Linearly rigid metric spaces and Kantorovich type norms.

J. Melleray[*], F. V. Petrov[†], A. M. Vershik[‡]

December 7, 2006

Abstract

We introduce and study the class of linearly rigid metric spaces; these are the spaces that admit a unique, up to isometry, linearly dense isometric embedding into a Banach space. The first nontrivial example of such a space was given by R. Holmes; he proved that the universal Urysohn space has this property. We give a criterion of linear rigidity of a metric space, which allows us to give a simple proof of the linear rigidity of the Urysohn space and other metric spaces. We relate these questions to the general theory of norms and metrics in spaces of measures on a metric space, and introduce the notion of a Banach norm compatible with a given metric; among these norms, the Kantorovich–Rubinshtein transportation metric is the maximal one, and the unit ball in this metric has a direct geometric description in the spirit of root polytopes.

Introduction

The goal of this paper is to describe the class of complete separable metric (=Polish) spaces which have the following property: there is a unique (up

[*]University of Illinois at Urbana-Champaign. E-mail: `melleray@math.uiuc.edu`.

[†]St. Petersburg Department of Steklov Institute of Mathematics. E-mail: `fedorpetrov@mail.ru`. Supported by the grants CRDF RUM1-2622,RFBR.05-01-00899

[‡]St. Petersburg Department of Steklov Institute of Mathematics. E-mail: `vershik@pdmi.ras.ru`. Supported by the grants CRDF RUM1-2622,NSh.4329.2006.1, INTAS 03-51-5018, Program SFB-701, of Math. Dept of University of Bielefeld, Germany.

to isometry) isometric embedding of this metric space (X, ρ) in a Banach space such that the linear span of the image of X is dense (in which case we say that X is *linearly dense*). One such embedding is well-known; it appears in Kantorovich's construction for the Kantorovich-Monge transport problem, we denote the Banach space generated by (X, ρ) when applying Kantorovich's method by $E_{(X,\rho)}$. So we must characterize the separable complete metric spaces which have no linearly dense isometric embedding in a Banach space besides Kantorovich's. We call such metric spaces *linearly rigid spaces*.

In order to study these spaces we must consider Kantorovich's construction from a new point of view: we introduce a "linear" geometry of metric spaces which is interesting in itself and nontrivial even for finite metric spaces. Namely, we define "root polytopes" for finite metric spaces (a generalization of root polytopes in the theory of simple Lie algebras) and study their geometry (section 1). We will return to this geometry elsewhere. The main result (section 3) is the criterion of linear rigidity in terms of *distance matrix* of the dense countable set in the Polish space. The formulation of the criterion uses the geometry of the space of distance matrices see [15] (section 2). The first nontrivial example of a linearly rigid space was Urysohn's universal space; this was proved by R.Holmes [8]. Remember that the Urysohn space is the unique (up to isometry) Polish space which is universal (in the class of all Polish spaces) and ultra-homogeneous (= any isometry between compact subspaces extends to an isometry of the whole space). Roughly speaking, our criterion states that a linearly rigid metric space must satisfy a weaker condition of universality; consequently there is more than one space with this property. We give two proofs of the main result. The formulation of the criterion and the first proof are due to the second and third author and are closer to the philosophy of the paper [15] and to the new view on the construction of the Kantorovich norm. The second proof (section 6), due to the first author, also uses duality for the transport problem but is more related to the paper of Holmes [8]. Besides the Urysohn space, an interesting example is the $\mathbb{N}$-analog of the Urysohn space - the integer valued ultra-homogeneous universal metric space, which is countable. We provide some other examples in section 4. In the short section 5 we discuss the extremal properties of the metric spaces which is also very intriguing for the general theory of the metric spaces. If (X, ρ) is a linearly rigid space, then the corresponding Kantorovich-Banach space $E_{X,\rho}$ is a remarkable space satisfying a weaker form of the universality property. At the same time (section 5), we claim that if (X, ρ) is the Urysohn space then $E_{X,\rho}$ is a universal Banach space, but is not isometric to the ϵ-homogeneous universal Gurariu Banach space.

1 Norms compatible with a metric and the Kantorovich space

Let (X, ρ) be a complete separable metric space. Consider the free vector space $V = \mathbb{R}(X)$ and the free affine space $V_0 = \mathbb{R}_0(X)$ generated by the space X (as a set) over the field of real numbers:

$$V(X) = \mathbb{R}(X) = \left\{ \sum a_x \cdot \delta_x,\ x \in X, a_x \in \mathbb{R} \right\}$$

$$\supset V_0(X) = \mathbb{R}_0(X) = \left\{ \sum a_x \cdot \delta_x,\ x \in X, a_x \in \mathbb{R}, \sum a_x = 0 \right\}$$

(all sums are finite). The space V_0 is a hyperplane in V. We omit the mention of the space (and also of the metric, see below) if no ambiguity is possible; we will mostly consider only the space V_0. One of the interpretations of the space $\mathbb{R}(X)$ (respectively, $\mathbb{R}_0(X)$) is that this is the *space of real measures* with finite support (respectively, the space of measures with total mass equal to zero: $\sum_x a_x = 0$). Let us fix a metric ρ on X and introduce a class of norms on V_0 *compatible with the metric*. For brevity, denote by $e_{x,y} = \delta_x - \delta_y$ the elementary signed measure corresponding to an ordered pair (x, y).

Definition 1. *We say that a norm $\|\cdot\|$ on the space V_0 is compatible with the metric ρ if $\|e_{x,y}\| = \rho(x, y)$ for all pairs $x, y \in X$.*

The rays $\{ce_{x,y},\ c > 0\}$ going through elementary signed measures will be called *fundamental rays* (the set of fundamental rays does not depend on the metric). If the metric is fixed, then a norm compatible with this metric determines, on each fundamental ray, a unique vector of unit norm; let us call these vectors (elementary signed measures) *fundamental vertices corresponding to a given metric.* They are given by the formula $(x \neq y)$:

$$\frac{e_{x,y}}{\rho(x, y)} \equiv \bar{e}_{x,y}.$$

Thus the set of norms compatible with a given metric ρ is the set of norms for which the fundamental vertices corresponding to this metric are of norm one.

Lemma 1. *Let us fix a Polish space (X, ρ). The unit ball with respect to every norm compatible with the metric ρ contains the convex hull of the set of fundamental vertices, hence among all the compatible norms there is a maximal norm $\|\cdot\|_{max}$; the unit ball in this norm is the convex hull of the set of fundamental vertices. This norm coincides with the classical Kantorovich transportation norm (see* [9]*). It is defined for every metric space (X, ρ).*

In the paper [9] the Kantorovich transportation metric on the space of probability Borel measures on the compact metric spaces was defined; the corresponding Banach norm was defined in the later paper [10]. Several authors rediscovered these metric and norm later (see short history in [16]). Nevertheless it seems that the definition of the Kantorovich norm in the lemma above as a maximal compatible norm is a new one. Remark by the way that during the last several years a great interest in this subject and its generalizations grew steadily (see [2, 17] for example).

