Geom. Funct. Anal. Vol. 28 (2018) 1641–1705
https://doi.org/10.1007/s00039-018-0467-6
Published online September 14, 2018

GAFA Geometric And Functional Analysis

SHARP FINITENESS PRINCIPLES FOR LIPSCHITZ SELECTIONS

Charles Fefferman and Pavel Shvartsman

Abstract. Let $(\mathcal{M}, \rho)$ be a metric space and let Y be a Banach space. Given a positive integer m, let F be a set-valued mapping from $\mathcal{M}$ into the family of all compact convex subsets of Y of dimension at most m. In this paper we prove a finiteness principle for the existence of a Lipschitz selection of F with the sharp value of the finiteness constant.

Contents

This research was supported by Grant No 2014055 from the United States-Israel Binational Science Foundation (BSF). Charles Fefferman was also supported in part by NSF Grant DMS-1265524 and AFOSR Grant FA9550-12-1-0425.

Keywords and phrases: Set-valued mapping, Lipschitz selection, Metric tree, Helly's theorem, Nagata dimension, Whitney partition, Steiner-type point

Mathematics Subject Classification: 46E35

1 Introduction

We prove a finiteness theorem for *Lipschitz selection problems*, conjectured by Yu. Brudnyi and Shvartsman [BS94, Shv02] and established in special cases by Fefferman, Israel and Luli [FIL16a, FIL17] and Shvartsman [Shv86, Shv92, Shv01, Shv02, Shv04].

In its simplest setting, our problem is as follows. We are given a metric space $(\mathcal{M}, \rho)$ and a positive integer m. For each point $x \in \mathcal{M}$, we are given a nonempty compact convex set $F(x) \subset \mathbb{R}^m$.

We want to find a Lipschitz map $f : \mathcal{M} \to \mathbb{R}^m$ such that $f(x) \in F(x)$ for all $x \in \mathcal{M}$. Such an f is called a *Lipschitz selection* of the set-valued map $F : \mathcal{M} \to \mathcal{K}(\mathbb{R}^m)$, where $\mathcal{K}(\mathbb{R}^m)$ denotes the family of all nonempty compact convex subsets of $\mathbb{R}^m$. If a Lipschitz selection f exists, then we ask how small we can take its Lipschitz seminorm.

In this setting, our main result implies the following:

Theorem 1.1. *Let $(\mathcal{M}, \rho)$ be a metric space, let $F : \mathcal{M} \to \mathcal{K}(\mathbb{R}^m)$, and let λ be a positive real number. Suppose that for every $\mathcal{M}' \subset \mathcal{M}$ consisting of at most 2^m points, the restriction $F|_{\mathcal{M}'}$ of F to $\mathcal{M}'$ has a Lipschitz selection $f_{\mathcal{M}'}$ with Lipschitz seminorm at most λ.*

Then F has a Lipschitz selection with Lipschitz seminorm at most $\gamma\lambda$. Here, γ depends only on the dimension m.

Equivalently, we may suppose that $\mathcal{M}$ contains at least 2^m points and take $\mathcal{M}'$ to contain *exactly* 2^m points.

Lipschitz selection problems are closely related to

Whitney's Extension Problem ([Whi34]) *Fix $m, n \geq 1$, and let f be a real-valued function defined on a given (arbitrary) closed set $E \subset \mathbb{R}^n$. Decide whether f extends to a function $F \in C^m(\mathbb{R}^n)$ with a finite C^m-norm.*

If such an extension F exists, then how small can we take its C^m-norm?

There is a finiteness principle for Whitney's Extension Problem, e.g. when $E \subset \mathbb{R}^n$ is a large finite set. See Brudnyi-Shvartsman [BS94, BS98, BS01, Shv82, Shv87, Shv02, Shv08] and the later papers of Fefferman, Israel, Klartag and Luli [Fef05a, Fef05b, Fef06, FK09a, FK09b, Fef09a, Fef09b, FIL16a, FIL17], as well as A. and Yu. Brudnyi [BB12] for that finiteness principle and several related results.

The idea of Lipschitz selection first arose in connection with Whitney's extension problem, see [BB12, BS98, BS94, BS01, Shv82, Shv84, Shv87]. In particular, a variant of a special case of Theorem 1.1 was the main ingredient in the proof [BS98, BS01, Shv82, Shv87, Shv02] of the finiteness principle for Whitney's Problem in the simplest non-trivial case, $m = 2$. The later papers [Fef05a, Fef05b, Fef06, FK09a, FK09b, Fef09a, FIL16a, FIL17] didn't explicitly mention Lipschitz selection, but they broadened Whitney's Problem by asking for functions $F \in C^m(\mathbb{R}^n)$ that agree with f on E to a given accuracy.

Of course, a Lipschitz selection problem may also be regarded as a search for a smooth function that agrees approximately with data.

Our main result is more general than Theorem 1.1. First of all, we allow $(\mathcal{M}, \rho)$ to be a *pseudometric space*, i.e., $\rho : \mathcal{M} \times \mathcal{M} \to [0, +\infty]$, $\rho(x, x) = 0$, $\rho(x, y) = \rho(y, x)$, $\rho(x, y) \le \rho(x, z) + \rho(z, y)$ for all $x, y, z \in \mathcal{M}$. Note that $\rho(x, y) = 0$ may hold with $x \ne y$, and $\rho(x, y)$ may be $+\infty$.

Secondly, the convex sets $F(x)$ needn't sit inside $\mathbb{R}^m$. Instead, we fix a Banach space $(Y, \|\cdot\|)$ and let $\mathcal{K}_m(Y)$ denote the family of all nonempty compact convex subsets $K \subset Y$ of dimension at most m. (We say that a convex subset of Y has dimension at most m if it is contained in an affine subspace of Y of dimension at most m.)

We write

$$N(m, Y) = \min\{2^{m+1}, 2^{\dim Y}\} \quad \text{if} \quad Y \quad \text{is finite-dimensional,} \tag{1.1}$$

and

$$N(m, Y) = 2^{m+1} \quad \text{if} \quad Y \quad \text{is infinite-dimensional.} \tag{1.2}$$

We define the Lipschitz seminorm of a map $f : \mathcal{M} \to Y$ for a Banach space Y and a pseudometric space $(\mathcal{M}, \rho)$ by setting

$$\|f\|_{\mathrm{Lip}(\mathcal{M},Y)} = \inf\{\, \lambda > 0 : \|f(x) - f(y)\| \le \lambda\, \rho(x, y) \quad \text{for all} \quad x, y \in \mathcal{M}\}.$$

In particular, $\|f\|_{\mathrm{Lip}(\mathcal{M},Y)} = +\infty$ if no such λ exists.

We can now state our main result in full generality; Theorem 1.1 will be a simple consequence.

Theorem 1.2. *Fix $m \ge 1$. Let $(\mathcal{M}, \rho)$ be a pseudometric space, and let $F : \mathcal{M} \to \mathcal{K}_m(Y)$ for a Banach space Y. Let λ be a positive real number.*

Suppose that for every $\mathcal{M}' \subset \mathcal{M}$ consisting of at most $N = N(m, Y)$ points, the restriction $F|_{\mathcal{M}'}$ of F to $\mathcal{M}'$ has a Lipschitz selection $f_{\mathcal{M}'}$ with Lipschitz seminorm $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}(\mathcal{M}',Y)} \le \lambda$.

Then F has a Lipschitz selection f with Lipschitz seminorm $\|f\|_{\mathrm{Lip}(\mathcal{M},Y)} \le \gamma\lambda$.

Here, γ depends only on m.

The "finiteness constants" 2^m in Theorem 1.1 and $N(m, Y)$ in Theorem 1.2 are optimal; see [Shv92] and[Shv02, Theorem 1.4]. We also refer the reader to the paper [FS, Section 8.1], which contains detailed proofs of this statement for $m = 1, 2$.

If the set $\mathcal{M}$ is *finite* in Theorem 1.1 or Theorem 1.2, then we can omit the assumption that the convex sets $F(x)$ $(x \in \mathcal{M})$ are compact. In this case, it is enough to assume that $F : \mathcal{M} \to \mathrm{Conv}_m(Y)$, where

$$\mathrm{Conv}_m(Y) = \{\text{ all nonempty convex subsets of } Y \text{ of dimension at most } m\}. \tag{1.3}$$

See Theorem 6.2.

For the case of the trivial distance function $\rho \equiv 0$, Theorems 1.2 and 6.2 agree with the classical Helly's Theorem [DGK63], except that the optimal finiteness constant for $\rho \equiv 0$ is

$$n(m, Y) = \min\{m + 2, \dim Y + 1\} \quad \text{in place of} \quad N(m, Y) = \min\{2^{m+1}, 2^{\dim Y}\}. \tag{1.4}$$

Thus, our results may be regarded as a generalization of Helly's Theorem. However, we make extensive use of Helly's Theorem in our proofs.

Theorem 1.2 and its variants were previously known in several special cases:

- $Y = \mathbb{R}^2$ [Shv02];
- Each $F(x)$ $(x \in \mathcal{M})$ is an affine subspace of Y of dimension at most m [Shv86, Shv92] $(Y = \mathbb{R}^m)$, [Shv01] (Y is a Hilbert space), [Shv04] (Y is a Banach space). Of course, all $F(x)$ are non-compact in this case;
- $(\mathcal{M}, \rho) = (\mathbb{R}^n, \|\cdot\|)$ and $Y = \mathbb{R}^m$ with the constant N and the constant γ depending on n as well as on m [FIL16a].

Let us recount how we arrived at our proof of Theorem 1.2. P. Shvartsman (unpublished) had already reduced Theorem 1.2 to the special case of a *metric tree* with nodes of bounded degree. We recall the relevant standard definitions.

Let $T = (X, E)$ be a finite (graph theoretic) tree, where X denotes the set of nodes of T, and E denotes the set of edges. The *degree* of a node $x \in X$ is the number of nodes y to which x is joined by an edge.

Suppose we assign a positive number $\Delta(e)$ to each edge $e \in E$. Then for $x, y \in X$ we can define their *distance* $d(x, y)$ to be the sum of $\Delta(e)$ over all the edges e in the "minimal path" joining x to y as in Figure 1.

We call d a *tree metric*; (X, d) is a *metric tree.*

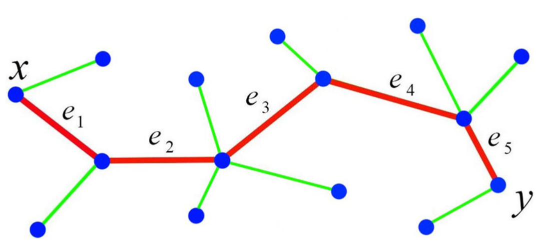

Figure 1: A minimal path joining nodes x and y in a tree. In this case, $d(x, y) = \Delta(e_1) + \Delta(e_2) + \cdots + \Delta(e_5)$.

Shvartsman's unpublished previous work reduced Theorem 1.2 to the following weakened form of a special case.

(Conjectured) Theorem 1.3. *Given $m \geq 1$, there exist $k^\sharp, \gamma$ depending only on m, for which the following holds.*

Let (X, d) be a metric tree in which each node has degree at most $m+1$.

Let $F : X \to \mathcal{K}_m(Y)$ for a Banach space Y, and let λ be a positive real number. Suppose that for every subset $X' \subset X$ consisting of at most $k^\sharp$ points, the restriction $F|_{X'}$ has a Lipschitz selection $f_{X'}$ with Lipschitz seminorm $\|f_{X'}\|_{\mathrm{Lip}(X',Y)} \leq \lambda$.

Then F has a Lipschitz selection f with Lipschitz seminorm $\|f\|_{\mathrm{Lip}(X,Y)} \leq \gamma\lambda$.

Remark 1.4. Note that here X' needn't be a subtree of X. Thus, in Figure 1, perhaps X' contains the nodes x and y but not the nodes that lie between them.

Note also that the optimal finiteness constant $N(m, Y)$ in Theorem 1.2 has been replaced in Theorem 1.3 by a sufficiently large constant $k^\sharp$ depending only on m.

On the other hand, the work of Fefferman, Israel and Luli [FIL16a] on "C^m Selection" implies a weakened version of Theorem 1.1, in which $(\mathcal{M}, \rho)$ is $\mathbb{R}^n$ with its standard Euclidean metric; the sharp finiteness constant 2^m in Theorem 1.1 is replaced by $k^\sharp$ as in Theorem 1.3; and the constant γ is allowed to depend on n as well as on m. See the web posting [FIL16b].

To prove Theorem 1.2, we set out to adapt the arguments in [FIL16a] from $\mathbb{R}^n$ to the setting of a metric tree. If we succeeded, Theorem 1.3 would follow, thus proving Theorem 1.2.

This attempt seemed highly unlikely to succeed; the geometry of a metric tree is of course radically different from that of $\mathbb{R}^n$. Nevertheless, we were able to adapt [FIL16a] and prove Theorem 1.3, thanks to one crucial similarity between $\mathbb{R}^n$ and metric trees - they have finite *Nagata dimension*. We recall the relevant definitions (see [Nag58, Ass82, LS05, BDHM09]).

Definition 1.5. *Let (X, d) be a metric space. Let D be a non-negative integer and let c be a positive real constant. We say that (X, d) satisfies Nagata (D, c) if for every real number $s > 0$ there exists a covering $(X_i)_{i \in I}$ of X by subsets X_i of diameter at most s, such that no ball of radius cs in X meets more than $D+1$ of the X_i. We call D, c the Nagata constants of (X, d).*

The least D for which (X, d) satisfies Nagata (D, c) for some $c > 0$ is the Nagata dimension (or Assouad-Nagata dimension) of (X, d).

Note that any *finite* metric space has Nagata dimension 0. The metric space $\mathbb{R}^n$ has Nagata dimension n ([LS05]). The space ℓ_∞ has infinite Nagata dimension because ℓ_∞ contains $\mathbb{R}^n$ for each $n \in \mathbb{N}$. Every *planar* connected graph whose nodes have finite degree has Nagata dimension at most $2^{10} - 1$ (see Ostrovskii, Rosenthal [OR15] for the precise statement and the proof).

Moreover, every *metric tree* satisfies Nagata $(1, c)$ for an absolute constant c. (See [LS05, Lemma 3.1 and Theorem 3.2]. For the reader's convenience, in Lemma 4.15

we prove that one can take $c = 1/16$.) This allows us to carry over arguments in [FIL16a] from $\mathbb{R}^n$ to an arbitrary metric tree.

More precisely, we prove the following result.

Theorem 1.6. *Given $m \ge 1$ there exists $k^\sharp$ depending only on m, for which the following holds.*

Let (X, d) be a finite metric space satisfying Nagata (D, c), and let $F : X \to \mathrm{Conv}_m(Y)$ for a Banach space Y. Let λ be a positive real number. Suppose that for every $X' \subset X$ consisting of at most $k^\sharp$ points, the restriction $F|_{X'}$ has a Lipschitz selection $f_{X'}$ with Lipschitz seminorm $\|f_{X'}\|_{\mathrm{Lip}(X',Y)} \le \lambda$.

Then F has a Lipschitz selection f with Lipschitz seminorm $\|f\|_{\mathrm{Lip}(X,Y)} \le \gamma\lambda$, where γ depends only on m and on the Nagata constants D, c.

Recall that $\mathrm{Conv}_m(Y)$ denotes the family of all nonempty convex subsets of Y of dimension at most m (see (1.3)).

As an immediate corollary, we obtain a stronger form of Theorem 1.3 in which we drop the assumption that each node has degree at most $m + 1$. See Corollary 4.16. So we have proven more than we need to establish Theorem 1.2. Because we needn't assume that the nodes of our metric tree have degree at most $m+1$, we can greatly simplify the earlier reduction of Theorem 1.2 to the case of metric trees.

This paper is organized as follows.

In Section 2 we construct "Whitney partitions of unity" associated to a "lengthscale" $r(x) > 0$ defined on a metric space of finite Nagata dimension. As in H. Whitney's classic paper [Whi34], such partitions are used to patch together functions defined in neighborhoods of varying sizes, while maintaining the smoothness of the functions being patched.

In Sections 3 and 4 we associate to a Lipschitz selection problem given by $F : \mathcal{M} \to \mathcal{K}_m(Y)$ a family of convex sets $\Gamma_\ell(x) \in \mathcal{K}_m(Y)$ parametrized by $x \in \mathcal{M}$ and $\ell \ge 0$. If for every subset $\mathcal{M}' \subset \mathcal{M}$ consisting of at most $k^\sharp(\ell, m)$ points there exists a Lipschitz selection $f_{\mathcal{M}'}$ of $F|_{\mathcal{M}'}$ with Lipschitz seminorm $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}(\mathcal{M}',Y)} \le \lambda$, then $\Gamma_\ell(x)$ is nonempty. That is how we use the hypothesis of Theorem 1.6.

As in [FIL16a, FIL16b], we use the $\Gamma_\ell(x)$ in Section 4 to prove Theorem 1.6. This is the most technically difficult part of our proof. The idea is to measure the difficulty of a Lipschitz selection problem by examining the size and shape of the $\Gamma_\ell(x)$. We proceed by induction on the difficulty of the problem, reducing hard cases to easier ones by first localizing to the correct lengthscale, then patching together local Lipschitz selections by a Whitney partition of unity. By the end of Section 4.10 we will have proven Theorem 1.6 and deduced Corollary 4.16, the strengthened version of Theorem 1.3 on metric trees (without any assumption of the degree of the nodes).

In Section 5 we return to the setting of a general metric space $(\mathcal{M}, \rho)$ and a map $F : \mathcal{M} \to \mathcal{K}_m(Y)$. We suppose that for every $\mathcal{M}' \subset \mathcal{M}$ consisting of at most $k^\sharp$ points, the restriction $F|_{\mathcal{M}'}$ has a Lipschitz selection with Lipschitz seminorm at most λ. Here, $k^\sharp$ is the same constant as in Theorem 1.6.

For each $x \in \mathcal{M}$, we define a nonempty compact convex "core"

$$G(x) \subset F(x), \tag{1.5}$$

with the following crucial property:

For every $x, y \in \mathcal{M}$, the *Hausdorff distance* from $G(x)$ to $G(y)$ is at most $\gamma_0 \lambda\, \rho(x, y)$. (1.6)

Here, γ_0 depends only on m.

Recall that the Hausdorff distance $\mathrm{d_H}(A, B)$ between two nonempty compact sets $A, B \subset Y$ is defined as the least $r \geq 0$ such that for each $x \in A$ there exists $y \in B$ such that $\|x - y\| \leq r$, and for each $x \in B$ there exists $y \in A$ such that $\|x - y\| \leq r$.

We define $G(x)$ by considering an arbitrary finite tree $T = (X, E)$ ($X = \{\text{nodes}\}$, $E = \{\text{edges}\}$) and an arbitrary map $\psi : X \to \mathcal{M}$ such that

$$\psi(x) \neq \psi(y) \quad \text{whenever} \quad x, y \in X \quad \text{are joined by an edge.} \tag{1.7}$$

We refer to ψ as an *admissible* mapping. (See Definition 5.4.) The map ψ induces a tree metric d on X by setting $d(x, y) = \rho(\psi(x), \psi(y))$ whenever x and y are nodes in X joined by an edge.

Moreover, we obtain a Lipschitz selection problem for the metric tree (X, d) by considering the map $F \circ \psi : X \to \mathcal{K}_m(Y)$. From Corollary 4.16 (i.e., Theorem 1.3 in its strengthened form), we learn that $F \circ \psi$ has a Lipschitz selection with Lipschitz seminorm at most $\gamma_0\, \lambda$. By considering all such Lipschitz selections for a fixed $T = (X, E)$, a node $a \in X$, and a map $\psi : X \to \mathcal{M}$ (satisfying (1.7)) such that $\psi(a) = x$, we define a nonempty compact convex set

$$O(x; [T, a, \psi]) \subset F(x) \quad \text{for each} \quad x \in \mathcal{M}.$$

(See Section 5.1 for the definition of the sets $O(x; [\cdot, \cdot, \cdot])$.)

The "core" $G(x)$ is then defined as the intersection of the sets $O(x; [T, a, \psi])$ over all finite trees $T = (X, E)$, all nodes $a \in X$, and all $\psi : X \to \mathcal{M}$ with $\psi(a) = x$ satisfying (1.7). The key properties (1.5), (1.6) of G follow easily once we know that $O(x; [T, a, \psi])$ is nonempty, and we easily deduce that key fact from Corollary 4.16.

Once we have produced a "core" $G(x)$ satisfying (1.5) and (1.6) (see Theorem 5.2), we can invoke a selection theorem of Shvartsman [Shv04], see Theorem 5.11. This result provides the existence of a Lipschitz (with respect to the Hausdorff distance $\mathrm{d_H}$) map $\mathrm{St} : \mathcal{K}_m(Y) \to Y$ such that $\mathrm{St}(K) \in K$ for all $K \in \mathcal{K}_m(Y)$. Furthermore, the $\mathrm{d_H}$-Lipschitz seminorm of St is bounded by a constant depending only on m. We refer to $\mathrm{St}(K)$ as "Steiner-type point" of K. See Section 5.2 for more detail.

We can now apply the Steiner-type point map St to the core G to establish the following weak form of Theorem 1.2.

Theorem 1.7. *Given $m \geq 1$ there exist constants $k^\sharp, \gamma_1$, depending only on m, for which the following holds.*

Let $(\mathcal{M}, \rho)$ be a metric space, let $F : \mathcal{M} \to \mathcal{K}_m(Y)$ for a Banach space Y, and let λ be a positive real number. Suppose that for every $\mathcal{M}' \subset \mathcal{M}$ consisting of at most $k^\sharp$ points, the restriction $F|_{\mathcal{M}'}$ has a Lipschitz selection $f_{\mathcal{M}'}$ with Lipschitz seminorm $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}(\mathcal{M}',Y)} \leq \lambda$.

Then F has a Lipschitz selection f with Lipschitz seminorm $\|f\|_{\mathrm{Lip}(\mathcal{M},Y)} \leq \gamma_1 \lambda$.

Note that we have here $k^\sharp$ instead of the sharp finiteness constant $N(m, Y)$, and that $(\mathcal{M}, \rho)$ is a metric space, rather than a pseudometric space.

To prove Theorem 1.2, it remains to pass from metric spaces to pseudometric spaces, and to pass from the large finiteness constant $k^\sharp$ to the optimal finiteness constant $N(m, Y)$.

We pass to pseudometric spaces in Section 6. In the context of Theorem 1.2, the task is easy. For Theorem 6.2, the variant of Theorem 1.2 in which $(\mathcal{M}, \rho)$ is finite but the sets $F(x)$ needn't be compact, it takes a bit more work.

Finally, we pass from $k^\sharp$ to $N(m, Y)$ by applying a result of Shvartsman [Shv02, Theorem 1.2].

Theorem 1.8. *Let $(\widetilde{\mathcal{M}}, \tilde{\rho})$ be a finite pseudometric space, let $\widetilde{F} : \widetilde{\mathcal{M}} \to \mathcal{K}_m(Y)$ for a Banach space Y, and let λ be a positive real number.*

Suppose that for every $S \subset \widetilde{\mathcal{M}}$ consisting of at most $N(m, Y)$ points, the restriction $\widetilde{F}|_S$ has a Lipschitz selection $\tilde{f}_S$ with Lipschitz seminorm $\|\tilde{f}_S\|_{\mathrm{Lip}(S,Y)} \leq \lambda$.

Then $\widetilde{F}$ has a Lipschitz selection $\tilde{f}$ with Lipschitz seminorm $\|\tilde{f}\|_{\mathrm{Lip}(\widetilde{\mathcal{M}},Y)} \leq C(\widetilde{\mathcal{M}})\,\lambda$, where $C(\widetilde{\mathcal{M}})$ depends only on m and on the number of points in $\widetilde{\mathcal{M}}$.

Note that Theorem 1.8 is the "bridge" between the existence of some finiteness constant $k^\sharp$ in Theorem 1.7 and the optimal finiteness constant $N(m, Y)$ in Theorem 1.2.

We combine Theorem 1.7 (for pseudometric spaces) with Theorem 1.8, to complete the proof of Theorem 1.2, our main result. The argument is simple: Using Theorem 1.8, we pass from $N(m, Y)$-point subsets to $k^\sharp$-point subsets; then, using Theorem 1.7, we pass from $k^\sharp$-point subsets to a full solution of our Lipschitz selection problem.

Finally, Section 7 states two variants of Theorem 1.2, and adds a few closing remarks.

As in [FIL16a], our present results lead to questions about efficient computation for Lipschitz selection problems on finite metric spaces. In connection with such issues, we ask whether the results of Har-Peled and Mendel [HM06] on the Well Separated Pairs Decomposition [CK95] can be extended from doubling metrics to metrics of bounded Nagata dimension.

Readers interested in checking details of our proofs may want to consult a much more detailed version of this paper posted on the arXiv [FS]. We mention also that

A. Brudnyi [Bru18] has advised us that he has an alternate proof of the passage from the finiteness principle for metric trees to the construction of the core.

2 Whitney Partitions and Patching Lemma

Let (X, d) be a metric space. We write $B(x, r)$ to denote the ball $\{y \in X : d(x, y) < r\}$ (strict inequality) in the metric space (X, d). We also write $\operatorname{diam} A = \sup\{d(a, b) : a, b \in A\}$ and

$$\operatorname{dist}(A', A'') = \inf\{d(a', a'') : a' \in A', a'' \in A''\}$$

to denote the diameter of a set $A \subset X$ and the distance between sets $A', A'' \subset X$ respectively.

2.1 Whitney partitions on metric spaces with finite Nagata dimension.

In this section, we prove the following result.

