Below $X, Y, \ldots$ always denote separable metrizable spaces.

1. For completeness, the diagonal principle used in Theorem 7.7 says that if h_n is a uniformly bounded (say by 1) sequence of functions on a countable set D the there is $N \subseteq \mathbb{N}$ such that $(h_n(d))_{n\in N}$ converges for every $d \in D$. Recall the usual diagonal proof or use a nonprincipal ultrafilter $\mathcal{U}$: define $h(d) = \lim_{n\to\mathcal{U}} h_n(d)$. Then h lies in the closure of $\{h_n : n \in \mathbb{N}\}$ in $[-1,1]^D$, and. . . .

2. Note that if K is compact and metrizable then $P(K)$ is a continuous image of $P(2^{\mathbb{N}})$ so to prove that $P(K)$ is compact it suffices to check compactness of $P(2^{\mathbb{N}})$. Here the Riesz representation theorem is much simpler.

3. Prove that convergence of measures is productive: If $\mu_n, \mu \in P(X)$ and $\nu_n, \nu \in P(Y)$ then $\mu_n \otimes \nu_n \longrightarrow \mu \otimes \nu$ if and only if $\mu_n \longrightarrow \mu$ and $\nu_n \longrightarrow \nu$

4. Let $\mu_n, \mu \in P(X)$ be absolutely continuous with respect to $\lambda \in P(X)$ and have densities (Radon-Nikodym derivatives) g_n and g, respectively.

Check that if $\lim_n g_n = g$ λ–almost everywhere then $\mu_n \longrightarrow \mu$; however, the inverse implication does not hold in general.

5. A sequence of $f_n : \Omega \longrightarrow \mathbb{R}$ of measurable functions (random variables) on some probability measure space (Ω, ν) is said to be convergent *in distribution* if $f_n[\nu] \in P(\mathbb{R})$ converge weakly.

Check that if $f_n \longrightarrow f$ **in probability** then $f_n \longrightarrow f$ in distribution, but not conversely.

6. Let $Y \subseteq X$; then we may think that $P(Y)$ is a subset of $P(X)$. Check that $P(Y)$ is a subspace of $P(X)$, i,e, the original topology on $P(Y)$ agrees with the one inherited from $P(X)$.

This may be compared with the fact that $C_b(Y)$ is, in general, **not** a subspace of $C_b(X)$.

7. Give a simple direct argument for the fact that if the sequence of $\mu_n \in P(\mathbb{R}^d)$ is converging then the family $\{\mu_n : n \in \mathbb{N}\}$ is uniformly tight.

8. Prove that a sequence $\mu_n \in P(\mathbb{R}^{\mathbb{N}})$ converges to μ if and only if $\pi_k[\mu_n] \longrightarrow \pi_k[\mu]$ for every k. Here $\pi_k : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^k$ denotes the projection onto the first k coordinates.

9. A sequence of $x_n \in X$ is said to be *uniformly distributed with respect to* $\mu \in P(X)$, if $1/n \sum_{k=1}^{n} \delta_{x_k} \longrightarrow \mu$.

Construct a uniformly distributed sequence in $[0,1]$ with respect to the Lebesgue measure on $[0,1]$.

10. For a given $\theta \in \mathbb{R}\setminus\mathbb{Q}$, set $x_n = n\cdot\theta - [n\cdot\theta]$. Try to prove that (x_n) is uniformly distributed. with respect to the Lebesgue measure on $[0,1]$.

Hint. An 'elementary' proof may be found in Billingsley, *Probability and measure* (25.1 in Polish 1987 edition). For an 'educated proof' test the convergence on (complex-valued) trygonometric polynomials $w_k(t) = \exp(2\pi i k t)$, and use the next item.

11. For $\mu_n, \mu \in P([0,1])$, $\mu_n \longrightarrow \mu$ if and only if $\lim_n \int_0^1 w_k \, \mathrm{d}\mu_n = \int_0^1 w_k \, \mathrm{d}\mu$ for every integer k (those trygonometric polynomials are defined above). This (Weyl's criterion) may be derived from the Stone-Weierstrass theorem.