Lifting measures

Let $\pi : X \to Y$ be a continuous map between compact metric spaces. If μ is a measure on X then $\pi\mu$ is the measure on Y satisfying $\pi\mu(E) = \mu(\pi^{-1}(E))$ for measurable $E \subseteq Y$ (this definition works also when X, Y are measurable spaces and π is measurable). Equivalently,

$$\forall g \in C(Y) \quad \int g \, d\pi\mu = \int g \circ \pi \, d\mu$$

(in the measurable case one requires this for measurable bounded functions, say). The measure $\pi\mu$ is called the push-forward of μ and is sometimes denotes $\pi_*\mu$ or $\pi_\#\mu$.

Proposition 1. *Let ν be a Borel probability measure on Y. Then there exists a Borel probability measure μ on X such that $\pi\mu = \nu$, i.e. $\mu(\pi^{-1}E) = \nu(E)$ for all Borel sets $E \subseteq Y$.*

Remark 2. μ need not be unique if π is not 1-1.

Remark 3. One can replace compactness by completeness, but then the theorem becomes much more technical (requires descriptive set theory).

Proof No. 1 (almost elementary). Start by constructing a sequence ν_n of atomic measures on Y with $\nu_n \to \nu$ weakly, i.e. $\int g \, d\nu_n \to \int g \, d\nu$ for all $g \in C(Y)$. To get such a sequence, given n choose a finite partition $\mathcal{E}_n$ of Y into measurable sets of diameter $< 1/n$ (for instance cover Y by balls B_i of radius $< 1/n$ and set $E_i = B_i \setminus \bigcup_{j<i} B_j$). For each $E \in \mathcal{E}_n$ choose x_E and set $\nu_n = \sum_{E \in \mathcal{E}_n} \nu(E) \cdot \delta_{x_E}$. One may verify that $\nu_n \to \nu$.

Now, each ν_n can be lifted to a probability measure μ_n on X such that $\pi\mu_n = \nu_n$: to see this, if $\nu_n = \sum w_i \cdot \delta_{y_i}$ choose $x_i \in \pi^{-1}(y_i)$ (there may be many choices, choose one), and set $\mu_n = \sum w_i \cdot \delta_{x_i}$.

Since the space of Borel probability measures on X is compact in the weak-* topology, by passing to a subsequence we can assume $\mu_n \to \mu$. Clearly μ is a probability measures; we claim $\pi\mu = \nu$. It is enough to show that $\int g \, d(\pi\mu) = \int g \, d\nu$ for every $g \in C(Y)$. Using the identity $\int g \, d\nu_n = \int g \circ \pi \, d\mu_n$ (which is equivalent to $\nu_n = \pi\mu_n$) we have

$$\int g \, d\nu = \lim \int g \, d\nu_n = \int g \circ \pi \, d\mu_n = \int g \circ \pi \, d\mu = \int g \, d(\pi\mu)$$

as claimed. □

Proof No. 2 (function-analytic). . First a few general remarks. A linear functional μ^* on $C(X)$ is positive if it takes non-negative values on non-negative functions. This property implies boundedness: to see this note that for any $f \in C(X)$ we have $\|f\|_\infty - f \geq 0$, hence by linearity and positivity $\mu^*(\|f\|_\infty) - \mu^*(f) \geq 0$, giving

$$\mu^*(f) \leq \mu^*(\|f\|_\infty) = \|f\|_\infty \cdot \mu^*(1)$$

Similarly, using $f + \|f\|_\infty \geq 0$ we get $\mu^*(f) \geq -\|f\|_\infty$. Combining the two we have $|\mu^*(f)| \leq C \|f\|_\infty$, where $C = \mu^*(1)$.

Since a positive functional μ^* is bounded it corresponds to integration against a regular signed Borel measure μ, and since $\int f \, d\mu = \mu^*(f) \geq 0$ for continuous $f \geq 0$, regularity implies that μ is a positive measure. Hence a linear functional $\mu^* \in C(X)^*$ corresponds to a probability measure if and only if it be positive and $\mu^*(1) = 1$ (this is the normalization condition $\int 1 \, d\mu = 1$).

We now begin the proof. Let $\nu^* : C(Y) \to \mathbb{R}$ be bounded positive the linear functional $g \mapsto \int g \, d\nu$. The map $\pi^* : C(Y) \to C(X)$, $g \mapsto g \circ \pi$, embeds $C(Y)$ isometrically as a subspace $V = \pi(C(Y)) < C(X)$, and lifts ν^* to a bounded linear functional $\mu_0^* : V \to \mathbb{R}$ (given by $\mu_0^*(g \circ \pi) = \nu^*(g)$).

Consider the positive cone $P = \{f \in C(X) : f \geq 0\}$, and let $s \in C(X)^*$ be the functional

$$s(f) = \sup\{0, -f(x) : x \in X\}$$

It is easy to check that s is a seminorm, that $s|_P \equiv 0$ and that $-\mu_0^*(f) \leq s(f)$ on V. Hence by Hahn-Banach we can extend $-\mu_0^*$ to a functional $-\mu^*$ on $C(X)$ satisfying $-\mu^* \leq s$, which for $f \in P$ implies $\mu^*(f) \geq -s(f) = 0$, so μ^* is positive. By the previous discussion there is a Borel probability measure μ such that $\int f\, d\mu = \mu^*(f)$; for $f = g \circ \pi$ this means that

$$\int g\, d\pi\mu = \int g \circ \pi\, d\mu = \mu^*(g \circ \pi) = \mu_0^*(g \circ \pi) = \nu^*(g) = \int g\, d\nu$$

so μ is the desired measure. $\square$