PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 130, Number 9, Pages 2687–2699
S 0002-9939(02)06434-1
Article electronically published on March 29, 2002

ON IMAGES OF BOREL MEASURES UNDER BOREL MAPPINGS

DIMITRIS GATZOURAS

(Communicated by David Preiss)

ABSTRACT. Let X and Y be metric spaces. We show that the tight images of a (fixed) tight Borel probability measure μ on X, under all Borel mappings $f\colon X \to Y$, form a closed set in the space of tight Borel probability measures on Y with the weak*-topology. In contrast, the set of images of μ under all continuous mappings from X to Y may not be closed. We also characterize completely the set of tight images of μ under Borel mappings. For example, if μ is non-atomic, then all tight Borel probability measures on Y can be obtained as images of μ, and as a matter of fact, one can always choose the corresponding Borel mapping to be of Baire class 2.

1. Main result

Given a complete and separable metric space Z, we denote by $M_1^+(Z)$ the space of all Borel probability measures on Z, with the topology inherited from the weak*-topology on $C_b(Z)^*$; here $C_b(Z)$ is the Banach space of bounded continuous functions $h\colon Z \to \mathbb{R}$ and $C_b(Z)^*$ is its dual. The space $M_1^+(Z)$ is metrizable, Z being assumed separable; the Prohorov metric is one metric that metrizes $M_1^+(Z)$ (cf. [1], Theorem 5 in Appendix III).

The following is the main result of this note, answering in the affirmative a question raised by S. Slijepcevic.

Theorem 1. *Let X and Y be complete, separable metric spaces and μ a Borel probability measure on X. Then the set $\{\mu \circ f^{-1} \mid f\colon X \to Y \text{ Borel}\}$ is closed in $M_1^+(Y)$.*

The rest of this note is organized as follows. In this section we prove Theorem 1. In Section 2 we give an example showing that the result of Theorem 1 is in general not true for images under continuous mappings. In Section 3, and in particular in Theorem 2 there, we characterize completely the set of images $\{\mu \circ f^{-1} \mid f\colon X \to Y \text{ Borel}, \mu \circ f^{-1} \text{ tight}\}$ for a tight Borel probability measure μ, when X and Y are arbitrary metric spaces (so we drop the assumptions of completeness and separability on X and Y, but require that the measures considered

Received by the editors November 15, 1999 and, in revised form, April 19, 2001.
2000 *Mathematics Subject Classification.* Primary 28A33, 46E27, 60B05, 60B10; Secondary 26A21, 28C15, 54E70, 54H05.
Key words and phrases. Convergence of a sequence of images of a measure, tight measure, Prohorov's theorem, characterization of images of a tight measure, Baire class 2 mapping.
This research was supported by the European Commission as part of the programmes ΕΠΕΤ and ΚΠΣ.

be tight); this characterization underlies the proof of Theorem 1. We also show in Section 3 that if ν is a tight Borel image of μ, then one can always write ν as $\mu \circ f^{-1}$ with f of Baire class 2 (Theorem 2 again), and that this is optimal with respect to the class of f. Section 3 also contains a generalization of the main result: restricting attention to tight measures only, allows one to drop the assumptions of completeness and separability on X and Y in Theorem 1 (Corollary 1). Finally, in Section 4 we discuss how the set $\{\mu \circ f^{-1} \mid f\colon X \to Y \text{ Borel}, \mu \circ f^{-1} \text{ tight}\}$ sits in $C_b(Y)^*$ (rather than $M_1^+(Y)$).

We now proceed with the proof of Theorem 1. We first fix some notation, which is to be retained throughout this section. Assume that:

(1) X and Y are complete, separable metric spaces, i.e., Polish spaces.
(2) μ is a Borel probability measure on X.
(3) $f_n\colon X \to Y$, $n \in \mathbb{N}$, is a sequence of Borel mappings, such that $\mu \circ f_n^{-1} \longrightarrow \nu$, in the weak*-topology, for some Borel probability measure ν on Y.

We will then show the following, which proves Theorem 1.

Proposition 1. *There exists a Borel mapping $f\colon X \to Y$, such that $\nu = \mu \circ f^{-1}$.*

We will call $x \in X$ an atom of μ if $\mu(\{x\}) > 0$. The set of all atoms of μ will be denoted by A; it is a countable set, finite or infinite and possibly empty. We will write μ_a for the atomic part of μ:

$$\mu_a(B) := \sum_{x \in B \cap A} \mu(\{x\});$$

μ_c will denote the non-atomic (or continuous) part of μ: $\mu_c(B) := \mu(B) - \mu_a(B)$. Note that $\mu = \mu_a + \mu_c$. We will call μ non-atomic if $\mu_a = 0$ and purely atomic if $\mu_c = 0$. Finally, for $x \in X$, δ_x will denote the measure

$$\delta_x(B) := \begin{cases} 1 & \text{if } x \in B, \\ 0 & \text{if } x \notin B. \end{cases}$$

Proposition 2 (Purely atomic case)**.** *If μ is purely atomic, then $\nu = \mu \circ f^{-1}$, for some Borel $f\colon X \to Y$.*

Proof. Suppose that μ is purely atomic, say $\mu = \sum_i p_i \delta_{x_i}$, where $p_i > 0$ for all i. So $A = \{x_i\}$. We first claim that, for each atom x_i of μ, there exists a compact set $C_i \subseteq Y$, such that

$$f_n(x_i) \in C_i \qquad \forall\, n \in \mathbb{N}. \tag{1.1}$$

Suppose not. Then for any compact $C \subseteq Y$, there exists an $n_C \geq 1$, such that

$$\mu \circ f_{n_C}^{-1}(C) \leq 1 - p_i. \tag{1.2}$$

Now note that since the sequence $\{\mu \circ f_n^{-1}\colon n \in \mathbb{N}\}$ converges weak* in $M_1^+(Y)$ it must be tight, by Prohorov's theorem ([1], Theorem 6.2); i.e., for each $\epsilon > 0$, there exists a compact set $C \subseteq Y$, such that

$$\mu \circ f_n^{-1}(C) \geq 1 - \epsilon \qquad \forall\, n \in \mathbb{N}.$$

But for $0 < \epsilon < p_i$, this is violated by (1.2).

