A characterisation of isometries
with respect to the Lévy-Prokhorov metric

GYÖRGY PÁL GEHÉR AND TAMÁS TITKOS

Abstract. According to the fundamental work of Yu.V. Prokhorov, the general theory of stochastic processes can be regarded as the theory of probability measures in complete separable metric spaces. Since stochastic processes depending upon a continuous parameter are basically probability measures on certain subspaces of the space of all functions of a real variable, a particularly important case of this theory is when the underlying metric space has a linear structure. Prokhorov also provided a concrete metrisation of the topology of weak convergence today known as the Lévy-Prokhorov distance. Motivated by these facts, the famous Banach-Stone theorem, and some recent works related to characterisations of onto isometries of spaces of Borel probability measures, here we give a complete description of surjective isometries with respect to the Lévy-Prokhorov metric in case when the underlying metric space is a separable Banach space. Our result can be considered as a generalisation of L. Molnár's earlier Banach-Stone-type result which characterises onto isometries of the space of all probability distribution functions on the real line wit respect to the Lévy distance. However, the present more general setting requires the development of an essentially new technique.

Mathematics Subject Classification (2010): 46B04 (primary); 46E27, 47B49, 54E40, 60B10, 28A33, 60A10, 60B05 (secondary).

1. Introduction

There is a long history and vast literature of isometries (*i.e.*, not necessarily surjective distance preserving maps) on different kind of metric spaces. Two classical results in the case of normed linear spaces are the Mazur-Ulam theorem which

every surjective isometry between real normed spaces is automatically affine (*i.e.*, linear up to translation), and the Banach-Stone theorem which provides the structure of onto linear isometries between Banach spaces of continuous scalar-valued functions on compact Hausdorff spaces. Since then several properties of surjective linear isometries on different types of normed spaces have been explored, see for instance the papers [2–5, 10, 13, 24] and the extensive books [16, 17]. The reader can find similar results on non-linear spaces for example in [6, 11, 18, 21, 27, 28].

The starting point of our investigation is Molnár's paper [25] where a complete description of surjective Lévy isometries of the non-linear space $\mathcal{D}(\mathbb{R})$ of all cumulative distribution functions was given. If $F, G \in \mathcal{D}(\mathbb{R})$, then their *Lévy distance* is defined by the following formula:

$$L(F, G) := \inf \left\{ \varepsilon > 0 \,\middle|\, \forall\, t \in \mathbb{R} \colon F(t - \varepsilon) - \varepsilon \le G(t) \le F(t + \varepsilon) + \varepsilon \right\}.$$

The importance of this metric lies in the fact that it metrises the topology of weak convergence on $\mathcal{D}(\mathbb{R})$. Molnár's result reads as follows (see [25, Theorem 1]): *let* $\Phi \colon \mathcal{D}(\mathbb{R}) \to \mathcal{D}(\mathbb{R})$ *be a surjective Lévy isometry,* i.e., *a bijective map satisfying*

$$L(F, G) = L(\Phi(F), \Phi(G)) \qquad (\forall\, F, G \in \mathcal{D}(\mathbb{R})).$$

Then there is a constant $c \in \mathbb{R}$ *such that* Φ *is one of the following two forms:*

$$\Phi(F)(t) = F(t + c) \qquad (\forall\, t \in \mathbb{R},\, F \in \mathcal{D}(\mathbb{R})),$$

or

$$\Phi(F)(t) = 1 - \lim_{s \to t-} F(-s + c) \qquad (\forall\, t \in \mathbb{R},\, F \in \mathcal{D}(\mathbb{R})).$$

In other words, every surjective Lévy isometry is induced by an isometry of $\mathbb{R}$ with respect to its usual norm (or equivalently, by a composition of a translation and a reflection on $\mathbb{R}$).

The investigation of surjective isometries on spaces of Borel probability measures was continued for example in [15, 26] for the Kolmogorov-Smirnov distance which is important in the Kolmogorov-Smirnov statistic and test, and in [7, 8, 20] with respect to the Wasserstein (or Kantorovich) metric which metrises the weak convergence.

Let (X, d) be a complete and separable metric space. We will denote the σ-*algebra of Borel sets on* X by $\mathscr{B}_X$ and the *set of all Borel probability measures* by $\mathcal{P}_X$. The Lévy distance gives a metrisation of weak convergence on $\mathcal{D}(\mathbb{R})$, or equivalently on $\mathcal{P}_{\mathbb{R}}$. In 1956 Prokhorov managed to metrise the weak convergence of $\mathcal{P}_X$ for general complete and separable metric spaces (X, d). The so-called *Lévy-Prokhorov distance* which was introduced by him in [30] is defined by

$$\pi(\mu, \nu) := \inf \left\{ \varepsilon > 0 \,\middle|\, \forall\, A \in \mathscr{B}_X \colon \mu(A) \le \nu(A^\varepsilon) + \varepsilon \right\}, \qquad (1.1)$$

where

$$A^\varepsilon := \bigcup_{x \in A} B_\varepsilon(x) \qquad \text{and} \qquad B_\varepsilon(x) := \{z \in X \mid d(x, z) < \varepsilon\}.$$

For the details and elementary properties see, *e.g.*, [19, page 27]. Let us point out that in the special case when $X = \mathbb{R}$ this metric differs from the original Lévy distance. Here arises the following very natural question:

What is the structure of onto isometries with respect to the Lévy-Prokhorov metric on $\mathcal{P}_X$ if X is a general separable real Banach space?

This paper is devoted to give an answer to this question. Namely, we will prove that every such transformation is induced by an affine isometry of the underlying space X.

There are some particularly important cases in our investigation which we emphasise now. Namely, since stochastic processes depending upon a continuous parameter are basically probability measures on certain subspaces of the space of all functions of a real variable (see, *e.g.*, [1, 14]), one particularly interesting case is when the underlying Banach space is $C([0, 1])$, *i.e.*, the space of all continuous real-valued functions on $[0, 1]$ endowed with the uniform norm $\| \cdot \|_\infty$. For details see [30, Chapter 2] or [9, Chapter 2]. Further two important cases are when X is a Euclidean space because of multivariate random variables, and when X is an infinite dimensional, separable real Hilbert space because of the theory of random elements in Hilbert spaces.

2. The setting and the statment of our main result

In this section we state the main result of the paper and collect some definitions and well-known facts about weak convergence of Borel probability measures. For more details the reader is referred to the textbooks of Billingsley [9], Huber [19] and Parthasarathy [29].

Let (X, d) be a complete, separable metric space and denote by $C_b(X; \mathbb{R})$ the Banach space of all real-valued bounded continuous functions. Recall that $\mathscr{B}_X$ is the smallest σ-algebra with respect to each $f \in C_b(X; \mathbb{R})$ is measurable. We say that an element of $\mathcal{P}_X$ is a *Dirac measure* if it is concentrated on one point, and for an $x \in X$ the symbol δ_x stands for the corresponding Dirac measure. The set of all Dirac measures on X is denoted by Δ_X. The collection of all *finitely supported measures* is

$$\mathcal{F}_X := \left\{ \sum_{i \in I} \lambda_i \delta_{x_i} \,\middle|\, \#I < \aleph_0, \ \sum_{i \in I} \lambda_i = 1, \ \lambda_i > 0, \ x_i \in X \ (\forall\, i \in I) \right\},$$

which is actually the convex hull of Δ_X. The *support* (or spectrum) of $\mu \in \mathcal{P}_X$ is the smallest d-closed set S_μ that satisfies $\mu(S_\mu) = 1$. Moreover, it is not hard to

verify the following equation:

$$S_\mu = \left\{ x \in X \mid \forall \, r > 0 \colon \mu(B_r(x)) > 0 \right\}.$$

The *closure* of a set $H \subseteq X$ will be denoted by $\overline{H}$.

We say that a sequence of measures $\{\mu_n\}_{n=1}^\infty \subset \mathcal{P}_X$ *converges weakly* to a $\mu \in \mathcal{P}_X$ if we have

$$\int f \, d\mu_n \to \int f \, d\mu \qquad (\forall \, f \in C_b(X; \mathbb{R})).$$

This type of convergence is metrised by the Lévy-Prokhorov metric given by (1.1). A map $\varphi \colon \mathcal{P}_X \to \mathcal{P}_X$ is called a π-*isometry* on $\mathcal{P}_X$ if

$$\pi(\mu, \nu) = \pi(\varphi(\mu), \varphi(\nu)) \qquad (\forall \, \mu, \nu \in \mathcal{P}_X)$$

is satisfied.

Now, we are in the position to state the main result of this paper.

Main Theorem. *Let $(X, \|\cdot\|)$ be a separable real Banach space and $\varphi \colon \mathcal{P}_X \to \mathcal{P}_X$ be a surjective π-isometry. Then there exists a surjective affine isometry $\psi \colon X \to X$ which induces φ,* i.e., *we have*

$$(\varphi(\mu))(A) = \mu\left(\psi^{-1}[A]\right) \qquad (\forall \, A \in \mathscr{B}_X), \tag{2.1}$$

where $\psi^{-1}[A]$ denotes the inverse-image set $\{\psi^{-1}(a) \mid a \in A\}$.