Proof. Temporarily denote the norm determined by the closed convex hull of the set of fundamental vertices by $\|\cdot\|'$. The Kantorovich norm $\|\cdot\|$ in the space of measures of the form $\phi = \phi_+ - \phi_-$, where $\phi_\pm$ are nonnegative finitely supported measures with equal mass ($\phi_-(X) = \phi_+(X)$), is defined as

$$\inf\left\{\sum a_{x,y}\rho(x,y) : \sum_y a_{x,y} = \phi_+(x), \sum_x a_{x,y} = \phi_-(y)\right\}.$$

Since the Kantorovich norms of all fundamental vertices are equal to one, their closed convex hull lies in the unit ball with respect to the Kantorovich norm, whence $\|\phi\|' \geq \|\phi\|$. On the other hand, by the duality theorem, there exists an optimal transportation plan $a = \{a_{x,y}\}$: $\|\phi\| = \sum(u(x) - u(y))a_{x,y}$, where u is a 1-Lipschitz function on X; moreover, if $a_{x,y} > 0$, then $u(x) - u(y) = \rho(x,y)$. Hence $\|\phi\| = \sum \rho(x,y)a_{x,y} = \sum \|e_{x,y}\|a_{x,y} = \sum \|e_{x,y}\|'a_{x,y} \geq \|\sum e_{x,y}a_{x,y}\|' = \|\sum(\delta_x - \delta_y)a_{x,y}\| = \|\phi_+ - \phi_-\|' = \|\phi\|'$. Thus the Kantorovich norm is the maximal norm $\|\cdot\|_{max}$ compatible with the metric ρ. □

Lemma 2. *Let X be a set. Consider the linear space $V_0(X)$ and specify some points $c(x,y) \cdot e_{x,y}$ on the fundamental rays $\mathbb{R}_+ \cdot e_{x,y}$, where the function $c(x,y)$ is defined for all pairs (x,y), $x \neq y$, positive, and symmetric: $c(x,y) = c(y,x)$. This set of points is the set of fundamental vertices of some metric on X if and only if no point lies in the relative interior of the convex hull of a set consisting of finitely many other fundamental vertices and the zero.*

Proof. Assume that we are given a set of fundamental vertices $c(x,y) \cdot e_{x,y}$. Consider the function defined by the formulas $\rho(x,y) = c(x,y)^{-1}$, $x \neq y$, and $\rho(x,x) = 0$. Let us check that the triangle inequality for this function is equivalent to the property of convex hulls mentioned in the lemma. Assume

$$\bar{e}_{a,b} = \sum_{i=1}^{n-1} \lambda_i \bar{e}_{x_i,x_{i+1}}, \qquad x_1 = a, \quad x_n = b, \qquad \sum_{i=1}^{n-1} \lambda_i < 1, \quad \lambda_i > 0,$$

(taking into account that $\bar{e}_{a,b} = \delta_a - \delta_b$ this is only way to represent $\bar{e}_{a,b}$ as a convex combination of other fundamental vectors). Equating the coefficients of $\delta_{x_i}, i = 1, 2, \dots n$ in both sides, we see that $\lambda_i/\rho(x_i, x_{i+1}) = 1/\rho(a,b); i = 1 \dots n-1$. So $\lambda_i = \rho(x_i, x_{i+1})/\rho(a,b)$. It follows that condition $\sum_{i=1}^{n-1} \lambda_i < 1$ contradicts the triangle inequality. Conversely, the latter relation is equivalent to the above decomposition of $\bar{e}_{a,b}$ into a convex combination. □

The completion of the space V_0 with respect to the norm $\| \cdot \|_{max}$ is the space defined by Kantorovich and Rubinshtein for compact metric spaces (see [10]) when studying the so-called transport problem. But we can complete V_0 with respect to a maximal norm $\| \cdot \|_{max}$ for an arbitrary metric space, denote it by $E_{X,\rho}$. In the paper [10] and in all subsequent papers, one usually considers only compact metric spaces; in this case, the completion contains all Borel measures of bounded variation, but is not exhausted by them. For an arbitrary metric space, this completion and the space of Borel measures are in general position.

The correspondence

$$(X, \rho) \mapsto E_{X,\rho}$$

is a functor from the category of Polish spaces (with Lipschitz maps as morphisms) to the category of Banach spaces (with linear bounded maps as morphisms).

Consider an arbitrary norm on V_0 compatible with the metric and extend it to the space V by setting $\|\delta_x\| = 0$ for some point $x \in X$ (there is no canonical extension, but the choice of the point x is not essential). Consider the completions $\bar{V}_0$ and $\bar{V}$ of the spaces V_0 and V with respect to these norms. Obviously, the metric space (X, ρ) has a canonical isometric embedding into $\bar{V}$, and, conversely, it is easy to see that if there exists an isometric embedding of the space (X, ρ) into some Banach space E, then the closure of the affine hull of the image of X in E is isometric to the completion of V with respect to some norm compatible with the metric. We will say that an isometric embedding of X into a Banach space E is *linearly dense* if the affine hull of the image of X coincides with E. Thus *every metric space has a linearly dense isometric embedding into the Banach space $E_{X,\rho}$.* It turns out that for some metric spaces (X, ρ) the Kantorovich norm is the *unique* norm on $V_0(X, \rho)$ compatible with the metric ρ; such spaces have a unique linearly dense isometric embedding into a Banach space (up to isometry). A trivial example of such a space is the metric space consisting of one or two points. The first nontrivial (and, as we will see below, necessarily infinite) example of such a space was discovered by R. Holmes [8], who showed that the universal Urysohn space has this property.

The main result of this paper is a description of metric spaces that have a unique (up to isometry) linearly dense isometric embedding into a Banach space. We call such metric spaces *linearly rigid*.

Remarks.

1. Let (X_n, ρ) be a finite metric space with n points in which the distance between any two points is equal to one: $\rho(i, j) = \delta_{i,j}$. In this case, the convex hull of the set of fundamental vertices in the space V_0 is a classical *root polytope* (= the convex hull of the set of all roots of a Lie algebra of series A_n; in this interpretation, V_0 is the conjugate space to the Cartan algebra). Thus the term “the root polytope of a finite metric space” will be used for the convex hull of the set of fundamental vertices of an arbitrary finite metric space, i.e., for the ball in the Kantorovich metric. The geometry of root polytopes, in particular, their convex type, provides invariants of metric spaces; we will return to this subject elsewhere.

2. The conjugate space of the Banach space $\bar{V}_0$ with the maximal (Kantorovich) norm is the quotient of the space $\mathrm{Lip}(X, \rho)$ of all Lipschitz functions on (X, ρ) by the subspace of constants. This fact is a basis for the duality theorem; it was obtained in the first works by L. V. Kantorovich and his coworkers on this subject.

3. In the modern literature (see, e.g., [17, 2]), there is an intensive study of *metrics on the simplex of Borel probability measures compatible with a metric*, i.e., metrics ϱ on this simplex such that $\varrho(\delta_x, \delta_y) = \rho(x, y)$. The Kantorovich metric is one of such metrics, and again it is maximal among all the compatible metrics. It is this metric that was first suggested in [9] and that allows one to construct the Kantorovich norm in the space of signed measures. Conversely, if we are given a norm on the space of signed measures, then the corresponding metric on measures can be recovered by the formula

$\varrho(\mu,\nu) = \|\mu - \nu\|$, where μ, ν are finitely supported positive measures on (X,ρ) with total mass equal to one. Moreover, *every norm compatible with the metric determines an admissible metric.* But, in general, not every compatible metric can be extended to an admissible norm. For example, in the case when a norm compatible with the metric ρ is unique (i.e., in the case when the space under consideration is linearly rigid), there are many distinct (nonisometric) metrics ϱ on the simplex of measures compatible with the metric ρ, but only one of them (the maximal one) generates a norm. The following problem then arises: which metrics generate a norm? In other words, when can the distance between the positive and the negative parts of a signed measure be taken as the norm of this signed measure? The solution of this problem will undoubtedly extend the possibilities of estimation of transportation-type metrics with the help of the more powerful machinery of norms in Banach spaces.