By (1.1) (and Cantor's diagonal method if A is infinite), we can find a subsequence $n_1 < n_2 < \dots$, such that, for each i, $f_{n_k}(x_i) \longrightarrow y_i$, as $k \longrightarrow \infty$, for some

$y_i \in Y$. Then

$$\mu \circ f_{n_k}^{-1} \longrightarrow \sum_i p_i \delta_{y_i} \qquad (k \longrightarrow \infty)$$

and so, by our assumption that $\mu \circ f_n^{-1} \longrightarrow \nu$,

$$\nu = \sum_i p_i \delta_{y_i};$$

we note at the outset that some of the y_i in this representation of ν may be the same. As the set A is countable, the mapping $f(x_i) = y_i$ and $f(x) = y$ for $x \in A^c$, where y is any fixed point in Y, is Borel-measurable and $\nu = \mu \circ f^{-1}$. □

We next turn to the non-atomic case. We begin with the case $X = Y = [0,1]$.

Lemma 1. *If P and Q are Borel probability measures on $[0,1]$, with P non-atomic, then $Q = P \circ f^{-1}$ for some Borel mapping $f\colon [0,1] \to [0,1]$.*

Proof. Let $F(x) := P\big([0,x]\big)$, $x \in [0,1]$, be the distribution function of P. Since P is non-atomic, F is continuous. It follows that if ξ is a random variable (on some probability space) with distribution P, then $F(\xi)$ has a uniform distribution on the unit interval; i.e., $P \circ F^{-1}$ is Lebesgue measure on the unit interval because F is continuous ([2], Exercise 14.3).

Now let $G\colon [0,1] \to [0,1]$ be the distribution function of Q and set $g(y) := \inf\{x \in [0,1]\colon G(x) \geq y\}$ for $y \in [0,1]$. Then, since $F(\xi)$ is uniform, $g(F(\xi))$ has distribution Q; i.e., since $P \circ F^{-1}$ is Lebesgue measure, $(P \circ F^{-1}) \circ g^{-1} = Q$ (cf. [2], 2nd Proof of Theorem 14.1, p. 189). Thus $Q = P \circ (g \circ F)^{-1}$, with $g \circ F$ Borel. □

The next result will allow us to pass from the unit interval $[0,1]$ to general spaces.

Theorem K (Kuratowski)**.** *Let Z be an uncountable, complete, separable metric space. Then there exists a bijection $\pi\colon Z \to [0,1]$, such that both π and π^{-1} are Borel mappings.*

Proof. [7], Chapter 15, Theorem 10. □

Proposition 3 (Non-atomic case)**.** *If μ is non-atomic, then $\nu = \mu \circ f^{-1}$ for some Borel mapping $f\colon X \to Y$.*

Proof. We shall in fact show that

$$\{\mu \circ f^{-1} \mid f\colon X \to Y \text{ Borel}\} = M_1^+(Y) \tag{1.3}$$

when μ is non-atomic. Proposition 3 is then obviously a consequence of this fact.

First note that since μ is non-atomic, any set of positive μ-measure must be uncountable. Thus X is uncountable. Suppose now that Y is also uncountable. Let π_X and π_Y be the bijections of Kuratowski's theorem (Theorem K), corresponding to X and Y respectively. Set $\tilde{\mu} := \mu \circ \pi_X^{-1}$ and $\tilde{\nu} := \nu \circ \pi_Y^{-1}$; $\tilde{\mu}$ and $\tilde{\nu}$ are Borel probability measures on $[0,1]$. Furthermore, $\tilde{\mu}$ is non-atomic, since π_X is injective and μ is non-atomic. It follows from Lemma 1 that $\tilde{\nu} = \tilde{\mu} \circ \tilde{f}^{-1}$ for some Borel $\tilde{f}\colon [0,1] \to [0,1]$. Now set $f := \pi_Y^{-1} \circ \tilde{f} \circ \pi_X$; then f is a Borel mapping of X into Y and $\nu = \mu \circ f^{-1}$.

To complete the proof we also need to consider the case where Y is countable, say $Y = \{y_j\colon j \in J\}$. Have disjoint Borel sets $X_j \subseteq X$, $j \in J$, such that $\mu(X_j) = \nu\big(\{y_j\}\big)$ for each j. This is possible because μ is non-atomic. (For example, use the intermediate value theorem for the function $\tilde{F}(x) := \tilde{\mu}\big([0,x]\big)$,

$x \in [0,1]$, which is continuous and with $\tilde{F}(0) = 0$ and $\tilde{F}(1) = 1$, to obtain points $x_0 := 0, x_1, x_2, \dots$ in $[0,1]$ such that $\tilde{\mu}(x_{j-1}, x_j) = \nu(\{y_j\})$ for all $j \in J$, and then set $X_j := \pi_X^{-1}(x_{j-1}, x_j)$. Here we identify J with $\{1, \dots, n\}$ for some $n \in \mathbb{N}$, or $\mathbb{N}$ itself, and the notation $\tilde{\mu}$, π_X is that of the previous paragraph.) Define $f(x) = y_j$ for $x \in X_j$, $j \in J$, and $f(x) = y$ for $x \in X \setminus \bigcup_{j \in J} X_j$, where y is any fixed point in Y. Then $\nu = \mu \circ f^{-1}$ and f is Borel because the sets X_j are Borel and J is countable. □

Proof of Proposition 1. If μ is either purely atomic or non-atomic, then we are done by the two preceding propositions. So assume that both $a := \mu_a(X) \neq 0$ and $c := \mu_c(X) \neq 0$. Then $a^{-1}\mu_a$ and $c^{-1}\mu_c$ are both probability measures.

