Charles University in Prague
Faculty of Mathematics and Physics

DOCTORAL THESIS

Michal Kraus

Some results in convexity and in Banach space theory

Department of Mathematical Analysis

Supervisor: prof. RNDr. Jaroslav Lukeš, DrSc.
Study programme: Mathematics
Specialization: Mathematical Analysis

Prague 2012

I would like to thank my supervisor, Jaroslav Lukeš, for his guidance and support during my studies, for many inspiring discussions, and for his help with preparing my first papers.

I am also grateful to Petr Hájek, with whom I worked on what now forms Chapter 3 of this thesis, for the time he dedicated to me, and for his help with looking for postdoc position. It was also he who suggested that I spend a year at Université de Franche-Comté in Besançon.

This brings me to Gilles Lancien, under whose guidance I worked in Besançon on what now forms Chapter 4 of this thesis. I would like to thank him for sharing his knowledge with me, and for many interesting conversations, which were always a great pleasure for me.

I thank the Department of Mathematical Analysis at Charles University in Prague, and also the Laboratoire de Mathématiques at Université de Franche-Comté in Besançon, for friendly and encouraging environment.

Last but not least, I would like to thank Eva, my family and my friends for their support and patience during my preoccupation with this thesis.

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In Besançon, 18th June 2012 Michal Kraus

Název práce: Několik výsledků v konvexitě a v teorii Banachových prostorů
Autor: Michal Kraus
Katedra: Katedra matematické analýzy
Vedoucí disertační práce: prof. RNDr. Jaroslav Lukeš, DrSc.
Abstrakt: Tato práce se skládá ze čtyř odborných článků. V prvním článku zkonstruujeme nemetrizovatelné kompaktní množiny s patologickými množinami simpliciality, čímž ukážeme, že vlastnosti množiny simpliciality, známé v metrizovatelném případě, neplatí bez předpokladu metrizovatelnosti. Ve druhém článku zkonstruujeme příklad týkající se remotal množin, čímž zodpovíme otázku Martína a Raa, a podáme nový důkaz tvrzení, že v každém nekonečně dimenzionálním Banachově prostoru existuje uzavřená konvexní omezená množina, která není remotal. Třetí článek je studie souvislostí mezi polynomy na Banachových prostorech a lineárními identitami. Zkoumáme za jakých podmínek je lineární identita splněná pouze polynomy, a popíšeme prostor polynomů splňujících takovou lineární identitu. V posledním článku studujeme existenci coarse a uniformních vnoření mezi Orliczovými prostory posloupností. Ukážeme, že existence vnoření mezi dvěma Orliczovými prostory posloupností je ve většině případů určena pouze hodnotami jejich horních Matuszewska-Orliczových indexů.
Klíčová slova: množina simpliciality, remotal množina, polynomy na Banachových prostorech, coarse vnoření, uniformní vnoření

Title: Some results in convexity and in Banach space theory
Author: Michal Kraus
Department: Department of Mathematical Analysis
Supervisor: prof. RNDr. Jaroslav Lukeš, DrSc.
Abstract: This thesis consists of four research papers. In the first paper we construct nonmetrizable compact convex sets with pathological sets of simpliciality, showing that the properties of the set of simpliciality known in the metrizable case do not hold without the assumption of metrizability. In the second paper we construct an example concerning remotal sets, answering thus a question of Martín and Rao, and present a new proof of the fact that in every infinite-dimensional Banach space there exists a closed convex bounded set which is not remotal. The third paper is a study of the relations between polynomials on Banach spaces and linear identities. We investigate under which conditions a linear identity is satisfied only by polynomials, and describe the space of polynomials satisfying such linear identity. In the last paper we study the coarse and uniform embeddability between Orlicz sequence spaces. We show that the embeddability between two Orlicz sequence spaces is in most cases determined only by the values of their upper Matuszewska-Orlicz indices.
Keywords: set of simpliciality, remotal set, polynomials on Banach spaces, coarse embedding, uniform embedding

Contents

Introduction

This thesis consists of the following four papers:

- *Point simpliciality in Choquet theory on nonmetrizable compact spaces*, Bull. Sci. Math. 135 (2011), 312–323.
- *Two remarks on remotality*, J. Approx. Theory 163 (2011), 307–310.
- (with P. Hájek) *Polynomials and identities on real Banach spaces*, J. Math. Anal. Appl. 385 (2012), 1015–1026.
- *Coarse and uniform embeddings between Orlicz sequence spaces*, submitted.

Each paper constitutes one chapter. Except for the third one, the papers are presented in their original form. The third paper, as presented here, differs very slightly from the published version. Let us now briefly introduce the topics treated in this thesis. Let us mention that all vector spaces considered here are over the real field.

In Chapter 1 we are concerned with Choquet theory of function spaces. We study the recent notion of the set of simpliciality, introduced by M. Bačák in [Bač09]. Suppose that $\mathcal{H}$ is a function space on a compact space K. The *set of simpliciality* of $\mathcal{H}$ is the set of all $x \in K$ for which there exists a unique maximal measure representing x. So we may say that the set of simpliciality of $\mathcal{H}$ is the set of all points of K at which the function space $\mathcal{H}$ is "locally simplicial".

Bačák in [Bač09] studied the set of simpliciality from various points of view. He was mainly interested in the case when K is metrizable. In that case, the set of simpliciality has some nice properties. Here we study the properties of the set of simpliciality for K nonmetrizable. We give examples showing that if K is nonmetrizable, then the set of simpliciality may behave quite pathologically, so Bačák's results are no longer true in this setting.

In Chapter 2 we present some results concerning the notion of remotality. Let X be a Banach space and $E \subset X$ be a bounded set. If $x \in X$, we define $D(x, E) := \sup\{\|x - z\| : z \in E\}$. We say that the set E is *remotal from a point* $x \in X$ if there exists a point $e \in E$ such that $\|x - e\| = D(x, E)$. In other words, E contains a farthest point from x. The set E is said to be *remotal* if it is remotal from all $x \in X$.

In recent years, remotal sets have received growing attention. Our work is a reaction to the paper [MaRa10] by Martín and Rao. They studied the following problem: characterize those Banach spaces in which every closed convex bounded set is remotal. Clearly in finite-dimensional spaces every closed bounded set is remotal. Martín and Rao proved that in every infinite-dimensional Banach space there exists a closed convex bounded set which is not remotal. In connection with the method of their proof, they asked whether the remotality of $\overline{\mathrm{co}}\,(E)$ from a point $x \in X$, where E is a weakly closed and bounded subset of a Banach space X, implies the remotality of E from x. We answer this question in the negative by finding a counterexample in c_0.

The second purpose of this chapter is to present an alternative proof of the fact that in every infinite-dimensional Banach space there exists a closed convex bounded set which is not remotal.

In Chapter 3 we study polynomials on Banach spaces. This is a joint work with Petr Hájek. We are interested in the relations between polynomials and linear identities. A classical example of theorems we are dealing with is a result due to Fréchet, Mazur and Orlicz, stating that if X, Y are Banach spaces, $f : X \to Y$ is a continuous mapping and $n \in \mathbb{N} \cup \{0\}$, then f is a polynomial of degree at most n if and only if

$$\sum_{k=0}^{n+1} (-1)^{n+1-k} \binom{n+1}{k} f(x + kh) = 0 \text{ for all } x, h \in X.$$

Similar identities were treated by other authors in their study of Banach spaces with polynomial norms. In our work we develop an abstract approach to linear identities, generalizing and unifying the aforementioned results. We study under which conditions a linear identity is satisfied only by polynomials, and describe the space of polynomials satisfying such linear identity. We also present a method for creating linear identities with prescribed properties based on the Lagrange interpolation theory.

As mentioned above, this chapter slightly differs from the published paper. The difference is in Theorems 3.3.8 and 3.4.1, which are slightly more general than the corresponding theorems in the published paper, and in comments between Corolary 3.3.5 and Theorem 3.3.7.

Chapter 4 presents some results in the nonlinear geometry of Banach spaces. These results were obtained during my stay at Université de Franche-Comté in Besançon, under the direction of Gilles Lancien.