2 The cone of distance matrices

A natural method of studying metric spaces is the method of distance matrices. It is considered in [15]. A distance matrix is a finite or infinite matrix $r = \{r_{i,j}\}$ that determines a metric or a semimetric on $\mathbb{N}$ or on $\mathbf{n}=\{1,2,\ldots,n\}$. Distance matrices form a convex weakly closed cone in the space of all real matrices of the corresponding finite or infinite order.

If in an (infinite) Polish metric space (X,ρ) there is a distinguished countable everywhere dense set $\{x_n\}_{n=1}^{\infty}$, then the distance matrix $\{\rho(x_i,x_j)\}$ associated with this set contains all the information on (X,ρ), and one can study the properties of the space (X,ρ), for example, universality or linear rigidity, using this distance matrix (see [15]). Denote by $X_n = \{x_1, x_2, \ldots, x_n\}$ the set of the first n points of the distinguished everywhere dense sequence chosen in X.

Definition 2. *We say that a vector $\{a_i\}$, $i = 1,2,\ldots,n$, is admissible for a finite distance matrix $\{r_{i,j}\}_{i,j=1}^{n}$ if*

$$|a_i - a_j| \le r_{i,j} \le a_i + a_j \text{ for all } i,j = 1,\ldots,n\,.$$

An admissible vector is a 1-Lipschitz function on the metric space (X_n, r). Not every 1-Lipschitz function f corresponds to an admissible vector, but there always exists a constant $C > 0$ (depending on f) such that the function

$f+C$ does correspond to an admissible vector. The set of admissible vectors (denote it by Adm_r) is a convex polyhedral set. It is unbounded: together with every vector v it contains the ray $v+\mathbb{R}_+\cdot(1,1,\dots)$. It is easy to see that Adm_r is the Minkowski sum of a convex polyhedron and the ray $\{(\lambda,\lambda,\dots),\, \lambda>0\}$ (see [15]). An extreme point of this set will be called an *extremal admissible vector*, and an extreme ray (i.e., a ray that is the intersection of Adm_r with a supporting plane to Adm_r), will be called an *extremal admissible ray.* These rays correspond to extreme points of the unit ball in the quotient of the space of Lipschitz functions by the one-dimensional space of constants, which is dual to the Kantorovich space $E_{X_n,r}$.

Recall (see [14, 15, 11, 13]) that P. S. Urysohn proved the existence and uniqueness up to isometry of a universal Polish space, i.e., a complete metric separable space such that every Polish space can be isometrically embedded into it (proper universality) and which is absolutely homogeneous, in the sense that every isometry between finite subsets can be extended to a global isometry. The following theorem gives a necessary and sufficient condition for a distance matrix to be the distance matrix of a countable everywhere dense system of points in the Urysohn space.

Theorem 1 (Universality criterion, [15]). *A space (X,r) is the universal Urysohn space if and only if the distance matrix $\{r_{i,j}\}$ of some (and hence every) countable everywhere dense system has the following universality property: for any $n\in\mathbb{N}$, $\epsilon>0$, and a vector $a\in\mathrm{Adm}(r^n)$ admissible for the distance matrix $r^n=\{r_{i,j}\}_{i,j=1}^n$, there exists m such that $\|a-\{r_{m,j}\}_{j=1}^n\|<\epsilon$.*

This criterion does not differ much from the condition used by Urysohn [14] or the conditions of Katětov [11] and Gromov [6]. However, the above formulation provides a more convenient method of constructing the universal space and allows one to prove the generic property of universal spaces, which is done in [15]. The following almost literal reformulation of the criteria above gives very useful characterization of the Urysohn space via extension of the Lipschitz functions from finite sets. It can be considered as a dual characterization of the Urysohn space.

The Polish space (X,r) is isometric to the universal Urysohn space iff for any $\epsilon>0$, any finite subset $F\subset X$ and any positive Lipschitz function u on the set F with induced metric r_F, there exist a point $x\in X$ such that $sup_{y\in F}\quad |r(x,y)-u(y)|<\epsilon$.

3 Linear rigidity

The following theorem provides a criterion of linear rigidity (cf. the criterion of universality).

Theorem 2. *Consider a Polish space* $(\bar{X}, \rho)$ *and an arbitrary everywhere dense sequence* $X = \{x_n\}$, $n = 1, 2, \dots,$ *of points of* $\bar{X}$. *The space* $(\bar{X}, \rho)$ *is linearly rigid if and only if the distance matrix* $M = \{r_{i,j}\}_{i,j=1}^{\infty}$ *of the sequence* X *satisfies the following condition: for every* $\epsilon > 0$, $n \in \mathbb{N}$, *and any extremal ray* L *of the set* $\mathrm{Adm}(r^n)$ *of admissible vectors for the distance matrix* $r^n \equiv \{r_{i,j}\}_{i,j=1}^{n}$ *of the first* n *points of the sequence* X, *there exists a vector* $v \in L$ *on this ray and a number* m *such that* $\|v - \{r_{m,j}\}_{j=1}^{n}\| < \epsilon$.

Remark. As a parallel to the reformulation of the criteria of universality which was given after theorem 1 we can reformulate also the Theorem 2 as follows: a Polish space (X, r) is linearly rigid iff for any $\epsilon > 0$, any finite subset $F \subset X$ and any *extremal* positive Lipschitz function u on the set F with the induced metric r_F there exist a point $x \in X$ and a constant a such that $sup_{y \in F}|r(x, y) - u(y) - a| < \epsilon$.

In this paper, we give two proofs of this theorem, corresponding to two possible points of view. The first proof, given below, is more measure-theoretic in nature, while the other one (given in the last section of the article), more elementary but less conceptual, is obtained by looking at the dual space to the space generated by X and using some known properties of Kantorovich spaces.

Proof. 1. The "only if" part. Let us first prove that if for some ϵ, a positive integer n, and an extremal ray $L \subset \mathrm{Adm}(r^n)$, the reverse inequality

$$\|v - \{r_{m,j}\}_{j=1}^{n}\| \geq \epsilon$$

holds for every vector $v \in L$ and every number m, then the space X is not linearly rigid. First of all we can assume that $n > 2$ because the extremal admissible rays $L \subset \mathbb{R}^2$ for two-point spaces are of the form $\{(r_{1,2} + \lambda, \lambda)\}_{\lambda \geq 0}$ and $\{(\lambda, r_{1,2} + \lambda)\}_{\lambda \geq 0}$, and in this case the vectors $(r_{1,2}, 0)$ and $(0, r_{1,2})$ already belong to the set of extremal rays. Thus, we can omit this extremal rays.