The converse of the above statement is trivial, namely, every transformation of the form (2.1) is obviously an onto π-isometry. Note that our theorem can be re-phrased in terms of push-forward measures. Namely, the action of φ is just the push-forward with respect to the isometry $\psi \colon X \to X$.

As we already mentioned, the Lévy-Prokhorov metric on $\mathcal{P}_\mathbb{R}$ differs from the Lévy distance on $\mathcal{P}_\mathbb{R}$. Therefore, in the special case when $X = \mathbb{R}$, our hypotheses are different from those given in [25, Theorem 1], although the conclusion is the same.

Our proof is given in the next section where we will have four major steps. This will be followed by some remarks in the final section, where we will also point out that our Main Theorem still holds if we replace $\mathcal{P}_X$ with an arbitrary weakly dense subset $\mathcal{S}$.

3. Proof

The proof is divided into four major steps. First, we will explore the action of φ on Δ_X. Then for finitely supported measures μ we will investigate the behaviour of its image $\varphi(\mu)$ near to the vertices of the convex hull of S_μ. This will be followed by

providing a procedure which will allow us to obtain important information about the "rest" of $\varphi(\mu)$. Finally, we close this section with the proof of the Main Theorem. Note that although our main result deals with Borel probability measures on separable Banach spaces, we state and prove some results in the context of complete and separable metric spaces.

3.1. First major step: the action on Dirac measures

Here we will investigate properties of the restricted map $\varphi|_{\mathcal{D}_X}$. Namely, we will prove that φ maps Δ_X onto Δ_X, furthermore, there is a surjective affine isometry of X which induces this restriction. In order to do this first, we formulate the metric phrase "distance one" by means of the supports of measures.

Proposition 3.1. *Let (X, d) be a complete, separable metric space and $\mu, \nu \in \mathcal{P}_X$. Then the following statements are equivalent:*

(i) $\pi(\mu, \nu) = 1$;
(ii) $\underline{d}(S_\mu, S_\nu) := \inf \{d(x, y) \mid x \in S_\mu, \ y \in S_\nu\} \geq 1$;
(iii) $S_\nu \cap S_\mu^1 = \emptyset$;
(iv) $S_\mu \cap S_\nu^1 = \emptyset$.

Proof. Observe that (ii) implies the following inequality for every $0 < \varepsilon < 1$:

$$1 = \mu(S_\mu) > \nu(S_\mu^\varepsilon) + \varepsilon = \varepsilon.$$

Consequently we have $\pi(\mu, \nu) \geq 1$. But on the other hand, $\pi(\mu, \nu) \leq 1$ holds for all $\mu, \nu \in \mathcal{P}_X$, and therefore the (ii)$\Rightarrow$(i) part is complete.

To prove (i)$\Rightarrow$(ii) assume that $\varrho := \underline{d}(S_\mu, S_\nu) < 1$. In this case one can fix two points $x^* \in S_\mu$ and $y^* \in S_\nu$, and a positive number $r > 0$ which satisfy both

$$\varrho \leq d(x^*, y^*) =: \varrho' < 1 \quad \text{and} \quad \varrho' + 2r < 1.$$

We also set $t := \min \{\mu(B_r(x^*)), \nu(B_r(y^*))\}$ which is clearly positive by the very definition of the support. We will show that $\hat{\varepsilon} := \max\{1 - t, \varrho' + 2r\} < 1$ is a suitable choice to guarantee

$$\mu(A) \leq \nu(A^{\hat{\varepsilon}}) + \hat{\varepsilon} \qquad (\forall \ A \in \mathscr{B}_X).$$

Indeed, if $A \in \mathscr{B}_X$ satisfies $\mu(A) \leq 1 - t$, then

$$\mu(A) \leq 1 - t \leq \nu(A^{1-t}) + 1 - t \leq \nu(A^{\hat{\varepsilon}}) + \hat{\varepsilon}.$$

On the other hand, if $\mu(A) > 1 - t$, then we observe that $\mu(A \cap B_r(x^*)) > 0$, and consequently $A \cap B_r(x^*)$ is not empty. Let us fix a point $z \in A \cap B_r(x^*)$. Using the triangle inequality we infer $d(y^*, z) \leq d(y^*, x^*) + d(x^*, z) < \varrho' + r$ and $B_r(y^*) \subseteq B_{\varrho'+2r}(z) \subseteq A^{\hat{\varepsilon}}$. Therefore we conclude

$$\mu(A) \leq 1 \leq t + \hat{\varepsilon} \leq \nu(B_r(y^*)) + \hat{\varepsilon} \leq \nu(A^{\hat{\varepsilon}}) + \hat{\varepsilon},$$

which implies $\pi(\mu, \nu) \leq \hat{\varepsilon} < 1$.

The equivalence of (ii), (iii) and (iv) follows from the definitions. $\qquad\square$

Next, let us define the *unit distance set* of a set of measures $\mathscr{A} \subseteq \mathcal{P}_X$ by

$$\mathscr{A}^{\mathrm{u}} = \left\{ \nu \in \mathcal{P}_X \,\middle|\, \forall\, \mu \in \mathscr{A} \colon \pi(\mu, \nu) = 1 \right\}.$$

(Remark that by definition we have $\emptyset^{\mathrm{u}} = \mathcal{P}_X$.) The following statement gives a metric characterisation of Dirac measures when X is a separable real Banach space. We point out that similar results were also crucial ideas in [15,25,26].

Proposition 3.2. *Let (X, d) be a complete, separable metric space and $\mu \in \mathcal{P}_X$ be an arbitrary Borel probability measure on it. Then the following three statements are equivalent:*

  (i) $(\{\mu\}^{\mathrm{u}})^{\mathrm{u}} = \{\mu\}$;
 (ii) *there exists an $x \in X$ such that*

  (ii/a) $\mu = \delta_x$;
  (ii/b) $B_1(y) \subseteq B_1(x)$ *implies $x = y$ for every $y \in X$;*

(iii) $\#(\{\mu\}^{\mathrm{u}})^{\mathrm{u}} = 1$.

Proof. First, let us characterise the elements of $(\{\mu\}^{\mathrm{u}})^{\mathrm{u}}$. It follows from Proposition 3.1 that

$$\{\mu\}^{\mathrm{u}} = \left\{ \nu \in \mathcal{P}_X \,\middle|\, S_\nu^1 \cap S_\mu = \emptyset \right\} = \left\{ \nu \in \mathcal{P}_X \,\middle|\, S_\nu \cap S_\mu^1 = \emptyset \right\}.$$

Applying this observation twice, we easily see that

$$(\{\mu\}^{\mathrm{u}})^{\mathrm{u}} = \bigcap_{\nu \in \{\mu\}^{\mathrm{u}}} \{\nu\}^{\mathrm{u}} = \bigcap_{\nu \in \mathcal{P}_X, S_\nu \cap S_\mu^1 = \emptyset} \left\{ \vartheta \in \mathcal{P}_X \,\middle|\, S_\nu \cap S_\vartheta^1 = \emptyset \right\}, \qquad (3.1)$$

and therefore we obtain the following equivalence:

$$\vartheta \in (\{\mu\}^{\mathrm{u}})^{\mathrm{u}} \qquad \Longleftrightarrow \qquad S_\vartheta^1 \subseteq S_\mu^1. \qquad (3.2)$$

(Note that if $\{\mu\}^{\mathrm{u}} = \emptyset$, then $\bigcap_{\nu \in \emptyset} \{\nu\}^{\mathrm{u}} = \mathcal{P}_X$ in (3.1) by definition.)
    Now, since $\mu \in (\{\mu\}^{\mathrm{u}})^{\mathrm{u}}$ always holds, the equivalence of (i) and (iii) is apparent. We continue with proving (i)$\Rightarrow$(ii). Observe that (3.2) implies

$$\left\{ \delta_z \,\middle|\, z \in S_\mu \right\} \subseteq (\{\mu\}^{\mathrm{u}})^{\mathrm{u}},$$

thus (ii/a) follows. On the other hand, if any $y \in X$ satisfies

$$S_{\delta_y}^1 = B_1(y) \subseteq B_1(x) = S_{\delta_x}^1,$$

then again by (3.2) we infer $x = y$.
    Finally, we show (ii)$\Rightarrow$(i). Assume that $\vartheta \in (\{\mu\}^{\mathrm{u}})^{\mathrm{u}}$. By (ii/b) we get that $S_\vartheta^1 \subseteq S_{\delta_x}^1$ holds if and only if $S_\vartheta \subseteq \{x\}$, which implies (i). $\qquad \square$

Remark 3.3. Note that if the diameter of the metric space X is less than 1, *i.e.*, there exists an $0 < r < 1$ such that $d(x, y) \le r$ ($\forall\, x, y \in X$), then $\pi(\mu, \nu) \le r$ holds for every $\mu, \nu \in \mathcal{P}_X$. In particular, $\{\mu\}^{\text{u}} = \emptyset$ and thus $(\{\mu\}^{\text{u}})^{\text{u}} = \mathcal{P}_X$ for every $\mu \in \mathcal{P}_X$.

The following lemma describes the action of φ on Dirac measures.