Let $(M, d_M), (N, d_N)$ be metric spaces and suppose that $f : M \to N$ is a mapping. Then f is called a *coarse embedding* if there exist nondecreasing functions $\rho_1, \rho_2 : [0, \infty) \to [0, \infty)$ such that $\lim_{t\to\infty} \rho_1(t) = \infty$ and

$$\rho_1(d_M(x, y)) \le d_N(f(x), f(y)) \le \rho_2(d_M(x, y)) \quad \text{for all } x, y \in M.$$

The mapping f is called a *uniform embedding* if f is injective and both f and $f^{-1} : f(M) \to M$ are uniformly continuous.

In the nonlinear geometry of Banach spaces, a considerable interest has been in the following general questions: when does a Banach space coarsely (uniformly) embed into another Banach space? Not much is known in general, but there are some results for special classes of Banach spaces. Due to the work of many mathematicians, the coarse and uniform embeddability between ℓ_p-spaces is now completely characterized. Our aim is to generalize this classification to a wider class of Banach spaces, namely to Orlicz sequence spaces. We give an almost complete classification of the coarse and uniform embeddability between these spaces. We show that the embeddability between two Orlicz sequence spaces is in most cases determined only by the values of their upper Matuszewska-Orlicz indices. On the other hand, we present examples showing that in some cases the embeddability is not determined by the values of the upper Matuszewska-Orlicz indices.

Chapter 1

Point simpliciality in Choquet theory on nonmetrizable compact spaces

1.1 Introduction

For notation and terminology we refer the reader to the next section. Let $\mathcal{H}$ be a function space on a compact space K. This paper is concerned with those probability measures μ on K for which there exists a unique maximal (with respect to the Choquet ordering $\preceq$) measure ν such that $\mu \preceq \nu$. A characterization of such measures was given by J. Köhn in [Köh70, Proposition 1] in the convex case, i.e. in the context of compact convex subsets of locally convex spaces, and extended to the general case of function spaces by M. Bačák in [Bač09, Theorem 5.1] (his proof is done only for metrizable compact spaces, but works with no change also without the assumption of metrizability). Let us present it here:

Theorem 1.1.1 (Köhn, Bačák)**.** *Let $\mathcal{H}$ be a function space on a compact space K and let $\mu \in \mathcal{M}^1(K)$. Then the following statements are equivalent:*

(i) *There exists a unique maximal measure $\nu \in \mathcal{M}^1(K)$ such that $\mu \preceq \nu$.*

(ii) *For every $f, g \in \mathcal{K}^c(\mathcal{H})$ we have $\mu((f+g)^*) = \mu(f^*) + \mu(g^*)$.*

(iii) *For every maximal $\nu \in \mathcal{M}^1(K)$, $\mu \preceq \nu$, and every $f \in \mathcal{K}^c(\mathcal{H})$, we have $\nu(f) = \mu(f^*)$.*

We denote by $\mathcal{M}^1_{PS}(\mathcal{H})$ the set of all measures from $\mathcal{M}^1(K)$ which satisfy some of the equivalent conditions of Theorem 1.1.1. If we take the Dirac measure ε_x for $x \in K$ and apply Theorem 1.1.1, we get a characterization of those points $x \in K$ for which there exists a unique maximal measure representing x (in the convex case, this result, with some other equivalent conditions, was proved also by S. Simons in [Sim70, Theorem 37]). In [Bač09], the set of all these points of K is called the *set of simpliciality* of $\mathcal{H}$ and denoted by $\mathrm{Sim}_{\mathcal{H}}(K)$. It turned out that if the space K is metrizable, the set $\mathrm{Sim}_{\mathcal{H}}(K)$ is closely related to measures from $\mathcal{M}^1_{PS}(\mathcal{H})$, and has some other nice properties. More precisely, we have the following (for proofs, see [Bač09, Theorems 4.5, 5.1 and 5.6]):

Theorem 1.1.2 (Bačák)**.** *Let $\mathcal{H}$ be a function space on a metrizable compact space K.*

(a) *The set* $\mathrm{Sim}_{\mathcal{H}}(K)$ *is Borel (in fact, a G_δ set).*

(b) *Let $\mu \in \mathcal{M}^1(K)$. Then $\mu \in \mathcal{M}^1_{PS}(\mathcal{H})$ if and only if $\mu(\mathrm{Sim}_{\mathcal{H}}(K)) = 1$.*

(c) *The set* $\mathrm{Sim}_{\mathcal{H}}(K)$ *is $\mathcal{H}$-extremal.*

The purpose of this paper is to study the validity of the statements of Theorem 1.1.2 without the assumption of metrizability. We will show that without this assumption, the statement (a) is false, and the statements (b) and (c) are false, even if the set $\mathrm{Sim}_{\mathcal{H}}(K)$ is Borel. The counterexamples in the general context of function spaces are presented in Section 1.4. In Section 1.5, we will show that the counterexamples may be constructed even in the convex case. Of course, it would be sufficient to present the examples only in the convex case, but we have decided to include the constructions also in the general context of function spaces, since these are much simpler and may be of some interest in themselves.

We will also show something more. Suppose that $\mu \in \mathcal{M}^1_{PS}(\mathcal{H})$. We know from Theorem 1.1.2(b), that this is equivalent to the fact that μ is carried by $\mathrm{Sim}_{\mathcal{H}}(K)$ if K is metrizable. If K is nonmetrizable, then, by a simple application of Theorem 1.1.1, we have at least that the atomic part of μ is carried by some (countable and therefore Borel) subset of $\mathrm{Sim}_{\mathcal{H}}(K)$. This is similar to the relation between maximal measures and the Choquet boundary $\mathrm{Ch}_{\mathcal{H}}(K)$ (if K is metrizable, then maximal probability measures are precisely those measures from $\mathcal{M}^1(K)$ which are carried by $\mathrm{Ch}_{\mathcal{H}}(K)$, see [LMNS10, Corollary 3.62], and in general, the atomic parts of maximal measures are carried by some subset of $\mathrm{Ch}_{\mathcal{H}}(K)$, see [LMNS10, Proposition 3.66]). Since maximal measures are always carried by $\overline{\mathrm{Ch}_{\mathcal{H}}(K)}$ (see [LMNS10, Proposition 3.64]), one may conjecture that μ is carried by $\overline{\mathrm{Sim}_{\mathcal{H}}(K)}$. Example 1.4.5 shows that even this statement is false. We do not know whether such an example may be found in the convex case.

In the construction of the examples we use the idea of the "porcupine" topology due to E. Bishop and K. de Leeuw, see [BiLe59, p. 327].

1.2 Preliminaries

Let us briefly summarize notation, terminology and basic facts used in this paper. For details and further information about Choquet theory see [LMNS10] or a classical book [Phe01]. All topological spaces throughout the paper are supposed to be Hausdorff. Let K be a compact space. We denote by $\mathcal{C}(K)$ the space of all real continuous functions on K equipped with the supremum norm. The symbol $\mathcal{M}^+(K)$ denotes the set of all nonnegative Radon measures (that is nonnegative regular Borel measures) on K. The symbol $\mathcal{M}(K)$ stands for the space of all signed Radon measures on K, while $\mathcal{M}^1(K)$ denotes the set of all probability Radon measures on K. All these sets of measures are equipped with the w^*-topology.

A linear subspace $\mathcal{H}$ of the space $\mathcal{C}(K)$ is called a *function space* if it contains all constant functions and separates points of K. If X is a compact convex subset of some locally convex space, then the space $A(X)$ of all continuous affine functions on X is a function space. If not stated otherwise, on a compact convex set X we will always consider the function space $A(X)$. We will refer to this setting as to the *convex case*. By a compact convex set we always mean a compact convex subset of a real locally convex space.

Let $\mathcal{H}$ be a function space on K. Let $\mu \in \mathcal{M}^1(K)$. We say that $x \in K$ is the *resultant* of μ (or μ *represents* x) if $f(x) = \mu(f)$ for every $f \in \mathcal{H}$; we denote the resultant of μ (which is unique if it exists) by $r(\mu)$. The set of all $\mu \in \mathcal{M}^1(K)$ which represent $x \in K$ is denoted by $\mathcal{M}_x(\mathcal{H})$. If X is a compact convex set, then every measure from $\mathcal{M}^1(X)$ has a resultant, see [LMNS10, Theorem 2.29]. If $\mu, \nu \in \mathcal{M}^+(K)$ and $\mu(f) = \nu(f)$ for every $f \in \mathcal{H}$, we write $\mu \sim \nu$.