Since $\{\mu \circ f_n^{-1}\}$ converges weak* in $M_1^+(Y)$, it is tight (Prohorov's theorem). It follows that each of $\{(a^{-1}\mu_a) \circ f_n^{-1}\}$ and $\{(c^{-1}\mu_c) \circ f_n^{-1}\}$ is tight. By Prohorov's theorem again (this time Theorem 6.1 of [1]), there exist probability measures ν^a and ν^c on Y, and a subsequence $n_1 < n_2 < \dots$, along which

$$(a^{-1}\mu_a) \circ f_{n_k}^{-1} \longrightarrow \nu^a \quad \text{and} \quad (c^{-1}\mu_c) \circ f_{n_k}^{-1} \longrightarrow \nu^c. \tag{1.4}$$

Propositions 2 and 3 now imply that $\nu^a = (a^{-1}\mu_a) \circ f_a^{-1}$ and $\nu^c = (c^{-1}\mu_c) \circ f_c^{-1}$, for some Borel mappings $f_a, f_c \colon X \to Y$. Since we are assuming that $\mu \circ f_n^{-1} \longrightarrow \nu$, (1.4) implies that

$$\nu(h) = \lim_{k\to\infty} \mu \circ f_{n_k}^{-1}(h) = \lim_{k\to\infty} \mu_a \circ f_{n_k}^{-1}(h) + \lim_{k\to\infty} \mu_c \circ f_{n_k}^{-1}(h) = a\nu^a(h) + c\nu^c(h),$$

for every bounded continuous function $h \colon Y \to \mathbb{R}$. It follows from this equality that $\nu = a\nu^a + c\nu^c$ (Theorem 1.3 of [1]), whence

$$\nu = \mu_a \circ f_a^{-1} + \mu_c \circ f_c^{-1}.$$

Recall that the set A of atoms of μ is countable, hence Borel. Thus the mapping $f(x) = f_a(x)$ if $x \in A$ and $f(x) = f_c(x)$ if $x \in A^c$ is Borel and $\nu = \mu \circ f^{-1}$. □

The proof of Theorem 1 is now complete.

2. Images under continuous mappings

It is natural to ask whether Theorem 1 remains true when one replaces the set $\{\mu \circ f^{-1} \mid f \text{ Borel}\}$ by $\{\mu \circ f^{-1} \mid f \text{ continuous}\}$. The answer is easily seen to be negative. In particular, it is easy to construct a sequence of continuous functions $f_n \colon [0,1] \to [0,1]$, and a Borel probability measure μ on $[0,1]$, such that $\mu \circ f_n^{-1} \longrightarrow \nu$ for some probability measure ν, which is not of the form $\mu \circ g^{-1}$ with g continuous. The following is an elementary example.

Example 1. Take μ to be Lebesgue measure on $[0,1]$, let

$$f_n(x) = \begin{cases} 2x/3 & \text{for } 0 \le x < x_n, \\ 2x/3 + 1/3 & \text{for } 1/2 \le x \le 1, \end{cases}$$

and define f_n to be linear in $[x_n, 1/2)$ and so that f_n is continuous on $[0,1]$; here $\{x_n\}$ is any sequence with $0 < x_1 < x_2 < \dots \uparrow 1/2$. Then $f_n \longrightarrow f$, where

$$f(x) = \begin{cases} 2x/3 & \text{for } 0 \le x < 1/2, \\ 2x/3 + 1/3 & \text{for } 1/2 \le x \le 1. \end{cases}$$

Hence $\mu \circ f_n^{-1} \longrightarrow \nu$, where $\nu = \mu \circ f^{-1}$, and ν cannot be written as $\mu \circ g^{-1}$ for some continuous g. For $\nu(a,b) = 0$ for any interval with $1/3 < a < b < 2/3$, and so if $\nu = \mu \circ g^{-1}$, then $g^{-1}(a,b)$ must be of Lebesgue measure zero. If g

is continuous however, $g^{-1}(a,b)$ is open and must therefore either have positive Lebesgue measure or be empty. But $g^{-1}(a,b)$ can also not be empty when g is continuous and $\nu = \mu \circ g^{-1}$, because we must have $g([0,1]) = [0,1]$ for such a g.

3. Further results and generalizations

Section 1 contains actually more information than is stated in Theorem 1. For example (1.3) shows that *any* Borel probability measure on Y can be obtained as an image of *any* non-atomic μ on X. Furthermore, the assumptions of completeness and separability on X and Y are superfluous throughout Section 1, provided one restricts attention to tight probability measures (see below for the definition). In the present section we drop these assumptions on X and Y altogether, give a complete characterization of the tight measures ν on Y that can be obtained as Borel images of a given tight measure μ on X (Theorem 2) and extend Theorem 1 to arbitrary metric spaces (Corollary 1).

Given metric spaces X and Y, write

$$B(\mu) := \{\mu \circ f^{-1} \mid f\colon X \to Y \text{ Borel}\}$$

for the set of images of a Borel measure μ on X under Borel mappings from X to Y. We also retain the notation of Section 1 pertaining to measures: for a measure κ on a metric space Z we write κ_a and κ_c for its atomic and continuous parts respectively and denote by A_κ its possibly empty and always countable set of atoms. Furthermore, having dropped the assumption that Z is Polish, we now denote by $M_1^+(Z)$ the space of *tight* Borel probability measures on Z. Recall that a Borel probability measure κ on a metric space Z is tight if, given any $\epsilon > 0$, there exists a compact set K such that $\kappa(K^c) < \epsilon$.

Note. A few remarks regarding terminology are perhaps in order. Recall that every finite Borel measure κ on a metric space Z is regular in the weak sense: for each Borel set B, and any $\epsilon > 0$, there exist an open set O and a closed set C, such that $C \subseteq B \subseteq O$ and $\kappa(O \smallsetminus C) < \epsilon$ ([7], Chapter 15, Proposition 11). A Borel probability measure κ is tight iff it is regular in the strong sense, i.e., for each $\epsilon > 0$, the closed set C above can be chosen to be compact. Recall that when Z is Polish, every Borel probability measure is strongly regular, hence tight (cf. [7], Chapter 15, Proposition 18, or [1], Theorem 1.4). Thus the notation $M_1^+(Z)$ introduced here agrees with the one already introduced for Polish spaces in Section 1 (i.e., for Polish spaces, $M_1^+(Z)$ contains *all* Borel probability measures on Z).

3.1. **Characterization of Borel images of a tight measure.**
Part (i) of the following theorem is in the heart of the proof of Theorem 1 (and Corollary 1).