Lemma 3.4. *Let $(X, \|\cdot\|)$ be a separable real Banach space, and let $\varphi \colon \mathcal{P}_X \to \mathcal{P}_X$ be a surjective π isometry. Then there exists a surjective affine isometry $\psi \colon X \to X$ such that*

$$\varphi(\delta_x) = \delta_{\psi(x)} \qquad (\forall\, x \in X). \tag{3.3}$$

Proof. Since φ is a bijective isometry, we have

$$\varphi\big((\{\mu\}^{\text{u}})^{\text{u}}\big) = \big(\{\varphi(\mu)\}^{\text{u}}\big)^{\text{u}} \qquad (\forall\, \mu \in \mathcal{P}_X).$$

Thus an easy application of the previous proposition yields $\varphi(\Delta_X) = \Delta_X$. This also means that there exists a bijective map $\psi \colon X \to X$ which induces the restriction $\varphi|_{\Delta_X}$, *i.e.*,

$$\varphi(\delta_x) := \delta_{\psi(x)} \qquad (\forall\, x \in X). \tag{3.4}$$

We will show that ψ is an isometry. Observe that

$$\pi(\delta_{x_1}, \delta_{x_2}) = \min\{1, \|x_1 - x_2\|\} \qquad (\forall\, x_1, x_2 \in X).$$

Therefore for all $\alpha \in (0, 1)$ we have

$$\|\psi(x_1) - \psi(x_2)\| = \alpha \iff \|x_1 - x_2\| = \alpha \qquad (\forall\, x_1, x_2 \in X). \tag{3.5}$$

If X is one-dimensional, then it is rather easy to see that (3.5) implies the isometriness of ψ. Now, assume that $\dim X \ge 2$. After suitable renorming (*i.e.*, considering the norm $\||\cdot\|| := \frac{1}{\alpha}\|\cdot\|$), from a result of T. M. Rassias and P. Šemrl [31, Theorem 1] we conclude that

$$\|\psi(x) - \psi(y)\| = n\alpha \iff \|x - y\| = n\alpha \qquad (\forall\, \alpha \in (0, 1), n \in \mathbb{N}),$$

and therefore ψ is indeed an isometry. Finally, by the famous Mazur-Ulam theorem we obtain that ψ is affine, which completes the proof. $\qquad\square$

We remark that the last step of the proof (using the Rassias-Šemrl theorem) can be also done by the extension theorem of Mankiewicz [23].

In light of the above lemma, from now on we may and do assume without loss of generality that φ acts identically on Δ_X, *i.e.*,

$$\varphi(\delta_x) = \delta_x \qquad (\forall\, x \in X), \tag{3.6}$$

and our aim will be to show that φ acts identically on the whole of $\mathcal{P}_X$. After we do so, to obtain the result of our Main Theorem for general surjective π-isometries will

be straightforward. Namely, if (3.3) is fulfilled, then we can consider the following modified transformation:

$$\varphi_\psi \colon \mathcal{P}_X \to \mathcal{P}_X, \quad \big(\varphi_\psi(\mu)\big)(A) := (\varphi(\mu))\,(\psi[A]) \qquad (\forall\,\mu \in \mathcal{P}_X,\, A \in \mathscr{B}). \quad (3.7)$$

By our assumption, φ_ψ fixes every element of Δ_X and thus also of $\mathcal{P}_X$, which implies (2.1).

Next, let us define the following continuous function for each $\mu \in \mathcal{P}_X$:

$$W_\mu \colon X \to [0, 1], \quad W_\mu(x) := \pi(\delta_x, \mu),$$

which will be called the *witness function* of μ. The main advantage of the assumption (3.6) is that the witness function becomes φ-invariant, *i.e.*,

$$W_\mu(x) = \pi(\delta_x, \mu) = \pi(\varphi(\delta_x), \varphi(\mu)) = \pi(\delta_x, \varphi(\mu)) = W_{\varphi(\mu)}(x) \quad (\forall\, x \in X). \quad (3.8)$$

It is natural to expect that the shape of the witness function carries some information about the measure. The last three major steps of the proof will be devoted to explore this for the φ-images of finitely supported measures in the setting of separable Banach spaces. However, as demonstrated by the next example, the witness function usually does not distinguish measures in general complete and separable metric spaces.

Example 3.5. Consider the complete and separable metric space (X, d) with

$$X := \{x_1, x_2, x_3\} \qquad \text{and} \qquad d(x_i, x_j) := \begin{cases} \frac{1}{3} & \text{if } i \neq j \\ 0 & \text{if } i = j. \end{cases}$$

Let $\mu := \frac{1}{2}\delta_{x_1} + \frac{1}{2}\delta_{x_2}$ and $\nu := \frac{1}{2}\delta_{x_2} + \frac{1}{2}\delta_{x_3}$. An easy calculation shows that we have $\pi(\delta_x, \mu) = \frac{1}{3} = \pi(\delta_x, \nu)$ for all $x \in X$ and hence $W_\mu \equiv W_\nu$.

3.2. Second major step: isolated atoms on the vertices of the convex hull of the support

Here we will prove that if μ is a finitely supported measure, and $\hat{x}$ is a vertex of the convex hull of S_μ, then $\hat{x}$ is an isolated atom of $\varphi(\mu)$ and

$$\mu(\{\hat{x}\}) = (\varphi(\mu))(\{\hat{x}\}).$$

We begin with a technical statement, which will be very useful in the sequel.

Proposition 3.6. *Let X be a separable real Banach space, and suppose that μ is a finitely supported measure. Then for every $\nu \in \mathcal{P}_X$, $\nu \neq \mu$ we have*

$$\pi(\mu, \nu) = \min\left\{\varepsilon > 0 \mid \forall\, A \subseteq S_\mu \colon \mu(A) \le \nu(\overline{A^\varepsilon}) + \varepsilon\right\}. \qquad (3.9)$$

Proof. Since one deals with finitely many balls, it is easy to see that the number ε obtained as the infimum without closures gives the minimum with closures. □

The following proposition plays a key role in the proof. But before stating it we introduce some notations. The convex hull of two points x and y will be denoted by $[x, y]$, and the symbol $]x, y[$ will stand for the set $[x, y] \setminus \{x, y\}$. If f is a real valued function on X and $c \in \mathbb{R}$, then the sets $\{x \in X \mid f(x) < c\}, \{x \in X \mid f(x) = c\}$, and $\{x \in X \mid f(x) \leq c\}$ will be denoted by $\{f < c\}, \{f = c\}$, and $\{f \leq c\}$, respectively.

Proposition 3.7. *Let $(X, \|\cdot\|)$ be a separable real Banach space, $\mu \in \mathcal{F}_X \setminus \Delta_X$ and K be the convex hull of S_μ. Assume that $\hat{x}$ is a vertex of K (which is a polytope) and set $\hat{\lambda} := \mu(\{\hat{x}\})$ (for which we obviously have $0 < \hat{\lambda} < 1$). Then for every $\vartheta \in \mathcal{P}_X$ with $S_\vartheta \subseteq K$ the following two conditions are equivalent:*

(i) *$\vartheta = \hat{\lambda}\delta_{\hat{x}} + (1 - \hat{\lambda})\tilde{\vartheta}$ where $\tilde{\vartheta} \in \mathcal{P}_X$ with $S_{\tilde{\vartheta}} \subseteq K \setminus B_r(\hat{x})$ for some $r > 0$;*

(ii) *there exist a number $0 < \rho \leq 1 - \hat{\lambda}$ and a half-line $\mathfrak{e}$ starting from $\hat{x}$ such that the restriction $W_\vartheta|_{\mathfrak{e}}$ is of the following form:*

$$W_\vartheta|_{\mathfrak{e}}(x) = \begin{cases} 1 & if \quad \|x - \hat{x}\| \geq 1 \\ \|x - \hat{x}\| & if \quad 1 - \hat{\lambda} < \|x - \hat{x}\| < 1 \\ 1 - \hat{\lambda} & if \quad 1 - \hat{\lambda} - \rho \leq \|x - \hat{x}\| \leq 1 - \hat{\lambda}. \end{cases} \tag{3.10}$$

Moreover, $S_{\varphi(\mu)} \subseteq K$ and $\hat{x}$ is an isolated atom of $\varphi(\mu)$ with $(\varphi(\mu))(\{\hat{x}\}) = \hat{\lambda}$.
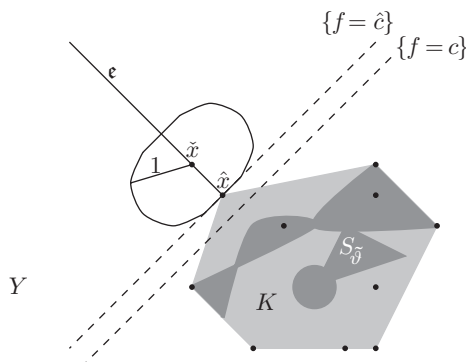

Figure 3.1. Illustration for Proposition 3.7 on the finite dimensional subspace Y. The support S_μ consists of the set of black points in K.