Whitney Partition Lemma 2.1. *Let (X, d) be a metric space, and let $r(x) > 0$ be a positive function on X. We assume the following, for constants $c_{\mathcal{N}} \in (0, 1]$, $D_{\mathcal{N}} \in \mathbb{N} \cup \{0\}$ and $C_{LS} \geq 1$:*

• (Nagata $(D_{\mathcal{N}}, c_{\mathcal{N}})$) *Given $s > 0$ there exists a covering of X by subsets X_i $(i \in I)$ of diameter at most s, such that every ball of radius $c_{\mathcal{N}} s$ in X meets at most $D_{\mathcal{N}} + 1$ of the X_i.*

• (Consistency of the Lengthscale) *Let $x, y \in X$. If $d(x, y) \leq r(x) + r(y)$, then*

$$C_{LS}^{-1} r(x) \leq r(y) \leq C_{LS} r(x). \tag{2.1}$$

Let $a > 0$.

Then there exist functions $\varphi_\nu : X \to \mathbb{R}$, and points $x_\nu \in X$, with the following properties:

- *Each $\varphi_\nu \geq 0$, and each $\varphi_\nu = 0$ outside $B(x_\nu, a r_\nu)$. Here and below, $r_\nu = r(x_\nu)$.*
- *Any given $x \in X$ satisfies $\varphi_\nu(x) \neq 0$ for at most D^* distinct ν.*
- *$\sum\limits_\nu \varphi_\nu = 1$ on X.*
- *For each ν and for all $x, y \in X$, we have*

$$|\varphi_\nu(x) - \varphi_\nu(y)| \leq C_{Wh}\, d(x, y)/r_\nu.$$

Here D^* and C_{Wh} are constants depending only on $c_{\mathcal{N}}$, $D_{\mathcal{N}}$, C_{LS} and a.

Proof. We write c, C to denote positive constants determined by $c_{\mathcal{N}}$, $D_{\mathcal{N}}$, C_{LS} and a. These symbols may denote different constants in different occurrences.

We introduce a large constant A to be fixed later. We make the following

Large A Assumption for Whitney Partitions 2.2. A exceeds a large enough constant determined by $c_{\mathcal{N}}$, $D_{\mathcal{N}}$, C_{LS}, a.

We write $c(A), C(A)$ to denote positive constants determined by A, c_N, D_N, C_{LS}, a. These symbols may denote different constants in different occurrences.

Let P denote the set of all integer powers of 2, including negative powers. For $s \in P$ let $(X(i,s))_{i\in I(s)}$ be a covering of X given by the Nagata (D_N, c_N) condition. Thus,

$$\operatorname{diam} X(i,s) \le s;$$

and, for fixed $s \in P$,

$$\text{any given} \quad x \in X \quad \text{lies in at most} \quad C \quad \text{of the sets} \quad X^{++}(i,s). \tag{2.2}$$

Here

$$X^{++}(i,s) = \{y \in X : d(y, X(i,s)) < c_N s/64\} \qquad (i \in I(s)).$$

We also define

$$X^{+}(i,s) = \{y \in X : d(y, X(i,s)) < c_N s/128\} \quad \text{for} \quad (i \in I(s)).$$

Let

$$\theta_{i,s}(x) = \max\{0, (1 - 256\, d(x, X(i,s))/(c_N s))\}$$

for $x \in X$, $i \in I(s)$, $s \in P$.

Then

$$0 \le \theta_{i,s} \le 1, \tag{2.3}$$

$$\|\theta_{i,s}\|_{\mathrm{Lip}(X,\mathbb{R})} \le C\, s^{-1}, \tag{2.4}$$

and

$$\theta_{i,s} = 0 \quad \text{outside} \quad X^{+}(i,s),$$

but

$$\theta_{i,s} = 1 \quad \text{on} \quad X(i,s). \tag{2.5}$$

For each $s \in P$ and $i \in I(s)$, we pick a representative point $x(i,s) \in X(i,s)$. (We may assume that the $X(i,s)$ are all nonempty.) We let Rel (relevant) denote the set of all (i,s) such that

$$A^{-3} r(x(i,s)) \le s \le A^{-1} r(x(i,s)). \tag{2.6}$$

We establish the basic properties of the set Rel .

Lemma 2.3. *Given $x_0 \in X$ there exists $(i,s) \in$ Rel such that $x_0 \in X(i,s)$ and therefore $\theta_{i,s}(x_0) = 1$.*

Proof. The "therefore" part of the lemma follows from (2.5).

Pick $s_0 \in P$ such that

$$s_0/2 \le r(x_0)/A^2 \le 2s_0.$$

Because the $X(i, s_0)$ $(i \in I(s_0))$ cover X, we may fix $i_0 \in I(s_0)$ such that $x_0 \in X(i_0, s_0)$. The points x_0 and $x(i_0, s_0)$ both belong to $X(i_0, s_0)$, hence

$$d(x_0, x(i_0, s_0)) \le \operatorname{diam} X(i_0, s_0) \le s_0 \le 2r(x_0)/A^2.$$

The Large A Assumption 2.2 and the CONSISTENCY OF THE LENGTHSCALE together now imply that

$$cr(x_0) \le r(x(i_0, s_0)) \le Cr(x_0),$$

and therefore

$$cs_0 \le r(x(i_0, s_0))/A^2 \le Cs_0.$$

Thanks to the Large A Assumption 2.2, we therefore have (2.6) for (i_0, s_0). Thus, $(i_0, s_0) \in$ REL and $x_0 \in X(i_0, s_0)$. □

LEMMA 2.4. *If $(i, s) \in$ REL and $x_0 \in X^{++}(i, s)$, then*

$$cA^{-3}r(x_0) \le s \le CA^{-1}r(x_0),$$

and therefore

$$\|\theta_{i,s}\|_{\mathrm{Lip}(X,\mathbb{R})} \le CA^3/r(x_0).$$

Proof. Both x_0 and $x(i, s)$ lie in $X^{++}(i, s)$, hence

$$d(x_0, x(i, s)) \le \operatorname{diam} X^{++}(i, s) \le 2c_{\mathcal{N}}s/64 + \operatorname{diam} X(i, s) \le Cs \le Cr(x(i, s))/A$$

thanks to (2.6).

The Large A Assumption 2.2 and CONSISTENCY OF THE LENGTHSCALE now tell us that

$$cr(x_0) \le r(x(i, s)) \le Cr(x_0),$$

and therefore (2.6) and (2.4) imply the conclusion of Lemma 2.4. □

COROLLARY 2.5. *Any given point $x_0 \in X$ lies in $X^{++}(i, s)$ for at most $C(A)$ distinct $(i, s) \in$ REL. Consequently, $\theta_{i,s}(x_0)$ is nonzero for at most $C(A)$ distinct $(i, s) \in$ REL.*

Proof. There are at most $C(A)$ distinct $s \in P$ satisfying the conclusion of Lemma 2.4. For each such s there are at most C distinct i such that $x_0 \in X^{++}(i, s)$; see (2.2). □

COROLLARY 2.6. *Suppose* $X^{++}(i,s) \cap X^{++}(i_0,s_0) \neq \emptyset$ *with* $(i,s),(i_0,s_0) \in \text{REL}$. *Then*

$$c(A)s_0 \le s \le C(A)s_0.$$

Proof. Pick $x_0 \in X^{++}(i,s) \cap X^{++}(i_0,s_0)$. Lemma 2.4 gives

$$c(A)r(x_0) \le s \le C(A)r(x_0) \quad \text{and} \quad c(A)r(x_0) \le s_0 \le C(A)r(x_0). \qquad \square$$

LEMMA 2.7. *Let* $(i_0,s_0),(i,s) \in \text{REL}$. *If* $x \in X^{+}(i_0,s_0)$, *then for any* $y \in X$

$$|\theta_{i,s}(x) - \theta_{i,s}(y)| \le C(A)\, d(x,y)/s_0. \tag{2.7}$$

Proof. We proceed by cases.

Case 1: $d(x,y) < c_N s_0/128$.

Then $x,y \in X^{++}(i_0,s_0)$. If x or y belongs to $X^{++}(i,s)$, then Corollary 2.6 tells us that

$$c(A)s_0 \le s \le C(A)s_0\,;$$

hence, (2.4) yields the desired estimate (2.7).

If instead neither x nor y belongs to $X^{++}(i,s)$, then $\theta_{i,s}(x) = \theta_{i,s}(y) = 0$, hence (2.7) holds trivially.

Case 2: $d(x,y) \ge c_N s_0/128$. Then (2.3) gives

$$|\theta_{i,s}(x) - \theta_{i,s}(y)| \le 1 \le C\, d(x,y)/s_0.$$

Thus, (2.7) holds in all cases. $\square$

Now define

$$\Theta(x) = \sum_{(i,s)\in\text{REL}} \theta_{i,s}(x) \quad \text{for all} \quad x \in X. \tag{2.8}$$

Corollary 2.5 shows that there are at most $C(A)$ nonzero summands in (2.8) for any fixed x. Moreover, each summand is between 0 and 1 [see (2.3)], and for each fixed x, at least one of the summands is equal to 1 (see Lemma 2.3). Therefore,

$$1 \le \Theta(x) \le C(A) \quad \text{for all} \quad x \in X. \tag{2.9}$$

LEMMA 2.8. *Let* $x,y \in X$ *and* $(i_0,s_0) \in \text{REL}$. *If* $x \in X^{+}(i_0,s_0)$, *then*

$$|\Theta(x) - \Theta(y)| \le C(A)\, d(x,y)/s_0.$$

Proof. There are at most $C(A)$ distinct $(i,s) \in \text{REL}$ for which $\theta_{i,s}(x)$ or $\theta_{i,s}(y)$ is nonzero. For each such (i,s) we apply Lemma 2.7, then sum over (i,s). $\square$

Now, for $(i_0, s_0) \in \text{REL}$, we set

$$\varphi_{i_0,s_0}(x) = \theta_{i_0,s_0}(x)/\Theta(x). \tag{2.10}$$

This function is defined on all of X, and it is zero outside $X^+(i_0, s_0)$. Moreover,

$$\varphi_{i_0,s_0} \geq 0 \quad \text{and} \quad \sum_{(i_0,s_0)\in\text{REL}} \varphi_{i_0,s_0} = 1 \quad \text{on} \quad X. \tag{2.11}$$

Note that because

$$\operatorname{diam} X^+(i_0, s_0) \leq C s_0 \leq C\, A^{-1} r(x(i_0, s_0))$$

[see (2.6)], the function φ_{i_0,s_0} is zero outside the ball $B(x(i_0, s_0), C\, A^{-1} r(x(i_0, s_0)))$. Thanks to our Large A Assumption 2.2, it follows that

$$\varphi_{i,s} \quad \text{is identically zero outside the ball} \quad B(x(i, s), ar(x(i, s))). \tag{2.12}$$

LEMMA 2.9. *For $x, y \in X$ and $(i_0, s_0) \in \text{REL}$, we have*

$$|\varphi_{i_0,s_0}(x) - \varphi_{i_0,s_0}(y)| \leq C(A)\, d(x, y)/s_0.$$

Proof. Suppose first that $x \in X^+(i_0, s_0)$. Then

$$\begin{aligned} |\varphi_{i_0,s_0}(x) - \varphi_{i_0,s_0}(y)| &= \left| \frac{\theta_{i_0,s_0}(x)}{\Theta(x)} - \frac{\theta_{i_0,s_0}(y)}{\Theta(y)} \right| \\ &\leq \frac{|\theta_{i_0,s_0}(x) - \theta_{i_0,s_0}(y)|}{\Theta(x)} + \theta_{i_0,s_0}(y)\, \frac{|\Theta(x) - \Theta(y)|}{\Theta(x)\Theta(y)}. \end{aligned}$$

The first term on the right is at most $C(A)\, d(x, y)/s_0$ by (2.4) and (2.9); the second term on the right is at most $C(A)\, d(x, y)/s_0$ thanks to (2.3), Lemma 2.8 and (2.9). Thus,

$$|\varphi_{i_0,s_0}(x) - \varphi_{i_0,s_0}(y)| \leq C(A)\, d(x, y)/s_0 \quad \text{if} \quad x \in X^+(i_0, s_0). \tag{2.13}$$

Similarly, (2.13) holds if $y \in X^+(i_0, s_0)$.

Finally, if neither x nor y belongs to $X^+(i_0, s_0)$, then

$$\varphi_{i_0,s_0}(x) = \varphi_{i_0,s_0}(y) = 0,$$

so (2.13) is obvious.

Thus, (2.13) holds in all cases. □

COROLLARY 2.10. *For $x, y \in X$ and $(i_0, s_0) \in \text{REL}$, we have*

$$|\varphi_{i_0,s_0}(x) - \varphi_{i_0,s_0}(y)| \leq C(A)\, d(x, y)/r(x(i_0, s_0)).$$

Proof. Immediate from Lemma 2.9 and inequalities (2.6). □

We can now finish the proof of the Whitney Partition Lemma 2.1. We pick A to be a constant determined by $c_{\scriptscriptstyle N}$, $D_{\scriptscriptstyle N}$, C_{LS}, a, taken large enough to satisfy the Large A Assumption 2.2. We then take our functions φ_ν to be the $\varphi_{(i,s)}$ $((i,s) \in \text{Rel})$, and we take our x_ν to be the points $x(i,s)$ $((i,s) \in \text{Rel})$. We set $r_\nu = r(x_\nu)$.

The following hold:

- Each $\varphi_\nu \geq 0$, and each $\varphi_\nu = 0$ outside $B(x_\nu, ar_\nu)$; see (2.11) and (2.12).
- Any given $x \in X$ satisfies $\varphi_\nu(x) \neq 0$ for at most C distinct ν. This follows from Corollary 2.5, definition (2.10), and the fact that A is now determined by $c_{\scriptscriptstyle N}$, $D_{\scriptscriptstyle N}$, C_{LS}, a.
- $\sum\limits_\nu \varphi_\nu = 1$ on X; see (2.11).
- For each ν and for all $x, y \in X$, we have

$$|\varphi_\nu(x) - \varphi_\nu(y)| \leq C\, d(x,y)/r_\nu;$$

see Corollary 2.10, and note that A is now determined by $c_{\scriptscriptstyle N}$, $D_{\scriptscriptstyle N}$, C_{LS} and a.

The proof of the Whitney Partition Lemma 2.1 is complete. □

Remark 2.11. Later on there will be another Large A Assumption different from that in this section.

2.2 Patching Lemma.

Patching Lemma 2.12. *Let (X,d) be a metric space, and let Y be a Banach space. For each ν in some index set, assume we are given the following objects:*

- *A point $x_\nu \in X$ and a positive number $r_\nu > 0$ (a "lengthscale").*
- *A function $\theta_\nu : X \to \mathbb{R}$.*
- *A vector $\eta_\nu \in Y$ and a vector-valued function $F_\nu : X \to Y$.*

We make the following assumptions: We are given positive constants $C_{LS} \geq 1$, C_{Wh}, C_η, $C^\#$, C_{Lip}, D^*, such that the following conditions are satisfied for each μ, ν

- (Consistency of the Lengthscale)

$$C_{LS}^{-1} \leq r_\nu / r_\mu \leq C_{LS} \quad \textit{whenever} \quad d(x_\mu, x_\nu) \leq r_\mu + r_\nu. \tag{2.14}$$

(Whitney Partition Assumptions)

- *$\theta_\nu \geq 0$ on X and $\theta_\nu = 0$ outside $B(x_\nu, a\, r_\nu)$, where*

$$a = (4\, C_{LS})^{-1}. \tag{2.15}$$

- *$|\theta_\nu(x) - \theta_\nu(y)| \leq C_{Wh} \cdot d(x,y)/r_\nu$ for $x, y \in X$.*
- *Any given $x \in X$ satisfies $\theta_\nu(x) \neq 0$ for at most D^* distinct ν.*
- *$\sum\limits_\nu \theta_\nu = 1$ on X.*
- (Consistency of the η_ν) $\|\eta_\mu - \eta_\nu\| \leq C_\eta \cdot [r_\nu + r_\nu + d(x_\mu, x_\nu)]$.
- (Agreement of F_ν with η_ν) $\|F_\nu(x) - \eta_\nu\| \leq C^\#\, r_\nu$ *for* $x \in B(x_\nu, r_\nu)$.

• (LIPSCHITZ CONTINUITY OF F_ν) $\|F_\nu(x) - F_\nu(y)\| \le C_{Lip} \cdot d(x,y)$ *for* $x, y \in B(x_\nu, r_\nu)$.

Define

$$F(x) = \sum_\nu \theta_\nu(x)\, F_\nu(x) \quad \textit{for} \;\; x \in X.$$

Then F *satisfies*

$$\|F(x) - F(y)\| \le C\, d(x,y) \quad \textit{for} \quad x, y \in X,$$

where C *is determined by* C_{LS}, C_{Wh}, C_η, $C^{\#}$, C_{Lip}, D^*.

To start the proof of the PATCHING LEMMA 2.12, we define a set of relevant ν by setting

$$\mathrm{RLV}(x) = \{\nu : \theta_\nu(x) \ne 0\}, \quad x \in X.$$

Then $1 \le \#(\mathrm{RLV}(x)) \le D^*$, and

$$d(x, x_\nu) \le a\, r_\nu \quad \text{for} \quad v \in \mathrm{RLV}(x). \tag{2.16}$$

We also recall that $C_{LS} \ge 1$ and $a = (4\, C_{LS})^{-1}$ so that

$$C_{LS} \cdot a = 1/4 \quad \text{and} \quad a \le 1/4. \tag{2.17}$$

We will use the following result.

LEMMA 2.13. *Let* $\nu, \nu_0 \in \mathrm{RLV}(x)$, $\mu_0 \in \mathrm{RLV}(y)$, *and suppose that* $d(x,y) \le a \cdot [r_{\nu_0} + r_{\mu_0}]$. *Then*

$$x, y \in B(x_\nu, r_\nu) \cap B(x_{\nu_0}, r_{\nu_0}) \cap B(x_{\mu_0}, r_{\mu_0})$$

and the ratios

$$r_{\nu_0}/r_{\mu_0},\; r_{\mu_0}/r_{\nu_0},\; r_\nu/r_{\nu_0},\; r_{\nu_0}/r_\nu,\; r_\nu/r_{\mu_0},\; r_{\mu_0}/r_\nu$$

are at most C_{LS}.

Proof. We have the following inequalities:

(★1) $d(x_\nu, x_{\nu_0}) \le d(x_\nu, x) + d(x, x_{\nu_0}) \le a\, r_\nu + a\, r_{\nu_0}$,

(★2) $d(x_{\nu_0}, x_{\mu_0}) \le d(x_{\nu_0}, x) + d(x,y) + d(y, x_{\mu_0}) \le a\, r_{\nu_0} + [a\, r_{\nu_0} + a\, r_{\mu_0}] + a\, r_{\mu_0}$,

(★3) $d(x_\nu, x_{\mu_0}) \le d(x_\nu, x) + d(x,y) + d(y, x_{\mu_0}) \le a\, r_\nu + [a\, r_{\nu_0} + a\, r_{\mu_0}] + a\, r_{\mu_0}$.

From (★1), (★2), (2.17), and CONSISTENCY OF THE LENGTHSCALE (2.14), we have

$$r_\nu/r_{\nu_0},\; r_{\nu_0}/r_\nu,\; r_{\nu_0}/r_{\mu_0},\; r_{\mu_0}/r_{\nu_0} \le C_{LS}.$$

Therefore, (★3) and (2.17) imply that

$$d(x_\nu, x_{\mu_0}) \le a\, r_\nu + C_{LS} a\, r_\nu + 2a\, r_{\mu_0} \le r_\nu + r_{\mu_0},$$

and, consequently, another application of CONSISTENCY OF THE LENGTHSCALE (2.14) gives

$$r_\nu/r_{\mu_0},\ r_{\mu_0}/r_\nu \le C_{LS}.$$

Next, note that, by (2.16) and (2.17),

$$d(x,x_\nu) \le a\,r_\nu < r_\nu$$

and

$$d(y,x_\nu) \le d(y,x) + d(x,x_\nu) \le [a\,r_{\nu_0} + a\,r_{\mu_0}] + a\,r_\nu \le (3C_{LS}\,a)r_\nu < r_\nu.$$

Hence,

$$x, y \in B(x_\nu, r_\nu).$$

Similarly,

$$d(x,x_{\nu_0}) \le a\,r_{\nu_0} < r_{\nu_0}$$

and

$$d(y,x_{\nu_0}) \le d(y,x) + d(x,x_{\nu_0}) \le [a\,r_{\mu_0} + a\,r_{\nu_0}] + a\,r_{\nu_0} \le (3C_{LS}\,a)r_{\nu_0} < r_{\nu_0}.$$

Hence,

$$x, y \in B(x_{\nu_0}, r_{\nu_0}).$$

Finally,

$$d(y,x_{\mu_0}) \le a\,r_{\mu_0} < r_{\mu_0}$$

and

$$d(x,x_{\mu_0}) \le d(x,y) + d(y,x_{\mu_0}) \le [a\,r_{\mu_0} + a\,r_{\nu_0}] + a\,r_{\mu_0} \le (3C_{LS}\,a)r_{\mu_0} < r_{\mu_0}.$$

Hence,

$$x, y \in B(x_{\mu_0}, r_{\mu_0}).$$

The proof of the lemma is complete. □

Proof of the Patching Lemma 2.12. We write c, C, C', etc. to denote positive constants determined by C_{LS}, C_{Wh}, C_η, $C^\#$, C_{Lip}, D^*. These symbols may denote different constants in different occurrences.

Let $x, y \in X$ be given. We must show that

$$\|F(x) - F(y)\| \le C\,d(x,y).$$

Fix μ_0, ν_0, with $\nu_0 \in \mathrm{RLV}(x)$ and $\mu_0 \in \mathrm{RLV}(y)$. We distinguish two cases.

CASE 1: Suppose

$$d(x,y) \le a \cdot [r_{\nu_0} + r_{\mu_0}] \quad \text{with} \quad a = (4\,C_{LS})^{-1}.$$

Then Lemma 2.13 yields

$$x, y \in B(x_\nu, r_\nu) \cap B(x_{\nu_0}, r_{\nu_0}) \cap B(x_{\mu_0}, r_{\mu_0}) \tag{2.18}$$

for all $\nu \in \mathrm{R\small{LV}}(x) \cup \mathrm{R\small{LV}}(y)$. (If $\nu \in \mathrm{R\small{LV}}(y)$, we apply Lemma 2.13 with y, x, μ_0, ν_0 in place of x, y, ν_0, μ_0.) Also, for such ν, Lemma 2.13 gives

$$c\,r_{\nu_0} \le r_\nu \le C\,r_{\nu_0} \quad \text{and} \quad c\,r_{\nu_0} \le r_{\mu_0} \le C\,r_{\nu_0}. \tag{2.19}$$

For $v \in \mathrm{R\small{LV}}(x)$, we have

$$\|F_\nu(y) - \eta_{\nu_0}\| \le \|F_\nu(y) - \eta_\nu\| + \|\eta_\nu - \eta_{\nu_0}\| \le C\,r_\nu + C\,[r_\nu + r_{\nu_0} + d(x_\nu, x_{\nu_0})]. \tag{2.20}$$

(Here, we may apply CONSISTENCY OF THE η_ν and AGREEMENT OF F_ν WITH η_ν, because $y \in B(x_\nu, r_\nu)$.) Also, by (2.18),

$$d(x_\nu, x_{\nu_0}) \le d(x_\nu, x) + d(x, x_{\nu_0}) \le r_\nu + r_{\nu_0} \quad \text{for} \quad \nu \in \mathrm{R\small{LV}}(x).$$

The above estimates and (2.19) tell us that

$$\|F_\nu(y) - \eta_{\nu_0}\| \le C\,r_{\nu_0} \quad \text{if} \quad \nu \in \mathrm{R\small{LV}}(x).$$

Similarly, suppose $v \in \mathrm{R\small{LV}}(y)$. Then (2.20) holds. (We may apply AGREEMENT OF F_ν WITH η_ν, because $y \in B(x_\nu, r_\nu)$.) Also, by (2.18),

$$d(x_\nu, x_{\nu_0}) \le d(x_\nu, y) + d(y, x_{\nu_0}) \le r_\nu + r_{\nu_0} \;\text{ for all }\; \nu \in \mathrm{R\small{LV}}(y).$$

The above estimates and (2.19) tell us that

$$\|F_\nu(y) - \eta_{\nu_0}\| \le C\,r_{\nu_0} \quad \text{for all} \quad \nu \in \mathrm{R\small{LV}}(y).$$

Thus,

$$\|F_\nu(y) - \eta_{\nu_0}\| \le C\,r_{\nu_0} \quad \text{for all} \quad \nu \in \mathrm{R\small{LV}}(x) \cup \mathrm{R\small{LV}}(y).$$

We now write

$$F(x) - F(y) = \sum_{\nu \in \mathrm{R\small{LV}}(x) \cup \mathrm{R\small{LV}}(y)} \theta_\nu(x) \cdot [F_\nu(x) - F_\nu(y)] + \sum_{\nu \in \mathrm{R\small{LV}}(x) \cup \mathrm{R\small{LV}}(y)} [\theta_\nu(x) - \theta_\nu(y)] \cdot [F_\nu(y) - \eta_{\nu_0}] \equiv I + II.$$

Here we have used the WHITNEY PARTITION ASSUMPTION that the sum of all θ_ν equals 1.