Theorem 2. *Let X and Y be metric spaces, μ a tight Borel probability measure on X and ν a tight Borel measure on Y. Write*

$$\mu_a = \sum_{i \in I} p_i \delta_{x_i} \qquad \big(A_\mu = \{x_i \colon i \in I\};\ p_i > 0\ \forall\, i \in I\big)$$

and

$$\nu_a = \sum_{j \in J} q_j \delta_{y_j} \qquad \big(A_\nu = \{y_j \colon j \in J\};\ q_j > 0\ \forall\, j \in J\big)$$

for the atomic parts of μ and ν respectively, with I and J possibly empty. Then:

(i) *$\nu \in B(\mu)$ iff ν is a probability measure, and there exists a partition $I = \bigcup_{j\in J} I_j$ of the atoms of μ, with some I_j possibly empty, such that*

$$q_j \geq \sum_{i\in I_j} p_i \qquad \forall j \in J.$$

(ii) *If $\nu \in B(\mu)$, then one can always write $\nu = \mu \circ f^{-1}$ with f of Baire class 2.*

Remark. Part (i) of this theorem

(A) simply reflects the obvious fact that the atoms of a measure cannot decompose when we take an image of it (but several atoms, or non-atomic mass, can aggregate to form a single atom), and

(B) also states that there are no further restrictions for a tight measure ν on Y to be a Borel image of μ (except of course that μ and ν must also have the same total mass).

Thus in particular, if μ is non-atomic, then $B(\mu) \cap M_1^+(Y) = M_1^+(Y)$, as already shown in (1.3) when X and Y are Polish, while when μ is purely atomic, then $B(\mu)$ consists precisely of all purely atomic measures of the form $\nu = \sum_{j\in J} q_j \delta_{y_j}$ with

$$q_j = \sum_{i\in I_j} p_i \qquad \forall j \in J,$$

where $\{I_j : j \in J\}$ ranges over all possible partitions $I = \bigcup_{j\in J} I_j$ of the atoms of μ.

Before giving the proof of Theorem 2 we briefly recall the terminology on Baire classes and collect a couple of elementary facts that will be used in the proof, for the convenience of the reader.

A Borel mapping $f\colon X \to Y$ is said to be of Baire class 0 if it is continuous; f is of class 1 if $f^{-1}(O)$ is an F_σ-set for every open set O in Y; it is of class 2 if, for every open O in Y, $f^{-1}(O)$ is a $G_{\delta\sigma}$-set, etc.; see Kuratowski [6] for these definitions. The composition of two mappings of classes a and b, respectively, is of class $a+b$. And if $X = X_1 \cup X_2$ with each X_i an $F_{\sigma\delta}$ and $f|_{X_i}$ is of class 2 for each $i = 1, 2$, then f is of class 2. These elementary facts will be used below and can be found in [6] (Theorem 2 in §31, III and Theorem 2 in §31, IV).

Proof of Theorem 2. One direction in (i) is entirely trivial. If $\nu = \mu \circ f^{-1}$ for some Borel $f\colon X \to Y$, then ν is obviously a probability measure, and $I_j := \{i \in I\colon x_i \in f^{-1}(\{y_j\})\}$ gives the desired partition of the atoms of μ.

For the other direction, consider a tight Borel probability measure ν on Y, for which a partition $I = \bigcup_{j\in J} I_j$ such that $q_j \geq \sum_{i\in I_j} p_i$ exists. We seek to find a mapping $f\colon X \to Y$, of Baire class 2, such that $\nu = \mu \circ f^{-1}$; this will then also prove (ii).

First observe that we may, without loss of generality, assume that μ is non-atomic. For if the theorem is true for non-atomic measures, have a Baire class 2 map $f_c\colon X \to Y$ such that $\nu' = \mu_c \circ f_c^{-1}$, where ν' is the measure on Y with $\nu'_c = \nu_c$ and $\nu'(\{y_j\}) = q_j - \sum_{i\in I_j} p_i$ for each $j \in J$, and then define $f\colon X \to Y$ by setting $f = f_c$ on $X \smallsetminus A_\mu$ and $f(x_i) := y_j$ for $i \in I_j$, $j \in J$. As $f|_{X\smallsetminus A_\mu}$ is of class 2 and $f|_{A_\mu}$ of class 1 (since A_μ is countable), and as each of the sets $X \smallsetminus A_\mu$ and A_μ is an $F_{\sigma\delta}$ (being G_δ and F_σ, respectively), the resulting mapping f is of class 2 on X, and clearly $\nu = \mu \circ f^{-1}$.

We may further assume that X and Y are σ-compact. For, since μ and ν are assumed to be tight, one can find σ-compact subsets $X' \subseteq X$ and $Y' \subseteq Y$ of full μ- and ν-measure, respectively. If $f\colon X' \to Y'$ is a Baire 2 mapping with $\nu|_{Y'} = (\mu|_{X'}) \circ f^{-1}$, extend f to all of X by letting it be constant on $X \setminus X'$; the resulting mapping will be of class 2 on all of X, for the same reasons as in the previous paragraph, and again one will have $\nu = \mu \circ f^{-1}$, by the very definition of X' and Y'.

So assume from now on that X and Y are σ-compact and that μ is non-atomic. The idea of course is to embed X and Y into $[0,1]$ and then use Lemma 1 to construct an f that carries μ to ν, as in the proof of Proposition 3. A little care is needed in doing this however, because here we also want to control the class of the resulting mapping f. We remark that a shorter and less involved argument for this part of the theorem can be given if one does not pay attention to the class of f, i.e., if one only cares to obtain part (i) of Theorem 2; see the Remark following this proof.

Let $e_X\colon X \to \mathbb{R}^\infty$ be the mapping given by $e_X(x) := \big(\rho_X(x,d_1), \rho_X(x,d_2), \dots\big)$, where ρ_X is a metric on X and $\{d_n\colon n \in \mathbb{N}\}$ is a countable dense subset of X. As is well known, e_X is a homeomorphism of X onto $e_X(X)$ (cf. [6], §22, II, Remark 3, or [1], Appendix I, p. 219). Define $e_Y\colon Y \to \mathbb{R}^\infty$ similarly.

Let G be a zero-dimensional, uncountable, G_δ-subset of $\mathbb{R}^\infty$ which has both full $\mu \circ e_X^{-1}$- and full $\nu \circ e_Y^{-1}$-measure. Such a set G may be constructed as follows. Let κ be the Borel measure $\kappa := \mu \circ e_X^{-1} + \nu \circ e_Y^{-1}$ on $\mathbb{R}^\infty$. Since, for each $n \geq 1$, $x \mapsto \kappa\left(\{(x_m)_{m\in\mathbb{N}} \in \mathbb{R}^\infty\colon x_n \leq x\}\right)$ is a non-decreasing function of $x \in \mathbb{R}$, there exists a countable, dense in $\mathbb{R}$ set $\{q_{n,k}\colon k \in \mathbb{N}\}$ such that $\kappa\left(\{(x_m)_{m\in\mathbb{N}} \in \mathbb{R}^\infty\colon x_n = q_{n,k}\}\right) = 0$ for each $k \in \mathbb{N}$. The set

$$G := \bigcap_n \bigcap_k \big\{(x_m)_{m\in\mathbb{N}} \in \mathbb{R}^\infty\colon x_n \neq q_{n,k}\big\}$$

then has the desired properties. It follows that G is homeomorphic to a closed subset of $\mathbb{I} := [0,1] \setminus \mathbb{Q}$ (cf. [6], §36, II, Theorem 2). Let $h\colon G \to h(G) \subseteq \mathbb{I}$ be a homeomorphism.