Proof. First, we construct a half-line $\mathfrak{e}$ which starts from $\hat{x}$ and satisfies

$$\underline{d}(\{x\}, K) = \|x - \hat{x}\| < \|x - k\| \qquad (\forall \, x \in \mathfrak{e}, k \in K \setminus \{\hat{x}\}). \tag{3.11}$$

Being the convex hull of a finite set, each vertex of K is strongly exposed, *i.e.*, there exists a continuous linear functional $f \in X^*$ with

$$\hat{c} := \max\{f(y) \mid y \in K\} = f(\hat{x}) \quad \text{and} \quad K \setminus \{\hat{x}\} \subset \{f < \hat{c}\}.$$

Let Y be the subspace generated by K. We fix an $\check{x} \in Y$ such that $\hat{x} \in \overline{B_1(\check{x})}$ and

$$\overline{B_1(\check{x})} \cap \{f \leq \hat{c}\} \cap Y \subseteq \{f = \hat{c}\}.$$

Note that as Y is finite dimensional, the existence of such an $\check{x}$ is guaranteed. Now, we define $\mathfrak{e}$ to be the half-line starting from $\hat{x}$ and going through $\check{x}$. It is straightforward that $\mathfrak{e}$ fulfils (3.11).

Next, we consider an arbitrary $\vartheta \in \mathcal{P}_X$ which satisfies (i). It is clear form the compactness of S_μ and S_ϑ, and the isolatedness of the point $\hat{x}$ in both S_μ and S_ϑ, that there is a number $c < \hat{c}$ such that

$$(S_\mu \cup S_\vartheta) \setminus \{\hat{x}\} \subseteq \{f < c\} \tag{3.12}$$

holds. Consequently, for every $x \in \mathfrak{e}$ we have

$$\underline{d}(\{x\}, K) = \|x - \hat{x}\| > \underline{d}(\{x\}, \{f \leq c\} \cap Y) > \underline{d}(\{x\}, (S_\mu \cup S_\vartheta) \setminus \{\hat{x}\}).$$

Therefore, if $x \in \mathfrak{e}$ and $\alpha := \|x - \hat{x}\| > 0$, then there exists a $y \in]x, \hat{x}[\subseteq \mathfrak{e}$ which satisfies the following equations:

$$\hat{\lambda} = \vartheta(\{\hat{x}\}) = \vartheta(\overline{B_\alpha(x)}) = \vartheta(\overline{B_\alpha(z)}) \qquad (\forall\, z \in [x, y]), \tag{3.13}$$

and

$$\hat{\lambda} = \mu(\{\hat{x}\}) = \mu(\overline{B_\alpha(x)}) = \mu(\overline{B_\alpha(z)}) \qquad (\forall\, z \in [x, y]). \tag{3.14}$$

In fact, y can be chosen to be any point on $]x, \hat{x}[$ such that

$$\|x - y\| \leq \underline{d}(\{f = \hat{c}\}, \{f = c\}).$$

We proceed to prove the equivalence of (i) and (ii). Trivially, both (i) and (ii) implies that $\vartheta \notin \Delta_X$, and therefore from now on we may and do assume that ϑ is not a Dirac measure. Recall that according to Proposition 3.6 we have

$$W_\vartheta(x) = \min\left\{\varepsilon > 0 \mid 1 \leq \vartheta(\overline{B_\varepsilon(x)}) + \varepsilon\right\} \qquad (\forall\, x \in X). \tag{3.15}$$

If (i) holds, then by combining (3.15) with (3.11) and (3.13) we obtain that $W_\vartheta|_\mathfrak{e}$ is of the form (3.10) with $\rho := \min\left(1 - \hat{\lambda}, \underline{d}(\{f = \hat{c}\}, \{f = c\})\right)$.

Conversely, we suppose that $\vartheta \in \mathcal{P}_X$, $S_\vartheta \subseteq K$ and ϑ satisfies (ii). Let x_1 and x_2 be the points on $\mathfrak{e}$ which satisfy $\|x_1 - \hat{x}\| = 1 - \hat{\lambda}$ and $\|x_2 - \hat{x}\| = 1 - \hat{\lambda} - \rho$. By (3.10) we have $W_\vartheta(x_1) = W_\vartheta(x_2) = 1 - \hat{\lambda}$. Therefore on one hand, we obtain

$$1 \leq \vartheta(\overline{B_{1-\hat{\lambda}}(x_1)}) + 1 - \hat{\lambda} = \vartheta(\{\hat{x}\}) + 1 - \hat{\lambda},$$

from which $\hat{\lambda} \leq \vartheta(\{\hat{x}\})$ follows. On the other hand,

$$1 > \vartheta(\overline{B_{1-\hat{\lambda}-\delta}(x_2)}) + 1 - \hat{\lambda} - \delta \geq \vartheta(\{\hat{x}\}) + 1 - \hat{\lambda} - \delta \quad (\forall\, 0 < \delta < \rho)$$

is satisfied. Hence we infer $\vartheta(\{\hat{x}\}) < \hat{\lambda} + \delta$ for every $\delta > 0$, and thus trivially $\vartheta(\{\hat{x}\}) = \hat{\lambda}$ holds. But we also observe the following:

$$1 > \vartheta\left(\overline{B_{1-\hat{\lambda}-\delta}(x_2)}\right) + 1 - \hat{\lambda} - \delta \geq \vartheta\left(\overline{B_{1-\hat{\lambda}-\frac{\rho}{2}}(x_2)}\right) + 1 - \hat{\lambda} - \delta \quad \left(\forall\, 0 < \delta < \frac{\rho}{2}\right)$$

which implies

$$\hat{\lambda} \geq \vartheta\left(\overline{B_{1-\hat{\lambda}-\frac{\rho}{2}}(x_2)}\right) \geq \vartheta\left(\{\hat{x}\}\right) = \hat{\lambda},$$

whence we conclude that $\hat{x}$ is indeed an isolated atom of ϑ.

For the last statement first, by Proposition 3.1 we infer $\pi(\delta_x, \mu) = 1$ for every $x \notin K^1$. The φ-invariance of the witness function gives

$$1 = \pi(\delta_x, \varphi(\mu)) \qquad \left(x \in X \setminus K^1\right),$$

and hence, again by Proposition 3.1, we conclude

$$S_{\varphi(\mu)} \cap B_1(x) = \emptyset \qquad \left(x \in X \setminus K^1\right).$$

Consequently, we obtain

$$S_{\varphi(\mu)} \subseteq X \setminus \left(X \setminus K^1\right)^1 \subseteq K.$$

Finally, an application of the equivalence of (i) and (ii) gives the rest. $\qquad \square$

We have the following consequence.

Corollary 3.8. *Let* $(X, \|\cdot\|)$ *be a separable real Banach space. If* $\mu \in \mathcal{F}_X$ *such that* $\#S_\mu \leq 2$, *then* $\varphi(\mu) = \mu$. *Moreover, if* $\nu \in \mathcal{P}_X$ *with* $W_\mu \equiv W_\nu$, *then* μ *and* ν *coincide.*

According to the above results, now we know that φ fixes every measure which has at most two points in its support. Although we are expecting the same for all $\mu \in \mathcal{F}_X$, right now we only have some information about the behaviour of $\varphi(\mu)$ near to the vertices of the convex hull of its support.

3.3. Third major step: the story beyond vertices

Here we show a procedure how the behaviour of $\varphi(\mu)$ can be completely explored in case when μ is a finitely supported measure. In order to do so, we need to introduce some technical notations. Let $s > 0$ be a positive parameter and define the *s-Lévy-Prokhorov distance* $\pi_s \colon \mathcal{P}_X \times \mathcal{P}_X \to [0, 1]$ by the following formula:

$$\pi_s(\mu, \nu) := \inf\left\{\varepsilon > 0 \mid \forall\, A \in \mathscr{B}_X : s \cdot \mu(A) \le s \cdot \nu(A^\varepsilon) + \varepsilon\right\} \quad (\forall\, \mu, \nu \in \mathcal{P}_X). \quad (3.16)$$

Note that although we do not know at this point whether π_s defines a metric on $\mathcal{P}_X$, we will see this later. The *s-witness function* (or modified witness function) of $\mu \in \mathcal{P}_X$ is defined by

$$W_{s,\mu} \colon X \to \mathbb{R}, \quad W_{s,\mu}(x) := \pi_s(\delta_x, \mu).$$

Obviously, if we set $s = 1$, then we get the original Lévy-Prokhorov metric and witness function.

In the next two lemmas we collect some properties of the s-Lévy-Prokhorov distance analogous to those provided in the previous major step.

