Lecture 18, Friday, May 6, 2002

Today's and next lecture:

(1) Lifting properties of X to $(\mathcal{P}_2(X), W_2)$: Polish, compact, but not "proper".

(2) Characterization of convergence in $(\mathcal{P}_2(X), W_2)$.

(3) Some examples.

Recall:

<u>Theorem</u> If X is a Polish space, then $(\mathcal{P}_2(X), W_2)$ is a metric space.

The β-metric of $\mathcal{P}(X)$:

$$\beta(\mu, \nu) = \sup\left\{ \left| \int_X f\,d\mu - \int_X f\,d\nu \right| : f \in BL(X),\ \|f\|_{BL} \le 1 \right\}$$

⊙ If X is separable, then $\mu_k \to \mu$ narrowly $\Longleftrightarrow$ $\int_X f\,d\mu_k \to \int_X f\,d\mu$ $\forall f \in BL(X)$ $\Longleftrightarrow$ $\beta(\mu_k, \mu) \to 0$.

⊙ If X is a Polish space, then $(\mathcal{P}(X), \beta)$ is complete.

<u>Theorem</u> Let X be a Polish space. Then $(\mathcal{P}_2(X), W_2)$ is a complete metric space.

<u>Proof</u> Let $\{\mu_n\}$ be a Cauchy sequence in $(\mathcal{P}_2(X), W_2)$. Since $\beta(\mu, \nu) \le W_2(\mu, \nu)$ $\forall \mu, \nu \in \mathcal{P}_2(X)$, $\{\mu_n\}$ is also a Cauchy sequence in $(\mathcal{P}(X), \beta)$, a complete metric space. Hence, $\beta(\mu_n, \mu) \to 0$ for some $\mu \in \mathcal{P}(X)$. We show that $\mu \in \mathcal{P}_2(X)$ and $W_2(\mu_n, \mu) \to 0$.

Let $\varepsilon > 0$. Let $N \ge 1$ be such that $W_2(\mu_m, \mu_n) \le \varepsilon$ for $m, n \ge N$. We claim for a given $\bar{x} \in X$ and $n \ge N$ that

$$\int_X d^2(x, \bar{x})\,d\mu(x) \le 2\varepsilon^2 + 2\int_X d^2(y, \bar{x})\,d\mu_n(y) < \infty, \qquad (*)$$

which would imply $\mu \in \mathcal{P}_2(X)$, and

$$W_2(\mu, \mu_n) \le \varepsilon, \tag{**}$$
which would imply $W_2(\mu_n, \mu) \to 0$.

Let $\gamma_{m,n} \in \mathcal{A}(\mu_m, \nu_n)$ be such that
$$W_2^2(\mu_m, \mu_n) = \int_{X\times X} d^2(x,y)\, d\gamma_{m,n}(x,y).$$
Since $\beta(\mu_m, \mu) \to 0$, $\mu_m \to \mu$ narrowly, and $\{\mu_m\}$ is precompact narrowly. Hence, by Prokhorov's Theorem, $\{\mu_m\}$ is equi-tight: $\forall \delta > 0$, $\exists$ compact $K \subseteq X$ such that $\mu_m(X \setminus K) < \delta$ for all $m \ge 1$. Thus, for any $n \ge 1$,
$$\pi_{m,n}(X\times X \setminus (K\times K)) \le \pi_{m,n}((X\setminus K)\times X) = \mu_m(X\setminus K) < \delta \ (m=1,2,\cdots).$$
Hence, $\{\pi_{m,n}\}_{m=1}^{\infty}$ is equi-tight. By Prokhorov's Theorem again, $\forall n \ge 1$, $\exists \gamma_n \in \mathcal{P}(X\times X)$ and a subseq. $\{\gamma_{m_k,n}\}$ of $\{\gamma_{m,n}\}$ such that $\gamma_{m_k,n} \to \gamma_n$ narrowly in $\mathcal{P}(X\times X)$. But $\pi_{\#}^{X}$ is continuous w.r.t. the narrow convergence, thus $\gamma_n \in \mathcal{A}(\mu, \mu_n)$.