Let us define a structure of a directed graph on X_n for $n > 2$ as follows: draw an edge $x_i \to x_j$ if $v_i - v_j = r_{i,j}$ for some $v \in L$ (the equality does not depend on the choice of $v \in L$, $v \neq 0$, since adding the same constant to all the v_i preserves the equalities and allows us to move the vector arbitrarily along the ray). Note that the constructed graph regarded as an undirected graph is connected. Indeed, if it is not connected, then we may assume without loss of generality that there are no edges between x_i and x_j for all indices i, j with $i \leq i_0$, $j > i_0$. Choose an arbitrary vector $w \in L$ and add a sufficiently large constant to its coordinates, $w := w + C \cdot (1, 1, \ldots, 1)$, so as to make all the coordinates greater than $\max_{i,j=1}^{n} r_{ij}$. Consider the vector $\delta_w = (1, 1, \ldots, 1, 0, 0, \ldots, 0)$ (with 1's in positions with indices $i \leq i_0$) and introduce the vectors $w_1, w_2 := v \pm \epsilon \cdot \delta_w$, where ϵ is a sufficiently small constant. We see from construction that $w_1, w_2 \in \mathrm{Adm}_{r^n} \setminus L$ and $w = (w_1 + w_2)/2$, which contradicts the extremality of L. Denote the edge set of the graph X_n by $e(X_n)$.

Let us define an element $\mu \in V_0(X_n)$ as the sum

$$\mu := \sum_{(a \to b) \in e(X_n)} e_{a,b},$$

and consider μ and other elements of $V_0(X_n)$ as measures on X_n. Vectors $v \in \mathrm{Adm}_{r^n}$ are Lipschitz functions on X_n, and we can integrate them with respect to these measures. By the definition of the measure μ, the vector v, regarded as a Lipschitz function on X_n, is the solution of the dual transportation problem with the measures μ_+ and μ_-, where $\mu = \mu_+ - \mu_-$; hence, by the duality theorem,

$$\|\mu\|_K = \int v d\mu = \sum_{(a \to b) \in e(X_n)} \rho(a, b).$$

Note that if the support of μ consists of two points (without loss of generality we may assume that these points are x_1, x_2 and $n = 2$), then, taking x_1 as x_m, we obtain a contradiction with the assumption $||v - \{r_{m,j}\}_{j=1}^{n}|| \geq \epsilon$; indeed, as the vector v lying on the extremal ray we can take the vector $(0, \rho(x_1, x_2))$, in other words, take $m = 1$.

Because we assume that $n > 2$ we have $\mu \neq e_{a,b}$, i.e., the support of μ consists of more than two points. In this case, the construction of a new norm $\| \cdot \|_n$ on $V_0(X_n)$ compatible with the metric uses a corrected measure μ. Namely, let us define the unit ball B_n of the new norm $\| \cdot \|_n$ as the convex hull in $V_0(X)$ of the unit ball B_K with respect to the Kantorovich norm, i.e., the convex hull of the elements $\bar{e}_{a,b} = e_{a,b} / \rho(a, b) \in V_0(X)$ and the point $\nu := \frac{\mu}{\|\mu\|_K - \epsilon}$.

Let us prove that the $\|\cdot\|_n$-norms of points of the form $\bar{e}_{a,b}$ are equal to one, as well as the $\|\cdot\|_K$-norms. Since $\|\mu\|_n \leq \|\mu\|_K - \epsilon$ (actually, the equality holds), this will prove that there exists another norm compatible with the metric apart from the maximal one, i.e., that the space X is not linearly rigid.

Assume that it is not true. This means that one of the points of the form $\bar{e}_{c,d}$ can be written as

$$\bar{e}_{c,d} = \lambda \cdot \nu + \sum_{k=1}^{N} \lambda_k \cdot \bar{e}_{a_k,b_k}, \qquad \lambda + \sum \lambda_k < 1$$

(and hence $\|\bar{e}_{c,d}\|_n < 1$). Note that the positivity of λ follows from our assumption that $\mu \neq e_{a,b}$ for any a, b from X. Let us integrate the admissible function $\rho(c,\cdot)$ with respect to the measures in the left- and right-hand sides.

We obtain

$$1 = \lambda \int \rho(c,\cdot)d\nu + \sum \lambda_k \frac{(\rho(c,a_k) - \rho(c,b_k))}{\rho(a_k,b_k)} < \lambda \int \rho(c,\cdot)d\nu + 1 - \lambda,$$

which implies that $\int \rho(c,\cdot)d\nu > 1$, and, taking into account the definition of the measure ν, we obtain

$$\sum_{(a\to b)\in e(X_n)} \big(\rho(c,a) - \rho(c,b) - \rho(a,b)\big) \geq -\epsilon$$

This implies the inequality

$$0 \leq \rho(a,b) + \rho(c,b) - \rho(a,c) \leq \epsilon.$$

Thus the differences of the distances from the point c to all (by the connectedness of the graph X_n) points of X_n are determined (up to ϵ), and the vector of distances from c to the points of X_n is close to the vector $v \in L$ chosen above. But the sequence $\{x_k\}$ is everywhere dense, hence it contains a point arbitrarily close to c whose vector of distances to the points of X_n coincides with the vector $v \in L$ up to ϵ, contradicting the original assumption.

We can say briefly that the essence of the proof was the link between the distance (ϵ) between the admissible vector (v) on the extremal ray (L) and functions ($c-->\rho(c,.)$), and the Kantorovich norm ($\|\mu\|_K$) of the measure (μ) which corresponds to that vector (v).

2. The "if" part. Let us prove that if for all $n \in \mathbb{N}$, $\epsilon > 0$, and an extremal ray $L \subset \mathrm{Adm}_{r^n}$ there exists a desired vector $v \in L$ and a point $x_m \in \{x_1, x_2, \dots\}$ whose vector of distances to $\{x_1, \dots, x_n\}$ coincides with v up to ϵ, then the space X is linearly rigid.

Namely, let us prove that for every signed measure $\mu \in V_0(X_n)$ and every norm $\|\cdot\|$ on $V_0(X)$ compatible with the metric, $\|\mu\| = \|\mu\|_K$. This will suffice, because $\cup_n X_n$ is dense in X. Let

$$\mu = \sum_{k=1}^{N} \alpha_k \bar{e}_{a_k,b_k}, \quad a_k, b_k \in X_n,\ \alpha_k \geq 0, \quad \|\mu\|_K = \sum \alpha_k.$$

The points $\bar{e}_{a_k,b_k}$ lie on some face of the unit ball of the space E_{X_n}. We may assume without loss of generality that it is a face of codimension 1. The corresponding supporting plane is determined by some linear functional of norm 1, i.e., a 1-Lipschitz function. Denote this function by f: $f(a_k) - f(b_k) = \rho(a_k, b_k)$. This is an admissible vector from $\mathrm{Adm}_{X_n,r}$ lying at an extremal ray L in $\mathrm{Adm}_{X_n,r}$.

First consider the case when all the a_k are equal: $a_k = a$.

By assumption, there is a point $c \in X$ such that $\rho(c, a) \geq \rho(c, b_k) + \rho(a, b_k) - \varepsilon$, its distance vector to the points $\{x_1, \dots, x_n\}$ will be close to a vector on the ray L).