Lemma 3.9. *Let* $(X, \|\cdot\|)$ *be a separable real Banach space and* $s > 0$. *Then* $(\mathcal{P}_X, \pi_s)$ *is a metric space. Furthermore, for every* $\mu \in \mathcal{F}_X$ *and* $\nu \in \mathcal{P}_X$ *with* $\nu \neq \mu$ *we have*

$$\pi_s(\mu, \nu) = \min\left\{\varepsilon > 0 \mid \forall\, A \subseteq S_\mu : s \cdot \mu(A) \le s \cdot \nu(\overline{A^\varepsilon}) + \varepsilon\right\}.$$

Proof. For the sake of clarity, let us use more detailed notations here. If $\|\cdot\|$ is a norm on X, then denote by $A^{\varepsilon, \|\cdot\|}$ and $\pi_{s, \|\cdot\|}$ the open ε-neighborhood of A and the s-Lévy-Prokhorov metric with respect to $\|\cdot\|$, respectively. Observe that the Borel σ-algebras of $(X, \|\cdot\|)$ and $\left(X, \frac{1}{s}\|\cdot\|\right)$ coincide as the norms are equivalent. By an elementary computation we have $A^{s\delta, \|\cdot\|} = A^{\delta, \frac{1}{s}\|\cdot\|}$, which yields

$$
\begin{aligned}
\pi_{s, \|\cdot\|}(\mu, \nu) &= \inf\left\{\varepsilon > 0 \;\middle|\; \forall\, A \in \mathscr{B}_X : s \cdot \mu(A) \le s \cdot \nu\left(A^{\varepsilon, \|\cdot\|}\right) + \varepsilon\right\} \\
&= \inf\left\{s\delta > 0 \;\middle|\; \forall\, A \in \mathscr{B}_X : s \cdot \mu(A) \le s \cdot \nu\left(A^{s\delta, \|\cdot\|}\right) + s\delta\right\} \\
&= s \cdot \inf\left\{\delta > 0 \;\middle|\; \forall\, A \in \mathscr{B}_X : \mu(A) \le \nu\left(A^{s\delta, \|\cdot\|}\right) + \delta\right\} \quad (3.17) \\
&= s \cdot \inf\left\{\delta > 0 \;\middle|\; \forall\, A \in \mathscr{B}_X : \mu(A) \le \nu\left(A^{\delta, \frac{1}{s}\|\cdot\|}\right) + \delta\right\} \\
&= s \cdot \pi_{\frac{1}{s}\|\cdot\|}(\mu, \nu)
\end{aligned}
$$

for every $\mu, \nu \in \mathcal{P}_X$. In particular, π_s is a metric on $\mathcal{P}_X$, and using the formula (3.9) completes the proof. $\qquad\square$

We omit the proof of the following lemma as it is a straightforward consequence of (3.17).

Lemma 3.10. *Let $(X, \| \cdot \|)$ be a separable real Banach space, $\mu \in \mathcal{F}_X \setminus \Delta_X$ and $s > 0$. Let us denote the convex hull of S_μ by K, and assume that $\hat{x}$ is a vertex of K. Set $\hat{\lambda} := \mu(\{\hat{x}\}) \in (0, 1)$. Then for every $\vartheta \in \mathcal{P}_X$ with $S_\vartheta \subseteq K$ the following two conditions are equivalent:*

(i) *$\vartheta = \hat{\lambda}\delta_{\hat{x}} + (1 - \hat{\lambda})\widetilde{\vartheta}$ where $\widetilde{\vartheta} \in \mathcal{P}_X$ with $S_{\widetilde{\vartheta}} \subseteq K \setminus B_r(\hat{x})$ for some $r > 0$;*

(ii) *there exist a number $0 < \rho \le s(1 - \hat{\lambda})$ and a half-line $\mathfrak{e}$ starting from $\hat{x}$ such that $W_{s,\vartheta}|_{\mathfrak{e}}$ has the following form:*

$$W_{s,\vartheta}|_{\mathfrak{e}}(x) = \begin{cases} s & \text{if} \quad \|x - \hat{x}\| \ge s \\ \|x - \hat{x}\| & \text{if} \quad s(1 - \hat{\lambda}) < \|x - \hat{x}\| < s \\ s(1 - \hat{\lambda}) & \text{if} \quad s(1 - \hat{\lambda}) - \rho \le \|x - \hat{x}\| \le s(1 - \hat{\lambda}). \end{cases}$$

As a consequence we have that if $\mu \in \mathcal{F}_X$, $\#S_\mu \le 2$, $\nu \in \mathcal{P}_X$ and $W_{s,\mu} \equiv W_{s,\nu}$, then $\mu = \nu$.

Next, let us suppose for a moment that $m \in \mathbb{N}$ pieces of atoms of $\vartheta \in \mathcal{P}_X$ have been already detected. (For instance by Lemma 3.7, if $\vartheta = \varphi(\mu)$ with $\mu \in \mathcal{F}_X$ then the atoms of ϑ in the vertices of the convex hull of S_ϑ can be detected.) Our aim with the forthcoming lemma is to describe a modified witness function of the remaining part of ϑ in terms of the (original) Lévy-Prokhorov distances between ϑ and some measures which are supported on at most $m + 1$ points. This will be later utilised in order to explore the action of φ on $\mathcal{F}_X$.

Lemma 3.11. *Let $(X, \| \cdot \|)$ be a separable real Banach space and $\vartheta \in \mathcal{P}_X$. Let $x \in X$ and $\{y_{j,l} \mid 1 \le j \le k, 1 \le l \le d_j\} \subset X$ be some pairwise different points such that*

$$\rho_j := \|x - y_{j,1}\| = \|x - y_{j,l}\| \qquad (\forall\, 1 \le l \le d_j)$$

holds for every $1 \le j \le k$,

$$\rho_j > \rho_{j+1} > 0 \qquad (\forall\, 1 \le j \le k - 1),$$

and

$$w_{j,l} := \vartheta(\{y_{j,l}\}) > 0 \qquad (\forall\, 1 \le j \le k, 1 \le l \le d_j).$$

We also set

$$w_j := \sum_{l=1}^{d_j} w_{j,l} = \vartheta(\{y_{j,1}, \dots, y_{j,d_j}\}) \qquad (\forall\, 1 \le j \le k),$$

$$\widetilde{w} := 1 - \sum_{j=1}^{k} w_j$$

and

$$\eta_r := \sum_{j=1}^{r} \sum_{l=1}^{d_j} w_{j,l} \cdot \delta_{y_{j,l}} + \left(1 - \sum_{j=1}^{r} w_j\right) \cdot \delta_x \in \mathcal{F}_X \qquad (\forall\, 0 \le r \le k). \quad (3.18)$$

Furthermore, denote by $\widetilde{\vartheta} \in \mathcal{P}_X$ the measure which satisfies

$$\vartheta = \sum_{j=1}^{k} \sum_{l=1}^{d_j} w_{j,l} \cdot \delta_{y_{j,l}} + \widetilde{w} \cdot \widetilde{\vartheta}. \tag{3.19}$$

Then the $\widetilde{w}$-witness function of $\widetilde{\vartheta}$ can be expressed in terms of the Lévy-Prokhorov distances of ϑ and η_r's in the following way:

$$W_{\widetilde{w}, \widetilde{\vartheta}}(x) = \begin{cases} \pi(\delta_x, \vartheta) & \text{if } x \text{ is not } (P_1) \\ \pi(\eta_r, \vartheta) & \text{if } x \text{ is } (P_r) \text{ but not } (P_{r+1}) \text{ with some } 1 \le r < k \\ \pi(\eta_k, \vartheta) & \text{if } x \text{ is } (P_k) \end{cases} \tag{3.20}$$

where for every $1 \le r \le k$ the property (P_r) means
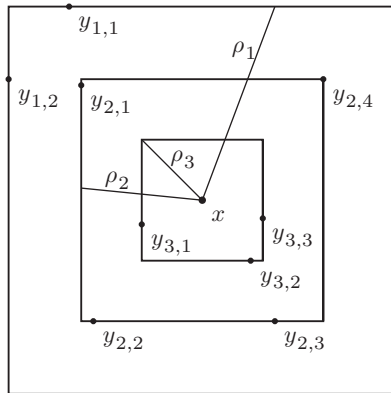
$$\pi(\eta_{r-1}, \vartheta) \le \rho_r. \tag{P_r}$$

Figure 3.2. An illustration when $X = \mathbb{R}^2$ with the ℓ^∞-norm.

Remark 3.12. It is extremely important to observe that the subscripts in the lemma above highly depend on the actual position of x. For instance on Figure 1 with that particular x we have $k = 3$. However, if x is moved slightly to the right, then k becomes 7. In particular, this changes (P_r) and therefore (3.20) as well.

Proof of Lemma 3.11. We split our proof into five parts.

Part 1. First, we prove that for each $1 \le r \le k$ and $0 < \varepsilon < \rho_r$ we have

$$\eta_r(A) \le \vartheta(\overline{A^\varepsilon}) + \varepsilon \quad \left(\forall\, A \subseteq S_{\eta_r}\right) \tag{3.21}$$

if and only if

$$\eta_r(\{x\}) \le \vartheta(\overline{B_\varepsilon(x)}) + \varepsilon \tag{3.22}$$

is satisfied. One direction is obvious. In order to see the reverse implication, observe that (3.21) holds trivially if $x \notin A$. On the other hand, if $x \in A$, then (3.22) yields (3.21) for this A by the following estimate:

$$\eta_r(A) = \eta_r(\{x\}) + \eta_r(A \setminus \{x\}) \leq \vartheta(\overline{B_\varepsilon(x)}) + \varepsilon + \vartheta(A \setminus \{x\}) \leq \vartheta(\overline{A^\varepsilon}) + \varepsilon.$$

Part 2. Here we show that the right-hand side of (3.20) is well defined. First, we observe that by Proposition 3.6 x is (P_r) if and only if

$$\eta_{r-1}(A) \leq \vartheta(\overline{A^{\rho_r}}) + \rho_r \quad (\forall\, A \subseteq S_{\eta_{r-1}}). \tag{3.23}$$

But by Part 1, this is equivalent to the following inequality:

$$1 - \sum_{j=1}^{r-1} w_j \leq \vartheta(\overline{B_{\rho_r}(x)}) + \rho_r. \tag{P_r'}$$

Next, let $2 \leq r \leq k$. In order to see the well-definedness, it is enough to show that if x is (P_r), then x is also (P_{r-1}). So assume that x is (P_r). Since $\rho_r < \rho_{r-1}$ and

$$w_{r-1} = \sum_{l=1}^{d_{r-1}} w_{r-1,l} = \vartheta(\{y_{r-1,1}, \ldots y_{r-1,d_{r-1}}\}) \leq \vartheta\left(\overline{B_{\rho_{r-1}}(x)} \setminus \overline{B_{\rho_r}(x)}\right),$$

we obtain

$$1 - \sum_{j=1}^{r-2} w_j \leq \vartheta(\overline{B_{\rho_{r-1}}(x)}) + \rho_{r-1}. \tag{P_{r-1}'}$$

Therefore x is indeed (P_{r-1}).

