3. Properties of measures on Polish spaces

We start with this observation.

LEMMA 3.1. *The Borel sets $\mathcal{B}(X)$ of (X, d'') are the same of (X, d), where d'' is the distance defined in Proposition 2.10.*

PROOF. From the point of view of the topology, it is easy to see that we are just adding the set $X \setminus B_n$ to the family of open sets, and these sets are already in the Borel family. □

COROLLARY 3.2. *The map $\phi : X \to (0, 1)$ is an isomorphism between the set $\mathcal{P}(X)$ of probability measures on X and the set $\mathcal{P}(\phi(X))$.*

As a consequence, we obtains the following.

PROPOSITION 3.3. *A probability measure μ on X is concentrated on a σ-compact set, i.e. a countable union of compact sets.*

PROOF. The measure $\phi_\sharp \mu \in \mathcal{P}(\phi(X))$ is a probability measure on $(0, 1)$, and by inner regularity there is a family of compact sets $\{K_n\}_n \in \mathbb{P}(\phi(X))$ such that

$$1 = \phi_\sharp \mu \left(\bigcup_n K_n \right) = \mu \left(\bigcup_n \phi^{-1}(K_n) \right).$$

Being ϕ^{-1} continuous, $\phi^{-1}(K_n)$ is compact. □

DEFINITION 3.4. A measure $\mu \in \mathcal{P}(X)$ is *tight* if it is concentrated on a σ-compact subset of X.
A family of measures $\{\mu_\alpha\}_\alpha \subset \mathcal{P}(X)$ is *tight* if

$$\forall \epsilon > 0 \ \exists K \text{ compact } \big(\sup_\alpha \mu_\alpha(X \setminus K) < \epsilon \big).$$

In particular Proposition 3.3 yields that every probability measure is tight. The tightness of a family of measure is equivalent to say that their mass cannot "escape to ∞", in a similar reasoning as the Lebesgue/Vitali Theorems for passing to the limit under the sing of integral. In Polish space the noncompactness appears also in bounded sets.

3.1. Weak convergence of measures and Prokhorov's Theorem. On $\phi(X) \subset (0, 1)$ the weak convergence of measure is the standard convergence by means of duality, i.e. tested with a continuous function ψ. By using the pullback map, $\psi \circ \phi$ is continuous and bounded in the d''-topology.

DEFINITION 3.5 (Weak/narrow convergence). We say that $\mathcal{P}(X) \ni \mu_n \rightharpoonup \mu \in \mathcal{P}(X)$ *in the weak/narrow sense* if for all continuous bounded function ψ on (X, d) it holds

$$\int \psi \mu_n \to \int \psi \mu.$$

REMARK 3.6. Note that as test functions we use the original topology: this implies that there are sequences of measures μ_n weakly convergence but such that $\phi_\sharp \mu_n$ is not weakly converging in $\mathbb{R}$. As an example, take the Dirac-delta measures on sequence converging for d but not for d''. For sequences this is not a problem: indeed we can chose as a base for the topology countably many balls B_n such that

$$\mu_k(\partial B_n), \mu(\partial B_n) = 0,$$

and in this case the weak convergence for (X, d) is equivalent to the weak convergence for (X, d'').

LEMMA 3.7. *The spaces $C_b(X)$ with the uniform topology is Polish.*

PROOF. Exercise. □

We first study Prokhorov's Theorem for sequences of measures on a G_δ-subset G of $[0, 1]$.

LEMMA 3.8. *Assume that $\nu_n \rightharpoonup \nu$ in $\mathcal{P}([0, 1])$, with $\nu_n \in \mathcal{P}(G)$. Then $\nu \in \mathcal{P}(G)$ iff the family $\{\nu_n\}_n$ is tight.*

PROOF. If the sequence is tight and converging, then its weak limit is in G: indeed for all ϵ by the u.s.c. of measures on closed sets one deduce that if $\nu_n(K) > 1 - \epsilon$ then also $\nu(K) > 1 - \epsilon$. Thus $\nu \in \mathcal{P}(G)$.

