_n}) + \int S_n f \, d\sigma_n + 1$

We need to wrestle the μ_n on the right hand side. We will get clever.

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

$\bar{q} = \left(\frac{\bar{n} - k}{m} \right)$ = number of copies of m. on each axis.



So we have

$$D_n = \{ \underset{\substack{\uparrow \\ \text{starting point}}}{k} + \overset{\substack{\text{which tile} \\ \downarrow}}{\bar{i}m} + \underset{\substack{\uparrow \\ \text{where within tile.}}}{\bar{j}} \mid \bar{i} \in D_{\bar{q}},\ \bar{j} \in D_m \} \cup \underset{\substack{\uparrow \\ \text{outside tiling.}}}{F}$$

F is a residual set with

$$|F| \le |D_n| - |D_{n-2m}|$$
$$= n^d - (n-2m)^d =: f_{n,m}$$

So we can write α^{D_m} as, for each $m < \frac{n}{2}$, $\bar{k} \in D_m$

$$\alpha^{D_m} = \bigvee_{\bar{i} \in D_{\bar{q}},\, \bar{j} \in D_m} T^{-(\bar{k} + \bar{i}m + \bar{j})} \alpha \vee \bigvee_{\bar{i} \in F} T^{-\bar{i}} \alpha$$

$$= \bigvee_{\bar{i} \in D_{\bar{q}}} T^{-(\bar{k} + \bar{i}m)} \underbrace{\bigvee_{\bar{j} \in D_m} T^{-\bar{j}} \alpha}_{\alpha^{D_m}} \vee \underbrace{\bigvee_{\bar{i} \in F} T^{-\bar{i}} \alpha}_{\alpha^F}$$

So now its a matter of putting the peices together.

$$\log P_{n,\delta}(T, f) \le H_{\sigma_n}(\alpha^{D_n}) + \int S_n f \, d\sigma_n + 1$$

$$\le \sum_{\bar{i} \in D_{\bar{q}}} H_{\sigma_n}(T^{-(\bar{k} + \bar{i}m)} \alpha^{D_m}) + \sum_{\bar{i} \in F} H_{\sigma_n}(T^{-\bar{i}} \alpha) + \int S_n f \, d\sigma_n + 1$$

$$\le \sum_{\bar{i} \in D_{\bar{q}}} H_{\sigma_n \circ T^{-(\bar{k} + \bar{i}m)}}(\alpha^{D_m}) + f_{n,m} \log|\alpha| + \int S_n f \, d\sigma_n + 1$$

We now sum over all $\bar{j} \in D_m$, and dividing by n^d.

$$\frac{m^d}{n^d}\log P_{n,\delta}(\tau,f) \le \frac{1}{n^d}\sum_{\tau\in D_n} H_{\sigma\circ T^{-\tau}}(\alpha^{D_m}) + \frac{m^d}{n^d}(f_{n,m}\log|\alpha|+1) + \frac{m^d}{n^d}\int S_n f\, d\sigma_n$$

$$= H_{\mu_n}(\alpha^{D_m}) + \frac{m^d}{n^d}(f_{n,m}\log|\alpha|+1) + m^d\int f\, d\mu_n$$

Since $\rho H_{\mu_1} + (1-\rho)H_{\mu_2} \le H_{\rho\mu_1+(1-\rho)\mu_2}$, $\rho\in(0,1)$, $\mu_1,\mu_2\in M(X)$.

Now by taking limits along n_j get

$$H_{\mu_{n_j}}(\alpha^{D_m}) \xrightarrow[j\to\infty]{} H_\mu(\alpha^{D_m}) \quad , \text{ since } \alpha \text{ is } \mu\text{-good}$$

$$\int f\, d\mu_{n_j} \to \int f\, d\mu \quad , \text{ since } f\in C_0(X).$$

and $\mu_n\to\mu$

Also $\lim_{n\to\infty}\frac{f_{n,m}}{n^d}$ $\lim_{n\to\infty}\frac{n^d-(n-2m)^d}{n^d} = 0$

So we have

$$m^d P_\delta(\tau,f) \le H_\mu(\alpha^{D_m}) + m^d\int f\, d\mu$$

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

$$P_\delta(\tau,f) \le h_\mu(\tau,\alpha) + \int f\, d\mu$$

$$\le h_\mu(\tau) + \int f\, d\mu.$$