Let $\bar{x} \in X$. For $m_k, n \ge N$, we have
$$\int_X d^2(x,\bar{x})\, d\mu_{m_k}(x) = \int_{X\times X} d^2(x,\bar{x})\, d\gamma_{m_k,n}(x,y)$$
$$\le 2\int_{X\times X} d^2(x,y)\, d\gamma_{m_k,n}(x,y) + 2\int_{X\times X} d^2(y,\bar{x})\, d\gamma_{m_k,n}(x,y)$$
$$= 2 W_2^2(\mu_{m_k}, \mu_n) + 2\int_{X\times X} d^2(y,\bar{x})\, d\gamma_{m_k,n}(x,y)$$
$$\le 2\varepsilon^2 + 2\int_X d^2(y,\bar{x})\, d\mu_n(y).$$
Sending $k \to \infty$, we get $(*)$. Since $\gamma_{m_k,n} \to \gamma_n$ narrowly,
$$\int_{X\times X} d^2(x,y)\, d\gamma_n(x,y) = \lim_{k\to\infty} \int_{X\times X} d^2(x,y)\, d\gamma_{m_k,n}(x,y)$$
$$= \lim_{k\to\infty} W_2^2(\mu_{m_k}, \mu_n) \le \varepsilon^2 \quad \text{for any } n \ge N.$$

Thus, since $\gamma_n \in \mathcal{A}(\mu, \mu_n)$,

$$W_2^2(\mu, \mu_n) \le \int_{X\times X} d^2(x,y)\, d\gamma_n(x,y) \le \varepsilon^2 \quad \forall n \ge N.$$

This is (**). <u>QED</u>.

<u>Theorem</u> (Characterization of W_2-convergence). Let X be a Polish space, and $\mu_n, \mu \in \mathcal{P}_2(X)$ $(n=1,2,\cdots)$.

(1) If $\mu_n \xrightarrow{W_2} \mu$ then $\mu_n \to \mu$ narrowly and for any $x_0 \in X$

$$\int_X d^2(x, x_0)\, d\mu_n(x) \to \int_X d^2(x, x_0)\, d\mu(x) \qquad (*)$$

(2) If $\mu_n \to \mu$ narrowly and $(*)$ holds true for some $x_0 \in X$, then $\mu_n \xrightarrow{W_2} \mu$.

<u>Remark</u>. The condition $(*)$ is necessary; cf. an example below.

<u>Corollary</u> If X is a compact metric space then $(\mathcal{P}_2(X), W_2) = (\mathcal{P}(X), W_2)$ is a compact metric space.

<u>Proof</u>. Fix $x_0 \in X$. Then $x \longmapsto d(x, x_0)$ is bounded and continuous on X, where d is the metric of X. Thus, $\mu \in \mathcal{P}(X) \Rightarrow \mu \in \mathcal{P}_2(X)$. Hence, $\mathcal{P}_2(X) = \mathcal{P}(X)$.

Since X is compact, $[C_b(X)]^* = [C(X)]^*$, and $\mathcal{M}_f(X) \cong [C(X)]^*$. $\mathcal{P}(X) \subseteq$ closed unit ball of $[C(X)]^*$, which is compact w.r.t. the weak-* topology, same as the narrow topology. Now, for any sequence $\mu_n \in \mathcal{P}_2(X) = \mathcal{P}(X)$ $(n=1,2,\cdots)$, there exists a subseq. $\{\mu_{n_j}\}$ and some $F \in [C(X)]^*$ such $\mu_{n_j} \to F$ weak-*. By Riesz's Theorem, $F \in \mathcal{M}_+(X)$. Moreover, $F(X) = 1$ as all $\mu_{n_j}(X) = 1$. Thus, $F \in \mathcal{P}(X)$. Hence, any $\{\mu_n\}$, a sequence in $(\mathcal{P}_2(X), W_2)$ has a convergent subseq. So, $(\mathcal{P}_2(X), W_2)$ is compact. <u>QED</u>

The following observation is interesting:

<u>Proposition</u> If X is a Polish space but unbounded, then $(\mathcal{P}_2(X), W_2)$ is not locally compact.

<u>Proof</u> We show that for any $\mu \in \mathcal{P}_2(X)$ and any $R>0$, the closed ball $\overline{B(\mu, R)}$ in $(\mathcal{P}_2(X), W_2)$ is not compact.

Pick up some $x_0 \in X$. Let $x_n \in X$ be such that $0 < d(x_n, x_0) \uparrow \infty$. Set for $n \in \mathbb{N}$

$$\varepsilon_n = R^2 \cdot \left[2 \int_X d^2(x, x_0)^2 d\mu(x) + 2 d^2(x_n, x_0)\right]^{-1} \downarrow 0.$$

$$\mu_n = (1-\varepsilon_n)\mu + \varepsilon_n \delta_{x_0} \in \mathcal{P}_2(X),$$

$$\hat{\mu}_n = (1-\varepsilon_n)\mu + \varepsilon_n \delta_{x_n} \in \mathcal{P}_2(X),$$

$$\gamma_n = ((1-\varepsilon_n)(Id \times Id)\#\mu + \varepsilon_n (Id \times x_0)\#\mu,$$

$$\hat{\gamma}_n = (1-\varepsilon_n)(Id \times Id)\#\mu + \varepsilon_n \delta_{x_n} \times \delta_{x_0}.$$

[Recall: $Id \times T: X \to X \times Y$: $x \mapsto (x, T(x))$.
If $T\#\nu_1 = \nu_2$ then $\gamma_T = (Id \times T)\#\nu_1 \in \mathcal{A}(\nu_1, \nu_2)$.]