It follows from Lipschitz continuity of F_ν that

$$\|I\| \le \sum_{\nu\in \mathrm{RLV}(x)\cup \mathrm{RLV}(y)} \theta_\nu(x)\cdot [C\,d(x,y)] = C\,d(x,y).$$

Each summand in II satisfies

$$|\theta_\nu(x)-\theta_\nu(y)| \le C\,d(x,y)/r_\nu \quad \text{and} \quad \|F_\nu(y)-\eta_{\nu_0}\| \le C\,r_{\nu_0} \le C\,C_{LS}\,r_\nu\,,$$

see (2.1). Hence

$$\|[\theta_\nu(x)-\theta_\nu(y)]\cdot[F_\nu(y)-\eta_{\nu_0}]\| \le C\,d(x,y).$$

Because there are at most $2D^*$ summands in II, it follows that

$$\|II\| \le C\,d(x,y).$$

Combining our estimates for terms I and II, we find that

$$\|F(x)-F(y)\| \le C\,d(x,y) \quad \text{in CASE 1.}$$

CASE 2: Suppose

$$d(x,y) > a\cdot[r_{\nu_0}+r_{\mu_0}] \quad \text{with} \quad a=(4\,C_{LS})^{-1}.$$

For $\nu\in \mathrm{RLV}(x)$, we have

$$d(x_\nu,x_{\nu_0}) \le d(x_\nu,x)+d(x,x_{\nu_0}) \le a\cdot r_\nu + a\cdot r_{\nu_0},$$

hence, by CONSISTENCY OF THE LENGTHSCALE (see (2.1)),

$$c\,r_{\nu_0} \le r_\nu \le C\,r_{\nu_0}$$

and

$$\begin{aligned}\|F_\nu(x)-\eta_{\nu_0}\| &\le \|F_\nu(x)-\eta_\nu\| + \|\eta_\nu-\eta_{\nu_0}\| \\ &\le C\,r_\nu + [C\,r_\nu + C\,r_{\nu_0} + Cd(x_\nu,x_{\nu_0})] \le C\,r_{\nu_0}.\end{aligned}$$

Here we use (2.16) and (2.17), and the fact that $x\in B(x_\nu,r_\nu)$.

Consequently,

$$\|F(x)-\eta_{\nu_0}\| = \left\|\sum_{v\in \mathrm{RLV}(x)} \theta_\nu(x)\cdot[F_\nu(x)-\eta_{\nu_0}]\right\| \le C\,r_{\nu_0}\sum_{v\in \mathrm{RLV}(x)}\theta_\nu(x) = C\,r_{\nu_0}.$$

Similarly,

$$\|F(y)-\eta_{\mu_0}\| \le C\,r_{\mu_0}.$$

Therefore,

$$\begin{aligned}\|F(x)-F(y)\| &\le C\,r_{\nu_0}+C\,r_{\mu_0}+\|\eta_{\nu_0}-\eta_{\mu_0}\| \le C'\,r_{\nu_0}+C'\,r_{\mu_0}+C'd(x_{\nu_0},x_{\mu_0})\\ &\le C'\,r_{\nu_0}+C'\,r_{\mu_0}+C'[d(x_{\nu_0},x)+d(x,y)+d(y,x_{\mu_0})]\\ &\le C''\,r_{\nu_0}+C''\,r_{\mu_0}+C''d(x,y).\end{aligned}$$

Moreover, because we are in CASE 2, we have

$$r_{\nu_0}+r_{\mu_0}\le \tfrac{1}{a}\,d(x,y)=4\,C_{LS}\,d(x,y).$$

It now follows that

$$\|F(x)-F(y)\|\le C'''d(x,y)\quad\text{in CASE 2.}$$

Thus, the conclusion of the PATCHING LEMMA holds in all cases. □

3 Sets Γ_ℓ, Labels and Bases

3.1 Main properties of Γ_ℓ. We recall that $(Y,\|\cdot\|)$ denotes a Banach space. Given a convex set $S\subset Y$ we let *affhull* (S) denote the affine hull of S, i.e., the smallest (with respect to inclusion) affine subspace of Y containing S. We define the affine dimension $\dim S$ of S as the dimension of its affine hull, i.e.,

$$\dim S=\dim \mathit{affhull}\,(S).$$

Given $y\in Y$ and $r>0$ we let

$$B_Y(y,r)=\{z\in Y:\ \|z-y\|\le r\}$$

denote a *closed* ball in Y with center y and radius r. By $B_Y=B_Y(0,1)$ we denote the unit ball in Y.

Given non-empty sets $A,B\subset Y$ we let $A+B=\{a+b: a\in A, b\in B\}$ denote the Minkowski sum of these sets. Given a positive real number λ by λA we denote the set $\lambda A=\{\lambda a: a\in A\}$.

We call a pseudometric space $(\mathcal{M},\rho)$ finite if $\mathcal{M}$ is finite, but we say that the pseudometric ρ is finite if $\rho(x,y)$ is finite for every $x,y\in\mathcal{M}$.

Let $(\mathcal{M},\rho)$ be a *finite* pseudometric space with a finite pseudometric ρ. Let us fix a constant $\lambda>0$, an integer $m\ge 0$, and a set-valued mapping $F:\mathcal{M}\to \mathrm{Conv}_m(Y)$. Recall that $\mathrm{Conv}_m(Y)$ denotes the family of all nonempty convex subsets of Y of dimension at most m.

In this section we introduce a family of convex sets $\Gamma_\ell(x)\subset Y$ parametrized by $x\in\mathcal{M}$ and a non-negative integer ℓ. To do so, we first define integers $k_0,k_1,k_2,\ldots$ by the formula

$$k_\ell=(m+2)^\ell\quad(\ell\ge 0). \tag{3.1}$$

Definition 3.1. *Let $x \in \mathcal{M}$ and let $S \subset \mathcal{M}$. A point $\xi \in Y$ belongs to the set $\Gamma(x,S)$ if there exists a mapping $f : S \cup \{x\} \to Y$ such that:*

(i) *$f(x) = \xi$ and $f(z) \in F(z)$ for all $z \in S \cup \{x\}$.*
(ii) *For every $z, w \in S \cup \{x\}$ the following inequality*

$$\|f(z) - f(w)\| \le \lambda\, \rho(z,w)$$

holds.

We then define

$$\Gamma_\ell(x) = \bigcap_{\substack{S \subset \mathcal{M} \\ \#S \le k_\ell}} \Gamma(x,S) \quad \text{for} \quad x \in \mathcal{M},\ \ell \ge 0. \tag{3.2}$$

For instance, given $x \in \mathcal{M}$ let us present an explicit formula for $\Gamma_0(x)$. By (3.2) for $\ell = 0$,

$$\Gamma_0(x) = \bigcap_{S \subset \mathcal{M},\ \#S \le 1} \Gamma(x,S).$$

Clearly, by Definition 3.1,

$$\Gamma(x,\{z\}) = F(x) \bigcap (F(z) + \lambda\, \rho(x,z) B_Y) \quad \text{for every} \quad z \in \mathcal{M},$$

and $\Gamma(x,\emptyset) = F(x)$, so that

$$\Gamma_0(x) = \bigcap_{z \in \mathcal{M}} (F(z) + \lambda\, \rho(x,z) B_Y)\,. \tag{3.3}$$

Remark 3.2. (i) Of course, the sets $\Gamma_\ell(x)$ also depend on the set-valued mapping F, the constant λ and m. However, we use Γ's only in this section, Sections 3-4 and Section 6.2 where these objects, i.e., F, λ and m, are clear from the context. Therefore we omit F, λ and m in the notation of Γ's.

(ii) As in the statement of Theorem 1.1, we may restrict attention to S containing exactly k_ℓ points in (3.2), provided $\mathcal{M}$ contains at least k_ℓ points.

The above Γ's are (possibly empty) *convex subsets* of Y. Note that

$$\Gamma(x,S) \subset F(x) \quad \text{for all} \quad x \in \mathcal{M} \quad \text{and} \quad S \subset \mathcal{M}. \tag{3.4}$$

Hence,

$$\Gamma(x,S) \subset \mathit{affhull}\,(F(x)) \qquad x \in \mathcal{M},\ S \subset \mathcal{M}. \tag{3.5}$$

From (3.4) and (3.2) we obtain

$$\Gamma_\ell(x) \subset F(x) \quad \text{for} \quad x \in \mathcal{M},\ \ell \ge 0. \tag{3.6}$$

Also, obviously,

$$\Gamma_\ell(x) \subset \Gamma_{\ell-1}(x) \quad \text{for} \quad x \in \mathcal{M},\ \ell \geq 1. \tag{3.7}$$

We describe main properties of the sets Γ_ℓ in Lemma 3.4 below. The proof of this lemma relies on Helly's intersection theorem [DGK63], a classical result from the Combinatorial Geometry of convex sets.

Helly's Theorem 3.3. *Let $\mathcal{K}$ be a finite family of nonempty convex subsets of Y lying in an affine subspace of Y of dimension m. Suppose that every subfamily of $\mathcal{K}$ consisting of at most $m+1$ elements has a common point. Then there exists a point common to all of the family $\mathcal{K}$.*

LEMMA 3.4. *Let $\ell \geq 0$. Suppose that the restriction $F|_{\mathcal{M}'}$ of F to an arbitrary subset $\mathcal{M}' \subset \mathcal{M}$ consisting of at most $k_{\ell+1}$ points has a Lipschitz selection $f_{\mathcal{M}'} : \mathcal{M}' \to Y$ with $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}(\mathcal{M}',Y)} \leq \lambda$. Then for all $x \in \mathcal{M}$*

(a) $\Gamma_\ell(x) \neq \emptyset$*;*
(b) $\Gamma_\ell(x) \subset \Gamma_{\ell-1}(y) + \lambda\, \rho(x,y)\, B_Y$ *for all $y \in \mathcal{M}$, provided $\ell \geq 1$.*

Proof. Thanks to (3.2), (3.5) and Helly's Theorem 3.3, conclusion (a) will follow if we can show that

$$\Gamma(x, S_1) \bigcap \cdots \bigcap \Gamma(x, S_{m+1}) \neq \emptyset \tag{3.8}$$

for every $S_1, \ldots, S_{m+1} \subset \mathcal{M}$ such that $\#S_i \leq k_\ell$ (each i). (We note that, by (3.5), each set $\Gamma(x,S)$ is a subset of the affine space *affhull* $(F(x))$ of dimension at most m. We also use the fact that there are only finitely many $S \subset \mathcal{M}$ because $\mathcal{M}$ is finite.)

However, $S_1 \cup \cdots \cup S_{m+1} \cup \{x\} \subset \mathcal{M}$ has cardinality at most

$$(m+1) \cdot k_\ell + 1 \leq k_{\ell+1}.$$

The lemma's hypothesis therefore produces a function $\tilde{f} : S_1 \cup \cdots \cup S_{m+1} \cup \{x\} \to Y$ such that $\tilde{f}(z) \in F(z)$ for all $z \in S_1 \cup \cdots \cup S_{m+1} \cup \{x\}$, and

$$\|\tilde{f}(z) - \tilde{f}(w)\| \leq \lambda\, \rho(z,w) \quad \text{for all} \quad z, w \in S_1 \cup \cdots \cup S_{m+1} \cup \{x\}.$$

Then $\tilde{f}(x)$ belongs to $\Gamma(x, S_i)$ for $i = 1, \ldots, m+1$, proving (3.8) and thus also proving (a).

To prove (b), let $x, y \in \mathcal{M}$, and let $\xi \in \Gamma_\ell(x)$ with $\ell \geq 1$. We must show that there exists $\eta \in \Gamma_{\ell-1}(y)$ such that $\|\xi - \eta\| \leq \lambda \cdot \rho(x,y)$. To produce such an η, we proceed as follows.

Given a set $S \subset \mathcal{M}$ we introduce a set $\hat{\Gamma}(x, y, \xi, S)$ consisting of all points $\eta \in Y$ such that there exists a mapping $f : S \cup \{x, y\} \to Y$ satisfying the followingconditions:

(i) $f(x)=\xi$, $f(y)=\eta$, and $f(z)\in F(z)$ for all $z\in S\cup\{x,y\}$;
(ii) For every $z,w\in S\cup\{x,y\}$ the following inequality

$$\|f(z)-f(w)\|\le\lambda\,\rho(z,w)$$

holds.

Clearly, $\hat\Gamma(x,y,\xi,S)$ is a *convex* subset of $F(y)$. Let us show that

$$\bigcap_{\substack{S\subset\mathcal{M}\\ \#S\le k_{\ell-1}}}\hat\Gamma(x,y,\xi,S)\ne\emptyset. \tag{3.9}$$

Thanks to Helly's Theorem 3.3, (3.9) will follow if we can show that

$$\hat\Gamma(x,y,\xi,S_1)\cap\cdots\cap\hat\Gamma(x,y,\xi,S_{m+1})\ne\emptyset \tag{3.10}$$

for all $S_1,\dots,S_{m+1}\subset\mathcal{M}$ with $\#S_i\le k_{\ell-1}$ (each i).

We set $\widetilde S=S_1\cup\cdots\cup S_{m+1}\cup\{y\}$. Then $\widetilde S\subset\mathcal{M}$ with

$$\#\widetilde S\le(m+1)\cdot k_{\ell-1}+1\le k_\ell.$$

Because $\xi\in\Gamma_\ell(x)\subset\Gamma(x,\widetilde S)$ (see (3.2)), there exists $\tilde f:S_1\cup\cdots\cup S_{m+1}\cup\{x,y\}\to Y$ such that

$$\tilde f(x)=\xi,\ \ \tilde f(z)\in F(z)\quad\text{for all}\quad z\in S_1\cup\cdots\cup S_{m+1}\cup\{x,y\},$$

and

$$\|\tilde f(z)-\tilde f(w)\|\le\lambda\,\rho(z,w)\quad\text{for}\quad z,w\in S_1\cup\cdots\cup S_{m+1}\cup\{x,y\}.$$

We then have $\tilde f(y)\in\hat\Gamma(x,y,\xi,S_i)$ for $i=1,\dots,m+1$, proving (3.10) and therefore also proving (3.9).

Let

$$\eta\in\bigcap_{\substack{S\subset\mathcal{M}\\ \#S\le k_{\ell-1}}}\hat\Gamma(x,y,\xi,S).$$

Taking $S=\emptyset$, we obtain a function $f:\{x,y\}\to Y$ with $f(x)=\xi$, $f(y)=\eta$ and

$$\|f(z)-f(w)\|\le\lambda\,\rho(z,w)\quad\text{for}\quad z,w\in\{x,y\}.$$

Therefore,

$$\|\eta-\xi\|\le\lambda\,\rho(z,w). \tag{3.11}$$

Moreover, because $\hat\Gamma(x,y,\xi,S)\subset\Gamma(y,S)$ for any $S\subset\mathcal{M}$ (see Definition 3.1), we have

$$\eta\in\bigcap_{\substack{S\subset\mathcal{M}\\ \#S\le k_{\ell-1}}}\Gamma(y,S)=\Gamma_{\ell-1}(y). \tag{3.12}$$

Our results (3.11), (3.12) complete the proof of (b). □

3.2 Statement of the Finiteness Theorem for bounded Nagata dimension. We place ourselves in the following setting.

- We fix a positive integer m.
- (X, d) is a finite metric space satisfying Nagata (D_N, c_N) (see Definition 1.5).
- Y is a Banach space. We write $\|\cdot\|$ for the norm in Y, and $\|\cdot\|_{Y^*}$ for the norm in the dual space Y^*. We write $\langle e, y\rangle$ to denote the natural pairing between vectors $y \in Y$ and dual vectors $e \in Y^*$.
- For each $x \in X$ we are given a convex set

$$F(x) \subset \text{Aff}_F(x) \subset Y,$$

where

$$\text{Aff}_F(x) \quad \text{is an affine subspace of} \quad Y, \quad \text{of dimension at most} \quad m.$$

Say, $\text{Aff}_F(x)$ is a translate of the vector subspace $\text{Vect}_F(x) \subset Y$.

- We make the following assumption for a large enough $k^\sharp$ determined by m.

Finiteness Assumption 3.5. *Given $S \subset X$ with $\#S \leq k^\sharp$, there exists $f^S : S \to Y$ with Lipschitz seminorm at most 1, such that $f^S(x) \in F(x)$ for all $x \in S$.*

The above assumption implies the existence of a Lipschitz selection with a controlled Lipschitz seminorm. More precisely, we have the following result.

Theorem 3.6 (Finiteness Theorem for bounded Nagata dimension). *Let (X, d) be a finite metric space satisfying Nagata (D_N, c_N).*

Given $m \in \mathbb{N}$ there exist a constant $k^\sharp \in \mathbb{N}$ depending only on m, and a constant $\gamma > 0$ depending only on m, c_N, D_N, for which the following holds: Let Y be a Banach space. For each $x \in X$, let $F(x) \subset Y$ be a convex set of (affine) dimension at most m.

Suppose that for each $S \subset X$ with $\#S \leq k^\sharp$ there exists $f^S : S \to Y$ with Lipschitz seminorm at most 1, such that $f^S(x) \in F(x)$ for all $x \in S$.

Then there exists $f : X \to Y$ with Lipschitz seminorm at most γ, such that $f(x) \in F(x)$ for all $x \in X$.

By applying Theorem 3.6 to the metric space $(X, \lambda\, d)$ we establish Theorem 1.6.

We place ourselves in the above setting until the end of the proof of Theorem 3.6 in the end of Section 4.9.

3.3 Labels and bases. A "label" is a finite sequence $\mathcal{A} = (e_1, e_2, \ldots, e_s)$ of functionals $e_a \in Y^*$, $a = 1, \ldots, s$, with $s \leq m$. Here, m is as in the hypothesis of Theorem 3.6.

We write $\#\mathcal{A}$ to denote the number s of functionals e_a appearing in $\mathcal{A}$. We allow the case $\#\mathcal{A} = 0$, in which case $\mathcal{A}$ is the empty sequence $\mathcal{A} = (\ )$.

Let $\Gamma \subset Y$ be a convex set, let $\mathcal{A} = (e_1, e_2, \ldots, e_s)$ be a label, and let r, C_B be positive real numbers. Finally, let $\zeta \in Y$.

Definition 3.7. *An* $(\mathcal{A}, r, C_B)$*-basis for* Γ *at* ζ *is a sequence of* s *vectors* $v_1, \ldots, v_s \in Y$*, with the following properties:*

(B0) $\zeta \in \Gamma$.
(B1) $\langle e_a, v_b \rangle = \delta_{ab}$ *(Kronecker delta) for* $a, b = 1, \ldots, s$.
(B2) $\|v_a\| \le C_B$ *and* $\|e_a\|_{Y^*} \le C_B$ *for* $a = 1, \ldots, s$.
(B3) $\zeta + \frac{r}{C_B} v_a$ *and* $\zeta - \frac{r}{C_B} v_a$ *belong to* Γ *for* $a = 1, \ldots, s$.

If $s \ge 1$, then of course (B3) implies (B0).

Let us note several elementary properties of $(\mathcal{A}, r, C_B)$-bases.

Remark 3.8. (i) If $s = 0$ then (B1), (B2), (B3) hold vacuously, so the assertion that Γ has an $((\ ), r, C_B)$-basis at ζ means simply that $\zeta \in \Gamma$;

(ii) If $r' \le r$ and $C'_B \ge C_B$, then any $(\mathcal{A}, r, C_B)$-basis for Γ at ζ is also an $(\mathcal{A}, r', C'_B)$-basis for Γ at ζ;

(iii) If $K \ge 1$, then any $(\mathcal{A}, r, C_B)$-basis for Γ at ζ is also an $(\mathcal{A}, Kr, KC_B)$-basis for Γ at ζ;

(iv) If $\Gamma \subset \Gamma'$, then every $(\mathcal{A}, r, C_B)$-basis for Γ at ζ is also an $(\mathcal{A}, r, C_B)$-basis for Γ' at ζ.

Lemma 3.9 ("Adding a vector"). *Suppose* $\Gamma \subset Y$ *(convex) has an* $(\mathcal{A}, r, C_B)$*-basis at* ξ*, where* $\mathcal{A} = (e_1, e_2, \ldots, e_s)$ *and* $s \le m - 1$.

Let $\eta \in \Gamma$*, and suppose that*

$$\|\eta - \xi\| \ge r$$

and

$$\langle e_a, \eta - \xi \rangle = 0 \quad \text{for} \quad a = 1, \ldots, s.$$

Then there exist $\zeta \in \Gamma$ *and* $e_{s+1} \in Y^*$ *with the following properties:*

- $\|\zeta - \xi\| = \frac{1}{2} r$.
- $\langle e_a, \zeta - \xi \rangle = 0$ *for* $a = 1, \ldots, s$ *(not necessarily for* $a = s + 1$*).*
- Γ *has an* $(\mathcal{A}^+, r, C'_B)$*-basis at* ζ*, where* $\mathcal{A}^+ = (e_1, \ldots, e_s, e_{s+1})$ *and* C'_B *is determined by* C_B *and* m.

Proof. In this proof, we write C to denote a positive constant determined by C_B and m. This symbol may denote different constants in different occurrences.

Let $(v_1, \ldots, v_s)$ be an $(\mathcal{A}, r, C_B)$-basis for Γ at ξ. Thus, $\xi \in \Gamma$,

$$\langle e_a, v_b \rangle = \delta_{ab} \quad \text{for} \quad a, b = 1, \ldots, s, \tag{3.13}$$

$$\|e_a\|_{Y^*} \le C_B, \quad \|v_a\| \le C_B \quad \text{for} \quad a = 1, \ldots, s, \tag{3.14}$$

$$\xi + \tfrac{r}{C_B} v_a, \quad \xi - \tfrac{r}{C_B} v_a \in \Gamma \quad \text{for} \quad a = 1, \ldots, s. \tag{3.15}$$

Let

$$\zeta = \tau\,\eta + (1 - \tau)\,\xi \quad \text{with} \quad \tau = \tfrac{1}{2}\, r\, \|\xi - \eta\|^{-1} \in (0, \tfrac{1}{2}].$$

(Note that, by the lemma's hypothesis, $\|\xi - \eta\|$ is non-zero so that τ and ζ are well defined.)

Our hypotheses on ξ and η tell us that

$$\zeta \in \Gamma, \quad \|\zeta - \xi\| = \tfrac{1}{2}r, \quad \langle e_a, \zeta - \xi\rangle = 0 \quad \text{for} \quad a = 1, \dots, s. \tag{3.16}$$

Because $\eta \in \Gamma$, Γ is convex, and $\tau \in (0, \frac{1}{2}]$, (3.15) implies

$$\zeta + \tfrac{1}{2}\tfrac{r}{C_B}\, v_a,\ \zeta - \tfrac{1}{2}\tfrac{r}{C_B}\, v_a \in \Gamma \quad \text{for} \quad a = 1, \dots, s. \tag{3.17}$$

Let

$$v_{s+1} = \frac{\zeta - \xi}{\|\zeta - \xi\|}. \tag{3.18}$$

(The denominator is nonzero, by (3.16).) Then

$$\zeta + \|\zeta - \xi\|\, v_{s+1} = \zeta + (\zeta - \xi) = 2\zeta - \xi = 2\tau\eta + (1 - 2\tau)\xi \in \Gamma$$

because $\xi, \eta \in \Gamma$, Γ is convex and $\tau \in (0, \frac{1}{2}]$.

Also,

$$\zeta - \|\zeta - \xi\|\, v_{s+1} = \zeta - (\zeta - \xi) = \xi \in \Gamma.$$

Recall that $\|\zeta - \xi\| = \frac{1}{2}\, r$, hence the above remarks and (3.17) together yield

$$\zeta + \tfrac{r}{C} v_a,\ \zeta - \tfrac{r}{C} v_a \in \Gamma \quad \text{for} \quad a = 1, \dots, s+1, \quad \text{and} \quad C = 2\max\{C_B, 1\}. \tag{3.19}$$

Also, because $\langle e_a, \zeta - \xi\rangle = 0$ for $a = 1, \dots, s$, (see (3.16)), the definition of v_{s+1}, together with (3.13), tells us that

$$\langle e_a, v_b\rangle = \delta_{ab} \quad \text{for} \quad a = 1, \dots, s \quad \text{and} \quad b = 1, \dots, s+1. \tag{3.20}$$

We prepare to define a functional $e_{s+1} \in Y^*$. To do so, we first prove the estimate

$$\sum_{a=1}^{s+1} |\lambda_a| \le C \left\| \sum_{a=1}^{s+1} \lambda_a\, v_a \right\| \quad \text{for all} \quad \lambda_1, \dots, \lambda_{s+1} \in \mathbb{R}. \tag{3.21}$$

To see this, we first note that for any $b = 1, \dots, s$, (3.20) yields the estimate

$$|\lambda_b| = \left| \langle e_b, \sum_{a=1}^{s+1} \lambda_a\, v_a\rangle \right| \le \|e_b\|_{Y^*} \cdot \left\| \sum_{a=1}^{s+1} \lambda_a\, v_a \right\| \le C_B \left\| \sum_{a=1}^{s+1} \lambda_a\, v_a \right\|. \tag{3.22}$$

Consequently,

$$\begin{aligned} |\lambda_{s+1}| = \|\lambda_{s+1} v_{s+1}\| &\le \left\| \sum_{a=1}^{s+1} \lambda_a\, v_a \right\| + \sum_{a=1}^{s} |\lambda_a|\, \|v_a\| \\ &\le \left\| \sum_{a=1}^{s+1} \lambda_a\, v_a \right\| + C_B \sum_{a=1}^{s} |\lambda_a| \le (1 + m\, C_B^2) \left\| \sum_{a=1}^{s+1} \lambda_a\, v_a \right\|. \end{aligned}$$

(Recall that $0 \le s \le m$.) Together with (3.22), this completes the proof of (3.21).

By (3.21) and the Hahn-Banach theorem, the linear functional

$$\sum_{a=1}^{s+1} \lambda_a v_a \to \lambda_{s+1}$$

on the span of $v_1, \dots, v_{s+1}$ extends to a linear functional $e_{s+1} \in Y^*$, with

$$\|e_{s+1}\|_{Y^*} \le C \tag{3.23}$$

and

$$\langle e_{s+1}, v_a \rangle = \delta_{s+1,a} \quad \text{for} \quad a = 1, \dots, s+1. \tag{3.24}$$

From (3.14), (3.16), (3.18), (3.20), (3.23), (3.24) we have

$$\zeta \in \Gamma, \tag{3.25}$$

$$\|e_a\|_{Y^*},\ \|v_a\| \le C \quad \text{for} \quad a = 1, \dots, s+1, \tag{3.26}$$

$$\langle e_a, v_b \rangle = \delta_{ab} \quad \text{for} \quad a, b = 1, \dots, s+1. \tag{3.27}$$

From (3.19), (3.25), (3.26), (3.27), we see that $v_1, \dots, v_{s+1}$ form an $((e_1, \dots, e_{s+1}), r, C)$-basis for Γ at ζ.