Next let $\pi_X\colon e_X^{-1}(G) \to [0,1]$ and $\pi_Y\colon e_Y^{-1}(G) \to [0,1]$ be the mappings defined by $\pi_X := h \circ e_X$ and $\pi_Y := h \circ e_Y$, respectively. π_X and π_Y are homeomorphisms of $e_X^{-1}(G)$ and $e_Y^{-1}(G)$ onto subsets $\tilde{X}$ and $\tilde{Y}$ of $\mathbb{I}$, and $\tilde{X} = h\big(e_X(X) \cap G\big)$ and $\tilde{Y} = h\big(e_Y(Y) \cap G\big)$ are F_σ in $\mathbb{I}$ because X and Y are assumed to be σ-compact and $h(G)$ is closed in $\mathbb{I}$.

Since μ is non-atomic, and since π_X is an injection, the measure $\tilde{\mu} := \mu \circ \pi_X^{-1}$ defines a non-atomic Borel probability measure on $[0,1]$. Consequently, by Lemma 1, there exists a mapping $\tilde{f}\colon [0,1] \to [0,1]$ such that $\tilde{\nu} = \tilde{\mu} \circ \tilde{f}^{-1}$, where $\tilde{\nu} := \nu \circ \pi_Y^{-1}$. Define $f := \pi_Y^{-1} \circ \tilde{f} \circ \pi_X$ on $\pi_X^{-1}\big(\tilde{X} \cap \tilde{f}^{-1}(\tilde{Y})\big)$ and then extend it to be constant on the complement of this set. Obviously, $\nu = \mu \circ f^{-1}$.

We now consider the class of f. One sees upon inspecting the proof of Lemma 1 that $\tilde{f}\colon [0,1] \to [0,1]$ is always a non-decreasing function. This implies that if I is an interval, then $\tilde{f}^{-1}(I)$ is also an interval, which may or may not contain its endpoints. In particular, if I is an interval which is open as a subset of $[0,1]$ (i.e., $I = (a,b)$, $I = [0,b)$, $I = (a,1]$, or $I = [0,1]$), then $\tilde{f}^{-1}(I)$ is an open interval union with a finite set, the latter consisting of one, both, or none of the endpoints of this interval.[1]

[1] As a matter of fact $\tilde{f}$ is also left-continuous, so the left endpoint is never included in the inverse image of an open interval, unless it equals 0.

It follows that:

(a) $\tilde{f}^{-1}(O)$ is an F_σ in $[0,1]$ whenever O is an open subset of [0,1];
(b) for any set C closed in $[0,1]$, $\tilde{f}^{-1}(C)$ is a closed $\smallsetminus$ countable set.

By (a), $\tilde{f}$ is of class 1 on $[0,1]$. By (b), $\tilde{f}^{-1}(\tilde{Y})$ is of the form $\tilde{f}^{-1}(\tilde{Y}) = F_1 \smallsetminus F_2$ with F_1 and F_2 both F_σ in $[0,1]$: $\tilde{Y} = F \smallsetminus \mathbb{Q}$, where F is an F_σ in $[0,1]$ (recall that $\tilde{Y}$ is F_σ in $\mathbb{I} = [0,1] \smallsetminus \mathbb{Q}$), whence by (b)

$$\tilde{f}^{-1}(\tilde{Y}) = \tilde{f}^{-1}(F) \smallsetminus \tilde{f}^{-1}(\mathbb{Q}) = (F_1 \smallsetminus C) \smallsetminus \tilde{f}^{-1}(\mathbb{Q})$$

with F_1 an F_σ in $[0,1]$ and C countable. Since $\tilde{f}^{-1}(\mathbb{Q})$ is also an F_σ, because by the monotonicity of $\tilde{f}$ again inverse images of singletons are either empty, single points, or intervals, $\tilde{f}^{-1}(\tilde{Y}) = F_1 \smallsetminus F_2$ with each F_i an F_σ in $[0,1]$, as asserted.

Since π_X is continuous, $\pi_X^{-1}\big(\tilde{X} \cap \tilde{f}^{-1}(\tilde{Y})\big)$ is of the form

$$\pi_X^{-1}\big(\tilde{X} \cap \tilde{f}^{-1}(\tilde{Y})\big) = F_1' \smallsetminus F_2'$$

with each F_i' an F_σ in $e_X^{-1}(G)$; since $e_X^{-1}(G)$ is itself a G_δ in X, it follows that $\pi_X^{-1}\big(\tilde{X} \cap \tilde{f}^{-1}(\tilde{Y})\big)$ is of the form

$$\pi_X^{-1}\big(\tilde{X} \cap \tilde{f}^{-1}(\tilde{Y})\big) = F' \cap G'$$

with F' an F_σ in X and G' a G_δ in X. Thus:

- both $\pi_X^{-1}\big(\tilde{X} \cap \tilde{f}^{-1}(\tilde{Y})\big)$ and its complement are $F_{\sigma\delta}$-subsets of X;
- f is of class 1 when restricted to $\pi_X^{-1}\big(\tilde{X} \cap \tilde{f}^{-1}(\tilde{Y})\big)$, since $f = \pi_Y^{-1} \circ \tilde{f} \circ \pi_X$ on this set, and $\tilde{f}$ is of class 1 while both π_X and π_Y^{-1} are continuous;
- f is of class 0 on the complement of $\pi_X^{-1}\big(\tilde{X} \cap \tilde{f}^{-1}(\tilde{Y})\big)$, being constant on that set.