Viceversa, we use the property that if $\nu_n \rightharpoonup \nu$ then for all $\epsilon > 0$ one considers a compact set $K \subset G$ such that
$$\nu(K) > 1 - \frac{\epsilon}{3}.$$
Then by the l.s.c. of the weak convergence on open sets, for all δ there exists N such that
$$\nu_n(G \cap (K + B_\delta(0))) > 1 - \frac{2\epsilon}{3} \quad \text{for all } n > N.$$
Thus there are compact sets $K_n \subset K + B_{2^{-n}}(0)$ such that $\nu_n(K_n) > 1 - \epsilon$. These compact sets converges to K in Hausdorff distance, so that
$$K \cup \bigcup_n K_n \quad \text{is compact.}$$
This is the tightness. $\qquad \square$

THEOREM 3.9 (Prokhorov). *A family of probability measure $\{\mu_\alpha\}_\alpha \subset \mathcal{P}(X)$ is sequentially precompact iff it is tight.*

Recall that sequentially precompact means that for every sequence $\mu_n = \mu_{\alpha_n}$ there is a converging subsequence.

PROOF. If the family is tight, then by a diagonal process we can find a sequence μ_n of measures and an increasing sequence K_n of compact sets such that
$$\forall n \geq m \big(\mu_n(K_m) > 1 - 2^{-m} \wedge \mu_n \llcorner K_m \rightharpoonup \mu_m \in \mathcal{M}^+(K_m)\big).$$
Clearly by u.s.c. it holds for $m < m'$
$$\mu_{m'} \llcorner K_m \geq \mu_m, \quad \|\mu_{m'} - \mu_m\| < 2^{-n}.$$
Hence the measure defined as
$$\mu = \sup_m \mu_m$$
satisfies $\mu(X) = 1$ and for $\psi \in C_b(X)$
$$\limsup_n \left| \int \psi \mu - \int \psi \mu_n \right| \leq \limsup_n \left| \int_{K_m} \psi \mu - \int_{K_m} \psi \mu_n \right| + 2^{-m+1} \|\psi\|_\infty$$
$$= \left| \int_{K_m} \psi(\mu - \mu_m) \right| + 2^{-m+1} \|\psi\|_\infty \leq 3 \cdot 2^{-m} \|\psi\|_\infty.$$

Assume that the family is not tight: we want to prove that there exists $\epsilon, \delta > 0$ such that for all compact set $K \subset X$ there is μ_α such that
$$\mu_\alpha(X \setminus K_\delta) > \epsilon, \quad \text{where } K_\delta = \big\{x \in X : \text{dist}(x, K) < \delta\big\}.$$
Assume conversely that for all $n \in \mathbb{N}$ there is K_n compact such that $\mu_\alpha(K_{n,2^{-n}}) > 1 - \epsilon 2^{-n}$ for all α. Define the family of compact sets
$$\tilde{K}_n = K_n \cap \bigcap_{m < n} \text{clos}(K_{m,2^{-m}}).$$
This family is compact in the Hausdorff topology, because (exercise)
$$\text{dist}_{\text{Haus}}(\tilde{K}_n, \tilde{K}_{n'}) \leq 2^{-n}, \quad n < n'.$$
Hence it converges to a limit compact set $\tilde{K}$. Moreover it holds
$$\mu_\alpha(\tilde{K}_{n,2^{-n}}) \geq \mu_\alpha(K_{n,2^{-n}}) - \sum_{m < n} \epsilon 2^{-m} > 1 - 2^{-m+1}\epsilon > 1 - \epsilon.$$
By the monotone convergence we thus conclude that
$$\mu_\alpha(\tilde{K}) = \lim_n \mu_\alpha(\tilde{K}_{n,2^{-n}}) > 1 - \epsilon,$$
i.e. $\{\mu_\alpha\}_\alpha$ is tight.