If f is a bounded function on K, its *upper envelope* f^* is defined by $f^* := \inf\{h \in \mathcal{H} : h \geq f\}$. If f is a bounded Borel function on K, then it is said to be *$\mathcal{H}$-convex* if $f(x) \leq \mu(f)$ for all $x \in K$ and $\mu \in \mathcal{M}_x(\mathcal{H})$. The set of all continuous $\mathcal{H}$-convex functions on K is denoted by $\mathcal{K}^c(\mathcal{H})$. The *Choquet ordering* on $\mathcal{M}^+(K)$ is defined as follows: if $\mu, \nu \in \mathcal{M}^+(K)$, then $\mu \preceq \nu$ provided $\mu(f) \leq \nu(f)$ for every $f \in \mathcal{K}^c(\mathcal{H})$. Clearly if $\mu \preceq \nu$ for $\mu, \nu \in \mathcal{M}^+(K)$, then $\mu \sim \nu$. Measures which are maximal in the Choquet ordering are called *maximal measures*. For every measure $\mu \in \mathcal{M}^1(K)$ there exists a maximal measure $\nu \in \mathcal{M}^1(K)$ such that $\mu \preceq \nu$, see [LMNS10, Theorem 3.65]. In particular, for every $x \in K$ there exists a maximal measure $\mu \in \mathcal{M}^1(K)$ representing x.

The *Choquet boundary* of $\mathcal{H}$ (denoted by $\mathrm{Ch}_{\mathcal{H}}(K)$) is the set of all $x \in K$ which have only one representing measure — Dirac measure concentrated at x, which we denote by ε_x. The set $\mathrm{Ch}_{\mathcal{H}}(K)$ is a G_δ set if K is metrizable, and may be non-Borel in general. If K is metrizable, then $\mu \in \mathcal{M}^+(K)$ is maximal if and only if it is carried by $\mathrm{Ch}_{\mathcal{H}}(K)$ (we say that μ is *carried* by a set $A \subset K$ if A is Borel and $\mu(K \setminus A) = 0$). A simple observation is that if a measure $\mu \in \mathcal{M}^+(K)$, where K is an arbitrary compact space, is carried by some subset of $\mathrm{Ch}_{\mathcal{H}}(K)$, then it is maximal, see [LMNS10, Corollary 3.60]. A point $x \in K$ is said to be *$\mathcal{H}$-exposed*, if there exists $f \in \mathcal{H}$ such that $f(x) = 0$ and $f > 0$ on $K \setminus \{x\}$. An important fact is that $\mathcal{H}$-exposed points of K belong to $\mathrm{Ch}_{\mathcal{H}}(K)$, see [LMNS10, Proposition 3.7]. If X is a compact convex set, then $\mathrm{Ch}_{A(X)}(X)$ equals the set of extreme points of X, denoted by $\mathrm{ext}(X)$.

As we have said in the Introduction, the *set of simpliciality* of $\mathcal{H}$, denoted by $\mathrm{Sim}_{\mathcal{H}}(K)$, is defined as the set of all $x \in K$ for which there exists a unique maximal measure representing x. If X is a compact convex set, we write simply $\mathrm{Sim}(X)$ instead of $\mathrm{Sim}_{A(X)}(X)$. Of course, we have $\mathrm{Ch}_{\mathcal{H}}(K) \subset \mathrm{Sim}_{\mathcal{H}}(K)$. Hence $\mathrm{Sim}_{\mathcal{H}}(K)$ is nonempty if $K \neq \emptyset$, since $\mathrm{Ch}_{\mathcal{H}}(K)$ is (see [LMNS10, Proposition 3.15]). We denote by $\mathcal{M}^1_{PS}(\mathcal{H})$ the set of all $\mu \in \mathcal{M}^1(K)$ for which there exists a unique maximal measure $\nu \in \mathcal{M}^1(K)$ such that $\mu \preceq \nu$ (in [Bač09], the set $\mathcal{M}^1_{PS}(\mathcal{H})$ was denoted simply by PS). Clearly, every maximal measure from $\mathcal{M}^1(K)$ belongs to $\mathcal{M}^1_{PS}(\mathcal{H})$.

A Borel subset B of K is called *$\mathcal{H}$-extremal* if for every $x \in B$ and every $\mu \in \mathcal{M}_x(\mathcal{H})$ we have $\mu(B) = 1$. If X is a compact convex set, then $A(X)$-extremal sets are called *measure extremal*. Further, a subset F of a compact convex set X is called *extremal*, if for every $x, y \in X$ and $\lambda \in (0, 1)$ such that $\lambda x + (1-\lambda)y \in F$, we have $x, y \in F$. Clearly, every measure extremal subset of X is extremal, but extremal subsets of X need not be measure extremal, even if they are Borel. See [DLS06] for a thorough discussion of the relation between extremal and measure extremal sets. The set $\mathrm{Sim}(X)$ is always extremal, see [Bač09, Theorem 4.1], and measure extremal if X is metrizable, as mentioned in Theorem 1.1.2(c), above. In Example 1.5.7 we will show that in general $\mathrm{Sim}(X)$ need not be measure extremal, even if it is Borel.

If X is a locally convex space, the topological dual of X is denoted by X^*. If X, Y are measurable spaces, $\varphi : X \to Y$ is a measurable mapping and μ is a measure on X, then we denote by $\varphi_\sharp\mu$ the image of the measure μ under the mapping φ.

Let us recall the notion of the state space, which will be the main tool to construct compact convex sets with desired properties. The *state space* $S(\mathcal{H})$ of the function space $\mathcal{H}$ is the set $\{\varphi \in \mathcal{H}^* : \varphi \geq 0, \varphi(1) = 1\}$. The set $S(\mathcal{H})$ is a compact convex subset of the space $\mathcal{H}^*$ endowed with the w^*-topology. On $S(\mathcal{H})$ we will always consider the w^*-topology.

Define a mapping $\phi : K \to S(\mathcal{H})$ by $\phi : x \mapsto \phi_x, x \in K$, where $\phi_x : f \mapsto f(x), f \in \mathcal{H}$. Then ϕ is a homeomorphism of K into $S(\mathcal{H})$ and carries $\mathrm{Ch}_{\mathcal{H}}(K)$ onto $\mathrm{ext}(S(\mathcal{H}))$. Let π be the quotient mapping from $\mathcal{M}(K)$ to $\mathcal{H}^*$, that is $\pi(\mu) := \mu|_{\mathcal{H}}, \mu \in \mathcal{M}(K)$. Then $S(\mathcal{H}) = \pi(\mathcal{M}^1(K))$. If $\varphi \in S(\mathcal{H})$ and $\mu \in \mathcal{M}^1(K)$ such that $\pi(\mu) = \varphi$, we write $\mu \sim \varphi$.

We will use the following properties of state spaces, for proofs see [LMNS10, Section 4.3].

Proposition 1.2.1. *Let $\mathcal{H}$ be a function space on a compact space K, and let $X := S(\mathcal{H})$ be its state space. Then we have the following.*

(a) *If $\mu \in \mathcal{M}^1(K)$, then $r(\phi_\sharp\mu) = \pi(\mu)$.*

(b) *A measure $\lambda \in \mathcal{M}^+(X)$ is maximal if and only if $\lambda = \phi_\sharp\mu$ for some $\mu \in \mathcal{M}^+(K)$ maximal.*

Let $\mu \in \mathcal{M}^+(K)$. Then μ is said to be *continuous* if $\mu(\{x\}) = 0$ for each $x \in K$. A point $x \in K$ is said to be an *atom* of μ if $\mu(\{x\}) > 0$. The measure μ is said to be *atomic* (or *discrete*) if there exists a set $M \subset K$ such that $\mu(K \setminus M) = 0$ and M consists of atoms of μ. A well known fact says that every $\mu \in \mathcal{M}^+(K)$ can be uniquely decomposed as $\mu = \mu_a + \mu_c$, where μ_a is atomic (the so-called *atomic part* of μ) and μ_c is continuous (the *continuous part* of μ). If M is a Borel subset of K, we denote by μ_M the measure defined by $\mu_M(A) := \mu(A \cap M)$ for $A \subset K$ Borel.