It follows that f is of class 2 on X. □

Remark. A more direct argument for part (i) of Theorem 2 can be given when one does not care about the class of the f carrying a given μ to a given ν. First reduce, as in the above proof, to the case where μ is non-atomic and X and Y are σ-compact. Define e_X and e_Y as before and then use Theorem K to obtain a Borel isomorphism $\pi\colon \mathbb{R}^\infty \to [0,1]$. Then, $\pi_X := \pi \circ e_X$ and $\pi_Y := \pi \circ e_Y$ are Borel isomorphisms from X and Y onto Borel subsets $\tilde{X}$ and $\tilde{Y}$ of $[0,1]$, respectively. Now use Lemma 1 as in the above proof to obtain a Borel $f\colon X \to Y$ with $\nu = \mu \circ f^{-1}$. The essential ingredients of this argument are of course (1) tightness, which allows one to restrict attention to σ-compact subsets of X and Y of full measure, and (2) the fact that any σ-compact metric space is Borel isomorphic to a Borel subset of $[0,1]$ (as the above argument shows), so that $\tilde{X}$ and $\tilde{Y}$ are Borel in $[0,1]$. Completeness and separability, which is what we assumed of X and Y in Proposition 3, are thus unnecessary, provided one restricts oneself to tight measures.

3.2. **Optimality in Theorem 2.** Theorem 2 is optimal with respect to the Baire class of f in a representation $\nu = \mu \circ f^{-1}$ of a $\nu \in B(\mu) \cap M_1^+(Y)$. Example 1 in Section 2 shows that it may be impossible to write a $\nu \in B(\mu) \cap M_1^+(Y)$ as $\nu = \mu \circ f^{-1}$ with f of class 0. The following example shows that it may also be impossible to choose f to be of class 1.

Example 2. Let κ be an arbitrary probability measure concentrated on the rationals in $[0,1]$ and such that $\kappa\big(\{q\}\big) > 0$ for each $q \in \mathbb{Q} \cap [0,1]$. Let $X = Y = [0,1]$

and λ be Lebesgue measure on $[0,1]$. Define

$$\mu = \frac{1}{2}\lambda + \frac{1}{2}\kappa,$$

$$f(x) = \begin{cases} 0 & \text{for } x \in \mathbb{Q} \cap [0,1], \\ x/2 + 1/2 & \text{for } x \in [0,1] \setminus \mathbb{Q}, \end{cases}$$

and let $\nu = \mu \circ f^{-1}$. Suppose now that $\nu = \mu \circ g^{-1}$ with g of class 1. Then $g^{-1}(\{0\})$ is a G_δ-set, G say, containing $\mathbb{Q} \cap [0,1]$. Hence G is also dense. Furthermore, $F := g^{-1}([0,1/2))$ is F_σ, contains G, and also has $\lambda(F) = 0$. But this is impossible by virtue of the Baire category theorem: If $F = \bigcup_n F_n$ with the F_n closed, each F_n is nowhere dense, because F^c is a set of full measure contained in F_n^c. If $G^c = \bigcup_n C_n$ with the C_n closed, then each C_n is also nowhere dense, because C_n^c contains G, which is dense. But then $X = \bigcup_n (F_n \cup C_n)$, a countable union of nowhere dense sets.

The following example shows that it may be impossible to choose f of class 1, even when μ is purely atomic. This example is due to S. Papadopoulou.

Example 3. Let p_n, $n \in \mathbb{N}$, be (strictly) positive numbers, which sum to 1, and such that no p_n can be written as a sum of other p_j's. This can be done, for example, by choosing $p_1 > 1/2$ and then $p_n > (1 - p_1 - \ldots - p_{n-1})/2$ for each $n > 1$. Let μ be the probability measure which puts mass p_n at q_n, where $\{q_n \colon n \in \mathbb{N}\}$ is an enumeration of the rationals in $[0,1]$. Let $Q_1 \cup Q_2$ be a partition of $\mathbb{Q} \cap [0,1]$ with each Q_i dense, and let $f \colon [0,1] \to [0,1]$ be any function which is one-to-one on the rationals and sends Q_1 into $[0,1/3]$ and Q_2 into $[2/3,1]$. The measure $\nu = \mu \circ f^{-1}$, on $[0,1]$, cannot be written as $\nu = \mu \circ g^{-1}$ with g of class 1. For we must have $f = g$ on $\mathbb{Q} \cap [0,1]$, by our choice of the p_n, and then $g^{-1}([0,1/3])$ and $g^{-1}([2/3,1])$ are both dense G_δ-sets, which are also disjoint. This is again impossible by Baire's theorem.

3.3. **Extension of Theorem 1 to arbitrary metric spaces.** Recall that, for an arbitrary metric space Z, we denote by $M_1^+(Z)$ the set of *tight* Borel probability measures on Z. Endowed with the weak*-topology inherited from $C_b(Z)^*$, this space is still metrizable by the Prohorov metric (cf. [1], Theorem 5 in Appendix III).

Corollary 1. *Let X and Y be metric spaces and $\mu \in M_+^1(X)$. Then $B(\mu) \cap M_+^1(Y)$ is closed in $M_+^1(Y)$.*

Proof. Recall the proof of Theorem 1. The assumptions of completeness and separability on X and Y were used

1. when applying Prohorov's theorem in the form of Theorem 6.2 of [1] (i.e., that weak*-convergence implies tightness) in the proofs of Propositions 1 and 2;
2. when applying Theorem K in the proof of Proposition 3, to get (1.3) for general X and Y from Lemma 1.

Prohorov's theorem remains true for arbitrary metric spaces when one restricts attention to tight Borel probability measures. In particular, one can replace Theorem 6.2 of [1] by Theorem 8 in Appendix III of the same reference. One can also replace (1.3) by

$$B(\mu) \cap M_1^+(Y) = M_1^+(Y) \tag{3.1}$$

for arbitrary X and Y and μ non-atomic, which is part of the conclusion of Theorem 2 (see Remark after Theorem 2; see also the Remark following the proof of Theorem 2 in this connection). Thus the proof of Theorem 1 goes through essentially verbatim to prove Corollary 1. □

4. Concluding remarks

4.1. Let Y be an arbitrary metric space. Since Y is a completely regular topological space, $C_b(Y)^*$ may be identified, via an isometric isomorphism, with the space $M(\check{Y})$ of regular, signed, Borel measures on the Stone-Čech compactification $\check{Y}$ of Y (cf. [3], Chapter V, Corollary 6.4). The set $M_b(Y)$ of finite, tight, signed, Borel measures on Y is obviously contained in $M(\check{Y})$ (a signed measure being tight iff both its positive and negative parts are).