Together with (3.16), this completes the proof of Lemma 3.9. □

Lemma 3.10 ("Transporting a Basis"). *Given $m \in \mathbb{N}$ and $C_B > 0$ there exists a constant $\varepsilon_0 \in (0, 1]$ depending only on m, C_B, for which the following holds:*

Suppose $\Gamma \subset Y$ (convex) has an $(\mathcal{A}, r, C_B)$-basis at ξ_0, where $\mathcal{A} = (e_1, e_2, \dots, e_s)$ and $s \le m$. Suppose $\Gamma' \subset Y$ (convex) satisfies:

(*) *Given any $\xi \in \Gamma$ there exists $\eta \in \Gamma'$ such that $\|\xi - \eta\| \le \varepsilon_0 r$.*

Then there exists $\eta_0 \in \Gamma'$ with the following properties:

- $\|\eta_0 - \xi_0\| \le C\, r$.
- $\langle e_a, \eta_0 - \xi_0 \rangle = 0$ *for $a = 1, \dots, s$.*
- *Γ' has an $(\mathcal{A}, r, C)$-basis at η_0.*

Here, C is determined by C_B and m.

Proof. In the trivial case $s = 0$ (see Remark 3.8 (i)), Lemma 3.10 holds because it simply asserts that there exists $\eta_0 \in \Gamma'$ such that $\|\eta_0 - \xi_0\| \le C\, r$, which is immediate from (*). We suppose $s \ge 1$.

We take

$$\varepsilon_0 \text{ to be less than a small enough positive constant determined by } C_B \text{ and } m. \tag{3.28}$$

We can take ε_0 to be, say, $\frac{1}{2}$ times that small positive constant.

We write c_1, c_2, c_3, C to denote positive constants determined by C_B and m. These symbols may denote different constants in different occurrences.

Let $(v_1, \dots, v_s)$ be an $(\mathcal{A}, r, C_B)$-basis for Γ at ξ_0. Thus, $\xi_0 \in \Gamma$,

$$\langle e_a, v_b \rangle = \delta_{ab} \quad \text{for} \quad a, b = 1, \dots, s, \tag{3.29}$$

$$\|e_a\|_{Y^*} \le C_B, \quad \|v_a\| \le C_B \quad \text{for} \quad a = 1, \dots, s, \tag{3.30}$$

and

$$\xi_0 + c_1 \sigma r\, v_a \in \Gamma \quad \text{for} \quad a = 1, \dots, s, \quad \sigma = \pm 1 \quad \text{and} \quad c_1 = 1/C_B. \tag{3.31}$$

Applying our hypothesis (*) to the vectors in (3.31), we obtain vectors

$$\zeta_{a,\sigma} \in Y \qquad (a = 1, \dots, s, \quad \sigma = \pm 1)$$

such that

$$\xi_0 + c_1 \sigma r\, v_a + \zeta_{a,\sigma} \in \Gamma' \quad \text{for} \quad a = 1, \dots, s, \quad \sigma = \pm 1, \tag{3.32}$$

and

$$\|\zeta_{a,\sigma}\| \le \varepsilon_0\, r \quad \text{for} \quad a = 1, \dots, s, \quad \sigma = \pm 1. \tag{3.33}$$

We define vectors

$$\eta_{00} = \frac{1}{2s} \sum_{a=1}^{s} \sum_{\sigma=\pm 1} (\xi_0 + c_1 \sigma r v_a + \zeta_{a,\sigma}) = \xi_0 + \frac{1}{2s} \sum_{a=1}^{s} \sum_{\sigma=\pm 1} \zeta_{a,\sigma} \tag{3.34}$$

and

$$\tilde{v}_a = \frac{[\xi_0 + c_1 r v_a + \zeta_{a,1}] - [\xi_0 - c_1 r v_a + \zeta_{a,-1}]}{2 c_1 r} = v_a + \left(\frac{\zeta_{a,1} - \zeta_{a,-1}}{2 c_1 r} \right) \tag{3.35}$$

for $a = 1, \dots, s$.

From (3.32) and the first equality in (3.34), we have $\eta_{00} \in \Gamma'$. From (3.33) and the second equality in (3.34), we have

$$\|\eta_{00} - \xi_0\| \le \varepsilon_0 r. \tag{3.36}$$

From (3.33) and the second equality in (3.35), we have

$$\|\tilde{v}_a - v_a\| \le C\, \varepsilon_0 \quad \text{for} \quad a = 1, \dots, s. \tag{3.37}$$

Also, for $b = 1, \dots, s$ and $\hat{\sigma} = \pm 1$, the first equalities in (3.34), (3.35) give

$$\eta_{00} + \frac{1}{s} c_1 r \hat{\sigma}\, \tilde{v}_b = \frac{1}{2s} \sum_{a=1}^{s} \sum_{\sigma=\pm 1} (\xi_0 + c_1 \sigma r v_a + \zeta_{a,\sigma})$$
$$+ \frac{\hat{\sigma}}{2s} [(\xi_0 + c_1 r v_b + \zeta_{b,1}) - (\xi_0 - c_1 r v_b + \zeta_{b,-1})],$$

which exhibits $\eta_{00} + \frac{1}{s} c_1 r \hat{\sigma}\, \tilde{v}_b$ as a convex combination of the vectors in (3.32). Consequently,

$$\eta_{00} + c_2 r\, \tilde{v}_b, \quad \eta_{00} - c_2 r\, \tilde{v}_b \in \Gamma' \quad \text{for} \quad b = 1, \ldots, s,$$

which implies that

$$\eta_{00} + c_2 r \sum_{a=1}^{s} \tau_a \tilde{v}_a \in \Gamma' \quad \text{for any} \quad \tau_1, \ldots, \tau_s \in \mathbb{R} \quad \text{with} \quad \sum_{a=1}^{s} |\tau_a| \le 1. \tag{3.38}$$

Here we use the following remark on convex sets: Suppose $\xi + \eta_i, \xi - \eta_i$, $(i = 1, \ldots, I)$ belong to a convex set Γ. Then

$$\xi + \sum_{i=1}^{I} \tau_i \eta_i \in \Gamma \quad \text{for all} \quad \tau_1, \ldots, \tau_I \in \mathbb{R} \quad \text{with} \quad \sum_{i=1}^{I} |\tau_i| \le 1.$$

From (3.29), (3.30), (3.37), we have

$$|\langle e_a, \tilde{v}_b \rangle - \delta_{ab}| \le C\varepsilon_0 \quad \text{for} \quad a, b = 1, \ldots, s. \tag{3.39}$$

We let A denote the $s \times s$ matrix $A = (\langle e_a, \tilde{v}_b \rangle)_{a,b=1}^{s}$. Let $I = (\delta_{ab})_{a,b=1}^{s}$ be the identity matrix. Given an $s \times s$ matrix T, we let $\|T\|_{op}$ denote the operator norm of T as an operator from ℓ_s^2 into ℓ_s^2. Clearly, $\|T\|_{op}$ is equivalent (with constants depending only on s) to $\max\{|t_{ab}| : 1 \le a, b \le s\}$ provided $T = (t_{ab})_{a,b=1}^{s}$.

Hence, by (3.39),

$$\|A - I\|_{op} \le C\varepsilon_0. \tag{3.40}$$

We recall the standard fact from matrix algebra which states that an $s \times s$ matrix T is invertible and the inequality $\|T^{-1} - I\|_{op} \le \|T - I\|_{op}/(1 - \|T - I\|_{op})$ is satisfied provided $\|T - I\|_{op} < 1$. Therefore, by (3.40), for ε_0 small enough, the matrix A is invertible, and the following inequality

$$\|A^{-1} - I\|_{op} \le 2\, \|A - I\|_{op} \tag{3.41}$$

holds.

Let $(A^{\mathbf{T}})^{-1} = (M_{gb})_{g,b=1,\ldots,s}$ where $A^{\mathbf{T}}$ denotes the transpose of A. Then

$$\left\langle e_a, \sum_{b=1}^{s} M_{gb}\, \tilde{v}_b \right\rangle = \delta_{ag} \quad \text{for} \quad a, g = 1, \ldots, s. \tag{3.42}$$

Moreover, by (3.40) and (3.41),

$$|M_{gb} - \delta_{gb}| \le C\, \varepsilon_0 \quad \text{for} \quad g, b = 1, \ldots, s. \tag{3.43}$$

We set

$$\hat{v}_g = \sum_{b=1}^{s} M_{gb}\, \tilde{v}_b \quad \text{for} \quad g = 1, \ldots, s. \tag{3.44}$$

Then (3.30), (3.37), (3.43), (3.44) yield

$$\|\hat{v}_g\| \leq C \quad \text{for} \quad g = 1, \dots, s, \tag{3.45}$$

while (3.42), (3.44) give

$$\langle e_a, \hat{v}_g \rangle = \delta_{ag} \quad \text{for} \quad a, g = 1, \dots, s. \tag{3.46}$$

Moreover, (3.38), (3.43), (3.44) together imply that

$$\eta_{00} + c_3 r \sum_{g=1}^{s} \tau_g \hat{v}_g \in \Gamma' \quad \text{for all} \quad \tau_1, \dots, \tau_s \quad \text{such that each} \quad |\tau_g| \leq 1. \tag{3.47}$$

To see this, we simply write the linear combination of the $\hat{v}_g$ in (3.47) as a linear combination of the $\tilde{v}_b$ using (3.44), and then recall (3.38).

From (3.30), (3.36) we have

$$|\langle e_a, \eta_{00} - \xi_0 \rangle| \leq C\varepsilon_0\, r \quad \text{for} \quad a = 1, \dots, s. \tag{3.48}$$

We set

$$\eta_0 = \eta_{00} - \sum_{g=1}^{s} \langle e_g, \eta_{00} - \xi_0 \rangle\, \hat{v}_g, \tag{3.49}$$

so that by (3.46),

$$\langle e_a, \eta_0 - \xi_0 \rangle = \langle e_a, \eta_{00} - \xi_0 \rangle - \sum_{g=1}^{s} \langle e_g, \eta_{00} - \xi_0 \rangle \langle e_a, \hat{v}_g \rangle = 0 \quad \text{for} \quad a = 1, \dots, s. \tag{3.50}$$

Also, by (3.36), (3.45), (3.48),

$$\|\eta_0 - \xi_0\| \leq \|\eta_{00} - \xi_0\| + \sum_{g=1}^{s} |\langle e_g, \eta_{00} - \xi_0 \rangle| \cdot \|\hat{v}_g\| \leq C\varepsilon_0 r. \tag{3.51}$$

From (3.48) and our *small* ε_0 *assumption* (3.28), we have

$$|\langle e_a, \eta_{00} - \xi_0 \rangle| \leq \tfrac{1}{2} c_3 r \quad \text{for} \quad a = 1, \dots, s,$$

with c_3 as in (3.47).

Therefore (3.47) and (3.49) tell us that

$$\eta_0 + c_3 r \sum_{g=1}^{s} \tau_g \hat{v}_g \in \Gamma' \quad \text{for any} \quad \tau_1, \dots, \tau_s \quad \text{such that} \quad |\tau_g| \leq \tfrac{1}{2} \quad \text{for each} \quad g.$$

In particular,

$$\eta_0 \in \Gamma' \tag{3.52}$$

and

$$\eta_0 + \tfrac{1}{2}c_3 r\,\hat{v}_g,\ \eta_0 - \tfrac{1}{2}c_3 r\,\hat{v}_g \in \Gamma' \quad \text{for} \quad g = 1, \dots, s.$$

Also, recalling (3.30), (3.45), (3.46), we note that

$$\|e_a\|_{Y^*},\ \|\hat{v}_a\| \le C \quad \text{for} \quad a = 1, \dots, s$$

and

$$\langle e_a, \hat{v}_g \rangle = \delta_{ag} \quad \text{for} \quad a, g = 1, \dots, s. \tag{3.53}$$

Our results (3.52),...,(3.53) tell us that $\hat{v}_1, \dots, \hat{v}_s$ form an $(\mathcal{A}, r, C)$-basis for Γ' at η_0, with $\mathcal{A} = (e_1, \dots, e_s)$. That's the third bullet point in the statement of Lemma 3.10. The other two bullet points are immediate from our results (3.51) and (3.50).

The proof of Lemma 3.10 is complete. □

4 The Main Lemma

4.1 Statement of the Main Lemma. Recall that (X, d) is a (finite) metric space satisfying Nagata (D_N, c_N).

For any label $\mathcal{A} = (e_1, \dots, e_s)$, we define

$$\ell(\mathcal{A}) = 2 + 3 \cdot (m - \#\mathcal{A}) = 2 + 3 \cdot (m - s). \tag{4.1}$$

Note that

$$\ell(\mathcal{A}) \ge \ell(\mathcal{A}^+) + 3 \quad \text{whenever} \quad \#\mathcal{A}^+ > \#\mathcal{A}.$$

We now choose the constant $k^\sharp$ in our Finiteness Assumption 3.5. We take

$$k^\sharp = k_{\ell^\# + 1} = (m+2)^{\ell^\# + 1} \tag{4.2}$$

as in equation (3.1), with

$$\ell^\# = 2 + 3m. \tag{4.3}$$

In this setting we define a family $\Gamma_\ell(x)$ of basic convex sets as in Section 3.1. More specifically, let $(\mathcal{M}, \rho) = (X, d)$, $\lambda = 1$ and let $F : X \to \mathrm{Conv}_m(Y)$ be the set-valued mapping from Theorem 3.6. We apply Definition 3.1 and formulae (3.1), (3.2) to these objects and obtain a family

$$\{\Gamma_\ell(x) : x \in X, \ell = 0, 1, \dots\}$$

of convex subsets of Y.

Finally, we apply Lemma 3.4 to the setting of this section. The Finiteness Assumption 3.5 enables us to replace the hypothesis of this lemma with the requirement $k^\sharp \ge k_{\ell+1}$, which together with definition (4.1) of $\ell(\mathcal{A})$ leads us to the following statement.

Lemma 4.1. *Let $\mathcal{A}$ be a label. Then*

(A) $\Gamma_\ell(x) \neq \emptyset$ *for any* $x \in X$ *and any* $\ell \le \ell(\mathcal{A})$.

(B) *Let* $1 \le \ell \le \ell(\mathcal{A})$, *let* $x, y \in X$, *and let* $\xi \in \Gamma_\ell(x)$. *Then there exists* $\eta \in \Gamma_{\ell-1}(y)$ *such that*

$$\|\xi - \eta\| \le d(x, y).$$

In Sections 4.2–4.9 we will prove the following result.

Main Lemma 4.2. *Let* $x_0 \in X$, $\xi_0 \in Y$, $r_0 > 0$, $C_B \ge 1$ *be given, and let* $\mathcal{A}$ *be a label.*

Suppose that $\Gamma_{\ell(\mathcal{A})}(x_0)$ *has an* $(\mathcal{A}, \varepsilon^{-1} r_0, C_B)$*-basis at* ξ_0, *where* $\varepsilon > 0$ *is less than a small enough constant* $\varepsilon^* > 0$ *determined by* m, C_B, c_N, D_N.

Then there exists $f : B(x_0, r_0) \to Y$ *with the following properties:*

$$\|f(z) - f(w)\| \le C(\varepsilon)\, d(z, w) \quad \text{for all} \quad z, w \in B(x_0, r_0), \tag{4.4}$$

$$\|f(z) - \xi_0\| \le C(\varepsilon)\, r_0 \quad \text{for all} \quad z \in B(x_0, r_0), \tag{4.5}$$

$$f(z) \in \Gamma_0(z) \quad \text{for all} \quad z \in B(x_0, r_0). \tag{4.6}$$

Here $C(\varepsilon)$ *is determined by* ε, m, C_B, c_N, D_N.

We will prove the Main Lemma 4.2 by downward induction on $\#\mathcal{A}$, starting with the case $\#\mathcal{A} = m$, and ending with the case $\#\mathcal{A} = 0$.

4.2 Proof of the Main Lemma in the base case. In this section, we assume the hypothesis of the Main Lemma 4.2 in the base case $\mathcal{A} = (e_1, \ldots, e_m)$. Thus, in this case $\#\mathcal{A} = m$ and $\ell(\mathcal{A}) = 2$, [see (4.1)].

We recall that for each $x \in X$ we have $\Gamma_\ell(x) \subset F(x) \subset \mathrm{Aff}_F(x)$ (all $\ell \ge 0$), where $\mathrm{Aff}_F(x)$ is a translate of the vector space $\mathrm{Vect}_F(x)$ of dimension $\le m$. We write C to denote a positive constant determined by m, C_B, c_N, D_N. This symbol may denote different constants in different occurrences.

Lemma 4.3. *For each* $z \in B(x_0, r_0)$, *there exists*

$$\eta^z \in \Gamma_1(z) \tag{4.7}$$

such that

$$\|\,\eta^z - \xi_0\,\| \le C\,\varepsilon^{-1} r_0, \tag{4.8}$$

$$\langle e_a, \eta^z - \xi_0 \rangle = 0 \quad \text{for} \quad a = 1, \ldots, m, \tag{4.9}$$

$$\Gamma_1(z) \quad \text{has an} \quad (\mathcal{A}, \varepsilon^{-1} r_0, C)\text{-basis at} \quad \eta^z. \tag{4.10}$$

Proof. We apply Lemma 3.10, taking Γ to be $\Gamma_2(x_0)$, Γ' to be $\Gamma_1(z)$, and r to be $\varepsilon^{-1} r_0$. To apply that lemma, we must check the key hypothesis (*), which asserts in the present case that

$$\text{Given } \ \xi \in \Gamma_2(x_0) \ \text{ there exists } \ \eta \in \Gamma_1(z) \ \text{ such that } \ \|\xi - \eta\| \le \varepsilon_0 \cdot (\varepsilon^{-1} r_0), \tag{4.11}$$

where ε_0 is a small enough constant determined by C_B and m.

To check (4.11), we recall Lemma 4.1 (B). Given $\xi \in \Gamma_2(x_0)$ there exists $\eta \in \Gamma_1(z)$ such that

$$\|\xi - \eta\| \le d(z, x_0) \le r_0 \quad (\text{because } z \in B(x_0, r_0)) < \varepsilon_0 \cdot (\varepsilon^{-1} r_0);$$

here, the last inequality holds thanks to our assumption that ε is less than a small enough constant determined by m, C_B, c_N, D_N.

Thus, (4.11) holds, and we may apply Lemma 3.10. That lemma provides a vector η^z satisfying (4.7),...,(4.10), completing the proof of Lemma 4.3. □

For each $z \in B(x_0, r_0)$, we fix a vector η^z as in Lemma 4.3. Repeating the idea of the proof of Lemma 4.3, we establish the following result.

Lemma 4.4. *Given $z, w \in B(x_0, r_0)$, there exists a vector*

$$\eta^{z,w} \in \Gamma_0(w) \tag{4.12}$$

such that

$$\|\eta^{z,w} - \eta^z\| \le C\,\varepsilon^{-1} d(z, w) \tag{4.13}$$

and

$$\langle e_a, \eta^{z,w} - \eta^z \rangle = 0 \quad \text{for} \quad a = 1, \ldots, m. \tag{4.14}$$

Proof. If $z = w$, we can just take $\eta^{z,w} = \eta^z$. Suppose $z \neq w$. Because $z, w \in B(x_0, r_0)$, we have $0 < d(z, w) \le 2r_0$. Therefore, (4.10) and Remark 3.8 (ii) tell us that

$$\Gamma_1(z) \quad \text{has an} \quad (\mathcal{A}, \tfrac{1}{2}\varepsilon^{-1} d(z, w), C)\text{-basis at} \quad \eta^z. \tag{4.15}$$

We prepare to apply Lemma 3.10, this time taking

$$\Gamma = \Gamma_1(z), \quad \Gamma' = \Gamma_0(w), \quad r = \tfrac{1}{2}\varepsilon^{-1} d(z, w).$$

We must verify the key hypothesis (*), which asserts in the present case that:

Given any $\xi \in \Gamma_1(z)$ there exists $\eta \in \Gamma_0(w)$ such that

$$\|\xi - \eta\| \le \varepsilon_0 \cdot (\tfrac{1}{2}\varepsilon^{-1} d(z, w)), \tag{4.16}$$

where ε_0 arises from the constant C in (4.15) as in Lemma 3.10. In particular, ε_0 depends only on m and C_B. Therefore, our assumption that ε is less than a small enough constant determined by m, C_B, c_N, D_N tells us that

$$d(z, w) < \varepsilon_0 \cdot (\tfrac{1}{2}\varepsilon^{-1} d(z, w)).$$

Consequently, Lemma 4.1 (B) produces for each $\xi \in \Gamma_1(z)$ an $\eta \in \Gamma_0(w)$ such that

$$\|\xi - \eta\| \le d(z, w) < \varepsilon_0 \cdot (\tfrac{1}{2}\varepsilon^{-1} d(z, w)),$$

which proves (4.16).

Therefore, we may apply Lemma 3.10. That lemma provides a vector $\eta^{z,w}$ satisfying (4.12), (4.13), (4.14), and additional properties that we don't need here.

The proof of Lemma 4.4 is complete. □

Lemma 4.5. *Let $w \in B(x_0, r_0)$. Then any vector $v \in \mathrm{Vect}_F(w)$ satisfying $\langle e_a, v\rangle = 0$ for $a = 1, \ldots, m$ must be the zero vector.*

Proof. Applying (4.10), we obtain an $(\mathcal{A}, \varepsilon^{-1}r_0, C)$-basis $(v_1, \ldots, v_m)$ for $\Gamma_1(w)$ at η^w. From the definition of an $(\mathcal{A}, \varepsilon^{-1}r_0, C)$-basis, see Definition 3.7, we have

$$\langle e_a, v_b\rangle = \delta_{ab} \quad \text{for} \quad a, b = 1, \ldots, m, \tag{4.17}$$

and

$$\eta^w + \tfrac{1}{C}\,\varepsilon^{-1}r_0 v_a,\ \eta^w - \tfrac{1}{C}\,\varepsilon^{-1}r_0 v_a \in \Gamma_1(w) \subset F(w) \subset \mathrm{Aff}_F(w) \quad \text{for} \quad a = 1, \ldots, m,$$

from which we deduce that

$$v_a \in \mathrm{Vect}_F(w) \quad \text{for} \quad a = 1, \ldots, m. \tag{4.18}$$

From (4.17), (4.18) we see that

$$v_1, \ldots, v_m \in \mathrm{Vect}_F(w)$$

are linearly independent. However, $\mathrm{Vect}_F(w)$ has dimension at most m. Therefore, $v_1, \ldots, v_m$ form a basis for $\mathrm{Vect}_F(w)$. Lemma 4.5 now follows at once from (4.17). □

Now let $z, w \in B(x_0, r_0)$. From Lemmas 4.3 and 4.4 we have

$$\eta^w,\ \eta^{z,w} \in \Gamma_0(w) \subset F(w) \subset \mathrm{Aff}_F(w),$$

and consequently

$$\eta^w - \eta^{z,w} \in \mathrm{Vect}_F(w). \tag{4.19}$$

On the other hand, (4.9) and (4.14) tell us that

$$\langle e_a, \eta^w - \xi_0\rangle = 0, \quad \langle e_a, \eta^z - \xi_0\rangle = 0, \quad \langle e_a, \eta^z - \eta^{z,w}\rangle = 0 \quad \text{for} \quad a = 1, \ldots, m.$$

Therefore,

$$\langle e_a, \eta^w - \eta^{z,w}\rangle = 0 \quad \text{for} \quad a = 1, \ldots, m. \tag{4.20}$$

From (4.19), (4.20) and Lemma 4.5, we conclude that $\eta^{z,w} = \eta^w$. Therefore, from (4.13), we obtain the estimate

$$\|\eta^z - \eta^w\| \le C\varepsilon^{-1}\, d(z, w) \quad \text{for} \quad z, w \in B(x_0, r_0). \tag{4.21}$$

We now define

$$f(z) = \eta^z \quad \text{for} \quad z \in B(x_0, r_0).$$

Then (4.7), (4.8), (4.21) tell us that

$$f(z) \in \Gamma_0(z) \quad \text{for all} \quad z \in B(x_0, r_0), \tag{4.22}$$

$$\|f(z) - \xi_0\| \le C\varepsilon^{-1}r_0 \quad \text{for} \quad z \in B(x_0, r_0), \tag{4.23}$$

and

$$\|f(z) - f(w)\| \le C\varepsilon^{-1}\, d(z, w) \quad \text{for} \quad z, w \in B(x_0, r_0). \tag{4.24}$$

Our results (4.22), (4.23), (4.24) immediately imply the conclusions of the Main Lemma 4.2.

This completes the proof of the Main Lemma 4.2 in the base case $\#\mathcal{A} = m$. □

4.3 Setup for the induction step. Fix a label $\mathcal{A} = (e_1, \ldots, e_s)$ with $0 \le s \le m-1$. We assume the

Inductive Hypothesis 4.6. Let $x_0^+ \in X$, $\xi_0^+ \in Y$, $r_0^+ > 0$, $C_B^+ \ge 1$ be given, and let $\mathcal{A}^+$ be a label such that $\#\mathcal{A}^+ > \#\mathcal{A}$.

Then the Main Lemma 4.2 holds, with x_0^+, ξ_0^+, r_0^+, C_B^+, $\mathcal{A}^+$, in place of x_0, ξ_0, r_0, C_B, $\mathcal{A}$, respectively.

We assume the

Hypotheses of the Main Lemma for the Label $\mathcal{A}$ 4.7. $x_0 \in X$, $\xi_0 \in Y$, $r_0 > 0$, $C_B \ge 1$, $\Gamma_{\ell(\mathcal{A})}(x_0)$has an $(\mathcal{A}, \varepsilon^{-1} r_0, C_B)$-basis at ξ_0.

We introduce a positive constant A, and we make the following assumptions.

Large A Assumption 4.8. A exceeds a large enough constant determined by m, C_B, $c_{\mathcal{N}}$, $D_{\mathcal{N}}$.

Small ε Assumption 4.9. ε is less than a small enough constant determined by A, m, C_B, $c_{\mathcal{N}}$, $D_{\mathcal{N}}$.