A characteristic function of a subset M of some set is denoted by χ_M, and for a characteristic function of a singleton $\{x\}$ we use an abbreviation χ_x.

1.3 Basic construction

First, we will construct a basic function space, which will be used later to construct the examples. It is a special case of the construction from [BiLe59, p. 327].

Definition 1.3.1. Let $M \subset [0,1]$ be arbitrary. We define sets $L_x \subset \mathbb{R}^3$, $x \in [0,1]$, by $L_x := \{(x,0,0)\}$ for $x \in [0,1] \setminus M$, and

$$L_x := \{(x,0,0),(x,1,1),(x,-1,1),(x,-1,-1),(x,1,-1)\}$$

for $x \in M$. Define

$$L := \bigcup_{x\in[0,1]} L_x.$$

Topologize L as follows: every point of $L \setminus \big([0,1] \times \{(0,0)\}\big)$ is an open set, and every point $(x,0,0)$, $x \in [0,1]$, has a base of neighbourhoods consisting of the sets

$$\{(x,0,0)\} \cup \bigcup_{y\in U\setminus\{x\}} L_y,$$

where U runs through all neighbourhoods of x in $[0,1]$. Then L is easily seen to be a compact space. Note that the relative topology on $[0,1] \times \{(0,0)\}$ inherited from L coincides with the Euclidean topology. For simplicity, we will write $[0,1]$ instead of $[0,1] \times \{(0,0)\}$, M instead of $M \times \{(0,0)\}$ and x instead of $(x,0,0)$, where no confusion is likely. Further, for $x \in M$, denote $a_x := (x,1,1), b_x := (x,-1,1), c_x := (x,-1,-1)$ and $d_x := (x,1,-1)$.

Define a function space $\mathcal{F}$ on L to be the set of all $f \in \mathcal{C}(L)$ which satisfy

$$f(x) = \frac{1}{2}f(a_x) + \frac{1}{2}f(c_x) = \frac{1}{2}f(b_x) + \frac{1}{2}f(d_x)$$

for every $x \in M$.

Further, define a mapping $\gamma : L \to [0,1]$ by $\gamma(y) := x$ if $y \in L_x$.

Let us now present some properties of the function space $\mathcal{F}$.

Claim 1.3.2. *We have* $\mathrm{Sim}_{\mathcal{F}}(L) = \mathrm{Ch}_{\mathcal{F}}(L) = L \setminus M$.

Proof. If $x \in M$, the functions $\chi_{a_x} - \chi_{c_x}$ and $\chi_{b_x} - \chi_{d_x}$ show that the points a_x, b_x, c_x, d_x are $\mathcal{H}$-exposed points of L, and therefore belong to $\mathrm{Ch}_{\mathcal{F}}(L)$. If $x \in [0,1] \setminus M$, define a function $f \in \mathcal{F}$ by $f(y) := |\gamma(y) - x|$, $y \in L$. This function shows that every $x \in [0,1] \setminus M$ is $\mathcal{H}$-exposed, hence belongs to $\mathrm{Ch}_{\mathcal{F}}(L)$. If $x \in M$, the point x does not belong to $\mathrm{Sim}_{\mathcal{F}}(L)$, since it has two maximal representing measures, $\frac{1}{2}\varepsilon_{a_x} + \frac{1}{2}\varepsilon_{c_x}$ and $\frac{1}{2}\varepsilon_{b_x} + \frac{1}{2}\varepsilon_{d_x}$. These measures are maximal since they are carried by a subset of $\mathrm{Ch}_{\mathcal{F}}(L)$. $\square$

Claim 1.3.3. *Let* $M = [0,1]$ *and let* $\mu \in \mathcal{M}^+(L)$ *be continuous and carried by* $[0,1]$ *(note that such a nontrivial measure clearly exists). If* $\nu \in \mathcal{M}^+(L)$ *is such that* $\nu \sim \mu$*, then* $\nu = \mu$*. In particular,* μ *is maximal.*

Proof. First we will show that ν is carried by $[0,1]$. Assume for the contradiction that this is not the case. Then ν has an atom, since $L \setminus [0,1]$ is discrete. Hence the measure $\gamma_\sharp\nu$, which is carried by $[0,1]$, also has an atom. Further, if $f \in \mathcal{C}([0,1])$, then $f \circ \gamma \in \mathcal{F}$, and consequently $\mu(f) = \mu(f \circ \gamma) = \nu(f \circ \gamma) = \gamma_\sharp\nu(f)$. Therefore $\mu = \gamma_\sharp\nu$, a contradiction with the continuity of μ. Hence ν is carried by $[0,1]$. But this means that $\nu = \gamma_\sharp\nu$ and therefore $\mu = \nu$. $\square$

Claim 1.3.4. *Let $M = [0,1]$. A measure $\mu \in \mathcal{M}^1(L)$ is maximal if and only if $\mu_{[0,1]}$ is continuous.*

Proof. Since $[0,1] = L \setminus \mathrm{Ch}_{\mathcal{F}}(L)$ by Claim 1.3.2, it follows that $\mu_{[0,1]}$ is continuous for μ maximal (see [LMNS10, Proposition 3.66]).

Let $\mu_{[0,1]}$ be continuous. Then $\mu_{[0,1]}$ is maximal by Claim 1.3.3. Further, $\mu_{L\setminus[0,1]}$ is maximal, since it is carried by $\mathrm{Ch}_{\mathcal{F}}(L)$. Hence $\mu = \mu_{[0,1]} + \mu_{L\setminus[0,1]}$ is maximal (since the sum of two maximal measures is again maximal, see [LMNS10, Theorem 3.70]). □

1.4 Function spaces

In this section, we present the promised counterexamples in the general context of function spaces. First, let us show that if we drop the assumption of metrizability, the set of simpliciality need not be Borel.

Example 1.4.1. *There exist a function space $\mathcal{H}$ on a compact space K such that $\mathrm{Sim}_{\mathcal{H}}(K)$ is not Borel.*

Proof. Let $K := L$ and $\mathcal{H} := \mathcal{F}$, where L and $\mathcal{F}$ are as in Definition 1.3.1, with M non-Borel in $[0,1]$. We know from Claim 1.3.2 that $\mathrm{Sim}_{\mathcal{H}}(K) = K \setminus M$. Since M is not Borel in $[0,1]$, we get that $\mathrm{Sim}_{\mathcal{H}}(K) \cap [0,1]$ is not Borel in $[0,1]$, and therefore $\mathrm{Sim}_{\mathcal{H}}(K)$ is not Borel in K. □

Remark 1.4.2. (a) We may, of course, take M in the construction of Example 1.4.1 much more bad than non-Borel. For example, if we take M which is not universally measurable in $[0,1]$ (a subset of some compact topological space is *universally measurable* if it is measurable with respect to the completion of any nonnegative Radon measure), for example M not Lebesgue measurable, then $\mathrm{Sim}_{\mathcal{H}}(K)$ is not universally measurable in K, as is easily seen.

(b) By a suitable modification of the construction in Definition 1.3.1 we may show that there is no connection between the complexity of the Choquet boundary and the set of simpliciality. Let $N \subset [0,1] \setminus M$. Define $L_x, x \in [0,1]$, as in Definition 1.3.1, with the exception that for $x \in N$ put $L_x := \{(x,0,0), (x,1,0), (x,-1,0)\}$, and topologize L similarly as before. The function space $\mathcal{F}$ will be defined as before, with the additional requirement that for every $f \in \mathcal{F}$ and $x \in N$ we have $f(x) = \frac{1}{2}f((x,1,0)) + \frac{1}{2}f((x,-1,0))$. Then we may show, similarly as in the proof of Claim 1.3.2, that

$$\mathrm{Ch}_{\mathcal{F}}(L) = L \setminus (M \cup N) \text{ and } \mathrm{Sim}_{\mathcal{F}}(L) = L \setminus M.$$

The next example shows that even if the set $\mathrm{Sim}_{\mathcal{H}}(K)$ is Borel and $\mu \in \mathcal{M}^1_{PS}(\mathcal{H})$, we cannot guarantee that μ is carried by $\mathrm{Sim}_{\mathcal{H}}(K)$.