We have

$$\|\mu\| = \left\| \sum \alpha_k \frac{(\delta_a - \delta_c) + (\delta_c - \delta_{b_k})}{\rho(a, b_k)} \right\|$$

$$\geq \sum \alpha_k \cdot \frac{\rho(a, b_k)}{\rho(a, c)} - \sum \alpha_k \cdot \frac{\rho(c, b_k)}{\rho(a, b_k)} \geq \sum \alpha_k \left(1 - \frac{\varepsilon}{\min_k \rho(a, b_k)}\right).$$

Letting $\varepsilon \to 0$, we obtain

$$\|\mu\| = \sum \alpha_k = \|\mu\|_K.$$

Now consider the general case. Find a point d such that $\rho(d, a_k) \geq \rho(d, b_k) + \rho(a_k, b_k) - \varepsilon$. Then we obtain

$$\begin{aligned}
\|\mu\| &= \left\| \sum \alpha_k \frac{(\delta_{a_k} - \delta_d) + (\delta_d - \delta_{b_k})}{\rho(a_k, b_k)} \right\| \\
&\geq \left\| \sum \alpha_k \frac{\delta_{a_k} - \delta_d}{\rho(a_k, b_k)} \right\| - \sum \alpha_k \left\| \frac{\delta_d - \delta_{b_k}}{\rho(a_k, b_k)} \right\| \\
&\geq \sum \alpha_k \frac{\rho(a_k, d)}{\rho(a_k, b_k)} - \sum \alpha_k \frac{\rho(d, b_k)}{\rho(a_k, b_k)} \geq \sum \alpha_k + o(1),
\end{aligned}$$

which completes the proof in this general case too. □

The next theorem shows how one can construct linearly rigid spaces by induction, successively increasing the number of points. Namely, it provides a universal procedure similar to the inductive construction of the Urysohn space (see [15]), and allows one to obtain (after completion) an arbitrary linearly rigid Polish space. This procedure should be described in the geometric terms of root polytopes in the space $V_0(X)$. A particular step of this procedure, adding new points $x_{n+1}, \dots, x_m$ to a finite metric space $X_n = \{x_1, \dots, x_n\}$ already constructed, consists in determining a line segment connecting the new fundamental vertices in $V_0(X)$ corresponding to any two points y, z being added that intersects a given face of the root polytope at an interior point. Then the whole face will be "rigid": in any norm compatible with the metric, all points of this face will be of norm one. Enumerating the sequences of faces of the root polytopes already constructed and "piercing" them by new line segments, we obtain a sequence of finite metric spaces for which all faces of all root polytopes are rigid; hence the completion of the constructed countable space will be linearly rigid.

Theorem 3 (Piercing theorem). *Let (X, r) be an arbitrary finite metric space, $\epsilon > 0$, and Γ be a face of the unit ball of the space E_X. Then the space (X, r) can be isometrically embedded into a finite metric space (Y, ρ) so that there exists a face Δ of the unit ball of the space $E_{Y,\rho}$ containing Γ and two vectors $\bar{e}_{z_1,z_2}$ and $\bar{e}_{u_1,u_2}$ such that the line segment connecting them intersects the face Δ at an interior point.*

The proof is similar to the proof of the "if" part of Theorem 1. The case of an arbitrary face reduces to considering a face that is the convex hull of vertices of the form $e_{a,x}$ with fixed a and different x. Two new fundamental vertices that are the endpoints of the line segment "piercing" such a face are given explicitly. Using this theorem, it is not difficult to justify the above procedure. Note that if an interior point of a face is of norm one, then all points of the face are also of norm one. Thus the recursive construction of a linearly rigid space consists in adding points satisfying the conditions of the piercing theorem so that eventually all faces become pierced.

The next theorem shows that a linearly rigid space cannot have a finite diameter if it has more than two points. This means (and is partially explained by the fact) that the unit sphere in the Banach space $E_{X,\rho}$ corresponding to a linearly rigid space is very degenerate: its finite-dimensional approximations, i.e., the root polytopes of finite metric spaces, have decreasing cross sections for growing dimensions $\bar{e}_{ab} = \dfrac{e_{a.b}}{\rho(a,b)}$, and hence the distances $\rho(a,b)$ are not bounded. This follows from the considerations of the previous theorem. Note that such a degeneracy of the unit sphere is typical for universal constructions (cf. the Poulsen simplex).

Theorem 4. *A linearly rigid metric space X containing more than two points is of infinite diameter and, in particular, noncompact.*

Proof. Assume the converse. Without loss of generality we assume that the space X is complete. Fix a point $a \in X$, denote by r_a the supremum of the distances $\rho(a,x)$ over $x \in X$, and choose a sequence of points (x_n) such that $\rho(a,x_n) \geq r_a - 1/n$. Then pick a countable dense subset $\{y_n\}$ of X, and define a sequence (z_n) by setting $z_{2n} = x_n$ and $z_{2n+1} = y_n$. Consider the points $\bar{e}_{a,z_k}$, $k = 1,\dots,N$. They lie on the same face of the unit ball of the space E_{X_N}, where $X_N = \{a, z_1, z_2, \dots, z_N\}$. Let us find the supporting ray in the set of admissible vectors Adm corresponding to this face. Applying Theorem 2 for this ray, we may find a point c_N such that

$$\rho(a, c_N) \geq \rho(a, z_k) + \rho(z_k, c_N) - 1/N, \qquad k \leq N.$$

In particular, $\rho(a,c_N) \geq \rho(a,x_k)+\rho(x_k,c_N)-1/N$, $2k \leq N$. Hence $\rho(x_k,c_N) \to 0$ as $k,N \to \infty$, so that the sequences (x_k), (c_k) are fundamental and have a common limit a'. The point a' satisfies the equalities $\rho(a,x) + \rho(x,a') = \rho(a,a')$ for all $x \in X$ (this is why we used the countable dense set $\{y_n\}$ in the definition of our sequence (z_n)).

Such a construction may be done for any point $a \in X$; note that for any $a, b \in X$ such that $a \neq b \neq a'$ (such a, b do exist if X has more than two points) we have $2\rho(a,a') = \rho(a,b) + \rho(a,b') + \rho(a',b) + \rho(a',b') = 2\rho(b,b')$, whence $\rho(a,a') \equiv D < \infty$. It also follows that $\rho(a,b) = \rho(a',b')$.

Without loss of generality, $\rho(a,b') = \rho(b,a') \geq \rho(a,b) = \rho(a',b') = 1$. Let $A = \{a,b,a',b'\}$. Define a function φ by the formulas $\varphi(a) = \varphi(a') = 0$, $\varphi(b) = \varphi(b') = 1$.

Such a function will be Lipschitz on $\{a,b,a',b'\}$; the corresponding face contains the points $\bar{e}_{a,b}$, $\bar{e}_{a',b'}$. Hence there exists a point c such that

$$\rho(c,a') \geq \rho(c,b') + 1/2, \qquad \rho(c,a) \geq \rho(c,b) + 1/2.$$

We have

$$\rho(c,c') = \rho(c,a') + \rho(a',c') = \rho(c,a') + \rho(c,a) \geq (\rho(c,b') + \rho(c,b) + 1) \geq D + 1.$$

The obtained contradiction proves the theorem. □

4 Corollaries

1. Theorem 5 (R. Holmes [8]). *The Urysohn space is linearly rigid.*

Proof. It suffices to compare the assumptions of the criterions of universality (Theorem 1) and linear rigidity (Theorem 2): the assumptions of the latter criterion require that the columns of the matrix should approximate only *extremal* admissible vectors rather than any admissible vectors as in the former one. □

The proof in [8] consists in a detailed study of embeddings of the Urysohn space into the Banach space $C([0,1])$ of continuous functions on the interval and proving that they are isometric; thus it does not allow one to study other examples (and even does not contain indications of their existence); some of the ideas in that proof led to the other proof of Theorem 2 that we give at the end of the paper. Roughly speaking our criterion of linear rigidity shows that linearly rigid spaces must "almost contain" up to isometry all extremely degenerated finite spaces and Urysohn space contains up to isometry all finite metric spaces (see next section).