Part 3. Next, we verify (3.20) in case when x is not (P_1), i.e., $\pi(\delta_x, \vartheta) > \rho_1$. Observe that since

$$1 > \vartheta(\overline{B_{\rho_1}(x)}) + \rho_1 = \widetilde{w} \cdot \widetilde{\vartheta}(\overline{B_{\rho_1}(x)}) + \sum_{i=1}^{k} w_i + \rho_1,$$

we have

$$\widetilde{w} > \widetilde{w} \cdot \widetilde{\vartheta}(\overline{B_{\rho_1}(x)}) + \rho_1,$$

and thus

$$\pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta}) > \rho_1$$

follows. Using this fact we obtain

$$\pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta}) = \min\left\{\varepsilon > \rho_1 \,\middle|\, \widetilde{w} \leq \widetilde{w} \cdot \widetilde{\vartheta}(\overline{B_\varepsilon(x)}) + \varepsilon\right\}$$

$$= \min\left\{\varepsilon > \rho_1 \,\middle|\, 1 \leq \sum_{j=1}^{k} w_j + \widetilde{w} \cdot \widetilde{\vartheta}(\overline{B_\varepsilon(x)}) + \varepsilon\right\}$$

$$= \min\left\{\varepsilon > \rho_1 \,\middle|\, 1 \leq \vartheta(\overline{B_\varepsilon(x)}) + \varepsilon\right\}$$

$$= \min\left\{\varepsilon > 0 \,\middle|\, 1 \leq \vartheta(\overline{B_\varepsilon(x)}) + \varepsilon\right\} = \pi(\delta_x, \vartheta),$$

which completes this part.

Part 4. We proceed to show (3.20) in the case when x is (P_r) but not (P_{r+1}) with some $1 \leq r < k$. As in the previous part, first we estimate the value of $\pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta})$. According to the re-phrasing (P'_r) and the assumption, we have

$$1 - \sum_{j=1}^{r-1} w_j \leq \vartheta(\overline{B_{\rho_r}(x)}) + \rho_r = \sum_{i=r}^{k} w_i + \widetilde{w} \cdot \widetilde{\vartheta}(\overline{B_{\rho_r}(x)}) + \rho_r$$

and

$$1 - \sum_{j=1}^{r} w_j > \vartheta(\overline{B_{\rho_{r+1}}(x)}) + \rho_{r+1} = \sum_{i=r+1}^{k} w_i + \widetilde{w} \cdot \widetilde{\vartheta}(\overline{B_{\rho_{r+1}}(x)}) + \rho_{r+1}.$$

Observe that these inequalities are equivalent to

$$\widetilde{w} \leq \widetilde{w} \cdot \widetilde{\vartheta}(\overline{B_{\rho_r}(x)}) + \rho_r$$

and

$$\widetilde{w} > \widetilde{w} \cdot \widetilde{\vartheta}(\overline{B_{\rho_{r+1}}(x)}) + \rho_{r+1},$$

respectively. Thus we conclude that

$$\rho_{r+1} < \pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta}) \leq \rho_r.$$

In particular, $\widetilde{\vartheta}$ is different from δ_x, and we have

$$\pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta}) = \min \left\{ \rho_{r+1} < \varepsilon \leq \rho_r \,\middle|\, \widetilde{w} \leq \widetilde{w} \cdot \widetilde{\vartheta}(\overline{B_\varepsilon(x)}) + \varepsilon \right\}. \tag{3.24}$$

From now on we consider two cases: (a) when $\rho_{r+1} < \pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta}) < \rho_r$, and (b) when $\pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta}) = \rho_r$. Assume first that (a) is fulfilled. Then (3.24) becomes

$$\pi_{\widetilde{w}}\left(\delta_x, \widetilde{\vartheta}\right) = \min \left\{ \rho_{r+1} < \varepsilon < \rho_r \,\middle|\, 1 - \sum_{i=1}^{r} w_i \leq \sum_{i=r+1}^{k} w_i + \widetilde{w} \cdot \widetilde{\vartheta}(\overline{B_\varepsilon(x)}) + \varepsilon \right\}$$

$$= \min \left\{ \rho_{r+1} < \varepsilon < \rho_r \,\middle|\, 1 - \sum_{i=1}^{r} w_i \leq \vartheta(\overline{B_\varepsilon(x)}) + \varepsilon \right\}$$

$$= \min \left\{ \rho_{r+1} < \varepsilon < \rho_r \,\middle|\, \eta_r(\{x\}) \leq \vartheta(\overline{\{x\}^\varepsilon}) + \varepsilon \right\}$$

$$= \min \left\{ \rho_{r+1} < \varepsilon < \rho_r \,\middle|\, \forall\, A \subseteq S_{\eta_r} : \eta_r(A) \leq \vartheta(\overline{A^\varepsilon}) + \varepsilon \right\} \quad \text{(by Part 1)}$$

$$= \min \left\{ \varepsilon > 0 \,\middle|\, \forall\, A \subseteq S_{\eta_r} : \eta_r(A) \leq \vartheta(\overline{A^\varepsilon}) + \varepsilon \right\}$$

$$= \pi(\eta_r, \vartheta),$$

which is exactly the desired equation. Second, suppose that (b) is satisfied. Consequently, we have

$$1 - \sum_{j=1}^{k} w_j = \widetilde{w} > \widetilde{w} \cdot \widetilde{\vartheta}\,(\overline{B_\varepsilon(x)}) + \varepsilon \quad (\forall\, \rho_{r+1} < \varepsilon < \rho_r),$$

whence

$$\eta_r(\{x\}) = 1 - \sum_{i=1}^{r} w_i > \sum_{i=r+1}^{k} w_k + \widetilde{w} \cdot \widetilde{\vartheta}\,(\overline{B_\varepsilon(x)}) + \varepsilon = \vartheta\,(\overline{B_\varepsilon(x)}) + \varepsilon$$

follows for every $\rho_{r+1} < \varepsilon < \rho_r$. In particular, we get

$$\pi(\eta_r, \vartheta) \geq \rho_r.$$

Finally, we verify that the converse inequality holds as well. Suppose indirectly that there exists an $A \subseteq S_{\eta_r}$ such that

$$\eta_r(A) > \vartheta\,(\overline{A^{\rho_r}}) + \rho_r.$$

Clearly, $x \notin A$ contradicts the above inequality, thus $x \in A$ follows. Therefore we have

$$
\begin{aligned}
\eta_r(A) > \vartheta\,(\overline{A^{\rho_r}}) + \rho_r &\geq \vartheta\,(A \cup \overline{B_{\rho_r}(x)}) + \rho_r \\
&\geq \vartheta\,(A \setminus \{x, y_{r,1}, \ldots, y_{r,d_r}\}) + \vartheta\,(\overline{B_{\rho_r}(x)}) + \rho_r \\
&= \eta_r(A \setminus \{x, y_{r,1}, \ldots, y_{r,d_r}\}) + \vartheta\,(\overline{B_{\rho_r}(x)}) + \rho_r.
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
1 - \sum_{j=1}^{r-1} w_j &= \eta_r(\{x, y_{r,1}, \ldots, y_{r,d_r}\}) \\
&\geq \eta_r(A) - \eta_r(A \setminus \{x, y_{r,1}, \ldots, y_{r,d_r}\}) > \vartheta\,(\overline{B_{\rho_r}(x)}) + \rho_r,
\end{aligned}
\tag{3.25}
$$

which contradicts (P_r). This completes the present part.