Example 1.4.3. *There exist a function space $\mathcal{H}$ on a compact space K and a measure $\mu \in \mathcal{M}^1_{PS}(\mathcal{H})$ (actually, μ is maximal), such that $\mathrm{Sim}_{\mathcal{H}}(K)$ is Borel (actually, open), and μ is carried by a compact set disjoint from $\mathrm{Sim}_{\mathcal{H}}(K)$.*

Proof. Let $K := L$ and $\mathcal{H} := \mathcal{F}$, where L and $\mathcal{F}$ are as in Definition 1.3.1, with $M = [0,1]$. Then, by Claim 1.3.2, we have $\mathrm{Sim}_{\mathcal{H}}(K) = K \setminus [0,1]$, which is an open set in K. Let $\mu \in \mathcal{M}^1(K)$ be continuous and carried by $[0,1]$. By Claim 1.3.3, μ is maximal, but μ is carried by the set $[0,1]$, which is compact and disjoint from $\mathrm{Sim}_{\mathcal{H}}(K)$. ☐

The following example shows that even if the set $\mathrm{Sim}_{\mathcal{H}}(K)$ is Borel, it need not be $\mathcal{H}$-extremal.

Example 1.4.4. *There exist a function space $\mathcal{H}$ on a compact space K such that* $\mathrm{Sim}_{\mathcal{H}}(K)$ *is Borel (actually, open), but not $\mathcal{H}$-extremal.*

Proof. Let L, $\mathcal{F}$ be as in Definition 1.3.1, with $M = [0,1]$. Define $K := L \cup \{a\}$, where $a \notin L$, and topologize K so that a is an isolated point of K and the topology on L remains the same. Let $\mu \in \mathcal{M}^1(K)$ be continuous and carried by $[0,1]$, and let

$$\mathcal{H} := \{f \in \mathcal{C}(K) : f|_L \in \mathcal{F}, f(a) = \mu(f)\}.$$

Clearly, $\mathcal{H}$ is a function space.

Now, $a \notin \mathrm{Ch}_{\mathcal{H}}(K)$, since μ represents a. If $x \in [0,1]$, the functions $\chi_{a_x} - \chi_{c_x}$ and $\chi_{b_x} - \chi_{d_x}$ show that the points a_x, b_x, c_x, d_x are $\mathcal{H}$-exposed, hence belong to $\mathrm{Ch}_{\mathcal{H}}(K)$. Further, each $x \in [0,1]$ does not belong to $\mathrm{Sim}_{\mathcal{H}}(K)$, since it has two maximal representing measures, $\frac{1}{2}\varepsilon_{a_x} + \frac{1}{2}\varepsilon_{c_x}$ and $\frac{1}{2}\varepsilon_{b_x} + \frac{1}{2}\varepsilon_{d_x}$ (which are maximal since they are carried by a subset of $\mathrm{Ch}_{\mathcal{H}}(K)$).

Let us show that $a \in \mathrm{Sim}_{\mathcal{H}}(K)$. Let $\nu \in \mathcal{M}_a(\mathcal{H})$ be maximal. Since $\mathrm{Ch}_{\mathcal{H}}(K) \subset L$ and L is a closed subset of K, the measure ν is carried by L. Since $\nu(f) = \mu(f)$ for every $f \in \mathcal{H}$, and both ν and μ are carried by L, we have $\nu(f) = \mu(f)$ for every $f \in \mathcal{F}$. By Claim 1.3.3, this entails $\nu = \mu$.

So we have $K \setminus \mathrm{Sim}_{\mathcal{H}}(K) = [0,1]$ and therefore $\mathrm{Sim}_{\mathcal{H}}(K)$ is an open set. But $\mu \in \mathcal{M}_a(\mathcal{H})$, $a \in \mathrm{Sim}_{\mathcal{H}}(K)$, and $\mu(\mathrm{Sim}_{\mathcal{H}}(K)) = 0$. Hence $\mathrm{Sim}_{\mathcal{H}}(K)$ is not $\mathcal{H}$-extremal. ☐

Finally, let us show that measures from $\mathcal{M}^1_{PS}(\mathcal{H})$ may even be carried by a compact set disjoint from the closure of the set of simpliciality. This of course makes Example 1.4.3 quite redundant, but the construction of Example 1.4.3 is easier than the construction of Example 1.4.5, and Example 1.4.3 may be of some interest in itself.

Example 1.4.5. *There exist a function space $\mathcal{H}$ on a compact space K and a measure $\mu \in \mathcal{M}^1_{PS}(\mathcal{H})$, such that* $\mathrm{Sim}_{\mathcal{H}}(K)$ *is Borel (actually, open), and μ is carried by a compact set disjoint from* $\overline{\mathrm{Sim}_{\mathcal{H}}(K)}$.

Proof. Let $L, \mathcal{F}$ be as in Definition 1.3.1, with $M = [0,1]$. Let

$$K := L \cup ([0,1] \times \{(1,0)\}) \cup ([0,1] \times \{(2,0)\}) \subset \mathbb{R}^3,$$

topologized so that the sets L and $K \setminus L$ are open, the topology on L remains the same as in Definition 1.3.1, and on $K \setminus L$ we have the Euclidean topology inherited from $\mathbb{R}^3$. The space K is clearly compact. Denote by $\mathcal{H}$ the function space

$$\left\{ f \in \mathcal{C}(K) : f|_L \in \mathcal{F} \text{ and } f((x,1,0)) = \frac{1}{2} f((x,0,0)) + \frac{1}{2} f((x,2,0)), x \in [0,1] \right\}.$$

Then $\operatorname{Sim}_{\mathcal{H}}(K) = \operatorname{Ch}_{\mathcal{H}}(K) = K \setminus \big([0,1] \times \{(0,0),(1,0)\}\big)$. Indeed, the functions $\chi_{a_x} - \chi_{c_x}$ and $\chi_{b_x} - \chi_{d_x}$ show that the points a_x, b_x, c_x, d_x are $\mathcal{H}$-exposed and therefore belong to $\operatorname{Ch}_{\mathcal{H}}(K)$. If $x \in [0,1]$, take a function $f \in \mathcal{H}$ such that $f((y,2,0)) = |y - x|$ for $y \in [0,1]$, and $f = 1$ on L. This function shows that $(x,2,0)$ is $\mathcal{H}$-exposed, hence $(x,2,0) \in \operatorname{Ch}_{\mathcal{H}}(K)$. If $x \in [0,1]$, then $(x,0,0)$ does not belong to $\operatorname{Sim}_{\mathcal{H}}(K)$, since it has two maximal representing measures, $\frac{1}{2}\varepsilon_{a_x} + \frac{1}{2}\varepsilon_{c_x}$ and $\frac{1}{2}\varepsilon_{b_x} + \frac{1}{2}\varepsilon_{d_x}$. These measures are maximal since they are carried by a subset of $\operatorname{Ch}_{\mathcal{H}}(K)$. Finally, if $x \in [0,1]$, then the point $(x,1,0)$ does not belong to $\operatorname{Sim}_{\mathcal{H}}(K)$, since it also has two maximal representing measures, $\frac{1}{4}\varepsilon_{a_x} + \frac{1}{4}\varepsilon_{c_x} + \frac{1}{2}\varepsilon_{(x,2,0)}$ and $\frac{1}{4}\varepsilon_{b_x} + \frac{1}{4}\varepsilon_{d_x} + \frac{1}{2}\varepsilon_{(x,2,0)}$. Again, these measures are maximal since they are carried by a subset of $\operatorname{Ch}_{\mathcal{H}}(K)$.