We write C to denote a positive constant determined by m, C_B, $c_{\mathcal{N}}$, $D_{\mathcal{N}}$; we write $C(\varepsilon, A)$ and $C'(\varepsilon, A)$ to denote positive constants determined by ε, m, A, C_B, $c_{\mathcal{N}}$, $D_{\mathcal{N}}$. These symbols may denote different constants in different occurrences.

Under the above assumptions, we will prove that there exists $f : B(x_0, r_0) \to Y$ satisfying

$$\|f(z) - f(w)\| \le C(\varepsilon, A)\, d(z, w) \quad \text{for all} \quad z, w \in B(x_0, r_0), \tag{4.25}$$

$$\|f(z) - \xi_0\| \le C(\varepsilon, A)\, r_0 \quad \text{for all} \quad z \in B(x_0, r_0), \tag{4.26}$$

$$f(z) \in \Gamma_0(z) \quad \text{for all} \quad z \in B(x_0, r_0). \tag{4.27}$$

These conclusions differ from the conclusions (4.4), (4.5), (4.6) of the Main Lemma 4.2 only in that here, $C(\varepsilon)$ is replaced by $C(\varepsilon, A)$.

Once we have proven the existence of such an f under the above assumptions, we then pick A to be a constant determined by m, C_B, $c_{\mathcal{N}}$, $D_{\mathcal{N}}$, taken large enough to satisfy the Large A Assumption 4.8.

Once we do so, our present Small ε Assumption 4.9 will follow from the small ε assumption made in the Main Lemma 4.2. Moreover, the conclusions (4.25), (4.26), (4.27) will then imply conclusions (4.4), (4.5), (4.6). Consequently, we will have proven the Main Lemma 4.2 for $\mathcal{A}$. That will complete our downward induction on $\#\mathcal{A}$, thereby proving the Main Lemma 4.2 for all labels.

To recapitulate:

We assume the Inductive Hypothesis 4.6 and the Hypotheses of the Main Lemma for the Label $\mathcal{A}$ 4.7, and we make the Large A Assumption 4.8 and the Small ε Assumption 4.9.

Under the above assumptions, our task is to prove that there exists $f : B(x_0, r_0) \to Y$ satisfying (4.25), (4.26), (4.27). Once we do that, the Main Lemma 4.2 will follow.

We keep the assumptions and notation of this section in force until the end of the proof of the Main Lemma 4.2.

4.4 A family of useful vectors. Recall that $\Gamma_{\ell(\mathcal{A})}(x_0)$ has an $(\mathcal{A}, \varepsilon^{-1}r_0, C_B)$-basis at ξ_0.

Let $z \in B(x_0, 10r_0)$. Then, thanks to our Small ε Assumption 4.9, we have

$$d(z, x_0) \le 10r_0 < \varepsilon_0 \cdot (\varepsilon^{-1}r_0), \tag{4.28}$$

where ε_0 arises from C_B, m as in Lemma 3.10.

We apply that lemma, taking $\Gamma = \Gamma_{\ell(\mathcal{A})}(x_0)$, $\Gamma' = \Gamma_{\ell(\mathcal{A})-1}(z)$, and $r = \varepsilon^{-1} r_0$, and using (4.28) and Lemma 4.1 (B) to verify the key hypothesis (*) in Lemma 3.10. Thus, we obtain a vector $\eta^z \in Y$, with the following properties:

$$\Gamma_{\ell(\mathcal{A})-1}(z) \quad \text{has an} \quad (\mathcal{A}, \varepsilon^{-1}r_0, C)\text{-basis at} \quad \eta^z, \tag{4.29}$$

$$\|\eta^z - \xi_0\| \le C\varepsilon^{-1}r_0, \tag{4.30}$$

and

$$\langle e_a, \eta^z - \xi_0 \rangle = 0 \quad \text{for} \quad a = 1, \dots, s. \tag{4.31}$$

We fix such a vector η^z for each $z \in B(x_0, 10r_0)$.

4.5 The basic lengthscales.

Definition 4.10. *Let $x \in B(x_0, 5r_0)$, and let $r > 0$. We say that (x, r) is OK if both conditions* (OK1) *and* (OK2) *below are satisfied.*

(OK1) $d(x_0, x) + 5r \le 5r_0$.

(OK2) *Either condition* (OK2A) *or condition* (OK2B) *below is satisfied.*

(OK2A) $\#B(x, 5r) \le 1$ *(i.e., $B(x, 5r)$ is the singleton $\{x\}$).*

(OK2B) *For some label $\mathcal{A}^+$ with $\#\mathcal{A}^+ > \#\mathcal{A}$, the following holds: For each $w \in B(x, 5r)$ there exists a vector $\zeta^w \in Y$ satisfying conditions* (OK2Bi), (OK2Bii), (OK2Biii) *below:*

(OK2Bi) $\Gamma_{\ell(\mathcal{A})-3}(w)$ *has an $(\mathcal{A}^+, \varepsilon^{-1}r, A)$-basis at ζ^w.*

(OK2Bii) $\|\zeta^w - \xi_0\| \le A\varepsilon^{-1}r_0$.

(OK2Biii) $\langle e_a, \zeta^w - \xi_0 \rangle = 0$ *for $a = 1, \dots, s$.*

Of course (OK1) guarantees that $B(x, 5r) \subset B(x_0, 5r_0)$.

Note that (x, r) cannot be OK if $r > r_0$, because then (OK1) cannot hold. On the other hand, if $x \in B(x_0, 5r_0)$, then $d(x_0, x) < 5r_0$, hence (OK1) holds for small enough r, and (OK2) holds as well (because $B(x, 5r) = \{x\}$ for small enough r; recall that (X, d) is a finite metric space). Thus, for fixed $x \in B(x_0, 5r_0)$, we find that (x, r) is OK if r is small enough, but not if r is too big.

For each $x \in B(x_0, 5r_0)$ we may therefore

$$\text{fix a } \textit{basic lengthscale} \quad r(x) > 0, \tag{4.32}$$

such that

$$(x, r(x)) \text{ is OK, but } (x, 2r(x)) \text{ is not OK.} \tag{4.33}$$

Indeed, we may just take $r(x)$ to be any r' such that (x, r') is OK and

$$r' > \tfrac{1}{2} \sup \{r : (x, r) \text{ is OK}\}.$$

We let RELX (relevant X) denote the set of all $x \in B(x_0, 5r_0)$ such that

$$B(x, r(x)) \cap B(x_0, r_0) \neq \emptyset. \tag{4.34}$$

Clearly,

$$B(x_0, r_0) \subset \text{RELX}. \tag{4.35}$$

From (4.33) and (OK1), we have

$$d(x_0, x) + 5r(x) \leq 5r_0 \quad \text{for each} \quad x \in B(x_0, 5r_0).$$

LEMMA 4.11. *Let $z_1, z_2 \in B(x_0, 5r_0)$. If*

$$d(z_1, z_2) \leq r(z_1) + r(z_2), \tag{4.36}$$

then

$$\tfrac{1}{4} r(z_1) \leq r(z_2) \leq 4r(z_1).$$

Proof. Suppose not. After possibly interchanging z_1 and z_2, we have

$$r(z_1) < \tfrac{1}{4} r(z_2). \tag{4.37}$$

Now $(z_2, r(z_2))$ is OK (see (4.33)). Therefore it satisfies (OK1), i.e.,

$$d(x_0, z_2) + 5r(z_2) \leq 5r_0.$$

Therefore, by (4.36),

$$\begin{aligned} &d(x_0, z_1) + 5 \cdot (2r(z_1)) \\ &\quad \leq d(x_0, z_2) + d(z_1, z_2) + 10r(z_1) \leq d(x_0, z_2) + r(z_1) + r(z_2) + 10r(z_1) \\ &\quad \leq d(x_0, z_2) + \tfrac{11}{4} r(z_2) + r(z_2) < d(x_0, z_2) + 5r(z_2) \leq 5r_0, \end{aligned}$$

i.e., $(z_1, 2r(z_1))$ satisfies (OK1).

Moreover,

$$B(z_1, 10r(z_1)) \subset B(z_2, 5r(z_2)). \tag{4.38}$$

Indeed, if $w \in B(z_1, 10r(z_1))$, then (4.37) and (4.36) give

$$d(w, z_2) \le d(w, z_1) + d(z_1, z_2) \le 10r(z_1) + r(z_1) + r(z_2) \le \tfrac{11}{4} r(z_2) + r(z_2) < 5r(z_2),$$

proving (4.38).

Because $(z_2, r(z_2))$ is OK, it satisfies (OK2A) or (OK2B). If $(z_2, r(z_2))$ satisfies (OK2A), then so does $(z_1, 2r(z_1))$, thanks to (4.38). In that case, $(z_1, 2r(z_1))$ satisfies (OK1) and (OK2A), hence $(z_1, 2r(z_1))$ is OK, contradicting (4.33).

On the other hand, suppose $(z_2, r(z_2))$ satisfies (OK2B). Fix $\mathcal{A}^+$ with $\#\mathcal{A}^+ > \#\mathcal{A}$ such that for every $w \in B(z_2, 5r(z_2))$ there exists ζ^w satisfying

- $\Gamma_{\ell(\mathcal{A})-3}(w)$ has an $(\mathcal{A}^+, \varepsilon^{-1} r(z_2), A)$-basis at ζ^w.
- $\|\zeta^w - \xi_0\| \le A\varepsilon^{-1} r_0$.
- $\langle e_a, \zeta^w - \xi_0 \rangle = 0 \quad$ for $a = 1, \dots, s$.

Thanks to (4.38) there exists such a ζ^w for every $w \in B(z_1, 5 \cdot (2r(z_1)))$.

Note that, by (4.37) and Remark 3.8 (ii), the $(\mathcal{A}^+, \varepsilon^{-1} r(z_2), A)$-basis in the first bullet point above is also an $(\mathcal{A}^+, \varepsilon^{-1} \cdot (2r(z_1)), A)$-basis.

It follows that $(z_1, 2r(z_1))$ satisfies (OK2B). We have seen that $(z_1, 2r(z_1))$ satisfies (OK1), so again $(z_1, 2r(z_1))$ is OK, contradicting (4.33).

Thus, in all cases, our assumption that Lemma 4.11 fails leads to a contradiction. □

4.6 Consistency of the useful vectors. Recall the useful vectors η^z ($z \in B(x_0, 10r_0)$), see (4.29), (4.30), (4.31), and the set RELX, see (4.34). In this section we establish the following result.

Lemma 4.12. *Let* $z_1, z_2 \in \mathrm{RELX}$. *Then*

$$\|\eta^{z_1} - \eta^{z_2}\| \le C\varepsilon^{-1}[r(z_1) + r(z_2) + d(z_1, z_2)].$$

Proof. If

$$r(z_1) + r(z_2) + d(z_1, z_2) \ge r_0/10,$$

then the lemma follows from (4.30) applied to $z = z_1$ and to $z = z_2$.

Suppose

$$r(z_1) + r(z_2) + d(z_1, z_2) < r_0/10. \tag{4.39}$$

Because $z_1 \in \mathrm{RELX}$, we have $d(z_1, x_0) \le r_0 + r(z_1)$, hence

$$d(z_1, x_0) + 5 \cdot (2r(z_1)) \le r_0 + 11r(z_1) < 5r_0.$$

Thus $(z_1, 2r(z_1))$ satisfies (OK1), and, in particular, $B(z_1, 10r(z_1)) \subset B(x_0, 5r_0)$.

Recall from (4.29) that $\Gamma_{\ell(\mathcal{A})-1}(z_2)$ has an $(\mathcal{A}, \varepsilon^{-1} r_0, C)$-basis at η^{z_2}. By (4.39) and Remark 3.8 (ii), it follows that

$$\Gamma_{\ell(\mathcal{A})-1}(z_2) \quad \text{has an} \quad (\mathcal{A}, \varepsilon^{-1}[r(z_1) + r(z_2) + d(z_1, z_2)], C)\text{-basis at} \quad \eta^{z_2}. \tag{4.40}$$

Our Small ε Assumption 4.9 shows that

$$d(z_1, z_2) \le \varepsilon_0 \cdot \varepsilon^{-1}[r(z_1) + r(z_2) + d(z_1, z_2)],$$

for the ε_0 arising from Lemma 3.10, where we use the constant C in (4.40) as the constant C_B in Lemma 3.10.

Therefore, by Lemmas 3.10 and 4.1 (B), with

$$\Gamma = \Gamma_{\ell(\mathcal{A})-1}(z_2), \quad \Gamma' = \Gamma_{\ell(\mathcal{A})-2}(z_1), \quad r = \varepsilon^{-1}[r(z_1) + r(z_2) + d(z_1, z_2)],$$

we obtain a vector $\zeta \in \Gamma_{\ell(\mathcal{A})-2}(z_1)$ such that

$$\|\zeta - \eta^{z_2}\| \le C\varepsilon^{-1}[r(z_1) + r(z_2) + d(z_1, z_2)] \tag{4.41}$$

and

$$\langle e_a, \zeta - \eta^{z_2}\rangle = 0 \quad \text{for} \quad a = 1, \dots, s,$$

hence

$$\langle e_a, \zeta - \eta^{z_1}\rangle = 0 \quad \text{for} \quad a = 1, \dots, s. \quad (\text{See } (4.31).) \tag{4.42}$$

We will prove that

$$\|\zeta - \eta^{z_1}\| \le \varepsilon^{-1} r(z_1);$$

(4.41) will then imply the conclusion of Lemma 4.12.

Suppose instead that

$$\|\zeta - \eta^{z_1}\| > \varepsilon^{-1} r(z_1). \tag{4.43}$$

We will derive a contradiction. By (4.29), Remark 3.8 (iv), and because $r(z_1) < r_0/10$ [see (4.39)], we know that

$$\Gamma_{\ell(\mathcal{A})-2}(z_1) \quad \text{has an} \quad (\mathcal{A}, \varepsilon^{-1} r(z_1), C)\text{-basis at} \quad \eta^{z_1}. \tag{4.44}$$

Our results (4.42), (4.44) and our assumption (4.43) are the hypotheses of Lemma 3.9 (“Adding a vector”). Applying that lemma, we obtain a vector $\hat{\zeta} \in \Gamma_{\ell(\mathcal{A})-2}(z_1)$, with the following properties:

$$\begin{aligned} &\|\hat{\zeta} - \eta^{z_1}\| = \tfrac{1}{2}\varepsilon^{-1} r(z_1), \\ &\langle e_a, \hat{\zeta} - \eta^{z_1}\rangle = 0 \quad \text{for} \quad a = 1, \dots, s; \end{aligned} \tag{4.45}$$

also

$$\Gamma_{\ell(\mathcal{A})-2}(z_1) \quad \text{has an} \quad (\mathcal{A}^+, \varepsilon^{-1} r(z_1), C)\text{-basis at} \quad \hat{\zeta}, \tag{4.46}$$

for a label of the form $\mathcal{A}^+ = (e_1, \dots, e_s, e_{s+1})$; and

$$\langle e_a, \hat{\zeta} - \xi_0\rangle = 0 \quad \text{for} \quad a = 1, \dots, s. \tag{4.47}$$

See (4.31).

In particular, $\#\mathcal{A}^+ = \#\mathcal{A} + 1$.

From (4.46) and Remark 3.8 (iii) we have, with a larger constant C,

$$\Gamma_{\ell(\mathcal{A})-2}(z_1) \quad \text{has an} \quad (\mathcal{A}^+, \varepsilon^{-1} \cdot (2r(z_1)), C)\text{-basis at} \ \hat{\zeta}. \tag{4.48}$$

Now let $w \in B(z_1, 5 \cdot (2r(z_1)))$. Let ε_0 arise from Lemma 3.10 where we use C from (4.48) as the constant C_B in Lemma 3.10. We have

$$d(z_1, w) < 10r(z_1) < \varepsilon_0 \cdot (\varepsilon^{-1} \cdot (2r(z_1))),$$

thanks to our Small ε Assumption 4.9. Therefore, Lemma 4.1 (B) allows us to verify the key hypothesis (*) in Lemma 3.10, with $\Gamma = \Gamma_{\ell(\mathcal{A})-2}(z_1)$, $\Gamma' = \Gamma_{\ell(\mathcal{A})-3}(w)$, $r = \varepsilon^{-1} \cdot (2r(z_1))$.

Applying Lemma 3.10, we obtain a vector $\zeta^w \in \Gamma_{\ell(\mathcal{A})-3}(w)$ with the following properties:

$$\begin{aligned} &\|\zeta^w - \hat{\zeta}\| \le C\varepsilon^{-1} \cdot (2r(z_1)), \\ &\langle e_a, \zeta^w - \hat{\zeta}\rangle = 0 \quad \text{for} \quad a = 1, \dots, s+1; \end{aligned} \tag{4.49}$$

hence by (4.47),

$$\langle e_a, \zeta^w - \xi_0\rangle = 0 \quad \text{for} \quad a = 1, \dots, s. \tag{4.50}$$

Also,

$$\Gamma_{\ell(\mathcal{A})-3}(w) \quad \text{has an} \quad (\mathcal{A}^+, \varepsilon^{-1} \cdot (2r(z_1)), C)\text{-basis at} \ \zeta^w. \tag{4.51}$$

We have

$$\|\zeta^w - \xi_0\| \le \|\zeta^w - \hat{\zeta}\| + \|\hat{\zeta} - \eta^{z_1}\| + \|\eta^{z_1} - \xi_0\| \le C\varepsilon^{-1} r(z_1) + \tfrac{1}{2}\varepsilon^{-1} r(z_1) + C\varepsilon^{-1} r_0$$

by (4.49), (4.45) and (4.30).

Recalling that $r(z_1) < r_0/10$, we conclude that

$$\|\zeta^w - \xi_0\| \le C\varepsilon^{-1} \cdot r_0. \tag{4.52}$$

Thus, for every $w \in B(z_1, 5 \cdot (2r(z_1)))$, our vector ζ^w satisfies (4.50), (4.51), (4.52). Comparing (4.51), (4.52), (4.50) with (OK2Bi), (OK2Bii), (OK2Biii), and recalling our Large A Assumption 4.8, we conclude that (OK2B) holds for $(z_1, 2r(z_1))$. We have already seen that (OK1) holds for $(z_1, 2r(z_1))$. Thus $(z_1, 2r(z_1))$ is OK, contradicting the defining property (4.33) of $r(z_1)$.

This contradiction proves that (4.43) cannot hold, completing the proof of Lemma 4.12. □

4.7 Additional useful vectors.

LEMMA 4.13. *Let $x \in B(x_0, 5r_0)$, and suppose that $\#B(x, 5r(x)) \geq 2$.*
Then there exist a vector $\zeta^x \in Y$ and a label $\mathcal{A}^+$ with the following properties:

$$\#\mathcal{A}^+ > \#\mathcal{A}, \tag{4.53}$$
$$\Gamma_{\ell(\mathcal{A})-3}(x) \quad \text{has an} \quad (\mathcal{A}^+, \varepsilon^{-1}r(x), A)\text{-basis at} \;\; \zeta^x, \tag{4.54}$$
$$\|\zeta^x - \eta^x\| \leq \varepsilon^{-1}r(x), \tag{4.55}$$
$$\langle e_a, \zeta^x - \eta^x\rangle = 0 \quad \text{for} \quad a = 1, \dots, s. \tag{4.56}$$

Proof. Recall that $(x, r(x))$ is OK. We are assuming that (OK2A) fails for $(x, r(x))$, hence (OK2B) holds. Fix $\mathcal{A}^+$ as in (OK2B), and let ζ^x be as in (OK2B) with $w = x$. Then (4.53), (4.54), (4.56) hold, thanks to (OK2B); however, (4.55) may fail in case $r(x)$ is much smaller than r_0. If (4.55) holds, we are done.

Suppose instead that (4.55) fails, i.e.,

$$\|\zeta^x - \eta^x\| > \varepsilon^{-1}r(x). \tag{4.57}$$

We recall from (4.29) that $\Gamma_{\ell(\mathcal{A})-1}(x)$ has an $(\mathcal{A}, \varepsilon^{-1}r_0, C)$-basis at η^x. We have also $r(x) \leq r_0$ because $(x, r(x))$ is OK; and

$$\Gamma_{\ell(\mathcal{A})-1}(x) \subset \Gamma_{\ell(\mathcal{A})-3}(x).$$

Therefore, by Remark 3.8 (iv),

$$\Gamma_{\ell(\mathcal{A})-3}(x) \quad \text{has an} \quad (\mathcal{A}, \varepsilon^{-1}r(x), C)\text{-basis at} \;\; \eta^x. \tag{4.58}$$

From (4.56), (4.57), (4.58) and Lemma 3.9 ("Adding a vector"), we obtain a vector $\hat{\zeta} \in Y$ and a label $\hat{\mathcal{A}}$ with the following properties:

$$\#\hat{\mathcal{A}} > \#\mathcal{A}, \tag{4.59}$$
$$\|\hat{\zeta} - \eta^x\| = \tfrac{1}{2}\varepsilon^{-1}r(x), \tag{4.60}$$
$$\langle e_a, \hat{\zeta} - \eta^x\rangle = 0 \quad \text{for} \quad a = 1, \dots, s, \tag{4.61}$$
$$\Gamma_{\ell(\mathcal{A})-3}(x) \quad \text{has an} \quad (\hat{\mathcal{A}}, \varepsilon^{-1}r(x), C')\text{-basis at} \;\; \hat{\zeta}. \tag{4.62}$$

Comparing (4.59),...,(4.62) with (4.53),...,(4.56), and recalling our Large A Assumption 4.8, we see that $\hat{\zeta}$ and $\hat{\mathcal{A}}$ have all the properties asserted for ζ^x and $\mathcal{A}^+$ in the statement of Lemma 4.13.

Thus, Lemma 4.13 holds in all cases. □

4.8 Local selections.

LEMMA 4.14. *Given* $x \in \mathrm{RELX}$, *there exists* $f : B(x, r(x)) \to Y$ *with the following properties:*

(I) $\|f(z) - f(w)\| \le C(\varepsilon, A)\, d(z, w)$ *for* $z, w \in B(x, r(x))$.
(II) $f(z) \in \Gamma_0(z)$ *for* $z \in B(x, r(x))$.
(III) $\|f(z) - \eta^x\| \le C(\varepsilon, A)\, r(x)$ *for* $z \in B(x, r(x))$.
(IV) $\|f(z) - \xi_0\| \le C(\varepsilon, A)\, r_0$ *for* $z \in B(x, r(x))$.

Proof. We proceed by cases.

Case 1. Suppose $\#B(x, 5r(x)) > 1$.

Then Lemma 4.13 applies. Let $\mathcal{A}^+$, ζ^x be as in that lemma. Thus,

$$\#\mathcal{A}^+ > \#\mathcal{A}, \tag{4.63}$$

$$\|\zeta^x - \eta^x\| \le \varepsilon^{-1} r(x) \tag{4.64}$$

and

$$\Gamma_{\ell(\mathcal{A})-3}(x) \quad \text{has an} \quad (\mathcal{A}^+, \varepsilon^{-1} r(x), A)\text{-basis at} \ \ \zeta^x;$$

hence, by Remark 3.8 (iv),

$$\Gamma_{\ell(\mathcal{A}^+)}(x) \quad \text{has an} \quad (\mathcal{A}^+, \varepsilon^{-1} r(x), A)\text{-basis at} \ \ \zeta^x, \tag{4.65}$$

because $\ell(\mathcal{A}) - 3 \ge \ell(\mathcal{A}^+)$ whenever $\#\mathcal{A}^+ > \#\mathcal{A}$.

We recall from our Small ε Assumption 4.9 that

$$\varepsilon \ \text{ is less than a small enough constant determined by } A, c_{\mathcal{N}}, D_{\mathcal{N}}, m. \tag{4.66}$$

Thanks to (4.65), (4.66), the Hypotheses of the Main Lemma 4.7 are satisfied, with $\mathcal{A}^+$, x, ζ^x, $r(x)$, A, in place of $\mathcal{A}$, x_0, ξ_0, r_0, C_B, respectively. Moreover, thanks to (4.63) and the Inductive Hypothesis 4.6, we are assuming the validity of the Main Lemma 4.2 for $\mathcal{A}^+, \dots, A$.

Therefore, we obtain a function $f : B(x, r(x)) \to Y$ satisfying (I), (II) and the inequality

$$\|f(z) - \zeta^x\| \le C(\varepsilon, A)\, r(x), \qquad z \in B(x, r(x)).$$

This inequality together with (4.64) implies (III).

Moreover, (IV) follows from (III) because, for $z \in B(x, r(x)) \subset B(x_0, 5r_0)$, we have

$$\|f(z) - \xi_0\| \le \|f(z) - \eta^x\| + \|\eta^x - \xi_0\| \le C(\varepsilon, A) r(x) + C\varepsilon^{-1} r_0 \le C'(\varepsilon, A) r_0;$$

here we use (4.30) and the fact that $(x, r(x))$ satisfies (OK1).

This completes the proof of Lemma 4.14 in Case 1.

Case 2. Suppose $\#B(x, 5r(x)) \le 1$.

Then, $B(x, 5r(x)) = \{x\}$ and $\eta^x \in \Gamma_{\ell(\mathcal{A})-1}(x) \subset \Gamma_0(x)$. Hence the function $f(x) = \eta^x$ satisfies (I),(II),(III), and also (IV) thanks to (4.30).

Thus, Lemma 4.14 holds in all cases. □

4.9 Proof of the Main Lemma: the final step. Let $\mathcal{B}_0$ be the metric space

$$\mathcal{B}_0 = \left(B(x_0, r_0), d|_{B(x_0,r_0)\times B(x_0,r_0)}\right),$$

i.e., the ball $B(x_0, r_0)$ supplied with the metric d.

For the rest of Section 4.9, we work in the metric space $\mathcal{B}_0$. Given $x \in B(x_0, r_0)$ and $r > 0$, we write $\widetilde{B}(x, r)$ to denote the ball in $\mathcal{B}_0$ with center x and radius r; thus $\widetilde{B}(x, r) = B(x, r) \cap B(x_0, r_0)$.