Part 5. Finally, we prove (3.20) when x is (P_k), *i.e.*, $\pi(\eta_{k-1}, \vartheta) \leq \rho_k$. We have to show that $\pi(\eta_k, \vartheta) = \pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta})$. Because of the assumption, we have

$$1 - \sum_{j=1}^{k-1} w_j \leq \vartheta\,(\overline{B_{\rho_k}(x)}) + \rho_k = w_k + \widetilde{w} \cdot \widetilde{\vartheta}\,(\overline{B_{\rho_k}(x)}) + \rho_k,$$

which implies $\widetilde{w} \leq \widetilde{w} \cdot \widetilde{\vartheta}\,(\overline{B_{\rho_k}(x)}) + \rho_k$, and hence,

$$\pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta}) \leq \rho_k.$$

We consider three cases: (a) when $0 < \pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta}) < \rho_k$, (b) when $\pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta}) = \rho_k$, and (c) when $\pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta}) = 0$. First, let us suppose (a). In this case we have

$$
\begin{aligned}
\pi_{\widetilde{w}}(\delta_x, \widetilde{\vartheta}) &= \min\left\{0 < \varepsilon < \rho_k \,\middle|\, \widetilde{w} \le \widetilde{w} \cdot \widetilde{\vartheta}\,(\overline{B_\varepsilon(x)}) + \varepsilon\right\} \\
&= \min\left\{0 < \varepsilon < \rho_k \,\middle|\, \eta_k(\{x\}) \le \vartheta\,(\overline{B_\varepsilon(x)}) + \varepsilon\right\} \\
&= \min\left\{0 < \varepsilon < \rho_k \,\middle|\, \forall\, A \subseteq S_{\eta_k} : \eta_k(A) \le \vartheta\,(\overline{A^\varepsilon}) + \varepsilon\right\} \quad \text{(by Part 1)} \\
&= \pi(\eta_k, \vartheta).
\end{aligned}
$$

Second, we assume (b). Let us observe the following for every $\varepsilon < \rho_k$:

$$
\eta_k(\{x\}) = \widetilde{w} > \widetilde{w} \cdot \widetilde{\vartheta}\,(\overline{B_\varepsilon(x)}) + \varepsilon = \vartheta\,(\overline{B_\varepsilon(x)}) + \varepsilon,
$$

which implies $\pi(\eta_k, \vartheta) \ge \rho_k$. To show the converse inequality, i.e., $\pi(\eta_k, \vartheta) \le \rho_k$, assume indirectly that there exists an $A \subseteq S_{\eta_k}$ such that

$$
\eta_k(A) > \vartheta\,(\overline{A^{\rho_k}}) + \rho_k.
$$

Very similarly, as in the verification of (3.25), we conclude that this inequality contradicts (P_k). Finally, the case (c) is trivial. $\square$

Since the modified witness function is obviously continuous, we also know the value of $W_{\widetilde{w},\widetilde{\vartheta}}(x)$ when $x \in \left\{y_{j,l} \,\middle|\, 1 \le j \le k, 1 \le l \le d_j\right\}$. Therefore if m pieces of atoms of $\vartheta \in \mathcal{P}_X$ have been already detected, then a modified witness function of the remaining part of ϑ can be calculated in terms of the Lévy-Prokhorov distances between ϑ and some measures supported on a set of at most $m + 1$ points.

3.4. Final major step: the action on $\mathcal{F}_X$ and $\mathcal{P}_X$

Now, we are in the position to verify our main result.

Proof of Main Theorem. Recall that we assumed (3.6) and that our aim is to show that φ is the identity map. Observe that it is enough to prove that φ acts identically on $\mathcal{F}_X$, as $\mathcal{F}_X$ is a weakly dense subset of $\mathcal{P}_X$ and φ is continuous. In order to do this we use induction on the cardinality of the support of $\mu \in \mathcal{F}_X$. By Corollary 3.8 our map φ fixes all measures with an at most two-element support. Let $k \in \mathbb{N}$, $k \ge 2$ and assume that we had already proved the following:

$$
\varphi(\nu) = \nu \qquad (\forall\, \nu \in \mathcal{P}_X, \#S_\nu \le k). \tag{3.26}
$$

Let us consider a measure

$$\mu = \sum_{i=1}^{k+1} \lambda_i \delta_{x_i} \in \mathcal{F}_X,$$

where the x_i's are pairwise different, $\sum_{i=1}^{k+1} \lambda_i = 1$ and each λ_i is positive. Assume also that for every $1 \le i \le k$ the point x_i lies outside of the convex hull of $\{x_j\}_{j=i+1}^{k+1}$. Let us use the following notations in the sequel:

$$\vartheta := \varphi(\mu) \quad \text{and} \quad \mu^{(i)} := \frac{1}{\sum\limits_{j=i+1}^{k+1} \lambda_j} \cdot \sum_{j=i+1}^{k+1} \lambda_j \delta_{x_j} \qquad (0 \le i \le k).$$

By Proposition 3.7 we observe that the support of ϑ is contained in the convex hull of $S_\mu = \{x_j\}_{j=1}^{k+1}$.

Now, we prove step by step that each x_i is an atom of $\varphi(\mu)$ with the same weight λ_i. By (3.8) we have $W_\mu \equiv W_\vartheta$, thus an application of Proposition 3.7 gives

$$\vartheta = \lambda_1 \cdot \delta_{x_1} + (1 - \lambda_1) \cdot \vartheta^{(1)},$$

with a measure $\vartheta^{(1)} \in \mathcal{P}_X$ such that $x_1 \notin S_{\vartheta^{(1)}}$ and $S_{\vartheta^{(1)}}$ lies in the convex hull of $S_\mu = \{x_j\}_{j=1}^{k+1}$. Utilising Lemma 3.11 and (3.26) for measures with supports of at most 2 elements we obtain

$$W_{1-\lambda_1, \mu^{(1)}} \equiv W_{1-\lambda_1, \vartheta^{(1)}}.$$

At this point, if k was 2, then $\#S_{\mu^{(1)}} = 2$, thus by Lemma 3.10 the measures $\mu^{(1)}$ and $\vartheta^{(1)}$ coincide, and therefore $\mu = \varphi(\mu)$ is yielded. Otherwise, applying Lemma 3.10 for the measures $\mu^{(1)}$ and $\vartheta^{(1)}$ gives

$$\vartheta = \lambda_1 \cdot \delta_{x_1} + \lambda_2 \cdot \delta_{x_2} + (1 - \lambda_1 - \lambda_2) \cdot \vartheta^{(2)},$$

with a measure $\vartheta^{(2)} \in \mathcal{P}_X$ such that $x_2 \notin S_{\vartheta^{(2)}}$ and $S_{\vartheta^{(2)}}$ lies in the convex hull of $S_{\mu^{(1)}} = \{x_j\}_{j=2}^{k+1}$. Using Lemma 3.11 and (3.26) for the case when the cardinality of the support is at most 3, we obtain

$$W_{1-\lambda_1-\lambda_2, \mu^{(2)}} \equiv W_{1-\lambda_1-\lambda_2, \vartheta^{(2)}}.$$

Iterating this procedure, the conclusion of the $(k-2)^{\mathrm{nd}}$ step is the following:

$$W_{1-\sum_{i=1}^{k-2}\lambda_i, \mu^{(k-2)}} \equiv W_{1-\sum_{i=1}^{k-2}\lambda_i, \vartheta^{(k-2)}} \tag{3.27}$$

where

$$\vartheta = \sum_{i=1}^{k-2}\lambda_i\delta_{x_i} + \left(1 - \sum_{i=1}^{k-2}\lambda_i\right)\cdot\vartheta^{(k-2)}$$

such that $\vartheta^{(k-2)} \in \mathcal{P}_X$, $x_{k-2} \notin S_{\vartheta^{(k-2)}}$ and $S_{\vartheta^{(k-2)}}$ lies in the convex hull of $S_{\mu^{(k-3)}} = \{x_{k-2}, x_{k-1}, x_k, x_{k+1}\}$. Utilising Lemma 3.10 for the measures $\mu^{(k-2)}$ and $\vartheta^{(k-2)}$ we get that

$$\vartheta = \sum_{i=1}^{k-1}\lambda_i\delta_{x_i} + \left(1 - \sum_{i=1}^{k-1}\lambda_i\right)\cdot\vartheta^{(k-1)}$$

with some $\vartheta^{(k-1)} \in \mathcal{P}_X$, $x_{k-1} \notin S_{\vartheta^{(k-1)}}$ and $S_{\vartheta^{(k-1)}}$ lies in the convex hull of $S_{\mu^{(k-2)}} = \{x_{k-1}, x_k, x_{k+1}\}$. Furthermore, by Lemma 3.11 and (3.26) we obtain

$$W_{1-\sum_{i=1}^{k-1}\lambda_i, \mu^{(k-1)}} \equiv W_{1-\sum_{i=1}^{k-1}\lambda_i, \vartheta^{(k-1)}}. \tag{3.28}$$

But since $\#S_{\mu^{(k-1)}} = 2$, Lemma 3.10 and (3.28) imply $\mu^{(k-1)} = \vartheta^{(k-1)}$, and therefore we conclude $\varphi(\mu) = \mu$, completing the proof. $\qquad\square$

4. Concluding remarks

We noted at the end of Section 2 that it is possible to give a characterisation of surjective π-isometries on certain subsets of $\mathcal{P}_X$. Namely, let $\mathcal{S} \subset \mathcal{P}_X$ be a weakly dense subset (possibly disjoint from Δ_X), and assume that $\phi\colon \mathcal{S} \to \mathcal{S}$ is onto and satisfies

$$\pi(\phi(\mu), \phi(\nu)) = \pi(\mu, \nu) \qquad (\forall\, \mu, \nu \in \mathcal{S}).$$

Since $(\mathcal{P}_X, \pi)$ is a complete metric space, there exists a unique isometric extension $\varphi\colon \mathcal{P}_X \to \mathcal{P}_X$, i.e., $\varphi|_{\mathcal{P}_X} = \phi$. Clearly, φ is a π-isometry which maps $\mathcal{P}_X$ into $\mathcal{P}_X$. Observe that $\varphi[\mathcal{P}_X]$ is closed in $\mathcal{P}_X$. On the other hand, as $\mathcal{S} = \varphi[\mathcal{S}] \subset \varphi[\mathcal{P}_X]$, we infer $\varphi[\mathcal{P}_X] = \mathcal{P}_X$. Therefore $\varphi\colon \mathcal{P}_X \to \mathcal{P}_X$ is induced by a surjective isometry $\psi\colon X \to X$, whence we conclude the same for $\phi\colon \mathcal{S} \to \mathcal{S}$, i.e.,

$$(\phi(\mu))(A) = \mu(\psi^{-1}[A]) \qquad (\forall\, \mu \in \mathcal{S}, A \in \mathscr{B}_X).$$