When Y is locally compact and separable, sequences in $M_1^+(Y)$ can only converge to elements in $M_1^+(Y)$. (Unfortunately, we could not locate a reference for this fact in the literature, so we give an argument below.) Thus in this case Corollary 1 also shows that the set of tight Borel images $B(\mu)\cap M_1^+(Y)$, of a tight Borel probability measure μ on X, is *sequentially* closed in $C_b(Y)^*$. When Y is actually compact, $M_1^+(Y)$ is a compact subset of $C_b(Y)^* = C(Y)^*$, and hence $B(\mu)$ is closed in $C_b(Y)^*$. (Thus $B(\mu)$ is compact when Y is compact.) When Y is not compact however, the set $B(\mu)\cap M_1^+(Y)$ may not be closed in $C_b(Y)^*$, as the following argument shows.

First observe that when Y is not compact, $M_1^+(Y)$ is not closed in $M(\check{Y})$: for if $y \in \check{Y} \setminus Y$, then $y_\alpha \longrightarrow y$ for some net $\{y_\alpha\} \subseteq Y$, as Y is dense in $\check{Y}$; consequently $\delta_{y_\alpha} \longrightarrow \delta_y$ in the weak*-topology, showing that $M_1^+(Y)$ is not closed in $M(\check{Y})$. But then, if μ is any non-atomic element of $M_1^+(X)$, $B(\mu)\cap M_1^+(Y) = M_1^+(Y)$, by Theorem 2, and so $B(\mu)\cap M_1^+(Y)$ is not closed in $C_b(Y)^* \simeq M(\check{Y})$.

One is naturally led to wonder whether the fact that $M_1^+(Z)$ is sequentially closed in $C_b(Z)^*$ remains true for arbitrary metric spaces Z. As above, this would of course automatically imply that $B(\mu)\cap M_1^+(Y)$ is sequentially closed in $C_b(Y)^*$ for arbitrary metric spaces X and Y and $\mu \in M_1^+(X)$, but this question is also of independent interest.

Note added in proof. It turns out that $M_1^+(Z)$ *is* sequentially closed in $C_b(Z)^*$ whenever the former contains *all* Borel probability measures on Z (cf. [4]); thus in particular, besides the case where Y is locally compact and separable, $B(\mu)\cap M_1^+(Y)$ is also sequentially closed in $C_b(Y)^*$ whenever Y is Polish. In general however, $M_1^+(Z)$ need *not* be sequentially closed in $C_b(Z)^*$ (cf. [1], comment following Theorem 8 in Appendix III), and neither does $B(\mu)\cap M_1^+(Y)$ need to be sequentially closed in $C_b(Y)^*$. For an example of the latter (taken from [1]), take Y to be a subset of the unit interval with inner Lebesgue measure 0 and outer Lebesgue measure 1, with the topology it inherits from $[0,1]$. If ν is outer Lebesgue measure on the Borel subsets of Y, every compact set in Y is Borel in $[0,1]$ and therefore has ν-measure 0, as it has inner measure 0; thus ν is non-tight. On the other hand, one can find measures ν_n of *finite* support on Y such that $\nu_n \to \nu$ (cf. [1], Theorem 4 in Appendix III). Then however, if μ is Lebesgue measure on $X := [0,1]$, one can easily find Borel functions $f_n : X \to Y$ such that $\nu_n = \mu \circ f_n^{-1}$, whence $\mu \circ f_n^{-1} \in M_1^+(Y), \mu \circ f_n^{-1} \to \nu$, yet $\nu \notin M_1^+(Y)$.

4.2. We now present an argument for the following fact, which was used above. *If Z is a separable, locally compact metric space and $\{\mu_n\}$ a sequence of measures in $M_1^+(Z)$ which converges weak* to some $\mu \in C_b(Z)^* \simeq M(\check{Z})$, then $\mu \in M_1^+(Z)$.*

Note. Since a locally compact and separable metric space is σ-compact, all Borel probability measures on such a space Z are tight; thus, as in the case of Polish spaces, $M_1^+(Z)$ again contains *all* Borel probability measures on Z.

In view of Prohorov's theorem and the preceding Note, it suffices to show that convergence of a sequence of Borel probability measures $\{\mu_n\}$ on Z in $C_b(Z)^*$ implies tightness of the sequence.

Assume $\{\mu_n\} \subseteq M_1^+(Z)$ is not tight. Then for some (and hence all sufficiently small) $\varepsilon > 0$ we have that

$$\limsup_{n\to\infty} \mu_n(K^c) \geq \varepsilon \qquad \text{for all compact } K \subseteq Z. \tag{4.1}$$

Fix ϵ to be the supremum of all those ε for which (4.1) holds and let $\epsilon' < \epsilon < \epsilon''$, with $\epsilon'' - \epsilon' \leq \epsilon'/2$.

Let O_n, $n \in \mathbb{N}$, be a countable collection of open sets with compact closures, whose union is all of Z; the existence of such a sequence is guaranteed by local compactness and separability. First have a compact set K_0 with $\mu_n(K_0^c) \leq \epsilon''$ for all sufficiently large n, say $n \geq N$. Next, use (4.1) to choose $n_1 \geq N$ with $\mu_{n_1}(K_0^c) > \epsilon'$, and then use tightness to choose a compact $K_1 \subseteq K_0^c$ with $\mu_{n_1}(K_1) \geq \epsilon'$. Having chosen $K_0, \ldots, K_{i-1}$, $i > 1$, use (4.1) to choose $n_i > n_{i-1}$, such that the μ_{n_i}-measure of the complement of the compact set $(K_0 \cup K_1 \cup \ldots \cup K_{i-1}) \cup (\overline{O}_1 \cup \ldots \cup \overline{O}_{i-1})$ is $> \epsilon'$, and then have a compact set K_i, with

$$K_i \subseteq [(K_0 \cup K_1 \cup \ldots \cup K_{i-1}) \cup (\overline{O}_1 \cup \ldots \cup \overline{O}_{i-1})]^c, \tag{4.2}$$

such that $\mu_{n_i}(K_i) \geq \epsilon'$.