Note that, since (X, d) satisfies Nagata $(D_{\mathcal{N}}, c_{\mathcal{N}})$, the metric space $\mathcal{B}_0$ satisfies Nagata $(D_{\mathcal{N}}, c_{\mathcal{N}})$ as well. See Definition 1.5.

Let $r : X \to \mathbb{R}_+$ be *the basic lengthscale* constructed in Section 4.5 (see (4.32)), and let

$$C_{LS} = 4 \quad \text{and} \quad a = (4\,C_{LS})^{-1}. \tag{4.67}$$

Note that, by Lemma 4.11, Consistency of the Lengthscale (see (2.1)) holds for the lengthscale $r(x)$ on $B(x_0, r_0)$ with the constant C_{LS} given by (4.67).

We apply the Whitney Partition Lemma 2.1 to the metric space $\mathcal{B}_0$, the lengthscale

$$\{r(x) : x \in B(x_0, r_0)\}$$

and the constants C_{LS}, a determined by (4.67), and obtain a partition of unity $\{\theta_\nu : B(x_0, r_0) \to \mathbb{R}_+\}$ and points

$$x_\nu \in B(x_0, r_0) \tag{4.68}$$

with the following properties.

- Each $\theta_\nu \geq 0$ and for each ν, $\theta_\nu = 0$ outside $\widetilde{B}(x_\nu, ar_\nu)$; here a is determined by (4.67), and $r_\nu = r(x_\nu)$.
- Any given x satisfies $\theta_\nu(x) \neq 0$ for at most D^* distinct ν, where D^* depends only on $c_{\mathcal{N}}$, $D_{\mathcal{N}}$.
- $\sum\limits_\nu \theta_\nu(x) = 1$ for all $x \in B(x_0, r_0)$.
- Each θ_ν satisfies

$$|\theta_\nu(x) - \theta_\nu(y)| \leq \frac{C}{r_\nu}\, d(x, y)$$

for all $x, y \in B(x_0, r_0)$; here again $r_\nu = r(x_\nu)$.

From Lemma 4.11, we know that

- For each μ, ν, if $d(x_\mu, x_\nu) \leq r_\mu + r_\nu$, then $\frac{1}{4} r_\nu \leq r_\mu \leq 4 r_\nu$.

Moreover, by (4.35) and (4.68),

$$x_\nu \in \text{RELX} \quad \text{for each} \quad \nu, \tag{4.69}$$

so that, by Lemma 4.14, there exists a function $\hat{f}_\nu : B(x_\nu, r_\nu) \to Y$ satisfying the following conditions

- $\|\hat{f}_\nu(z) - \hat{f}_\nu(w)\| \le C(\varepsilon, A)\, d(z,w) \quad$ for $\; z, w \in B(x_\nu, r_\nu)$.
- $\hat{f}_\nu(z) \in \Gamma_0(z) \quad$ for $\; z \in B(x_\nu, r_\nu)$.
- $\|\hat{f}_\nu(z) - \eta_\nu\| \le C(\varepsilon, A)\, r_\nu \quad$ for $\; z \in B(x_\nu, r_\nu)$, where $\eta_\nu \equiv \eta^{x_\nu}$.
- $\|\hat{f}_\nu(z) - \xi_0\| \le C(\varepsilon, A)\, r_0 \quad$ for $\; z \in B(x_\nu, r_\nu)$.

Let $f_\nu = \hat{f}_\nu|_{\widetilde{B}(x_\nu, r_\nu)}$. We extend f_ν from $\widetilde{B}(x_\nu, r_\nu) = B(x_\nu, r_\nu) \cap B(x_0, r_0)$ to all of $B(x_0, r_0)$ by setting $f_\nu = 0$ outside $\widetilde{B}(x_\nu, r_\nu)$.

Since each $x_\nu \in$ RELX (see (4.69)), from Lemma 4.12, we have

- $\|\eta_\nu - \eta_\mu\| \le C(\varepsilon, A) \cdot [r_\nu + r_\mu + d(x_\nu, x_\mu)]$ for each μ, ν.

The above conditions on the θ_ν, η_ν, $\hat{f}_\nu$, f_ν, r_ν and a (cf. (2.15) with (4.67)) allow us to apply the Patching Lemma 2.12 on $\mathcal{B}_0$. We conclude that

$$f(x) = \sum_\nu \theta_\nu(x)\, f_\nu(x) \quad (\text{all} \;\; x \in B(x_0, r_0))$$

satisfies

$$\|f(x) - f(y)\| \le C(\varepsilon, A)\, d(x,y) \quad \text{for} \;\; x, y \in B(x_0, r_0).$$

Moreover, for fixed $x \in B(x_0, r_0)$, we know that $f(x)$ is a convex combination of finitely many values $f_\nu(x)$ with $\widetilde{B}(x_\nu, ar_\nu) \ni x$; for those ν we have $f_\nu(x) \in \Gamma_0(x)$ and $\|f_\nu(x) - \xi_0\| \le C(\varepsilon, A)\, r_0$. Therefore, $f(x) \in \Gamma_0(x)$ and $\|f(x) - \xi_0\| \le C(\varepsilon, A)\, r_0$ for all $x \in B(x_0, r_0)$.

Thus, f satisfies (4.25), (4.26) and (4.27), completing the proof of the Main Lemma 4.2. □

Proof of the Finiteness Theorem 3.6 for bounded Nagata dimension. Let $x_0 \in X$, $r_0 = \operatorname{diam} X + 1$, $C_B = 1$, and $\mathcal{A} = (\;)$. Let $\varepsilon = \frac{1}{2}\varepsilon^*$ where ε^* is as in the Main Lemma 4.2 for m, $C_B = 1$, $c_{\mathcal{N}}$ and $D_{\mathcal{N}}$. Thus, ε depends only on m, $c_{\mathcal{N}}$ and $D_{\mathcal{N}}$.

By Lemma 4.1 (A), $\Gamma_{\ell(\mathcal{A})}(x_0) \ne \emptyset$ so that there exists $\xi_0 \in \Gamma_{\ell(\mathcal{A})}(x_0)$. Since $\#\mathcal{A} = 0$, the set $\Gamma_{\ell(\mathcal{A})}(x_0)$ has an $(\mathcal{A}, \varepsilon^{-1} r_0, C_B)$-basis at ξ_0. See Remark 3.8, (i).

Hence, by the Main Lemma 4.2, there exists a mapping $f : B(x_0, r_0) \to Y$ such that

$$\|f(z) - f(w)\| \le C\, d(z,w) \quad \text{for all} \quad z, w \in B(x_0, r_0),$$

and

$$f(z) \in \Gamma_0(z) \quad \text{for all} \quad z \in B(x_0, r_0).$$

Here C is a constant determined by ε, m, C_B, $c_{\mathcal{N}}$, $D_{\mathcal{N}}$. Thus, C depends only on m, $c_{\mathcal{N}}$, $D_{\mathcal{N}}$.

Clearly, $B(x_0, r_0) = X$. Furthermore, $\Gamma_0(z) \subset F(z)$ for every $z \in X$ (see (3.6)), so that $f(z) \in F(z)$, $z \in X$. Thus, f is a Lipschitz selection of F on X with Lipschitz seminorm at most a certain constant depending only on m, $c_{\mathcal{N}}$, $D_{\mathcal{N}}$.

The proof of Theorem 3.6 is complete. □

Recall that Theorem 3.6 immediately implies Theorem 1.6.

4.10 The Finiteness Principle on metric trees. Let us consider an important example of a metric space with finite Nagata dimension.

Let $T = (X, E)$ be a *finite tree*. Here X denotes the set of nodes and E denotes the set of edges of T. We write $x \leftrightarrow y$ to indicate that nodes $x, y \in X$, $x \neq y$, are joined by an edge; we denote that edge by $[xy]$.

Suppose we assign a positive number $\Delta(e)$ to each edge $e \in E$. Then we obtain a notion of distance $d(x, y)$ for any $x, y \in X$, as follows.

We set

$$d(x,x) = 0 \quad \text{for every} \quad x \in X. \tag{4.70}$$

Because T is a tree, any two distinct nodes $x, y \in X$ are joined by one and only "path"

$$x = x_0 \leftrightarrow x_1 \leftrightarrow \cdots \leftrightarrow x_L = y \quad \text{with all the} \ \ x_i \ \ \text{distinct.}$$

We define

$$d(x,y) = \sum_{i=1}^{L} \Delta([x_{i-1}x_i]). \tag{4.71}$$

We call the resulting metric space (X, d) a *metric tree*.

For the reader's convenience we prove the following slight variant of a result from [LS05].

LEMMA 4.15. *Every metric tree satisfies Nagata* $(1, c)$ *with* $c = 1/16$. *(See Definition 1.5).*

Proof. Given a metric tree (X, d), we fix an origin $0 \in X$ and make the following definition:

Every point $x \in X$ is joined to the origin by one and only one "path"

$$0 = x_0 \leftrightarrow x_1 \leftrightarrow \cdots \leftrightarrow x_L = x, \quad \text{with all the} \ \ x_i \ \ \textit{distinct.}$$

We call $x_0, x_1, \ldots, x_L$ the *ancestors* of x. We define the *distinguished ancestor* of x, denoted $\mathrm{DA}(x)$, to be x_i for the smallest $i \in \{0, \ldots, L\}$ for which

$$d(0, x_i) > \lfloor d(0,x) \rfloor - 1, \tag{4.72}$$

where $\lfloor \cdot \rfloor$ denotes the greatest integer function. (Note that there is at least one x_i satisfying (4.72), namely $x_L = x$. Thus, every $x \in X$ has a distinguished ancestor.)

We note two simple properties of $\mathrm{DA}(x)$, namely,

(1) $d(x, \mathrm{DA}(x)) \leq 2$;

(2) $\mathrm{DA}(x)$ is an ancestor of any ancestor y of x that satisfies $d(0,y) > \lfloor d(0,x) \rfloor - 1$.

We now exhibit a Nagata covering of X for the lengthscale $s = 4$.

For $q = 0, 1$ and $z \in X$, let

$$X_q(z) = \{x \in X : z = \mathrm{DA}(x) \ \ \text{and} \ \ \lfloor d(0,x) \rfloor \equiv q \mod 2\}.$$

Clearly, the $X_q(z)$ cover X. Moreover, (1) tells us that each $X_q(z)$ has diameter at most 4.

We assert the following

CLAIM: If $z \neq z'$ and $q = q'$, then the distance from $X_q(z)$ to $X_{q'}(z')$ is at least 1/2.

The CLAIM immediately implies that any given ball $B \subset X$ of radius 1/4 meets at most one of the $X_0(z)$ and at most one of the $X_1(z)$, hence at most two of the $X_q(z)$.

Let us establish the CLAIM; if it were false, then we could find

$$z \neq z', \; q \in \{0,1\}, \; x \in X_q(z), \; x' \in X_q(z') \quad \text{with} \quad d(x,x') \leq 1/2.$$

We will derive a contradiction from these conditions as follows.

Because $d(x,x') \leq 1/2$, we have

$$|\,\lfloor d(0,x)\rfloor - \lfloor d(0,x')\rfloor\,| \leq 1.$$

On the other hand, $\lfloor d(0,x)\rfloor \equiv \lfloor d(0,x')\rfloor \mod 2$. Hence, $\lfloor d(0,x)\rfloor = \lfloor d(0,x')\rfloor$.

Next, let $\tilde{z}$ be the closest common ancestor of x, x'. Because $d(x,x') \leq 1/2$, we have $d(x,\tilde{z}) \leq 1/2$ and $d(x',\tilde{z}) \leq 1/2$, and therefore the ancestor $\tilde{z}$ of x satisfies

$$d(0,\tilde{z}) > \lfloor d(0,x)\rfloor - 1.$$

Hence, (2) implies that z is an ancestor of $\tilde{z}$. Similarly, z' is an ancestor of $\tilde{z}$.

It follows that either z is an ancestor of z', or z' is an ancestor of z. Without loss of generality, we may suppose that z is an ancestor of z'. Consequently, z is an ancestor of x'; moreover,

$$d(0,z) > \lfloor d(0,x)\rfloor - 1 = \lfloor d(0,x')\rfloor - 1.$$

Thanks to (2), we now know that z' is an ancestor of z. Thus, each of the points z, z' is an ancestor of the other, and therefore $z = z'$, contradicting an assumption that the CLAIM is false.

We have produced a covering of an arbitrary metric tree by subsets X_i of diameter at most 4, such that no ball of radius 1/4 intersects more than two of the X_i.

Applying the above result to the metric tree $(X, \frac{4}{s}\, d)$ for given $s > 0$, we produce a covering of X by X_i such that, with respect to d, each X_i has diameter at most s, and no ball of radius $s/16$ meets more than two of the X_i. Thus, we have verified the Nagata condition for metric trees. □

Let us apply Theorem 1.6 to *metric trees*. Thus, we obtain the following

COROLLARY 4.16. *Let $m \in \mathbb{N}$, let (X,d) be a metric tree and let λ be a positive constant.*

Let $F : X \to \mathrm{Conv}_m(Y)$ be a set-valued mapping such that, for every subset $X' \subset X$ with $\#X' \leq k^\sharp$, the restriction $F|_{X'}$ has a Lipschitz selection $f_{X'} : X' \to Y$ with $\|f_{X'}\|_{\mathrm{Lip}(X',Y)} \leq \lambda$.

Then F has a Lipschitz selection $f : X \to Y$ with $\|f\|_{\mathrm{Lip}(X,Y)} \le \gamma_0\,\lambda$.

Here $k^\sharp = k^\sharp(m)$ is the constant from Theorem 1.6, and $\gamma_0 = \gamma_0(m)$ is a constant depending only on m.

5 Metric Trees and Lipschitz Selections with Respect to the Hausdorff Distance

We recall that $(Y, \|\cdot\|)$ denotes a Banach space, and $\mathcal{K}_m(Y)$ denotes the family of all nonempty compact convex subsets $K \subset Y$ of dimension at most m. Recall also that $\mathrm{d_H}(A, B)$ denotes the Hausdorff distance between $A, B \in \mathcal{K}_m(Y)$.

In this section we work with finite trees $T = (X, E)$, where X denotes the set of nodes and E denotes the set of edges of T. As in Section 4.10, we write $u \leftrightarrow v$ to indicate that $u, v \in X$ are distinct nodes joined by an edge in T.

We supply X with a metric d defined by formulae (4.70) and (4.71), and we refer to the metric space (X, d) as a *metric tree (with respect to the tree $T = (X, E)$)*.

Remark 5.1. Sometimes we will be looking simultaneously at two different pseudometrics, say ρ and $\tilde{\rho}$, on a pseudometric space, say on $\mathcal{M}$. In this case we will speak of a ρ-Lipschitz selection and ρ-Lipschitz seminorm, or a $\tilde{\rho}$-Lipschitz selection and $\tilde{\rho}$-Lipschitz seminorm to make clear which pseudometric we are using. Furthermore, sometimes given a mapping $f : \mathcal{M} \to Y$ we will write $\|f\|_{\mathrm{Lip}((\mathcal{M},\rho),Y)}$ to denote the Lipschitz seminorm of f with respect to the pseudometric ρ.

5.1 The "core" of a set-valued mapping and the Finiteness Principle.

Until the end of Section 6 we write $k^\sharp$ and γ_0 to denote the constants from Corollary 4.16. Recall that these constants depend only on m.

In this and the next subsection we prove the following result.

Theorem 5.2. *Let $(\mathcal{M}, \rho)$ be a metric space, and let $F : \mathcal{M} \to \mathcal{K}_m(Y)$ for a Banach space Y. Let λ be a positive real number.*

Suppose that for every subset $\mathcal{M}' \subset \mathcal{M}$ consisting of at most $k^\sharp$ points, the restriction $F|_{\mathcal{M}'}$ has a Lipschitz selection $f_{\mathcal{M}'}$ with Lipschitz seminorm $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}(\mathcal{M}',Y)} \le \lambda$.

Then there exists a mapping $G : \mathcal{M} \to \mathcal{K}_m(Y)$ satisfying the following conditions:

(i) *$G(x) \subset F(x)$ for every $x \in \mathcal{M}$;*

(ii) *For every $x, y \in \mathcal{M}$ the following inequality*

$$\mathrm{d_H}(G(x), G(y)) \le \gamma_0\,\lambda\,\rho(x, y)$$

holds.

Let $(\mathcal{M}, \rho)$ be a metric space and let $F : \mathcal{M} \to \mathcal{K}_m(Y)$ be a set-valued mapping. We suppose that the following assumption is satisfied.

Assumption 5.3. *For every subset $\mathcal{M}' \subset \mathcal{M}$ consisting of at most $k^\sharp$ points, the restriction $F|_{\mathcal{M}'}$ of F to $\mathcal{M}'$ has a ρ-Lipschitz selection $f_{\mathcal{M}'} : \mathcal{M}' \to Y$ with $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}((\mathcal{M}',\rho),Y)} \le \lambda$.*

Our aim is to prove the existence of a mapping $G : \mathcal{M} \to \mathcal{K}_m(Y)$ satisfying conditions (i) and (ii) of Theorem 5.2. We refer to G as a "core" of the set-valued mapping F.

Let $T = (X, E)$ be an arbitrary finite tree. We introduce the following

DEFINITION 5.4. *A mapping* $\psi : X \to \mathcal{M}$ *is said to be admissible with respect to* T *if for every two distinct nodes* $u, v \in X$ *with* $u \leftrightarrow v$ *(i.e.,* u *is joined by an edge to* v*), we have* $\psi(u) \neq \psi(v)$.

Let $\psi : X \to \mathcal{M}$ be an admissible mapping. Then ψ gives rise a *tree metric* $d_{T,\psi} : X \times X \to \mathbb{R}_+$ defined by

$$d_{T,\psi}(u, v) = \rho(\psi(u), \psi(v)) \quad \text{for every} \;\; u, v \in X, \;\; u \leftrightarrow v. \tag{5.1}$$

See (4.71).

Clearly, by the triangle inequality,

$$\rho(\psi(u), \psi(v)) \le d_{T,\psi}(u, v) \quad \text{for every} \;\; u, v \in X. \tag{5.2}$$

Now define a set-valued mapping $F_{T,\psi} : X \to \mathcal{K}_m(Y)$ by the formula

$$F_{T,\psi}(u) = F(\psi(u)), \quad u \in X.$$

LEMMA 5.5. *The set-valued mapping* $F_{T,\psi} = F \circ \psi$ *has a* $d_{T,\psi}$*-Lipschitz selection* $f : X \to Y$ *such that*

$$\|f\|_{\mathrm{Lip}((X,\, d_{T,\psi}),Y)} \le \gamma_0\, \lambda. \tag{5.3}$$

Proof. Let $X' \subset X$ be an arbitrary subset of X with $\#X' \le k^\sharp$, and let $\mathcal{M}' = \psi(X')$. Then

$$\#\mathcal{M}' \le \#X' \le k^\sharp$$

so that, by Assumption 5.3, the restriction $F|_{\mathcal{M}'}$ has a ρ-Lipschitz selection $f_{\mathcal{M}'} : \mathcal{M}' \to Y$ with $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}((\mathcal{M}',\rho),Y)} \le \lambda$.

Let $g_{X'} : X' \to Y$ be defined by

$$g_{X'}(u) = f_{\mathcal{M}'}(\psi(u)), \quad u \in X'.$$

Then $g_{X'}$ is a *selection* of the restriction $F_{T,\psi}|_{X'}$, i.e., $g_{X'}(u) \in F_{T,\psi}(u)$ for all $u \in X'$. Furthermore, for every $u, v \in X'$

$$\|g_{X'}(u) - g_{X'}(v)\| = \|f_{\mathcal{M}'}(\psi(u)) - f_{\mathcal{M}'}(\psi(v))\| \le \lambda\, \rho(\psi(u), \psi(v))$$

so that, by (5.2),

$$\|g_{X'}(u) - g_{X'}(v)\| \le \lambda\, d_{T,\psi}(u, v)$$

proving that the $d_{T,\psi}$-Lipschitz seminorm of $g_{X'}$ is bounded by λ.

Hence, by Corollary 4.16, the set-valued mapping $F_{T,\psi}$ has a $d_{T,\psi}$-Lipschitz selection $f : X \to Y$ satisfying inequality (5.3). □

We will need the following two definitions.

Definition 5.6. *Let* $x \in \mathcal{M}$. *The family* $\mathrm{APT}(x)$ *consists of all triples* $L = [T, a, \psi]$ *where*

- $T = (X, E)$ *is a finite tree with the family of nodes* X *and the family of edges* E*;*
- $a \in X$ *is a node of* T*;*
- $\psi : X \to \mathcal{M}$ *is an admissible mapping with respect to* T *such that* $\psi(a) = x$.

We refer to each triple $L = [T, a, \psi] \in \mathrm{APT}(x)$ *as an admissibly placed tree rooted at* a. *We call* $\mathrm{APT}(x)$ *the family of all Admissibly Placed Trees associated with* x.

Definition 5.7. *Let* $x \in \mathcal{M}$. *Given a finite tree* $T = (X, E)$ *and a triple* $L = [T, a, \psi] \in \mathrm{APT}(x)$ *we let* $O(x; L)$ *denote the subset of* Y *defined by*

$$O(x; L) = \{f(a) : f \text{ is a } d_{T,\psi}\text{-Lipschitz selection of } F_{T,\psi} \text{ with } \|f\|_{\mathrm{Lip}((X,d_{T,\psi}),Y)} \le \gamma_0\, \lambda\}.$$

We recall that a convex subset of Y has dimension at most m if it is contained in an affine subspace of Y of dimension at most m.

Lemma 5.8. *Let* $x \in \mathcal{M}$ *and let* $L = [T, a, \psi] \in \mathrm{APT}(x)$. *Then* $O(x; L)$ *is a nonempty compact convex subset of* $F(x)$ *of dimension at most* m.

Proof. By Lemma 5.5, the mapping $F_{T,\psi} = F \circ \psi$ has a $d_{T,\psi}$-Lipschitz selection $f : X \to Y$ with $\|f\|_{\mathrm{Lip}((X,d_{T,\psi}),Y)} \le \gamma_0\, \lambda$. Therefore, by Definition 5.7, $f(a) \in O(x; L)$ proving that $O(x; L) \ne \emptyset$.

The convexity of $O(x; L)$ directly follows from the convexity of sets $F(y)$ ($y \in \mathcal{M}$) and Definition 5.7. Furthermore, if $f : X \to Y$ is a selection of $F_{T,\psi} = F \circ \psi$, then $f(a) \in F(\psi(a)) = F(x)$ (recall that $x = \psi(a)$, see Definition 5.6).

Hence, $O(x; L) \subset F(x)$. This also proves that $\dim O(x; L) \le \dim F(x) \le m$.

Let us prove that $O(x; L)$ is compact whenever each set $F(y), y \in \mathcal{M}$, is. Since $O(x; L) \subset F(x)$ and $F(x)$ is a compact set, $O(x; L)$ is a bounded set. We prove that $O(x; L)$ is closed.

Let $h \in Y$, and a let $h_n \in O(x; L), n = 1, 2, \ldots$ be a sequence of points converging to h:

$$h = \lim_{n\to\infty} h_n. \tag{5.4}$$

We will prove that $h \in O(x; L)$.

By Definition 5.7, there exists a sequence of mappings $f_n \in \mathrm{Lip}((X, d_{T,\psi}), Y)$ such that

$$f_n(u) \in F(\psi(u)) \quad \text{and} \quad \|f_n\|_{\mathrm{Lip}((X,d_{T,\psi}),Y)} \le \gamma_0\, \lambda \tag{5.5}$$

for every $u \in X$ and $n \in \mathbb{N}$, and

$$h_n = f_n(a), \quad n = 1, 2, \ldots. \tag{5.6}$$

Note that $(X, d_{T,\psi})$ is a *finite* metric space, and each set $F(\psi(u)), u \in X$, is a finite dimensional compact subset of Y. Therefore, there exists a subsequence $n_k \in \mathbb{N}$, $k = 1, 2, \ldots$, such that $(f_{n_k}(u))_{k=1}^{\infty}$ converges in Y for every $u \in X$. Let

$$\tilde{f}(u) = \lim_{k\to\infty} f_{n_k}(u), \quad u \in X. \tag{5.7}$$

Then, by (5.4) and (5.6),

$$h = \lim_{k\to\infty} h_{n_k} = \lim_{k\to\infty} f_{n_k}(a) = \tilde{f}(a). \tag{5.8}$$

Since each set $F(\psi(u)), u \in X$, is closed, by (5.5) and (5.7), $\tilde{f}(u) \in F(\psi(u))$ for every $u \in X$, proving that $\tilde{f}$ is a *selection* of the set-valued mapping $F_{T,\psi} = F \circ \psi$ on X. Since each mapping $f_n : X \to Y$ is $d_{T,\psi}$-Lipschitz with $\|f_n\|_{\mathrm{Lip}((X,d_{T,\psi}),Y)} \le \gamma_0\,\lambda$, by (5.7), $\tilde{f}$ is $d_{T,\psi}$-Lipschitz as well, with $\|\tilde{f}\|_{\mathrm{Lip}((X,d_{T,\psi}),Y)} \le \gamma_0\,\lambda$.

Thus, by (5.8) and Definition 5.7, $h \in O(x;L)$ proving the lemma. □

Given $x \in \mathcal{M}$ let

$$G(x) = \bigcap_{L\,\in\,\mathrm{APT}(x)} O(x;L). \tag{5.9}$$

Clearly, by Lemma 5.8, for every $x \in \mathcal{M}$ the set

$$G(x) \;\; \textit{is a convex compact subset of} \;\; F(x). \tag{5.10}$$

In the next section, we will prove that $G(x) \ne \emptyset$ for each $x \in \mathcal{M}$ and that

$$\mathrm{d_H}(G(x), G(y)) \le \gamma_0\,\lambda\,\rho(x,y) \quad \text{for every} \;\; x, y \in \mathcal{M}. \tag{5.11}$$

Recall that $\mathrm{d_H}$ denotes the Hausdorff distance between subsets of Y.