Consider another two examples of linearly rigid universal spaces.

2. Let us consider the countable metric space denoted by $\mathbb{QU}_{\geq 1}$. It is a universal and absolutely homogeneous space in the class of countable metric spaces with rational distances not smaller than one. Such a space can be constructed in exactly the same way as the Urysohn space.

Theorem 6. *The space* $\mathbb{QU}_{\geq 1}$ *is linearly rigid.*

Indeed, the assumptions of the criterion of linear rigidity (Theorem 2) are obviously satisfied.

This example, as well as the next one, is of interest because it is an example of a countable linearly rigid space. Thus the corresponding Banach space $E_{\mathbb{QU}_{\geq 1}}$ has a basis. It is not known whether the space $E_{\mathbb{U}}$ has a basis.

3. The following example is of special interest also for another reason. Consider the space $\mathbb{Z}\mathbb{U}$, the universal and absolutely homogeneous space in the class of metric spaces with integer distances between points. Let us show that it is also linearly rigid. For this, let us check the condition of the criterion of linear rigidity. Fix X_n and a ray L of the admissible set Adm. Note that the differences of the coordinates of every vector from X_n are integers; this follows from the connectivity of the graph in the proof of Theorem 1. Hence on this ray there is a vector with integer coordinates, which is realized as the vector of distances from some point $x \in X$ to X_n.

Let us introduce a structure of a graph on this space by assuming that pairs of points at distance one are neighbors. This graph has remarkable properties: it is universal but not homogeneous (as a graph), its group of isomorphisms coincides with the group of isometries of this space regarded as a metric space. As follows from [3, 4], there exists an isometry that acts transitively on this space; hence the incidence matrix has a Toeplitz realization.

4. Another example of a linearly rigid space is any of the three above spaces with an arbitrary open bounded set removed. It easily follows from the universality criterion that if we remove the unit ball, then the obtained space will be *isometric to the original one*.

5 Extremality and the properties of the Banach–Kantorovich space $E_{X,\rho}$

Recall (see [15]) that the set of infinite distance matrices is a convex weakly closed cone. Its extreme rays correspond to metrics on $\mathbb{N}$, and hence to metrics on the completion, that cannot be written as the half sums of any other nonproportional metrics. Note that the universal real (Urysohn) space $\mathbb{U}$ *is extremal* in this sense (see [15]). It follows that the distance matrices of everywhere dense systems of points of extremal metric spaces form an everywhere dense G_δ-set in the space of distance matrices. The integer space $\mathbb{Z}\mathbb{U}$ is also extremal; the extremality of both spaces follows from a result of Avis [1], which states that every finite metric space with commensurable distances can be embedded into a finite extremal metric space, and hence the assumptions of the criterion of linear rigidity are satisfied. It is not known whether any linearly rigid space is extremal.

Apparently, the Banach–Kantorovich spaces $E_{X,\rho}$ corresponding to linearly rigid metric spaces (X,ρ) have not been studied, and they are undoubtedly of interest. For example, if $(\mathbb{U},\rho)$ is the universal Urysohn space, then, as was observed in [13], $E_{\mathbb{U},\rho}$ is a universal Banach space, i.e., every separable space can be linearly isometrically embedded into it. This follows from a strong theorem of Godefroy and Kalton [5], which states that if some separable Banach space F has an isometric embedding into a Banach space B, then it also has a linear isometric embedding into B. However, $E_{\mathbb{U},\rho}$ is not a homogeneous universal space — linearly isometric finite-dimensional subspaces in this space must not necessarily be sent to each other by a linear isometry, and even linear ϵ-isometry, of the whole space, as is the case for the Gurariy space [7, 12]. Hence the space $E_{\mathbb{U},\rho}$ is not isometric to the Gurariy space. The authors do not know any characterization of $E_{\mathbb{U},\rho}$ as a Banach space. The same holds for the spaces $E_{\mathbb{QU}_{\geq 1}}$ and $E_{\mathbb{ZU}}$. Such a characterization is undoubtedly of interest.

6 Proof of the main result from a dual point of view.

As promised, we give another proof of Theorem 2; to explain it, we need to set some notation. First, if (X,x_0) is a pointed metric space and B is a Banach space, we say that (X,x_0) is *embedded in* $(B,0)$ if X is isometrically embedded in B in such a way that x_0 is mapped to 0. The norm on the linear span of (X,x_0) is then said to be *compatible with* the metric on (X,x_0). This is the same definition as before, except that now we specify which δ_x has norm equal to 0; the reason why we have to consider pointed metric spaces here is that we want to use the dual space of the Kantorovich space of X, and it depends on the choice of a constant. Different choices of that constant lead to isometric linear structures, but one needs it to write down formulas. If (X,x_0) is a pointed metric space then we denote by $(X,x_0)'$ the set of all 1-Lipschitz maps f on X such that $f(x_0)=0$ (this is the unit ball in the dual to the Kantorovich space of (X,x_0)). Recall that $f\in(X,x_0)'$ is *extremal* if it is an extreme point of that convex set. In that case, one can see that f is extremal if, and only if, one may change the indices $X=\{x_0,x_1,\dots,x_n\}$ in such a way that there exists $j\leq n$ such that one of the following things happen:

$f(x_i)=\rho(x_0,x_i)$ for all $i\le j$, and $f(x_i)=\sup\{\rho(x_0,x_k)-\rho(x_i,x_k)\colon k\le j\}$, or
$f(x_i)=-\rho(x_0,x_i)$ for all $i\le j$, and $f(x_i)=\inf\{\rho(x_0,x_k)+\rho(x_i,x_k)\colon k\le j\}$. The first line means that f takes values as big as possible on $x_1,\dots,x_j$, then values as small as possible (knowing the first j values) on $x_{j+1},\dots,x_n$; the second line means that f takes values as small as possible on $x_1,\dots,x_j$, then values as large as possible (knowing the first j values) on $x_{j+1},\dots,x_n$.