It can be verified that $\gamma_n \in \mathcal{A}(\mu, \hat{\mu}_n)$ and $\hat{\gamma}_n \in \mathcal{A}(\mu_n, \hat{\mu}_n)$. Thus,

$$W_2^2(\mu_n, \mu) \le 2 W_2^2(\mu_n, \hat{\mu}_n) + 2 W_2^2(\mu, \hat{\mu}_n)$$

$$\le 2 \int_{X \times X} d^2(x, y) d\hat{\gamma}_n(x, y) + 2 \int_{X \times X} d^2(x, y) d\gamma_n(x, y)$$

$$= 2 \varepsilon_n d^2(x_n, x_0) + 2 \varepsilon_n \int_X d^2(x, x_0) d\mu(x)$$

$$= R^2 \quad (n = 1, 2, \cdots).$$

Thus, $\mu_n \in \overline{B(\mu, R)}$ w.r.t. W_2 $(n = 1, 2, \cdots)$. Moreover,

$$\liminf_{n\to\infty} \int_X d^2(x,x_0)\, d\mu_n(x)$$
$$= \liminf_{n\to\infty} \left[(1-\varepsilon_n) \int_X d^2(x,x_0)\, d\mu + \varepsilon_n d^2(x_n, x_0) \right]$$
$$= \int_X d^2(x,x_0)\, d\mu + \frac{R^2}{2}. \qquad (**)$$

Now, if there exists a subseq. of μ_n, $\mu_{n_j} \xrightarrow{W_2} \tilde{\mu} \in \mathcal{P}_2(X)$. Then, $\mu_{n_j} \to \tilde{\mu}$ narrowly. But clearly $\mu_n \to \mu$ narrowly. Thus $\tilde{\mu} = \mu$. Then, $\mu_{n_j} \xrightarrow{W_2} \mu$, which should imply that

$$\int_X d^2(x,x_0)\, d\mu_{n_j}(x) \to \int_X d^2(x,x_0)\, d\mu(x).$$

This contradicts $(**)$. <u>QED</u>

The example below demonstrates two points.
(1) The condition $(*)$ for some x_0 in the main theorem of convergence is necessary. If $(*)$ fails for some $x_0 \in X$, then $\mu_n \not\to \mu$ in W_2, but it is still possible that $\mu_n \to \mu$ narrowly.
(2) If X is a Polish space, and is also proper, then $(\mathcal{P}_2(X), W_2)$ may not be proper. Recall that a metric space is proper, if bounded and closed sets are compact. In a proper space, any bounded sequence has a convergent subsequence.

<u>Example</u> $X = \mathbb{R}$, $\mu_n = (1-\frac{1}{n})\delta_0 + \frac{1}{n^2}\delta_n \in \mathcal{P}(\mathbb{R})$ $(n=1,2,\cdots)$, and $\mu = \delta_0$. Then, $\mu \in \mathcal{P}_2(X)$ as

$$\int_X d^2(x,0)\,d\mu(x) = \int_{\mathbb{R}} |x|^2 \, d\delta_0(x) = 0 .$$

Also, each $\mu_n \in \mathcal{P}_2(X)$, since

$$\int_X d^2(x,0)\,d\mu_n(x) = \int_{\mathbb{R}} |x|^2 \, d\left(\left(1-\tfrac{1}{n}\right)\delta_0 + \tfrac{1}{n^2}\delta_n\right) = \tfrac{1}{n^2} n^2 = 1 .$$

By the condition (*) in the convergence theorem, no subseq. of $\{\mu_n\}$ is W_2-convergent. In particular, $\mu_n \not\to \mu$ in W_2. But, $\mu_n \to \mu$ narrowly, as

$$\int_{\mathbb{R}} f\,d\mu_n = \left(1-\tfrac{1}{n}\right)f(0) + \tfrac{1}{n^2}f(n) \to f(0) = \int_{\mathbb{R}} f\,d\mu \quad \forall f \in C_b(\mathbb{R}).$$

If $\gamma_n \in \mathcal{A}(\mu_n, \mu)$ then $\gamma_n(\mathbb{R}\times\{0\}) = \mu(\{0\}) = 1$, and hence,

$$\int_{\mathbb{R}\times\mathbb{R}} |x-y|^2 \, d\gamma(x,y) = \int_{\mathbb{R}\times\mathbb{R}} x^2 \, d\gamma(x,y) = \int_{\mathbb{R}} x^2 \, d\mu_n(x) = \frac{n^2}{n^2} = 1.$$

Hence, $W_2(\mu_n, \mu) = 1$ $(n = 1, 2, \dots)$, and $\{\mu_n\}$ is W_2-bounded.