5.2 Lipschitz continuity of the "core" with respect to the Hausdorff distance.

LEMMA 5.9. *For every $x \in \mathcal{M}$, the set $G(x) \ne \emptyset$.*

We must show that

$$\bigcap_{L\,\in\,\mathrm{APT}(x)} O(x;L) \ne \emptyset.$$

See (5.9). By Lemma 5.8, each $O(x;L)$ is a nonempty compact subset of the compact set $F(x)$. Therefore, it is enough to show that

$$O(x;L_1) \cap \cdots \cap O(x;L_N) \ne \emptyset \tag{5.12}$$

for every finite subcollection $\{L_1, \ldots, L_N\} \subset \mathrm{APT}(x)$.

Let $L_1, \ldots, L_N \in \mathrm{APT}(x)$ with $L_i = [T_i, a_i, \psi_i]$, $i = 1, \ldots, N$, where each $T_i = (X_i, E_i)$ is a finite tree.

We introduce a procedure for gluing the finite trees $T_i = (X_i, E_i)$, $i = 1, \dots, N$, together. Recall that X_i here denotes the set of nodes of T_i, and E_i denotes the set of edges of T_i. By passing to isomorphic copies of the T_i, we may assume that the sets X_i are pairwise disjoint. Then we form a finite tree $T^+ = (X^+, E^+)$ from $T_1, \dots, T_N$ by identifying together all the nodes $a_1, \dots, a_N$. We spell out details below.

For each i, we write J_i to denote the set of all the neighbors of a_i in T_i. Also, we write X'_i to denote the set $X_i \setminus \{a_i\}$, and we write E'_i to denote all the edges in T_i that join together points of X'_i (i.e. not including a_i as an endpoint).

We introduce a new node a^+ distinct from all the nodes of all the T_i.

The finite tree $T^+ = (X^+, E^+)$ is then defined as follows. The nodes in X^+ are all the nodes in all the X'_i, together with the single node a^+. The edges in E^+ are all the edges belonging to any of the E'_i, together with edges joining a^+ to all the nodes in all the J_i. One checks easily that T^+ is a finite tree. We say that T^+ arises by "gluing together the T_i by identifying the a_i".

Note that T^+ contains an isomorphic copy of each T_i as a subtree; the relevant isomorphism φ_i carries the node a_i of T_i to the node a^+ of T^+, and φ_i is the identity on all other nodes of T_i. Each edge $[ab]$ of the tree T_i is carried to the edge $[\varphi_i(a)\,\varphi_i(b)]$ of T^+.

This concludes our discussion of the gluing of trees T_i.

We define a map $\psi^+ : X^+ \to \mathcal{M}$ by setting

$$\psi^+(a^+) = x \tag{5.13}$$

and

$$\psi^+(b) = \psi_i(b) \quad \text{for all} \quad b \in X'_i = X_i \setminus \{a_i\},\ i = 1, \dots, N. \tag{5.14}$$

One checks that ψ^+ is an admissible map, and $\psi^+(a^+) = x$. Thus, $L^+ = [T^+, a^+, \psi^+]$ belongs to $\mathrm{APT}(x)$. Consequently, by Lemma 5.5, there exists a d_{T^+,ψ^+}-Lipschitz selection f^+ of $F \circ \psi^+$ with d_{T^+,ψ^+}-Lipschitz seminorm $\le \gamma_0\,\lambda$. (We recall that the metric d_{T^+,ψ^+} is determined by formula (5.1).)

The map

$$f_i(b) = \begin{cases} f^+(b), & \text{if } \ b \in X_i \setminus \{a_i\}, \\ f^+(a^+), & \text{if } \ b = a_i, \end{cases}$$

is a d_{T_i,ψ_i}-Lipschitz selection of $F \circ \psi_i$ with d_{T_i,ψ_i}-Lipschitz seminorm $\le \gamma_0\,\lambda$, therefore

$$f^+(a^+) \in O(x; L_i) \quad \text{for each} \quad i = 1, \dots, N.$$

Thus, (5.12) holds, completing the proof of Lemma 5.9. □

We know that the affine dimension of each set $F(x)$ is at most m. Since $G(x) \subset F(x)$, the same is true for each set $G(x)$, $x \in \mathcal{M}$. This observation, Lemma 5.9 and statement (5.10) imply that G maps the metric space $\mathcal{M}$ into the family $\mathcal{K}_m(Y)$.

We are in a position to prove inequality (5.11).

Lemma 5.10. *For every $x, y \in \mathcal{M}$ the following inequality*

$$\mathrm{d_H}(G(x), G(y)) \le \gamma_0\, \lambda\, \rho(x, y)$$

holds.

Proof. We may suppose $x \ne y$, else the desired conclusion is obvious. Let us prove that

$$I = G(x) + \gamma_0\, \lambda\, \rho(x, y)\, B_Y \supset G(y). \tag{5.15}$$

Recall that by $B_Y = B_Y(0, 1)$ we denote the closed unit ball in Y.

If we can prove that, then by interchanging the roles of x and y we obtain also

$$G(y) + \gamma_0\, \lambda\, \rho(x, y)\, B_Y \supset G(x).$$

These two inclusions tell us that $\mathrm{d_H}(G(x), G(y)) \le \gamma_0\, \lambda\, \rho(x, y)$, proving the lemma.

Let us prove (5.15). By definition,

$$I = \left[\bigcap_{L \,\in\, \mathrm{APT}(x)} O(x; L)\right] + \gamma_0\, \lambda\, \rho(x, y)\, B_Y.$$

See (5.9). We will check that

$$\left[\bigcap_{L \,\in\, \mathrm{APT}(x)} O(x; L)\right] + \gamma_0\, \lambda\, \rho(x, y)\, B_Y$$
$$= \bigcap \left\{ [O(x; L_1) \cap \cdots \cap O(x; L_N)] + \gamma_0\, \lambda\, \rho(x, y)\, B_Y \right\}, \tag{5.16}$$

where the first intersection of the right-hand side is taken over all finite sequences $L_1, \dots, L_N$ of elements of $\mathrm{APT}(x)$.

Indeed, the left-hand side of (5.16) is obviously contained in the right-hand side. Conversely, let ξ belong to the right-hand side of (5.16). Then any finite subcollection of the compact sets

$$K_L = \{\eta \in B_Y : \xi - \gamma_0\, \lambda\, \rho(x, y)\, \eta \in O(x; L)\}$$

has nonempty intersection. (The above sets are compact because $O(x; L)$ is compact.)

Therefore,

$$\bigcap_{L \in \mathrm{APT}(x)} K_L \ne \emptyset,$$

proving that ξ belongs to the left-hand side of (5.16). The proof of (5.16) is complete.

Thanks to (5.16), our desired inclusion (5.15) will follow if we can show that

$$[O(x; L_1) \cap \cdots \cap O(x; L_N)] + \gamma_0\, \lambda\, \rho(x, y)\, B_Y \supset G(y) \tag{5.17}$$

for any $L_1, \dots, L_N \in \mathrm{APT}(x)$. Then the proof of Lemma 5.10 is reduced to the task of proving (5.17).

Let $L_i = [T_i, a_i, \psi_i]$ where $T_i = (X_i, E_i)$. Then a_i is a node of the tree T_i, $i = 1, \dots, N$. We introduce a new node a^+ and form the tree $T^+ = (X^+, E^+)$ as in the proof of Lemma 5.9. Thus T^+ arises by *gluing together the trees* T_i *by identifying the* a_i.

We also introduce an admissible map $\psi^+ : X^+ \to \mathcal{M}$ as in the proof of Lemma 5.9, see (5.13) and (5.14).

We now introduce a new node $\tilde{a}$ not present in T^+. We define a new tree $\widetilde{T} = (\widetilde{X}, \widetilde{E})$ as follows.

- The nodes in $\widetilde{X}$ are the nodes in X^+, together with the new node $\tilde{a}$.
- The edges in $\widetilde{E}$ are the edges in E^+, together with a single edge joining $\tilde{a}$ to a^+.

We define a map $\widetilde{\psi} : \widetilde{T} \to \mathcal{M}$ by setting

$$\widetilde{\psi} = \psi^+ \quad \text{on} \quad T^+, \quad \widetilde{\psi}(\tilde{a}) = y.$$

Then one checks that $\widetilde{T} = (\widetilde{X}, \widetilde{E})$ is a tree and $\widetilde{\psi}(\tilde{a}) = y$. Furthermore, since ψ^+ is admissible on T^+ and $x \neq y$, the mapping $\widetilde{\psi}$ is admissible on $\widetilde{T}$.

Let $\widetilde{L} = [\widetilde{T}, \tilde{a}, \widetilde{\psi}]$, and let $\eta \in G(y)$. Then, by definition (5.9), $\eta \in O(y; \widetilde{L})$ so that there exists a $d_{\widetilde{T},\widetilde{\psi}}$-Lipschitz selection $\tilde{f}$ of $F \circ \widetilde{\psi}$, with $d_{\widetilde{T},\widetilde{\psi}}$-Lipschitz seminorm $\le \gamma_0\, \lambda$, satisfying $\tilde{f}(\tilde{a}) = \eta$. See Definition 5.7. [We also recall that the metric $d_{\widetilde{T},\widetilde{\psi}}$ is defined by formulae (5.1) and (4.71).]

Restricting this $\tilde{f}$ to T^+ and arguing as in the proof of Lemma 5.9, we see that

$$\tilde{f}(a^+) \in O(x; L_1) \cap \dots \cap O(x; L_N).$$

On the other hand, our Lipschitz bound for $\tilde{f}$ gives

$$\|\tilde{f}(a^+) - \eta\| = \|\tilde{f}(a^+) - \tilde{f}(\tilde{a})\| \le \gamma_0\, \lambda\, \rho(\widetilde{\psi}(a^+), \widetilde{\psi}(\tilde{a})) = \gamma_0\, \lambda\, \rho(x, y).$$

Then,

$$\eta \in [O(x; L_1) \cap \dots \cap O(x; L_N)] + \gamma_0\, \lambda\, \rho(x, y)\, B_Y$$

proving (5.17). □

The proof of Theorem 5.2 is complete. □

We turn to the final step of the proof of Theorem 1.7. The following selection theorem is a special case of [Shv04, Theorem 1.2].

Theorem 5.11. *Let Y be a Banach space, and let $m \ge 1$. Then there exists a map* $\mathrm{St} : \mathcal{K}_m(Y) \to Y$ *such that*

(α) $\operatorname{St}(K) \in K$ *for all* $K \in \mathcal{K}_m(Y)$

and

(β) $\|\operatorname{St}(K) - \operatorname{St}(K')\| \le C(m) \cdot \mathrm{d_H}(K, K')$ *for all* $K, K' \in \mathcal{K}_m(Y)$.

Here $C(m)$ *depends only on* m.

We refer to $\operatorname{St}(K)$ as the *"Steiner-type point" of* K. In the special case $Y = \mathbb{R}^m$, we can take $\operatorname{St}(K)$ to be the *Steiner point* of K. Recall that the Steiner point of K may be defined as the limit as $R \to \infty$ of the barycenter of $K + B(R)$, where "+" denotes Minkowski sum, and $B(R)$ is the standard Euclidean ball of radius R about 0. For general Y, there is no simple description of the "Steiner-type point" $\operatorname{St}(K)$ in [Shv04].

To construct the Lipschitz selection f and establish Theorem 1.7, we just set

$$f(x) = \operatorname{St}(G(x)) \quad \text{for} \quad x \in \mathcal{M},$$

where G is the core defined by (5.9). Since $G(x) \in \mathcal{K}_m(Y)$ for each $x \in \mathcal{M}$, the function f is well defined on $\mathcal{M}$.

By part (i) of Theorem 5.2 and part (α) of Theorem 5.11,

$$f(x) = \operatorname{St}(G(x)) \in G(x) \subset F(x) \quad \text{for} \quad x \in \mathcal{M}.$$

On the other hand, part (ii) of Theorem 5.2 and part (β) of Theorem 5.11 imply that

$$\begin{aligned}\|f(x) - f(y)\| = \|\operatorname{St}(G(x)) - \operatorname{St}(G(y))\| &\le C(m) \cdot \mathrm{d_H}(G(x), G(y)) \\ &\le C(m) \cdot \gamma_0\, \lambda\, \rho(x, y)\end{aligned}$$

for all $x, y \in \mathcal{M}$. Thus, f is a Lipschitz selection of F with Lipschitz seminorm at most $C(m) \cdot \gamma_0\, \lambda$. Recalling that $C(m)$ and γ_0 depend only on m, we conclude that Theorem 1.7 holds. □

6 Pseudometric Spaces

In this section we prove Theorem 1.2, the Finiteness Principle for Lipschitz Selections, and Theorem 6.2, a variant of Theorem 1.2 for finite pseudometric spaces.

Until the end of Section 6 we write γ_1 to denote the constant $\gamma_1 = \gamma_1(m)$ from Theorem 1.7. Everywhere in Section 6 we write $k^\sharp = k^\sharp(m)$ and $\gamma_0 = \gamma_0(m)$ to denote the constants from Corollary 4.16.

Let $(\mathcal{M}, \rho)$ be a pseudometric space. Recall that we say that the *pseudometric* ρ *is finite* if

$$\rho(x, y) < \infty \quad \text{for all} \;\; x, y \in \mathcal{M}. \tag{6.1}$$

On the other hand, we say that $(\mathcal{M}, \rho)$ is a *finite pseudometric space* if $\mathcal{M}$ contains only finitely many points.

Given a set-valued mapping $F : \mathcal{M} \to \operatorname{Conv}_m(Y)$, by a *selection* of F (not necessarily Lipschitz) we mean simply a map $f : \mathcal{M} \to Y$ such that $f(x) \in F(x)$ for all $x \in \mathcal{M}$.

6.1 The final step of the proof of the Finiteness Principle. In this section we prove an analog of Theorem 1.7 for pseudometric spaces.

PROPOSITION 6.1. *Let $(\mathcal{M},\rho)$ be a pseudometric space satisfying (6.1), and let $\lambda > 0$. Let $F:\mathcal{M}\to\mathcal{K}_m(Y)$ be a set-valued mapping such that for every subset $\mathcal{M}'\subset\mathcal{M}$ consisting of at most $k^\sharp$ points, the restriction $F|_{\mathcal{M}'}$ of F to $\mathcal{M}'$ has a Lipschitz selection $f_{\mathcal{M}'}:\mathcal{M}'\to Y$ with $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}(\mathcal{M}',Y)}\le\lambda$.*

Then F has a Lipschitz selection $f:\mathcal{M}\to Y$ with $\|f\|_{\mathrm{Lip}(\mathcal{M},Y)}\le\gamma_1\lambda$.

Proof. A selection of F may be regarded as a point of the Cartesian product

$$\mathcal{F}=\prod_{x\in\mathcal{M}}F(x).$$

We endow $\mathcal{F}$ with the product topology. Then, by Tychonoff's theorem, $\mathcal{F}$ is compact because each $F(x)$ is compact.

For $\varepsilon>0$ and $x,y\in\mathcal{M}$, let

$$\rho_\varepsilon(x,y)=\begin{cases}\rho(x,y)+\varepsilon, & \text{if } \ x\ne y,\\ 0, & \text{if } \ x=y.\end{cases}$$

Then $(\mathcal{M},\rho_\varepsilon)$ is a metric space. For any $\mathcal{M}'\subset\mathcal{M}$ with $\#\mathcal{M}'\le k^\sharp$ there exists a selection of $F|_{\mathcal{M}'}$ with ρ-Lipschitz seminorm $\le\lambda$, hence with ρ_ε-Lipschitz seminorm $\le\lambda$. By Theorem 1.7, F has a selection with ρ_ε-Lipschitz seminorm $\le\gamma_1\lambda$.

Let Selec(ε) be the set of all selections of F with ρ_ε-Lipschitz seminorm at most $\gamma_1\lambda$. Then Selec(ε) is a closed subset of $\mathcal{F}$. We have just seen that Selec(ε) is nonempty. Because

$$\mathrm{Selec}(\varepsilon)\subset\mathrm{Selec}(\varepsilon')\quad\text{for}\quad\varepsilon<\varepsilon',$$

it follows that

$$\mathrm{Selec}(\varepsilon_1)\cap\mathrm{Selec}(\varepsilon_2)\cap\cdots\cap\mathrm{Selec}(\varepsilon_N)\ne\emptyset$$

for any $\varepsilon_1,\varepsilon_2,\dots,\varepsilon_N>0$.

Because $\mathcal{F}$ is compact and each Selec(ε) is closed in $\mathcal{F}$, it follows that

$$\bigcap_{\varepsilon>0}\mathrm{Selec}(\varepsilon)\ne\emptyset.$$

Furthermore, any $f\in\cap\{\mathrm{Selec}(\varepsilon):\varepsilon>0\}$ is a selection of F with ρ-Lipschitz seminorm $\le\gamma_1\lambda$.

The proof of Proposition 6.1 is complete. □

Proof of Theorem 1.2. Suppose that ρ is a finite pseudometric, i.e., condition (6.1) holds.

Let $\mathcal{M}'$ be an arbitrary subset of $\mathcal{M}$ consisting of at most $k^\sharp$ points. (Note that, by definitions (1.1), (1.2) and (4.2), (4.3), the constant $k^\sharp > N(m,Y)$ for every $m \in \mathbb{N}$.) Then, by the theorem's hypothesis, for every set $S \subset \mathcal{M}'$ with $\#S \le N(m,Y)$, the restriction $F|_S$ has a Lipschitz selection $f_S : S \to Y$ with $\|f_S\|_{\mathrm{Lip}(S,Y)} \le \lambda$. Hence, by Theorem 1.8, the restriction $F|_{\mathcal{M}'}$ of F to $\mathcal{M}'$ has a Lipschitz selection $f_{\mathcal{M}'} : \mathcal{M}' \to Y$ whose seminorm satisfies $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}(\mathcal{M}',Y)} \le \gamma\lambda$ where γ is a constant depending only on m and $\#\mathcal{M}'$. Since $\#\mathcal{M}' \le k^\sharp$ and $k^\sharp$ depends only on m, the constant γ depends only on m as well.

Hence, by Proposition 6.1, F has a Lipschitz selection $f : \mathcal{M} \to Y$ with $\|f\|_{\mathrm{Lip}(\mathcal{M},Y)} \le \gamma_1\gamma\lambda$. Recall that γ_1 is a constant depending only on m.

This completes the proof of Theorem 1.2 for the case of a *finite* pseudometric ρ.

To pass to the general case in which $\rho(x,y)$ may take the value $+\infty$ is an easy exercise. We define an equivalence relation on $\mathcal{M}$ by calling x and y equivalent when $\rho(x,y)$ is finite. On each equivalence class we produce a Lipschitz selection of F, with controlled Lipschitz seminorm, by invoking the known case of Theorem 1.2 in which all distances are finite. By combining those Lipschitz selections into a single function defined on the union of all the equivalence classes, we obtain the desired Lipschitz selection of F. Details are spelled out in [FS].

The proof of Theorem 1.2 is complete. □

6.2 Finite pseudometric spaces. In this section we prove a variant of our main result, Theorem 1.2, related to the case of *finite* pseudometric spaces. As we have noted in the Introduction, for the case of the trivial distance function $\rho \equiv 0$ defined on a finite pseudometric space, Theorem 6.2 below agrees with the classical Helly's Theorem [DGK63] [up to the values of $N(m,Y)$ and the optimal finiteness constant for $\rho \equiv 0$ (see 1.4)].

Theorem 6.2. *Let $(\mathcal{M},\rho)$ be a finite pseudometric space, and let $F : \mathcal{M} \to \mathrm{Conv}_m(Y)$ be a set-valued mapping from $\mathcal{M}$ into the family $\mathrm{Conv}_m(Y)$ of all convex subsets of Y of affine dimension at most m. Let λ be a positive real number.*

Suppose that for every subset $\mathcal{M}' \subset \mathcal{M}$ consisting of at most $N(m,Y)$ points, the restriction $F|_{\mathcal{M}'}$ of F to $\mathcal{M}'$ has a Lipschitz selection $f_{\mathcal{M}'}$ with Lipschitz seminorm $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}(\mathcal{M}',Y)} \le \lambda$.

Then F has a Lipschitz selection f with Lipschitz seminorm $\|f\|_{\mathrm{Lip}(\mathcal{M},Y)} \le \gamma\,\lambda$.

Here, γ depends only on m.

Our proof of this result relies on an analog of Proposition 6.1 for a *finite* pseudometric space $(\mathcal{M},\rho)$ and a set-valued mapping $F : \mathcal{M} \to \mathrm{Conv}_m(Y)$. See Proposition 6.6 below.

We will need three auxiliary lemmas.

LEMMA 6.3. *Let $\lambda > 0$ and let $(\mathcal{M}, \rho)$ be a finite metric space. Let F be a set-valued mapping on $\mathcal{M}$ which to every $x \in \mathcal{M}$ assigns a nonempty convex bounded subset of Y of dimension at most m.*

Suppose that for every subset $\mathcal{M}' \subset \mathcal{M}$ with $\#\mathcal{M}' \le k^\sharp$, the restriction $F|_{\mathcal{M}'}$ of F to $\mathcal{M}'$ has a Lipschitz selection $f_{\mathcal{M}'} : \mathcal{M}' \to Y$ with $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}(\mathcal{M}',Y)} \le \lambda$.

Then F has a Lipschitz selection $f : \mathcal{M} \to Y$ with $\|f\|_{\mathrm{Lip}(\mathcal{M},Y)} \le 2\gamma_1\lambda$. Here, γ_1 is as in Proposition 6.1.

Proof. We introduce a new set-valued mapping on $\mathcal{M}$ defined by

$$\widetilde{F}(x) = (F(x))^{\,\mathrm{cl}} \quad \text{for all} \quad x \in \mathcal{M}.$$

Here the sign cl denotes the closure of a set in Y.

Since the sets $F(x)$, $x \in \mathcal{M}$, are finite dimensional and bounded, each set $\widetilde{F}(x)$ is compact so that $\widetilde{F} : \mathcal{M} \to \mathcal{K}_m(Y)$. Furthermore, since $F(x) \subset \widetilde{F}(x)$ on $\mathcal{M}$, the mapping $\widetilde{F}$ satisfies the hypothesis of Proposition 6.1.

By this proposition, there exists a mapping $\tilde{f} : \mathcal{M} \to Y$ such that

$$\tilde{f}(x) \in \widetilde{F}(x) = (F(x))^{\,\mathrm{cl}} \quad \text{for all} \quad x \in \mathcal{M}, \tag{6.2}$$

and

$$\|\tilde{f}(x) - \tilde{f}(y)\| \le \gamma_1 \,\lambda \,\rho(x, y) \quad \text{for all} \quad x, y \in \mathcal{M}. \tag{6.3}$$

Since $\mathcal{M}$ is a *finite* metric space, the following quantity

$$\delta = \gamma_1 \,\lambda \min_{x,y \in \mathcal{M},\, x \ne y} \rho(x, y) \tag{6.4}$$

is positive. Therefore, by (6.2), for each $x \in \mathcal{M}$ there exists a point $f(x) \in F(x)$ such that

$$\|f(x) - \tilde{f}(x)\| \le \delta/2.$$

Thus $f : \mathcal{M} \to Y$ is a selection of F on $\mathcal{M}$. Let us estimate its Lipschitz seminorm. For every $x, y \in \mathcal{M}$ (distinct), by (6.3) and (6.4),

$$\begin{aligned} \|f(x) - f(y)\| &\le \|f(x) - \tilde{f}(x)\| + \|\tilde{f}(x) - \tilde{f}(y)\| + \|\tilde{f}(y) - f(y)\| \\ &\le \delta/2 + \gamma_1\lambda \,\rho(x, y) + \delta/2 \le 2\gamma_1\lambda \,\rho(x, y). \end{aligned}$$

Hence, $\|f\|_{\mathrm{Lip}(\mathcal{M},Y)} \le 2\gamma_1\lambda$, and the proof of the lemma is complete. □

The second auxiliary lemma provides additional properties of sets Γ_ℓ defined in Section 3.1 [see (3.2) and Definition 3.1]. We will need these properties in the proof of Lemma 6.5 below.

LEMMA 6.4. *Let $(\mathcal{M},\rho)$ be a finite pseudometric space satisfying (6.1). Let $\ell \geq 0$ and let $F : \mathcal{M} \to \mathrm{Conv}_m(Y)$. Suppose that for every subset $\mathcal{M}' \subset \mathcal{M}$ with $\#\mathcal{M}' \leq k_{\ell+1}$ the restriction $F|_{\mathcal{M}'}$ of F to $\mathcal{M}'$ has a Lipschitz selection $f_{\mathcal{M}'} : \mathcal{M}' \to Y$ with $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}(\mathcal{M}',Y)} \leq \lambda$.*

Let $x_0 \in \mathcal{M}$, $\xi_0 \in \Gamma_\ell(x_0)$, and let $1 \leq k \leq \ell+1$. Let S be a subset of $\mathcal{M}$ with $\#S = k$ containing x_0.

Then there exists a mapping $f^S : S \to Y$ such that

(a) $f^S(x_0) = \xi_0$.
(b) $f^S(y) \in \Gamma_{\ell+1-k}(y)$ *for all* $y \in S$.
(c) $\|f^S\|_{\mathrm{Lip}(S,Y)} \leq 3^k\lambda$.

Proof. We recall that the sequence of positive integers k_ℓ is defined by the formula (3.1).

We proceed by induction on k. For $k = 1$, we have $S = \{x_0\}$, and we can just set $f^S(x_0) = \xi_0$.

For the induction step, we fix $k \geq 2$ and suppose the lemma holds for $k-1$; we then prove it for k. Thus, let $\xi_0 \in \Gamma_\ell(x_0)$, $x_0 \in S$, $\#S = k \leq \ell+1$.