We proceed to mention some typical examples of weakly dense subsets of $\mathcal{P}_X$ (for which the above statement holds). 1) The set of all *discrete Borel probability measures*, which is the collection of those $\mu \in \mathcal{P}_X$ that are concentrated on a countable subset of X. 2) The class of all *continuous Borel probability measures*, i.e., those

$\mu \in \mathcal{P}_X$ such that $\mu(\{x\}) = 0$ for every $x \in X$. 3) Let $n \in \mathbb{N}$, $X = \mathbb{R}^n$ and $\| \cdot \|$ be an arbitrary norm on $\mathbb{R}^n$. Since any two norms on $\mathbb{R}^n$ are equivalent, the Borel σ-algebra $\mathscr{B}_{\mathbb{R}^n}$ does not depend on $\| \cdot \|$. We say that $\mu \in \mathcal{P}_{\mathbb{R}^n}$ is an *absolutely continuous Borel probability measure* if it is absolutely continuous with respect to the usual Lebesgue measure on $\mathbb{R}^n$. This set is clearly weakly dense in $\mathcal{P}_{\mathbb{R}^n}$, as every element of $\mathcal{F}_X$ can be approximated.

Next, as we have mentioned in the introduction, the most important special cases of our result are the following: 1) when X is an infinite dimensional, separable real Hilbert space; 2) when X is the real Banach space $C([0, 1])$; and 3) when X is an n-dimensional Euclidean space ($n \in \mathbb{N}$). We make some comments on how our proof could be modified in these cases. In the first two cases the underlying Banach spaces are of infinite dimension, hence the support of any $\mu \in \mathcal{F}_X$ lies in a finite dimensional affine subspace. In case of 1) the equivalence in Proposition 3.7 can be done for every element $\hat{x}$ in S_μ by choosing a half-line $\mathfrak{e}$ orthogonal to that affine subspace. Therefore the proof becomes much simpler as we immediately obtain that every $\mu \in \mathcal{F}_X$ is fixed by φ. A similar argument simplifies the proof for general strictly convex infinite dimensional separable Banach spaces. In case of 2) the space is of infinite dimension but the norm is not strictly convex. Despite of this obstacle the proof still can be shortened by utilising the Lindenstrauss-Troyansky theorem [12, 22, 32]. Namely, if $\mu \in \mathcal{F}_X$ and S_μ is contained in the kernel of a strongly exposing functional (for the definition see, *e.g.*, [12]), then the equivalence part of Proposition 3.7 can be verified for every element $\hat{x}$ in S_μ. Since by the Lindenstrauss-Troyansky theorem it is easy to see that every $\mu \in \mathcal{F}_X$ can be weakly approximated by such measures, we easily complete the proof of the Main Theorem in this case too. It seems that for finite dimensional spaces, even for the case of 3), we really have to do the whole procedure presented in Section 3, or at least we are not aware of any shortening possibilities.

Finally, we note that throughout Section 3 there were some parts where we considered general complete and separable metric spaces. But later on most of our techniques required that the underlying space had a linear structure. In our opinion it would be interesting to find a characterisation of all surjective π-isometries in the setting of other special (but still general enough) kinds of complete separable metric spaces.

References

[1] W. AMBROSE, *On measurable stochastic processes*, Trans. Amer. Math. Soc. **47** (1940), 66–79.

[2] J. ARAUJO and J. J. FONT, *Linear isometries between subspaces of continuous functions*, Trans. Amer. Math. Soc. **349** (1997), 413–428.

[3] J. ARAUJO and J. J. FONT, *Linear isometries on subalgebras of uniformly continuous functions*, Proc. Edinburgh Math. Soc. (2) **43** (2000), 139–147.

[4] J. ARAUJO, *The noncompact Banach-Stone theorem*, J. Operator Theory **55** (2006), 285–294.

[5] M. Bachir, *Remarks on isometries of products of linear spaces*, Extracta Math. **30** (2015), 1–13.

[6] Ben Yaacov I. and J. Melleray, *Isometrisable group actions*, Proc. Amer. Math. Soc. **144** (2016), 4081–4088.

[7] J. Bertrand and B. R. Kloeckner, *A geometric study of Wasserstein spaces: Hadamard spaces*, J. Topol. Anal. **4** (2012), 515–542.

[8] J. Bertrand and B. R. Kloeckner, *A geometric study of Wasserstein spaces: isometric rigidity in negative curvature*, Int. Math. Res. Not. IMRN (2016), 1368–1386.

[9] P. Billingsley, "Convergence of Probability Measures", Second edition, John Wiley & Sons, Inc., New York, 1999.

[10] D. P. Blecher and L. E. Labuschagne, *Logmodularity and isometries of operator algebras*, Trans. Amer. Math. Soc. **355** (2003), 1621–1646 (electronic).

[11] F. Botelho, J. Jamison and L. Molnár, *Surjective isometries on Grassmann spaces*, J. Funct. Anal. **265** (2013), 2226–2238.

[12] J. Bourgain, *Strongly exposed points in weakly compact convex sets in Banach spaces*, Proc. Amer. Math. Soc. **58** (1976), 197–200.

[13] L. Cheng, Q. Cheng, K. Tu and J. Zhang, *A universal theorem for stability of ε-isometries of Banach spaces*, J. Funct. Anal. **269** (2015), 199–214.

[14] J. L. Doob, *Stochastic processes depending on a continuous parameter*, Trans. Amer. Math. Soc. **42** (1937), 107–140.

[15] G. Dolinar and L. Molnár, *Isometries of the space of distribution functions with respect to the Kolmogorov–Smirnov metric*, J. Math. Anal. Appl. **348** (2008), 494–498.

[16] R. J. Fleming and J. E. Jamison, "Isometries on Banach Spaces: Function Spaces", Chapman & Hall/CRC, Boca Raton, FL, 2003.

[17] R. J. Fleming and J. E. Jamison, "Isometries on Banach Spaces. Vector-valued Function Spaces", Vol. 2, Chapman & Hall/CRC, Boca Raton, FL, 2008.

[18] Gy. P. Gehér and P. Šemrl, *Isometries of Grassmann spaces*, J. Funct. Anal. **270** (2016), 1585–1601.

[19] P. J. Huber, "Robust Statistics", John Wiley & Sons, Inc., New York, 1981.

[20] B. Kloeckner, *A geometric study of Wasserstein spaces: Euclidean spaces*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5) **9** (2010), 297–323.

[21] E. Le Donne and A. Ottazzi, *Isometries of Carnot groups and sub-Finsler homogeneous manifolds*, J. Geom. Anal. **26** (2016), 330–345.

[22] J. Lindenstrauss, *On operators which attain their norm*, Israel J. Math. **1** (1963), 139–148.

[23] P. Mankiewicz, *On extension of isometries in normed linear spaces*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **20** (1972), 367–371.

[24] M. Martín, *The group of isometries of a Banach space and duality*, J. Funct. Anal. **255** (2008), 2966–2976.

[25] L. Molnár, *Lévy isometries of the space of probability distribution functions*, J. Math. Anal. Appl. **380** (2011), 847–852.

[26] L. Molnár, *Kolmogorov–Smirnov isometries and affine automorphisms of spaces of distribution functions*, Cent. Eur. J. Math. **9** (2011), 789–796.

[27] D. Monclair, *Isometries of Lorentz surfaces and convergence groups*, Math. Ann. **363** (2015), 101–141.

[28] P. Niemiec, *Isometry groups of proper metric spaces*, Trans. Amer. Math. Soc. **366** (2014), 2597–2623.

[29] K. R. Parthasarathy, "Probability Measures on Metric Spaces", Academic Press, Inc., New York-London 1967.

[30] Yu. V. Prokhorov, *Convergence of random processes and limit theorems in probability theory*, Theory Probab. Appl. **1** (1956), 157–214.

[31] T. M. Rassias and P. Šemrl, *On the Mazur–Ulam theorem and the Aleksandrov problem for unit distance preserving mappings*, Proc. Amer. Math. Soc. **118** (1993), 919–925.

[32] S. L. TROYANSKI, *On locally uniformly convex and differentiable norms in certain non-separable Banach spaces*, Studia Math. **37** (1970/71), 173–180.

[33] C. VILLANI, "Optimal Transport: Old and New", Springer-Verlag, Berlin, 2009.

MTA-SZTE Analysis and Stochastics Research Group
Bolyai Institute
University of Szeged
Aradi vértanúk tere 1
Szeged H-6720, Hungary
and
MTA-DE "Lendület" Functional Analysis Research Group
Institute of Mathematics
University of Debrecen
P.O. Box 12
Debrecen H-4010, Hungary
gehergyuri@gmail.com
gehergy@math.u-szeged.hu

Alfréd Rényi Institute of Mathematics
Hungarian Academy of Sciences
Reáltanoda utca 13-15
Budapest H-1053, Hungary
titkos.tamas@renyi.mta.hu