Let $\lambda \in \mathcal{M}^1([0,1])$ be continuous and let $\lambda_i, i = 0,1,2$, be the copy of the measure λ on the line segment $[0,1] \times \{(i,0)\} \subset K$. Let us show that $\lambda_1 \in \mathcal{M}^1_{PS}(\mathcal{H})$. To this end, we will show that in fact if $\nu \in \mathcal{M}^1(K)$ is a maximal measure such that $\nu \sim \lambda_1$, then $\nu = \frac{1}{2}\lambda_0 + \frac{1}{2}\lambda_2$. It is clear that $\lambda_1 \sim \frac{1}{2}\lambda_0 + \frac{1}{2}\lambda_2$. So, let $\nu \in \mathcal{M}^1(K)$ be maximal such that $\nu \sim \lambda_1$. Then ν is carried by $\overline{\operatorname{Ch}_{\mathcal{H}}(K)} = K \setminus \big([0,1] \times \{(1,0)\}\big)$.

Let $f \in \mathcal{C}\big([0,1] \times \{(2,0)\}\big)$. Let $g \in \mathcal{H}$ be such that $g = 0$ on L and $g = f$ on $[0,1] \times \{(2,0)\}$. Then

$$\nu_{[0,1]\times\{(2,0)\}}(f) = \nu_{[0,1]\times\{(2,0)\}}(g) = \nu(g) = \lambda_1(g) = \frac{1}{2}\lambda_2(g) = \frac{1}{2}\lambda_2(f),$$

and therefore $\nu_{[0,1]\times\{(2,0)\}} = \frac{1}{2}\lambda_2$.

If $f \in \mathcal{F}$, let $g \in \mathcal{H}$ be such that $g = f$ on L and $g = 0$ on $[0,1] \times \{(2,0)\}$. Then

$$\nu_L(f) = \nu_L(g) = \nu(g) = \lambda_1(g) = \frac{1}{2}\lambda_0(g) = \frac{1}{2}\lambda_0(f),$$

and therefore, by Claim 1.3.3, we have $\nu_L = \frac{1}{2}\lambda_0$.

Hence $\nu = \nu_L + \nu_{[0,1]\times\{(2,0)\}} = \frac{1}{2}\lambda_0 + \frac{1}{2}\lambda_2$. Therefore, if we denote $\mu := \lambda_1$, we have $\mu \in \mathcal{M}^1_{PS}(\mathcal{H})$, but μ is carried by a compact set $[0,1] \times \{(1,0)\}$, which is disjoint from the set $\overline{\operatorname{Sim}_{\mathcal{H}}(K)} = K \setminus ([0,1] \times \{(1,0)\})$. $\square$

1.5 The convex case

In this section we will show that the pathologies from Examples 1.4.1, 1.4.3 and 1.4.4 may occur even in the convex case. The method is standard — we will take the state space of an appropriate function space. However, the constructions are not as straightforward as the constructions of compact convex sets whose sets of extreme points have pathological properties. In that case, one may use the fact that if $\mathcal{H}$ is a function space on a compact space K, and X is the state space of $\mathcal{H}$, then $\phi(\mathrm{Ch}_{\mathcal{H}}(K)) = \mathrm{ext}(X)$ (see Preliminaries for explanation), and simply transfer the properties of $\mathrm{Ch}_{\mathcal{H}}(K)$ to $\mathrm{ext}(X)$. This need not hold (and in our examples it does not) for $\mathrm{Sim}_{\mathcal{H}}(K)$ and $\mathrm{Sim}(X)$. However, we have at least the following simple fact, which will be useful in our constructions.

Proposition 1.5.1. *Let $\mathcal{H}$ be a function space on a compact space K and let $X := S(\mathcal{H})$ be its state space. Then $\phi(\mathrm{Sim}_{\mathcal{H}}(K)) = \mathrm{Sim}(X) \cap \phi(K)$.*

Proof. A simple application of Proposition 1.2.1. □

Example 1.5.2. *There exists a compact convex set X such that* $\mathrm{Sim}(X)$ *is not Borel.*

Proof. Let $L, \mathcal{F}$ be as in Definition 1.3.1, with M non-Borel in $[0, 1]$. Let $X := S(\mathcal{F})$ be the state space of the function space $\mathcal{F}$. As we have shown in the construction of Example 1.4.1, the set $\mathrm{Sim}_{\mathcal{F}}(L)$ is not Borel in L. By Proposition 1.5.1, we have $\phi(\mathrm{Sim}_{\mathcal{F}}(L)) = \mathrm{Sim}(X) \cap \phi(L)$. Since ϕ is a homeomorphism, the set $\mathrm{Sim}(X)$ is non-Borel. □

Before we proceed to the next example, let us prove the following statement.

Proposition 1.5.3. *Let $L, \mathcal{F}$ be as in Definition 1.3.1, with $M = [0, 1]$, and let $X := S(\mathcal{F})$ be the state space of the function space $\mathcal{F}$. Then the set* $\mathrm{Sim}(X)$ *is a G_δ set.*

Proof. Let $\varphi \in X$ and let $\mu \in \mathcal{M}^1(L)$ be such that $\mu \sim \varphi$. Then the measure $\gamma_\sharp\mu$ on $[0, 1]$ (γ was defined in Definition 1.3.1) is uniquely determined by φ, as easily follows from the fact that $f \circ \gamma \in \mathcal{F}$ if $f \in \mathcal{C}([0, 1])$. Denote this measure by $\gamma_\sharp\varphi$. Further, denote by N_φ the set of atoms of $\gamma_\sharp\varphi$. Let $x \in N_\varphi$ and let μ^x be the image of the measure μ_{L_x} under the mapping $\omega_x : L_x \to \mathbb{R}^2$ defined by $\omega_x(x) = (0, 0)$, $\omega_x(a_x) = (1, 1)$, $\omega_x(b_x) = (-1, 1)$, $\omega_x(c_x) = (-1, -1)$, $\omega_x(d_x) = (1, -1)$. Denote by r^μ_x the resultant of the measure $\frac{\mu^x}{\gamma_\sharp\varphi(\{x\})}$ (which is a probability measure on the unit square in $\mathbb{R}^2$, that is, a square with vertices $(1, 1), (-1, 1), (-1, -1), (1, -1)$). If $f \in (\mathbb{R}^2)^*$, then the function defined by $f \circ \omega_x$ on L_x and 0 on $L \setminus L_x$ belongs to $\mathcal{F}$, and an easy computation then shows that r^μ_x does not depend on the choice of $\mu \sim \varphi$. Hence we may denote the point r^μ_x by r^φ_x. Further, denote by C the unit square in $\mathbb{R}^2$ and E the union of its edges. For $n \in \mathbb{N}$, let $C_n := \left(1 - \frac{1}{n}\right) C$.

Let us now describe members of $\mathrm{Sim}(X)$.

Claim 1.5.4. *Let $\varphi \in X$. Then $\varphi \in$* $\mathrm{Sim}(X)$ *if and only if $r^\varphi_x \in E$ for every $x \in N_\varphi$.*

Proof of Claim 1.5.4. First, note that $\mathrm{Sim}(C) = E$, which is quite easy to prove.

Let $r^\varphi_x \notin E$ for some $x \in N_\varphi$. Since $r^\varphi_x \notin \mathrm{Sim}(C)$, there are $\nu_1, \nu_2 \in \mathcal{M}^1(C)$, $\nu_1 \neq \nu_2$, which represent r^φ_x, and are maximal, that is, they are carried by the set of vertices of C. Let $\mu \in \mathcal{M}^1(L)$ be maximal such that $\mu \sim \varphi$. Then $\mu_1 := \mu - \mu_{L_x} + \gamma_\sharp\varphi(\{x\})(\omega_x^{-1})_\sharp\nu_1$ and $\mu_2 := \mu - \mu_{L_x} + \gamma_\sharp\varphi(\{x\})(\omega_x^{-1})_\sharp\nu_2$ are two different measures such that $\mu_1, \mu_2 \sim \varphi$, which can be easily proved using the fact that for $f \in \mathcal{F}$ the function $f \circ \omega_x^{-1}$ is the restriction of a continuous affine function on C to the vertices of C and to $(0, 0)$. Further, these two measures are maximal by

Claim 1.3.4, since they are continuous on $[0,1]$. Hence, by Proposition 1.2.1, we have that $\phi_\sharp\mu_1, \phi_\sharp\mu_2$ are two different maximal measures representing a point φ.