Set $\hat{S} = S \setminus \{x_0\}$. We pick $\hat{x}_0 \in \hat{S}$ to minimize $\rho(\hat{x}_0, x_0)$, and we pick $\hat{\xi}_0 \in \Gamma_{\ell-1}(\hat{x}_0)$ such that $\|\hat{\xi}_0 - \xi_0\| \leq \lambda\,\rho(\hat{x}_0, x_0)$. (See Lemma 3.4 (b).) For $y \in \hat{S}$ we have $\rho(y, x_0) \geq \rho(\hat{x}_0, x_0)$, hence

$$\rho(y,\hat{x}_0) + \rho(\hat{x}_0, x_0) \leq [\rho(y,x_0) + \rho(x_0,\hat{x}_0)] + \rho(\hat{x}_0,x_0) \leq 3\rho(y,x_0). \tag{6.5}$$

By the induction hypothesis, there exists $\hat{f} : \hat{S} \to Y$ such that

($\hat{a}$) $\hat{f}(\hat{x}_0) = \hat{\xi}_0$.
($\hat{b}$) $\hat{f}(y) \in \Gamma_{(\ell-1)+1-(k-1)}(y) = \Gamma_{\ell+1-k}(y)$ for all $y \in \hat{S}$.
($\hat{c}$) $\|\hat{f}\|_{\mathrm{Lip}(\hat{S},Y)} \leq 3^{k-1}\lambda$.

We now define $f : S \to Y$ by setting

$$f(y) = \hat{f}(y) \quad \text{for} \quad y \in \hat{S}; \qquad f(x_0) = \xi_0.$$

Then f obviously satisfies (a) and (b). To see that f satisfies (c), we first recall ($\hat{c}$); thus it is enough to check that

$$\|f(y) - f(x_0)\| \leq 3^k\lambda\,\rho(y,x_0)$$

for $y \in \hat{S}$, i.e.,

$$\|\hat{f}(y) - \xi_0\| \leq 3^k\lambda\,\rho(y,x_0) \quad \text{for} \quad y \in \hat{S}.$$

However, for $y \in \hat{S}$ we have

$$\begin{aligned}\|\hat{f}(y) - \xi_0\| &\leq \|\hat{f}(y) - \hat{\xi}_0\| + \|\hat{\xi}_0 - \xi_0\| = \|\hat{f}(y) - \hat{f}(\hat{x}_0)\| + \|\hat{\xi}_0 - \xi_0\| \\ &\leq 3^{k-1}\lambda\,\rho(y,\hat{x}_0) + \lambda\,\rho(\hat{x}_0,x_0),\end{aligned}$$

thanks to ($\hat{c}$) and the definition of $\hat{\xi}_0$.

Therefore,

$$\|\hat{f}(y)-\xi_0\| \le 3^{k-1}\lambda\,[\rho(y,\hat{x}_0)+\rho(\hat{x}_0,x_0)] \le 3^k\lambda\,\rho(y,x_0),$$

by (6.5).

Thus, f satisfies (a), (b), (c), completing our induction. □

We turn to the last auxiliary lemma. Let

$$\tilde{\ell}=k^\sharp \quad \text{and let} \quad k^*=k_{\tilde{\ell}+1} \tag{6.6}$$

where $k_\ell=(m+2)^\ell$, see (3.1).

LEMMA 6.5. *Let $(\mathcal{M},\rho)$ be a finite pseudometric space satisfying (6.1), and let $x_0\in\mathcal{M}$ and $\lambda>0$.*

Let $F:\mathcal{M}\to \operatorname{Conv}_m(Y)$ be a set-valued mapping such that for every subset $\mathcal{M}'\subset\mathcal{M}$ consisting of at most k^* points, the restriction $F|_{\mathcal{M}'}$ of F to $\mathcal{M}'$ has a Lipschitz selection $f_{\mathcal{M}'}:\mathcal{M}'\to Y$ with $\|f_{\mathcal{M}'}\|_{\operatorname{Lip}(\mathcal{M}',Y)}\le\lambda$.

Then there exists a point $\xi_0\in F(x_0)$ such that the following statement holds: For every subset $S\subset\mathcal{M}$ with $\#S\le k^\sharp$, there exists a mapping $f_S:S\to Y$ with $\|f_S\|_{\operatorname{Lip}(S,Y)}\le C\lambda$ such that

$$\|f_S(x)-\xi_0\|\le C\lambda\,\rho(x,x_0) \quad \textit{for every} \quad x\in S, \tag{6.7}$$

and

$$f_S(x)\in F(y)+\lambda\,\rho(x,y)\,B_Y \quad \textit{for every} \quad x\in S, y\in\mathcal{M}. \tag{6.8}$$

Here C is a constant depending only on m.

Proof. By the lemma's hypothesis, (6.6) and by Lemma 3.4 *(a)*,

$$\Gamma_{\tilde{\ell}}(x)\ne\emptyset \quad \text{for every} \quad x\in\mathcal{M}.$$

Let $\xi_0\in\Gamma_{\tilde{\ell}}(x_0)$. By (3.6),

$$\xi_0\in\Gamma_{\tilde{\ell}}(x_0)\subset F(x_0).$$

Let $S\subset\mathcal{M}$, $\#S\le k^\sharp$. Let $\widetilde{S}=S\cup\{x_0\}$ and let $k=\#\widetilde{S}=\#(S\cup\{x_0\})$. Then

$$1\le k\le \#S+1\le k^\sharp+1=\tilde{\ell}+1.$$

Therefore, by Lemma 6.4, there exists a mapping $f^{\widetilde{S}}:\widetilde{S}\to Y$ with $\|f^{\widetilde{S}}\|_{\operatorname{Lip}(\widetilde{S},Y)}\le 3^k\lambda$ such that $f^{\widetilde{S}}(x_0)=\xi_0$ and

$$f^{\widetilde{S}}(x)\in\Gamma_{\tilde{\ell}+1-k}(x) \quad \text{for all} \quad x\in\widetilde{S}.$$

Recall that $k\le\tilde{\ell}+1=k^\sharp+1$ so that

$$\|f^{\widetilde{S}}\|_{\operatorname{Lip}(\widetilde{S},Y)}\le C\lambda$$

with $C = 3^{k^\sharp+1}$. Since $k^\sharp$ depends only on m, the constant C depends only on m as well.

Hence, by (3.7),

$$f^{\widetilde{S}}(x) \in \Gamma_{\tilde{\ell}+1-k}(x) \subset \Gamma_0(x) \quad \text{for every} \quad x \in \widetilde{S}. \tag{6.9}$$

Let

$$f_S = f^{\widetilde{S}}|_S.$$

Then $\|f_S\|_{\mathrm{Lip}(S,Y)} \le \|f^{\widetilde{S}}\|_{\mathrm{Lip}(\widetilde{S},Y)} \le C\lambda$. Moreover, by (6.9),

$$f_S(x) \in \Gamma_0(x) \quad \text{for all} \quad x \in S. \tag{6.10}$$

Since $\|f^{\widetilde{S}}\|_{\mathrm{Lip}(\widetilde{S},Y)} \le C\lambda$ and $x_0 \in \widetilde{S}$,

$$\|f_S(x) - \xi_0\| = \|f^{\widetilde{S}}(x) - f^{\widetilde{S}}(x_0)\| \le C\lambda\,\rho(x,x_0) \quad \text{for every} \quad x \in S.$$

Furthermore, by (3.3) and (6.10), for every $x \in S$

$$f_S(x) \in \Gamma_0(x) = \bigcap_{y\in\mathcal{M}} (F(y) + \lambda\,\rho(x,y)\,B_Y)$$

so that

$$f_S(x) \in F(y) + \lambda\,\rho(x,y)\,B_Y \quad \text{for every} \quad x \in S, y \in \mathcal{M}.$$

The proof of the lemma is complete. □

Proposition 6.6. *Let $(\mathcal{M},\rho)$ be a finite pseudometric space satisfying (6.1), and let $\lambda > 0$.*

Let $F : \mathcal{M} \to \mathrm{Conv}_m(Y)$ be a set-valued mapping such that for every subset $\mathcal{M}' \subset \mathcal{M}$ with $\#\mathcal{M}' \le k^*$, the restriction $F|_{\mathcal{M}'}$ of F to $\mathcal{M}'$ has a Lipschitz selection $f_{\mathcal{M}'} : \mathcal{M}' \to Y$ with $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}(\mathcal{M}',Y)} \le \lambda$.

Then F has a Lipschitz selection $f : \mathcal{M} \to Y$ with $\|f\|_{\mathrm{Lip}(\mathcal{M},Y)} \le \gamma_2\lambda$ where γ_2 is a constant depending only on m.

Proof. Let $x_0 \in \mathcal{M}$. By Lemma 6.5, there exists a point $\xi_0 \in F(x_0)$ such that for every set $S \subset \mathcal{M}$ with $\#S \le k^\sharp$ there exists a mapping $f_S : S \to Y$ with $\|f_S\|_{\mathrm{Lip}(S,Y)} \le C\lambda$ such that (6.7) and (6.8) hold. Here C is a constant depending only on m.

We introduce a new set-valued mapping $\widetilde{F} : \mathcal{M} \to \mathrm{Conv}_m(Y)$ by letting

$$\widetilde{F}(x) = \left(\bigcap_{y\in\mathcal{M}} [F(y) + \lambda\,\rho(x,y)\,B_Y]\right) \bigcap B_Y(\xi_0, C\lambda\,\rho(x,x_0)), \quad x \in \mathcal{M}. \tag{6.11}$$

By Lemma 6.5 and definition (6.11), for every set $S \subset \mathcal{M}$ consisting of at most $k^\sharp$ points the restriction $\widetilde{F}|_S$ of $\widetilde{F}$ to S has a Lipschitz selection $f_S : S \to Y$ with $\|f_S\|_{\mathrm{Lip}(S,Y)} \le C\lambda$. In particular, $\widetilde{F}(x) \ne \emptyset$ for every $x \in \mathcal{M}$.

Let us introduce a binary relation "$\sim$" on $\mathcal{M}$ by letting

$$x \sim y \quad \Longleftrightarrow \quad \rho(x,y) = 0.$$

Clearly, "$\sim$" satisfies the axioms of an equivalence relation, i.e., it is reflexive, symmetric and transitive. Given $x \in \mathcal{M}$, by $[x] = \{y \in \mathcal{M} : \ y \sim x\}$ we denote the equivalence class of x. Let

$$[\mathcal{M}] = \mathcal{M}/\sim \; = \{\,[x] : x \in \mathcal{M}\,\}$$

be the corresponding quotient set of $\mathcal{M}$ by "$\sim$", i.e., the family of all equivalence classes of $\mathcal{M}$ by "$\sim$". Finally, given an equivalence class $U \in [\mathcal{M}]$ let us choose a point $w_U \in U$ and put

$$W = \{w_U : U \in [\mathcal{M}]\}.$$

Clearly, (W, ρ) is a *finite metric space*. Let

$$\hat{F} = \widetilde{F}|_W. \tag{6.12}$$

Then, by (6.11) and (6.12), $\hat{F}$ is a set-valued mapping defined on a finite metric space which takes values in the family of all nonempty convex *bounded* subsets of Y of dimension at most m. Furthermore, this mapping satisfies the hypothesis of Lemma 6.3 with $C\lambda$ in place of λ.

Therefore, by this lemma, there exists a Lipschitz selection $\hat{f} : W \to Y$ of $\hat{F}$ on W with

$$\|\hat{f}\|_{\mathrm{Lip}(W,Y)} \le 2\gamma_1\, C\lambda = \gamma_2\lambda.$$

Here $\gamma_2 = 2\gamma_1 C$ is a constant depending only on m (because γ_1 and C depend on m only).

We define a mapping $f : \mathcal{M} \to Y$ by letting

$$f(x) = \hat{f}(w_{[x]}), \qquad x \in \mathcal{M}.$$

Then f is a *selection of* F on $\mathcal{M}$. Indeed, let $x \in \mathcal{M}$. Since $\hat{f}$ is a selection of $\hat{F} = \widetilde{F}|_W$, and $w_{[x]} \in W$,

$$f(x) = \hat{f}(w_{[x]}) \in \widetilde{F}(w_{[x]})$$

so that, by (6.11),

$$f(x) \in \widetilde{F}(w_{[x]}) \subset F(x) + \lambda\, \rho(w_{[x]}, x)\, B_Y.$$

But $w_{[x]} \sim x$ so that $\rho(w_{[x]}, x) = 0$, proving that $f(x) \in F(x)$.

Let us prove that $\|f\|_{\mathrm{Lip}(\mathcal{M},Y)} \le \gamma_2\lambda$, i.e.,

$$\|f(x) - f(y)\| \le \gamma_2\lambda\,\rho(x,y) \quad \text{for all} \quad x,y \in \mathcal{M}. \tag{6.13}$$

In fact, since $\|\hat{f}\|_{\mathrm{Lip}(W,Y)} \le \gamma_2\lambda$,

$$\begin{aligned}\|f(x) - f(y)\| &= \|\hat{f}(w_{[x]}) - \hat{f}(w_{[y]})\| \le \gamma_2\lambda\,\rho(w_{[x]}, w_{[y]})\\ &\le \gamma_2\lambda\,(\rho(w_{[x]},x) + \rho(x,y) + \rho(y,w_{[y]})) = \gamma_2\lambda\,\rho(x,y),\end{aligned}$$

proving (6.13).

The proof of Proposition 6.6 is complete. □

Proof of Theorem 6.2. We prove this theorem following the scheme of the proof of Theorem 1.2. In particular, to study a pseudometric ρ that takes only finite values, we use Proposition 6.6 and the constant k^* rather than Proposition 6.1 and $k^\sharp$ respectively.

We note that [Shv02, Remark 1.3] implies a variant of Theorem 1.8 for the case of a finite pseudometric space $(\widetilde{\mathcal{M}}, \tilde{\rho})$ and a set-valued mapping $\widetilde{F}$ with convex (not necessarily compact) images $\widetilde{F}(x)$, $x \in \widetilde{\mathcal{M}}$, of dimension at most m.

As in the proof of Theorem 1.2, the passage from the case of finite pseudometrics $\rho : \mathcal{M} \times \mathcal{M} \to \mathbb{R}_+$ to the general case of an arbitrary pseudometric $\rho : \mathcal{M} \times \mathcal{M} \to \mathbb{R}_+ \cup \{+\infty\}$ is an easy exercise. □

7 Further Results and Comments

• **Generalization of the finiteness principle: set-valued mappings with closed images.**

In Theorem 1.2 we prove the finiteness principle for set-valued mappings F whose values are convex *compact* sets with affine dimension bounded by m. The following result shows that this family of sets can be slightly extended.

Theorem 7.1. *Theorem 1.2 holds provided the requirement $F : \mathcal{M} \to \mathcal{K}_m(Y)$ in its formulation is replaced with the following one: for every $x \in \mathcal{M}$ the set $F(x)$ is a closed convex subset of Y of dimension at most m, and there exists $x_0 \in \mathcal{M}$ such that $F(x_0)$ is bounded.*

For the proof of this statement we refer the reader to [FS, p. 74].

Theorem 7.1 implies the following result.

Theorem 7.2. *Theorem 1.2 holds provided the requirement $F : \mathcal{M} \to \mathcal{K}_m(Y)$ in its formulation is replaced with $F : \mathcal{M} \to \mathcal{K}_m(Y) \cup \mathit{Aff}_m(Y)$.*

Here $\mathit{Aff}_m(Y)$ denotes the family of all affine subspaces of Y of dimension at most m.

Proof. The result follows from [Shv04] whenever $F : \mathcal{M} \to \mathit{Aff}_m(Y)$, and from Theorem 7.1 whenever there exists $x_0 \in \mathcal{M}$ such that $F(x_0) \in \mathcal{K}_m(Y)$. □

• **Steiner-type points as a special case of the finiteness principle for Lipschitz selections.**

Let Y be a Banach space. Given $m \in \mathbb{N}$ let $\mathcal{M} = \mathcal{K}_m(Y)$ be the family of all nonempty convex compact subsets of Y of affine dimension at most m equipped with the Hausdorff distance $\rho = \mathrm{d_H}$.

Let $F : \mathcal{M} \to \mathcal{K}_m(Y)$ be the identity mapping on $\mathcal{K}_m(Y)$, i.e.,

$$F(K) = K \quad \text{for every} \quad K \in \mathcal{K}_m(Y).$$

By Theorem 5.11, this mapping has a selection $S_Y : \mathcal{M} \to Y$ whose $\mathrm{d_H}$-Lipschitz seminorm is bounded by a constant $\gamma = \gamma(m)$ depending only on m.

The following claim asserts that the mapping F satisfies the hypothesis of Theorem 1.2.

Claim 7.3. For every subset $\mathcal{M}' \subset \mathcal{M}$ with $\#\mathcal{M}' \le N(m, Y)$ the restriction $F|_{\mathcal{M}'}$ has a $\mathrm{d_H}$-Lipschitz selection $f_{\mathcal{M}'} : \mathcal{M}' \to Y$ with $\|f_{\mathcal{M}'}\|_{\mathrm{Lip}((\mathcal{M}',\mathrm{d_H}),Y)} \le \theta$ where $\theta = \theta(m)$ is a constant depending only on m.

For a simple proof of this claim we refer the reader to [FS, p. 73].

Claim 7.3 shows that Theorem 5.11 can be considered as a particular case of our main result, Theorem 1.2, which is applied to the metric space $(\mathcal{K}_m(Y), \mathrm{d_H})$. (Note that the proof of Theorem 1.2 uses Theorem 5.11.) In general, this metric space has the same complexity as an L_∞-space, and may have infinite Nagata dimension. For example, if $Y = \ell_\infty$ then $\mathcal{K}_m(Y)$ contains the set of one point subsets of ℓ_∞. As we have noted in the Introduction, ℓ_∞ has infinite Nagata dimension so that $\mathcal{K}_m(Y)$ has infinite Nagata dimension as well.

In this case we are unable to prove Theorem 5.11 using the ideas and methods developed in Sections 2–4.

Thus, analyzing the scheme of the proof of Theorem 1.2, we observe that this proof is actually based on solutions of the Lipschitz selection problem for two independent particular cases of this problem, namely, for metric trees, see Corollary 4.16, and for the metric space $(\mathcal{K}_m(Y), \mathrm{d_H})$, see Theorem 5.11. Theorem 5.2 proven in Section 5 provides a certain "bridge" between these two independent results (i.e., Corollary 4.16 and Theorem 5.11). Combining all these results, we finally obtain a proof of Theorem 1.2 in the general case.

Acknowledgements

We are grateful to Alexander Brudnyi, Arie Israel, Bo'az Klartag, Garving (Kevin) Luli and the participants of the 10th Whitney Problems Conference, Williamsburg, VA, for valuable conversations. We thank the referee for very careful reading and numerous suggestions, which led to improvements in our exposition. We are grateful also to the College of William and Mary, Williamsburg, VA, USA, the American Institute of Mathematics, San Jose, CA, USA, the Fields Institute, Toronto, Canada, the University of Arkansas, AR, USA, the Banff International Research Station, Banff, Canada, the Centre International de Rencontres Mathématiques (CIRM), Luminy, Marseille, France, and the Technion, Haifa, Israel, for hosting and supporting workshops on the topic of this paper and closely related problems. Finally, we thank the US-Israel Binational Science Foundation, the US National Science Foundation, the Office of Naval Research and the Air Force Office of Scientific Research for generous support.

References

[Ass82] P. Assouad. Sur la distance de Nagata. *C. R. Acad. Sci. Paris Sér. I Math.*, (1)294 (1982), 31–34.

[BDHM09] N. Brodskiy, J. Dydak, J. Higes, and A. Mitra. Assouad-Nagata dimension via Lipschitz extensions. *Israel J. Math.*, 171 (2009), 405–423.

[Bru18] A. Brudnyi. A note on the Lipschitz selection. *C. R. Math. Rep. Acad. Sci. Canada*, (1)40 (2018), 29–32.

[BB12] A. Brudnyi, and Yu. Brudnyi. *Methods of Geometric Analysis in extension and trace problems.* Vols. I, II, Monographs in Mathematics, Vol. 102. Springer, Basel (2012).

[BS94] Y. Brudnyi, and P. Shvartsman. Generalizations of Whitney's extension theorem. *Internat. Math. Res. Notices* (3) (1994), 129–139.

[BS98] Yu. Brudnyi, and P. Shvartsman. The trace of jet space $J^k\Lambda_\omega$ to an arbitrary closed subset of $\mathbb{R}^n$. *Trans. Amer. Math. Soc.*, (4)350 (1998), 1519–1553.

[BS01] Yu. Brudnyi, and P. Shvartsman. Whitney Extension Problem for Multivariate $C^{1,\omega}$-functions. *Trans. Amer. Math. Soc.*, (6)353 (2001), 2487–2512.

[CK95] P. B. Callahan, and S. R. Kosaraju. A decomposition of multidimensional point sets with applications to k-nearest-neighbors and n-body potential fields. *J. Assoc. for Computing Machinery*, 42 (1995), 67–90.

[DGK63] L. Danzer, B. Grünbaum, and V. Klee. Helly's Theorem and its relatives. In: *AMS Symposium on Convexity, Seattle, Proceedings of Symposium on Pure Mathematics*, Vol. 7. Amer. Math. Soc., Providence, RI (1963), pp. 101–180.

[Fef05a] C. Fefferman. A sharp form of Whitney extension theorem. *Annals of Math.*, (1)161 (2005), 509–577.

[Fef05b] C. Fefferman. A Generalized Sharp Whitney Theorem for Jets. *Rev. Mat. Iberoamericana*, (2)21 (2005) 577–688.

[Fef06] C. Fefferman. Whitney extension problem for C^m. *Annals of Math.*, (1)164 (2006), 313–359.

[FK09a] C. Fefferman, and B. Klartag. Fitting a C^m-smooth function to data I. *Annals of Math.*, (1)169 (2009), 315–346.

[FK09b] C. Fefferman, and B. Klartag. Fitting a C^m-smooth function to data II. *Revista Mat. Iberoamericana*, (1)25 (2009), 49–273.

[Fef09a] C. Fefferman. Fitting a C^m-smooth function to data III. *Annals of Math.*, (1)170 (2009), 427–441.

[Fef09b] C. Fefferman. Whitney extension problems and interpolation of data. *Bulletin A.M.S.*, (2)46 (2009), 207–220.

[FIL16a] C. Fefferman, A. Israel, and G. K. Luli. Finiteness principles for smooth selection. *Geom. Funct. Anal.*, (2)26 (2016), 422–477.

[FIL16b] C. Fefferman, A. Israel, and G. K. Luli. Finiteness principles for C^m and Lipschitz selection. 82 pp., Whitney Extension Problems: C^m and Sobolev functions on subsets of $\mathbb{R}^n$. The 9th Whitney Problems Workshop, May–June 2016, Haifa, Israel. http://www.math.technion.ac.il/Site/events/EvntOffices/event.php?eid=eo108. Accessed June 2016.

[FIL17] C. Fefferman, A. Israel, and G. K. Luli. Interpolation of data by smooth non-negative functions. *Rev. Mat. Iberoam.*, (1)33 (2017), 305–324.

[FS] C. Fefferman, and P. Shvartsman. Sharp finiteness principles for Lipschitz selections: long version. arXiv:1708.00811v2.

[HM06] S. Har-Peled, and M. Mendel. Fast Construction of Nets in Low-Dimensional Metrics and Their Applications. *SIAM J. Comput.*, (5)35 (2006), 1148–1184.

[LS05] U. Lang, and T. Schlichenmaier. Nagata dimension, quasisymmetric embeddings, and Lipschitz extensions. *Int. Math. Res. Not.*, (58) (2005), 3625–3655.

[Nag58] J. Nagata. Note on dimension theory for metric spaces. *Fund. Math.*, 45 (1958), 143–181.

[OR15] M. Ostrovskii, D. Rosenthal. Metric dimensions of minor excluded graphs and minor exclusion in groups. *Internat. J. Algebra Comput.*, (4)25 (2015), 541–554.

[Shv82] P. Shvartsman. The traces of functions of two variables satisfying the Zygmund condition. (Russian) *Studies in the Theory of Functions of Several Real Variables,* 145–168, Yaroslav. Gos. Univ., Yaroslavl' (1982).

[Shv84] P. Shvartsman. Lipschitz sections of set-valued mappings and traces of functions from the Zygmund class on an arbitrary compactum (Russian). *Dokl. Akad. Nauk SSSR*, (3)276 (1984), 559–562; English transl. in *Soviet Math. Dokl.* 29 (1984), no. 3, 565–568.

[Shv86] P. Shvartsman. Lipschitz sections of multivalued mappings. (Russian) *Studies in the theory of functions of several real variables,* 121–132, 149, Yaroslav. Gos. Univ., Yaroslavl' (1986).

[Shv87] P. Shvartsman. On the traces of functions of the Zygmund class. *Sib. Mat. Zh.*, (5)28, (1987), 203–215; English transl. in *Sib. Math. J.* 28 (1987), 853–863.

[Shv92] P. Shvartsman. K-functionals of weighted Lipschitz spaces and Lipschitz selections of multivalued mappings. *Interpolation spaces and related topics*, (Haifa, 1990), 245–268, Israel Math. Conf. Proc., 5, Bar-Ilan Univ., Ramat Gan (1992).

[Shv01] P. Shvartsman. On Lipschitz selections of affine-set valued mappings. *Geom. Funct. Anal.*, (4)11, (2001), 840–868.

[Shv02] P. Shvartsman. Lipschitz selections of set-valued mappings and Helly's theorem. *J. Geom. Anal.*, (2)12 (2002), 289–324.

[Shv04] P. Shvartsman. Barycentric selectors and a Steiner-type point of a convex body in a Banach space. *J. Funct. Anal.*, (1)210 (2004), 1–42.

[Shv08] P. Shvartsman. The Whitney extension problem and Lipschitz selections of set-valued mappings in jet-spaces. *Trans. Amer. Math. Soc.*, (10)360 (2008), 5529–5550.

[Whi34] H. Whitney. Analytic extension of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.*, 36 (1934), 63–89.

Charles Fefferman, Department of Mathematics, Princeton University, Fine Hall Washington Road, Princeton, NJ 08544, USA cf@math.princeton.edu

Pavel Shvartsman, Department of Mathematics, Technion - Israel Institute of Technology, 32000 Haifa, Israel pshv@technion.ac.il

Received: February 11, 2018
Accepted: June 25, 2018