Let $\bar{\epsilon}, \bar{\delta}$ be the constants such that for every compact sets K there exists μ_α such that $\mu_\alpha(K_{\bar{\delta}}) \leq 1 - \bar{\epsilon}$. Starting with μ_1 and K_1 such that $\mu_1(K_1) > 1 - \frac{\bar{\epsilon}}{2}$ (because μ_1 is tight, Proposition 3.3), we can find a probability μ_2 such that $\mu_2(K_{1,\bar{\delta}}) < 1 - \bar{\epsilon}$. Hence there is a compact set $K_2 \subset X \setminus K_{1,\bar{\delta}}$ such that $\mu_2(K_2) > \frac{\bar{\epsilon}}{2}$. Proceeding in this way, we obtain a sequence of probabilities $\mu_n \in \{\mu_\alpha\}_\alpha$ and compact sets K_n such that

$$\operatorname{dist}(K_n, K_m) > \bar{\delta}, \quad \mu_n(K_n) > \frac{\bar{\epsilon}}{2}.$$

Assume that there is a subsequence not relabeled weakly converging to μ. Then for the Lipschitz function

$$\psi_m(x) = \min\left\{1, \frac{\operatorname{dist}(x, \bigcup_{n>m} K_n)}{\delta}\right\} \tag{3.1}$$

one has

$$\int \psi_m \mu = \lim_n \int \psi_m \mu_n < 1 - \frac{\bar{\epsilon}}{2}.$$

Hence since pointwise

$$\psi_k \nearrow 1$$

one concludes that with the monotone convergence theorem that

$$\mu(X) = \lim_k \int \psi_k \mu < 1 - \frac{\bar{\epsilon}}{2},$$

which is a contradiction being μ a probability. $\square$

REMARK 3.10. For a given sequence $\mu_n \in \mathcal{P}(X)$ weakly converging to μ, by Remark 3.6 we can assume that μ_n weakly converges also w.r.t. the continuous functions in the stronger metric d''. In this case Lemma 3.8 gives the tightness of the sequence μ_n in the stronger topology, and then also in the original one.

COROLLARY 3.11. *The space $\mathcal{P}(X)$ with the weak topology is Polish spaces.*

PROOF. Consider the metric

$$\tilde{d}(\mu, \mu') = \sum_n 10^{-n} \sup\left\{\int \psi(\mu - \mu'), 0 \leq \psi \leq 1, \psi \ n\text{-Lipschitz}\right\}.$$

It is fairly easy to see that this is a metric and also separable, by taking the convex combinations of Dirac deltas centered on a dense subsequence of X (exercise).

If μ_n is not tight, then in the proof of Prokhorov Theorem we have constructed a family of $\frac{1}{\delta}$-Lipschitz functions ψ_m given by (3.1) such that

$$\sup_k\left\{\int \psi_k(\mu - \mu')\right\} > \frac{\bar{\epsilon}}{2},$$

thus it is not a Cauchy sequence for $\tilde{d}$. If there is ψ continuous such that

$$n \mapsto \int \psi \mu_n$$

is not converging, then there is a Lipschitz function $\tilde{\psi}$ such that $\int \tilde{\psi} \mu_n$ is not Cauchy, thus concluding that it is not converging in the distance $\tilde{d}$.

Conversely, if μ_n is a weakly convergent sequence, then it is immediate to see that it is Cauchy for $\tilde{d}$. $\square$

3.2. Exercises.

(1) Let $(X, \mathcal{B})$ be a measurable space, and assume that $\mathcal{B}$ is countably generated by the family of sets $\{B_n\}_n$ which is separating: for all $x \neq y$ there is $B \in \mathcal{B}$ such that $x \in B$, $y \in X \setminus B$. Define a separable metric d such that the Borel set of the topology are $\mathcal{B}$. (This exercise is just to show that one can play around with a countable family of sets in order to build suitable metric/Polish/Borel structures.)

(2) Construct a countable dense family of Lipschitz function in $C(X)$.

(3) Prove the finitely atomic measures are dense in $\mathcal{P}(X)$ w.r.t. the weak convergence.

(4) Prove that an equivalent definition for the distance on $\mathcal{P}(X)$ is
$$\tilde{d}(\mu, \mu') = \inf \left\{ \epsilon : \forall K \text{compact} \big(\mu(K) \leq \mu'(K_\epsilon) \ \wedge \ \mu(K) \leq \mu'(K_\epsilon) \big) \right\}.$$

(5) Prove that $\mathcal{P}(X)$ as a subset of the dual space of $C_b(X)$ is a G_δ-set.