Let A be the union of the K_i over even $i \geq 0$ and B the union over odd i. The sets A and B are disjoint, by (4.2), and, using (4.2) again, it is easy to see that $A \cup B = \bigcup_{i\geq 0} K_i$ is closed (the O_n are open and cover Z). Define a function $f \in C_b(Z)$ by setting $f = 0$ on A, $f = 1/\epsilon'$ on B, and extending f continuously onto all of Z and so that $0 \leq f \leq 1/\epsilon'$. Then $\mu_{n_i}(f) \geq 1$ for all odd i, and $\mu_{n_i}(f) \leq (\epsilon'' - \epsilon') \cdot (1/\epsilon') \leq 1/2$ for i even and positive, so that $\{\mu_{n_i}\}$ cannot converge in $C_b(Z)^*$.

Note. As pointed out by the referee, the above argument is really a variant of the argument proving Shur's theorem, namely that strong and weak convergence in ℓ^1 are equivalent ([5], Theorem 99), and one can, as a matter of fact, derive the above statement—of sequential closedness of $M_1^+(Z)$ for locally compact separable Z—as a consequence of Shur's theorem.

4.3. We conclude with one more observation. Let $M_b^+(Z)$ denote the space of finite, non-negative, tight, Borel measures on an arbitrary metric space Z, with the weak*-topology. Then, part of the argument in the Proof of Proposition 1 provides a proof of the following general fact (which, as we learned, seems to be part of the folklore); this fact was essentially used (and proved) implicitly in the proof of Proposition 1. *If Λ and K are closed sets in $M_b^+(Z)$, then their algebraic sum $K + \Lambda$ is also closed.*

The complete argument is as follows. First note that $M_b^+(Z)$ is metrizable: If ρ_1 denotes the Prohorov metric on $M_1^+(Z)$, the metric defined by

$$\rho_b(\kappa, 0) := \kappa(Z)$$

for all $\kappa \in M_b^+(Z)$, where 0 is the measure identically equal to 0, and by

$$\rho_b(\kappa, \lambda) := |\kappa(Z) - \lambda(Z)| + \rho_1(\kappa', \lambda') \cdot \min\{\kappa(Z), \lambda(Z)\}$$

for $\kappa, \lambda \in M_b^+(Z) \setminus \{0\}$, where κ' and λ' are the measures κ and λ respectively normalized to be probability measures, provides a metric on $M_b^+(Z)$ compatible with the weak*-topology. Thus sequences are enough to characterize closed sets in $M_b^+(Z)$.

Let $\{\kappa_n + \lambda_n\}$ be a sequence in $K + \Lambda$ that converges to some element μ in $M_b^+(Z)$. We consider two cases:

Case 1: One (or both) of the numerical sequences $\{\kappa_n(Z)\}$ and $\{\lambda_n(Z)\}$, say $\{\kappa_n(Z)\}$, has a subsequence $n_1 < n_2 < \dots$ that converges to 0. Then $\kappa_{n_j} \to 0$ in $M_b^+(Z)$, whence also $\lambda_{n_j} \to \mu$. Thus $\mu = 0 + \mu$ with $0 \in K$ and $\mu \in \Lambda$, because we have assumed K and Λ to be closed.

Case 2: Both $\inf_n \kappa_n(Z) > 0$ and $\inf_n \lambda_n(Z) > 0$. Since $\kappa_n + \lambda_n \to \mu$ in $M_b^+(Z)$, the normalized measures $\mu_n := (\kappa_n + \lambda_n)/[\kappa_n(Z) + \lambda_n(Z)]$ converge in $M_1^+(Z)$, and therefore form a tight sequence, by Prohorov's theorem. It follows that each of the normalized sequences $\{\kappa'_n\}$ and $\{\lambda'_n\}$ is tight and therefore has a convergent subsequence, again by Prohorov's theorem. Thus $\kappa'_{n_j} \longrightarrow \kappa'$ and $\lambda'_{n_j} \longrightarrow \lambda'$, for some subsequence $\{n_j\}$ and probability measures κ' and λ' on Z. Now choose a further subsequence $\{n_{j(i)}\}$, such that the limits $\lim_i \kappa_{n_{j(i)}}(Z)$ and $\lim_i \lambda_{n_{j(i)}}(Z)$ exist; this is possible because the convergence of $\{\kappa_n + \lambda_n\}$ to an element $\mu \in M_b^+(Z)$ implies that $\{\kappa_n(Z)\}$ and $\{\lambda_n(Z)\}$ are bounded sequences of real numbers. Then $\kappa_{n_{j(i)}} \longrightarrow \kappa$ and $\lambda_{n_{j(i)}} \longrightarrow \lambda$, where $\kappa := \kappa' \cdot \lim_i \kappa_{n_{j(i)}}(Z)$ and λ is defined similarly. It follows that $\mu = \kappa + \lambda$ with $\kappa \in K$ and $\lambda \in \Lambda$, because we have assumed K and Λ to be closed.

Note. Since Z is an arbitrary metric space in the above argument, one is using one direction of Prohorov's theorem in the form of Theorem 8 in Appendix III of [1] rather than Theorem 6.2 of the same reference, just as in the proof of Corollary 1.

Acknowledgments

The author expresses his gratitude to S. Lalley and S. Papadopoulou for enlightening discussions while writing this note. The author is also indebted to an anonymous referee, whose comments and suggestions led to dramatic improvements over earlier versions of this article.

References

1. P. Billingsley (1968) *Convergence of Probability Measures*, Wiley, New York. MR **38:**1718
2. P. Billingsley (1995) *Probability and Measure*, 3rd ed., Wiley, New York. MR **95k:**60001
3. J. B. Conway (1990) *A Course in Functional Analysis*, 2nd ed., Springer, New York. MR **91e:**46001
4. D. Gatzouras (2001) *On Weak Convergence of Probability Measures in Metric Spaces*, preprint.
5. P. Habala, P. Hájek, V. Zizler (1996) *Introduction to Banach Spaces* [I], Matfyzpress, vydavatelstvi Matematicko-fyzikálni fakulty Univerzity Karlovy.

6. K. Kuratowski (1966) *Topology*, Vol. 1, Academic Press, New York. MR **36:**840
7. H. L. Royden (1988) *Real Analysis*, 3rd ed., Macmillan Publishing Company, New York. MR **90g:**00004

Department of Mathematics, University of Crete, Leoforos Knossou, 714 09 Iraklion, Crete, Greece

E-mail address: gatzoura@math.uoc.gr

Current address: Department of Mathematics, Agricultural University of Athens, Iera Odos 75, 118 55 Athens, Greece

E-mail address: gatzoura@aua.gr