Let $r_x^\varphi \in E$ for every $x \in N_\varphi$. Let $\lambda_1, \lambda_2 \in \mathcal{M}^1(X)$ be maximal measures representing a point φ. By Proposition 1.2.1, we have $\lambda_1 = \phi_\sharp\mu_1$ and $\lambda_2 = \phi_\sharp\mu_2$, where $\mu_1, \mu_2 \in \mathcal{M}^1(L)$ are maximal, and $\mu_1, \mu_2 \sim \varphi$. Then $\gamma_\sharp\mu_1 = \gamma_\sharp\mu_2 = \gamma_\sharp\varphi$, and therefore, since μ_1, μ_2 are continuous on $[0,1]$ by Claim 1.3.4, and atomic on $L \setminus [0,1]$ by the discreteness of $L \setminus [0,1]$, we have that both $(\mu_1)_{[0,1]}$ and $(\mu_2)_{[0,1]}$ are equal to the continuous part of $\gamma_\sharp\varphi$. Let us show that if $x \in N_\varphi$, then $(\mu_1)_{L_x} = (\mu_2)_{L_x}$. To this end it suffices to show that $\frac{(\mu_1)^x}{\gamma_\sharp\varphi(\{x\})} = \frac{(\mu_2)^x}{\gamma_\sharp\varphi(\{x\})}$. But this is clear since these two measures are maximal in C (they are carried by vertices of C) and have the same resultant $r_x^\varphi \in E = \operatorname{Sim}(C)$. Hence $\mu_1 = \mu_2$ and therefore $\lambda_1 = \lambda_2$. □

Hence, by Claim 1.5.4, we have

$$\begin{aligned} X \setminus \operatorname{Sim}(X) &= \{\varphi \in X : r_x^\varphi \notin E \text{ for some } x \in N_\varphi\} \\ &= \bigcup_{n\in\mathbb{N}} \{\varphi \in X : r_x^\varphi \in C_n \text{ for some } x \in N_\varphi\} \\ &= \bigcup_{n,m\in\mathbb{N}} \left\{\varphi \in X : r_x^\varphi \in C_n \text{ and } \gamma_\sharp\varphi(\{x\}) \geq \frac{1}{m} \text{ for some } x \in N_\varphi\right\}. \end{aligned}$$

Denote the sets from the last union depending on $n, m \in \mathbb{N}$ by F_{nm}.

Claim 1.5.5. *For every $n, m \in \mathbb{N}$ we have $\overline{F_{nm}} \subset X \setminus \operatorname{Sim}(X)$.*

Proof of Claim 1.5.5. Let $n, m \in \mathbb{N}$ and let $\{\varphi_\alpha\}$ be a net in F_{nm} such that $\varphi_\alpha \to \varphi \in X$. We have to show that $r_x^\varphi \notin E$ for some $x \in N_\varphi$. First, let $x_\alpha \in N_{\varphi_\alpha}$ be witnesses of the fact that $\varphi_\alpha \in F_{nm}$ and let $\mu_\alpha \in \mathcal{M}^1(L)$ be such that $\mu_\alpha \sim \varphi_\alpha$. By passing to a subnet if necessary, we may suppose that $x_\alpha \to x \in [0,1]$ and $\mu_\alpha \to \mu \in \mathcal{M}^1(L)$ (since $\mathcal{M}^1(L)$ is compact). Clearly $\mu \sim \varphi$. Further, we may suppose that either $x_\alpha \neq x$ for every α or $x_\alpha = x$ for every α.

If $x_\alpha \neq x$ for every α, then $\mu(\{x\}) \geq \frac{1}{m}$. Indeed, if $g_k \in \mathcal{C}([0,1])$ is such that $0 \leq g_k \leq 1$, $g_k(x) = 1$ and $g_k(y) = 0$ for $|y-x| \geq \frac{1}{k}$, then a function f_k on L defined by $f_k := g_k \circ \gamma$ on $L \setminus L_x$, $f_k(x) := 1$ and $f_k := 0$ on $L_x \setminus \{x\}$ belongs to $\mathcal{C}(L)$, and clearly $\liminf_\alpha \mu_\alpha(f_k) \geq \frac{1}{m}$. Hence $\mu(f_k) = \lim_\alpha \mu_\alpha(f_k) \geq \frac{1}{m}$. Since $f_k \to \chi_{\{x\}}$, the Lebesgue dominated convergence theorem shows that $\mu(\{x\}) = \lim_k \mu(f_k) \geq \frac{1}{m}$. Consequently $x \in N_\varphi$. To show that $r_x^\varphi \notin E$ it clearly suffices to show that $\frac{\mu^x(g)}{\gamma_\sharp\varphi(\{x\})} < 1$ for every $g \in (\mathbb{R}^2)^*$ such that $g \leq 1$ on C. So take such a g. Define a function f on L by $f := g \circ \omega_x$ on L_x and $f := 0$ on $L \setminus L_x$. Then $f \in \mathcal{F}$ and we have

$$\frac{\mu^x(g)}{\gamma_\sharp\varphi(\{x\})} = \frac{\mu(f)}{\gamma_\sharp\varphi(\{x\})} = \frac{\mu_{L_x\setminus\{x\}}(f)}{\gamma_\sharp\varphi(\{x\})} \leq \frac{\mu(L_x \setminus \{x\})}{\gamma_\sharp\varphi(\{x\})} = \frac{\mu(L_x \setminus \{x\})}{\mu(\{x\}) + \mu(L_x \setminus \{x\})} < 1$$

(the last inequality follows from $\mu(\{x\}) \geq \frac{1}{m} > 0$). Hence $r_x^\varphi \notin E$.

Let $x_\alpha = x$ for every α. Since $\mu_\alpha \to \mu$ and L_x is a compact subset of L, we have $\limsup_\alpha \mu_\alpha(L_x) \leq \mu(L_x)$ (see [LMNS10, Theorem A.85(b)]). Hence $\mu(L_x) \geq \frac{1}{m}$

and consequently $x \in N_\varphi$. Further, there exists an index α_0 such that $\frac{\mu_\alpha(L_x)}{\mu(L_x)} < 1 + \frac{1}{n}$ for $\alpha \succeq \alpha_0$. Take again $g \in (\mathbb{R}^2)^*$ such that $g \leq 1$ on C, and define $f \in \mathcal{F}$ by $f := g \circ \omega_x$ on L_x and $f := 0$ on $L \setminus L_x$. Then we have $\frac{\mu_\alpha(f)}{\mu_\alpha(L_x)} = \frac{(\mu_\alpha)^x(g)}{\gamma_\sharp\varphi_\alpha(\{x\})} = g(r_x^{\varphi_\alpha}) \leq 1 - \frac{1}{n}$, since $r_x^{\varphi_\alpha} \in C_n$ (x witnesses the fact that $\varphi_\alpha \in F_{nm}$). Then for $\alpha \succeq \alpha_0$ we have

$$\frac{\mu_\alpha(f)}{\gamma_\sharp\varphi(\{x\})} = \frac{\mu_\alpha(f)}{\mu(L_x)} = \frac{\mu_\alpha(L_x)}{\mu(L_x)} \frac{\mu_\alpha(f)}{\mu_\alpha(L_x)} < \left(1 + \frac{1}{n}\right)\left(1 - \frac{1}{n}\right) = 1 - \frac{1}{n^2}.$$

Since

$$\frac{\mu_\alpha(f)}{\gamma_\sharp\varphi(\{x\})} \to \frac{\mu(f)}{\gamma_\sharp\varphi(\{x\})},$$

we have $\frac{\mu^x(g)}{\gamma_\sharp\varphi(\{x\})} = \frac{\mu(f)}{\gamma_\sharp\varphi(\{x\})} \leq 1 - \frac{1}{n^2} < 1$. Hence $r_x^\varphi \notin E$. $\square$

Now it is easy to finish the proof of the proposition. By Claim 1.5.5 we have

$$X \setminus \operatorname{Sim}(X) = \bigcup_{n,m \in \mathbb{N}} F_{nm} \subset \bigcup_{n,m \in \mathbb{N}} \overline{F_{nm}} \subset X \setminus \operatorname{Sim}(X),$$

and therefore

$$X \setminus \operatorname{Sim}(X) = \bigcup_{n,m \in \mathbb{N}} \overline{F_{nm}}.$$

Hence the set $\operatorname{Sim}(X)$ is a G_δ set. $\square$

We are now ready to present the remaining examples.