We may now proceed with the proof; its principle is to look at the dual formula for the Kantorovich norm. Given a pointed metric space (X,x_0), and a vector $v=\sum a_x\cdot\delta_x\in V(X)$, that formula is

$$\|v\|=\sup\{|\sum a_x f(x)|\colon f\in(X,x_0)'\}\ .$$

As explained before, the choice of the point $x_0\in X$ is not important (different choices of x_0 yield isometric linear structures). Furthermore, it is easy to see (because of the Hahn-Banach theorem) that any norm $\|\cdot\|'$ compatible with the metric on (X,x_0) is defined by a similar formula, namely

$$\|v\|'=\sup\{|\sum a_x f(x)|\colon f\in N_{\|\cdot\|'}\}\,,$$

where $N_{\|\cdot\|'}$ is some subset of $(X,x_0)'$ such that for any $x,y\in X$ there exists $f\in N_{\|\cdot\|'}$ satisfying $|f(x)-f(y)|=\rho(x,y)$; one can think of $N_{\|\cdot\|'}$ as being the unit ball of the dual space to the Banach space in which X is embedded. Then, one possible way of understanding linear rigidity is to look at what one might say about such subsets $N_{\|\cdot\|'}$ of $(X,x_0)'$, which we call below *norming sets* of X. In particular, it is natural to wonder which maps come close to realizing the supremum which appears in the definition of the Kantorovich norm on V; this is where extremal maps come into the picture.

Indeed, a consequence of the duality theorem is that, for any extremal map on a finite subset $\{x_1,\dots,x_n\}$ of X, there exists a linear combination of $x_1,\dots,x_n$ such that the only way of coming close to realizing the sup in the definition of the Kantorovich norm of that combination is to pick a map g close to $\pm f$. In turn, this means that extremal maps have to somehow "appear" in any norming set of X if X is to be linearly rigid.

Given a pointed metric space (X,x_0) define $f_{x,y}\colon X\to\mathbb{R}$ $(x,y\in X^2)$ by

$$f_{x,y}(z)=\frac{\rho(x,z)-\rho(y,z)}{2}+\frac{\rho(y,x_0)-\rho(x,x_0)}{2}\ .$$

Then $|f_{x,y}(x) - f_{x,y}(y)| = \rho(x,y)$ for any $x, y \in X$, so the discussion above shows that the norm $\|\cdot\|'$ on V defined by

$$\|v\|' = \sup\{|\sum a_x f_{y,z}(x)| \colon (y,z) \in X^2\}$$

is compatible with the metric on (X, x_0). Pick now $x_1, \dots, x_n \in X$.
Assume now that X is linearly rigid; then we know that, because of the property of extremal maps explained above, for any $\varepsilon > 0$ and any extremal map f on $\{x_0, x_1, \dots, x_n\}$ there must exist some $(x,y) \in X^2$ and $\delta = \pm 1$ such that $|f_{x,y}(x_i) + \delta f(x_i)| \leq \varepsilon$ for all $i = 1, \dots, n$.
Reindexing $x_1, \dots, x_n$ and replacing f by $-f$ if necessary, we may assume that there exists $j \leq n$ such that $f(x_i) = \rho(x_0, x_i)$ for all $i \leq j$ and $f(x_i) = \sup\{\rho(x_0, x_k) - \rho(x_k, x_i) \colon k \leq j\}$ for all $i > j$. Exchanging x and y if necessary, we assume that $\delta = -1$ above.

Claim. Pick $\varepsilon > 0$, and let $x_1, \dots, x_n$, f and $f_{x,y}$ be as above. Then there exist constants c, c' such that $|\rho(x, x_i) - (c + f(x_i))| \leq 6\varepsilon$ and $|\rho(y, x_i) - (c' - f(x_i))| \leq 6\varepsilon$ for all $i = 1, \dots, n$.

The inequalities above are not optimal but are sufficient for our proof to work; given the discussion before the claim, it is clear that proving it is enough to conclude the proof of the "only if" part of Theorem 2. Indeed, it shows that (X, x_0) can only be linearly rigid if for all $x_1, \dots, x_n \in X$, all extremal $f \in (\{x_1, \dots, x_n\}, x_0)'$ and all $\varepsilon > 0$ there exist x, y and constants c, c' such that $|c + f(x_i) - \rho(x, x_i)| \leq \varepsilon$ and $|f(x_i) - c' + \rho(y, x_i)| \leq \varepsilon$ for all $i = 1, \dots, n$. We are only interested in proving the existence of x; it turns out that for the proof to work we have to prove the existence of x and y at the same time (we say more on this after the proof of the Claim).

Proof of the Claim. For $i \leq j$ we have $f_{x,y}(x_i) \geq \rho(x_0, x_i) - \varepsilon$. Given the definition of $f_{x,y}$, this means that

$$\rho(x, x_i) + \rho(y, x_0) \geq 2\rho(x_0, x_i) + \rho(y, x_i) + \rho(x, x_0) - 2\varepsilon, \text{ so}$$

$$\rho(x, x_i) + \rho(y, x_0) \geq (\rho(x_0, x_i) + \rho(x, x_0)) + (\rho(x_0, x_i) + \rho(y, x_i)) - 2\varepsilon \;.$$

The triangle inequality then implies that $|\rho(x, x_i) - (\rho(x, x_0) + \rho(x_0, x_i))| \leq 2\varepsilon$, and $|\rho(y, x_i) - (\rho(y, x_0) - \rho(x_0, x_i))| \leq 2\varepsilon$. Thus, setting $c = \rho(x, x_0)$ and $c' = \rho(y, x_0)$, we get a better inequality than what we wanted for $i \leq j$ (recall

that $f(x_i) = \rho(x_0, x_i)$ for all $i \leq j$).
Now, if $i \geq j+1$, one has $f(x_i) = \rho(x_0, x_k) - \rho(x_k, x_i)$ for some $k \leq j$. This implies that $f_{x,y}(x_i) \leq \rho(x_0, x_k) - \rho(x_k, x_i) + \varepsilon$. Because of the definition of $f_{x,y}$, this means that

$$\rho(x_k, x_i) - \rho(y, x_i) + \rho(y, x_0) - \rho(x, x_0) \leq 2\rho(x_0, x_k) - 2\rho(x_i, x_k) + 2\varepsilon \ .$$

An equivalent form of this inequality is

$$(\rho(x, x_i)+\rho(x_i, x_k))+(\rho(y, x_0)-\rho(x_0, x_k)+\rho(x_i, x_k)) \leq (\rho(x_0, x_k)+\rho(x, x_0))+\rho(y, x_i)+2\varepsilon \ .$$

Using the inequalities that we proved above for $k \leq j$, we finally obtain that

$$(\rho(x, x_i) + \rho(x_i, x_k)) + (\rho(y, x_k) + \rho(x_k, x_i)) \leq \rho(x, x_k) + \rho(y, x_i) + 6\varepsilon$$

This yields $\rho(x, x_i) + \rho(x_i, x_k) \leq \rho(x, x_k) + 6\varepsilon$ and $\rho(y, x_k) + \rho(x_k, x_i) \leq \rho(y, x_i) + 6\varepsilon$. This implies that $\rho(x, x_i) \leq \rho(x, x_k) - \rho(x_i, x_k) + 6\varepsilon \leq \rho(x, x_0) + \rho(x_0, x_k) - \rho(x_0, x_i) + 6\varepsilon = c + f(x_i) - 6\varepsilon$; also, $\rho(x, x_i) \geq \rho(x, x_k) - \rho(x_k, x_i) \geq \rho(x, x_0) + \rho(x_0, x_k) - 2\varepsilon - \rho(x_k, x_i) \geq c + f(x_i) - 2\varepsilon$. Put together, these two inequalities give us what we wanted for x; a similar proof works for y. $\square$

Notice that if we only had used functions $f_x(z) = \rho(x, z) - \rho(x, 0)$, then we would have obtained, using a similar proof, that if X is linearly rigid and f is any extremal map in $(\{x_0, x_1, \ldots, x_n\}, x_0)'$ then there has to exist some x and some constant c such that $\rho(x, x_i)$ is close to $c + \delta f(x_i)$ for all i, where $\delta = \pm 1$. The problem is that this condition is (at least formally) weaker than the one in Theorem 2; this is why we need to use two points to define the functions in our norming set.

We have now finished the proof of the "only if" part of Theorem 2; one can prove the "if" part using similar ideas, which we do with help from the two lemmas below.

Lemma 3. *Assume that X satisfies the hypothesis of Theorem 2 and that (X, x_0) is embedded in some Banach space $(B, 0)$. Then for any $\varepsilon > 0$ and any $x_1, \ldots, x_n \in X$ there exists a linear functional φ on B such that $\|\varphi\| = 1$ and $\varphi(x_i) \geq \varphi(x_1) + \rho(x_1, x_i) - \varepsilon$, for all $i = 1, \ldots, n$. This can also be written as $|\varphi(x_i) - (\varphi(x_1) + \rho(x_1, x_i))| \leq \varepsilon$ for all $i = 1, \ldots, n$.*

Proof. Since X satisfies the hypothesis of Theorem 2, there exists $z \in X$ such that $\rho(z, x_1) \geq \rho(z, x_i) + \rho(x_i, x_1) - \varepsilon$ for all $i = 1, \ldots, n$. The Hahn-Banach theorem ensures that there exists some linear functional φ such that $\|\varphi\| = 1$ and $\varphi(z) = \varphi(x_1) + \rho(x_1, z)$. Hence, one has $\varphi(x_i) \geq \varphi(z) - \rho(z, x_i) = \varphi(x_1) + \rho(x_1, z) - \rho(z, x_i) \geq \varphi(x_1) + \rho(x_1, x_i) - \varepsilon$. $\square$

Lemma 4. *Assume that X satisfies the hypothesis of Theorem 2, and that (X, x_0) is embedded in some Banach space $(B, 0)$. Then for any $x_1, \ldots, x_n \in X$ and any extremal map f on $\{x_0, x_1, \ldots, x_n\}$ there exists a linear functional φ such that $\|\varphi\| = 1$ and $|\varphi(x_i) - f(x_i)| \leq \varepsilon$ for all $i = 1, \ldots, n$.*

Proof. Pick $\varepsilon > 0$; set $x_0 = 0$ and pick some $y \in X$ such that $|\rho(y, x_i) - c + f(x_i)| \leq \varepsilon$ for some $c \in \mathbb{R}$ and all $i = 0, \ldots, n$. Then apply Lemma 3 to $y, 0, x_1, \ldots, x_n$ (in that order). This yields φ such that $\|\varphi\| = 1$, $|\varphi(0) - (\varphi(y) + \rho(y, 0))| \leq \varepsilon$, and $|\varphi(x_i) - (\varphi(y) + \rho(y, x_i))| \leq \varepsilon$ for all $i = 1, \ldots, n$. The first inequality gives $|\varphi(y) + \rho(y, 0)| \leq \varepsilon$, hence $|\varphi(y) + c| \leq 2\varepsilon$. Then the second inequality yields that $|\varphi(x_i) - (-c + \rho(y, x_i)| \leq 3\varepsilon$; the definition of y shows that this implies $|\varphi(x_i) - f(x_i)| \leq 4\varepsilon$ for all $i = 1, \ldots, n$. Since $\varepsilon > 0$ was arbitrary, this concludes the proof. $\square$

Lemma 4 proves the "if" part in Theorem 2: assume that X satisfies the hypothesis of Theorem 2, that (X, x_0) is embedded in $(B, 0)$, and pick $x_1, \ldots, x_n \in X$. Then $\|\sum a_i x_i\| = \sup\{\sum a_i \varphi(x_i) \colon \|\varphi\| = 1, \varphi \in B^*\}$, where B^* is the dual space of B. Lemma 4 then gives $\|\sum a_i x_i\| \geq \sup\{\sum a_i f(x_i) \colon f \in (\{x_0, x_1, \ldots, x_n\}, x_0)' \colon f \text{ is extremal } \big)$, and the right-hand sup is equal to the Kantorovich norm or $\sum a_i x_i$. So any norm compatible with the metric on (X, x_0) has to be bigger than the Kantorovich norm; we know that the Kantorovich norm is maximal among compatible norms, hence this proves that all norms compatible with the metric are equal.

Acknowledgment. The authors are grateful to V.Pestov for useful discussions.

References

[1] D. Avis, "On the extreme rays of the metric cone," *Canad. J. Math.*, **32**, No. 1, 126–144 (1980).

[2] L. A. Caffarelli and S. Salsa (eds.), *Optimal Transportation and Applications*, Lectures from the C.I.M.E. Summer School held in Martina

Franca, September 2–8, 2001. *Lecture Notes in Math.*, **1813**. Springer-Verlag, Berlin, 2003.

[3] P. Cameron, "Homogeneous Cayley objects," *European J. Combin.*, **21**, 745–760 (2000).

[4] P. Cameron and A. M. Vershik, "Some isometry groups of the Urysohn space," *Ann. Pure Appl. Logic*, **143**, Nos. 1–3, 70–78 (2006).

[5] G. Godefroy and N. Kalton, "Lipschitz-free Banach spaces," *Studia Math.*, **159**, No. 1, 121–141 (2003).

[6] M. Gromov, *Metric Structures for Riemannian and Non-Riemannian Spaces*, Birkhäuser Boston, Boston, 1999.

[7] V. I. Gurariy, "Spaces of universal placement, isotropic spaces and a problem of Mazur on rotations of Banach spaces," *Sibirsk. Mat. Zh.*, **7**, 1002–1013 (1966).

[8] R. Holmes, "The universal separable metric space of Urysohn and isometric embeddings thereof in Banach spaces," *Fund. Math.*, **140**, No. 3, 199–223 (1992).

[9] L. V. Kantorovich, "On the translocation of masses," *Dokl. Akad. Nauk SSSR*, **37**, Nos. 7–8, 227–229 (1942).

[10] L. V. Kantorovich and G. Sh. Rubinshtein, "On a space of totally additive functions," *Vestn. Leningr. Univ.*, **13**, No. 7, 52–59 (1958).

[11] M. Katětov, "On universal metric spaces," in: *General Topology and Its Relations to Modern Analysis and Algebra*, VI (Prague, 1986), Helderman, Berlin, 1988, pp. 323–330.

[12] W. Lusku, "The Gurarij spaces are unique," *Arch der Mathematik*, **27**, No. 6, 627–635 (1976).

[13] J. Melleray, "Géométrie de l'espace d'Urysohn et théorie descriptive des ensembles", Ph.D Thesis, Paris, 2005.

[14] P. S. Urysohn, "Sur un espace metrique universel," *Bull. Sci. Math.*, **51**, 1–28 (1927).

[15] A. M. Vershik, “Random metric spaces and universality,” *Russian Math. Surveys*, **59**, No. 2, 259–295 (2004).

[16] A. M. Vershik, “Kantorovich metric: initial history and little-known applications”, *J. Math. Sci.*, **133**, No. 4, 1410–1417 (2006).

[17] C. Villani, *Topics in Optimal Transportation*, Amer. Math. Soc., Providence, RI (2003).