Example 1.5.6. *There exist a compact convex set X and $\lambda \in \mathcal{M}^1_{PS}(A(X))$ (actually, λ is maximal), such that* $\operatorname{Sim}(X)$ *is Borel (actually, a G_δ set), and λ is carried by a compact set disjoint from* $\operatorname{Sim}(X)$.

Proof. Let $L, \mathcal{F}$ be as in Definition 1.3.1, with $M = [0,1]$, and let $\mu \in \mathcal{M}^1(L)$ be continuous and carried by $[0,1]$. Let $X := \mathrm{S}(\mathcal{F})$ be the state space of $\mathcal{F}$, and let $\lambda := \phi_\sharp\mu$. The set $\operatorname{Sim}(X)$ is a G_δ set by Proposition 1.5.3. By Claim 1.3.3, the measure μ is maximal, and therefore λ is maximal by Proposition 1.2.1(b). Since $\operatorname{Sim}_{\mathcal{F}}(L) = L \setminus [0,1]$ by Claim 1.3.2, and $\phi(\operatorname{Sim}_{\mathcal{F}}(L)) = \operatorname{Sim}(X) \cap \phi(L)$ by Proposition 1.5.1, the measure λ is carried by a compact set disjoint from $\operatorname{Sim}(X)$. $\square$

Example 1.5.7. *There exists a compact convex set X such that* $\operatorname{Sim}(X)$ *is Borel (actually, a G_δ set), but not measure extremal.*

Proof. Let again $L, \mathcal{F}$ be as in Definition 1.3.1, with $M = [0,1]$, and let $\mu \in \mathcal{M}^1(L)$ be continuous and carried by $[0,1]$. Let $X := \mathrm{S}(\mathcal{F})$ be the state space of $\mathcal{F}$, and $\lambda := \phi_\sharp\mu$. By Proposition 1.5.3, the set $\operatorname{Sim}(X)$ is a G_δ set. Further, by the same argument as in the construction of Example 1.5.6, the measure λ is carried by a compact set disjoint from $\operatorname{Sim}(X)$. Let $\varphi \in X$ be the resultant of the measure λ. Then $\varphi \in \operatorname{Sim}(X)$. Indeed, let $\Gamma \in \mathcal{M}^1(X)$ be a maximal measure representing φ. By Proposition 1.2.1 we have $\Gamma = \phi_\sharp\nu$ for some $\nu \in \mathcal{M}^1(L)$, and $\nu \sim \mu$. By Claim 1.3.3 we have $\nu = \mu$, and therefore $\Gamma = \lambda$. Hence $\varphi \in \operatorname{Sim}(X)$ (this follows also from Claim 1.5.4, but this statement is unnecessarily strong for this purpose), and $\operatorname{Sim}(X)$ is not measure extremal. $\square$

1.6 Final remarks

Let $\mathcal{H}$ be a function space on a compact space K. Denote by $\mathcal{M}^1_{QS}(\mathcal{H})$ the set of all $\mu \in \mathcal{M}^1(K)$ for which there exists a unique maximal measure ν such that $\mu \sim \nu$ (in [Bač09], the set $\mathcal{M}^1_{QS}(\mathcal{H})$ was denoted by QS). Some characterizations of measures from $\mathcal{M}^1_{QS}(\mathcal{H})$ was given in [Bač09, Theorem 5.3]. Since clearly $\mathcal{M}^1_{QS}(\mathcal{H}) \subset \mathcal{M}^1_{PS}(\mathcal{H})$, we have that measures from $\mathcal{M}^1_{QS}(\mathcal{H})$ are carried by $\mathrm{Sim}_{\mathcal{H}}(K)$ if K is metrizable. However, the measure μ from the construction of Example 1.4.3 belongs to $\mathcal{M}^1_{QS}(\mathcal{H})$ by Claim 1.3.3, but it is carried by a compact set disjoint from the (Borel) set of simpliciality. This may happen also in the convex case. Indeed, the measure λ from the construction of Example 1.5.6 has the same properties. To see that λ really belongs to $\mathcal{M}^1_{QS}(A(X))$, we may use the simple fact that if Y is a compact convex set, then $\nu \in \mathcal{M}^1(Y)$ belongs to $\mathcal{M}^1_{QS}(A(Y))$ if and only if $r(\nu) \in \mathrm{Sim}(Y)$, cf. [Bač09, Remark 5.4]. That $r(\lambda) \in \mathrm{Sim}(X)$ was shown in the construction of Example 1.5.7.

Since the measure μ from the construction of Example 1.4.5 belongs to the set $\mathcal{M}^1_{QS}(\mathcal{H})$, we see that a measure from $\mathcal{M}^1_{QS}(\mathcal{H})$ may even be carried by a compact set disjoint from the closure of the set of simpliciality. However, this cannot happen in the convex case. Indeed, if X is a compact convex set, then the set $\mathrm{Sim}(X)$ is extremal (see [Bač09, Theorem 4.1]). Hence, if $r(\mu) \in \mathrm{Sim}(X)$ for some measure $\mu \in \mathcal{M}^1(X)$, then $\mathrm{supt}(\mu) \subset \overline{\mathrm{Sim}(X)}$ (where $\mathrm{supt}(\mu)$ denotes the support of μ). To prove this, assume for the contradiction that $\mathrm{supt}(\mu) \not\subset \overline{\mathrm{Sim}(X)}$. Then there exists a compact convex set $Y \subset X \setminus \overline{\mathrm{Sim}(X)}$ such that $\mu(Y) > 0$. Further, we have $\mu(X \setminus Y) > 0$, since otherwise we would have $r(\mu) \in Y$. If we denote $x_1 := r\left(\frac{\mu_Y}{\mu(Y)}\right)$ and $x_2 := r\left(\frac{\mu_{X\setminus Y}}{\mu(X\setminus Y)}\right)$, then we have $r(\mu) = \mu(Y)x_1 + (1-\mu(Y))x_2$, as is easily checked. But $r(\mu) \neq x_1$, since $x_1 \in Y$, which is a contradiction with the extremality of the set $\mathrm{Sim}(X)$.

Acknowledgments. The author would like to thank Miroslav Bačák for suggesting the topic and Jaroslav Lukeš for many useful remarks on preliminary versions of this paper.

This research was partially supported by the grants GAČR 201/07/0388 and SVV-2010-261316.

Chapter 2

Two remarks on remotality

Let X be a Banach space (all spaces throughout the paper are considered to be real) and $E \subset X$ be a bounded set. If $x \in X$, we define $D(x, E) := \sup\{\|x - z\| : z \in E\}$. We say that the set E is *remotal* from a point $x \in X$ if there exists a point $e \in E$ such that $\|x - e\| = D(x, E)$. The set E is said to be *remotal* if it is remotal from all $x \in X$.

Consider the following problem: characterize those Banach spaces in which every closed convex bounded set is remotal. Clearly in finite-dimensional spaces every closed bounded set is remotal. M. Sababheh and R. Khalil claimed in [SaKh08, Theorem A] that among reflexive spaces, those spaces in which every closed convex bounded set is remotal are precisely the finite-dimensional ones. However, their proof was not entirely correct. Later, T.S.S.R.K. Rao in [Rao09, Theorem 2.3] proved the assertion of [SaKh08, Theorem A] by showing that even in every Banach space which fails the Schur property, there exists a closed convex bounded set which is not remotal. M. Martín and T.S.S.R.K. Rao in [MaRa10, Theorem 7] then solved the problem completely by showing that in every infinite-dimensional Banach space there exists a closed convex bounded set which is not remotal. Their method was (as well as the method of the previous works [Rao09] and [SaKh08]), roughly speaking, the following. First, they proved that if E is a bounded subset of a Banach space, then, under some additional assumptions on the set E, the remotality of $\overline{\text{co}}\,(E)$ from a point $x \in X$ implies the remotality of E from x. Then they constructed an appropriate bounded set E (considering separately the spaces which fail the Schur property, reproving [Rao09, Theorem 2.3], and the others) which is not remotal from 0, and therefore also $\overline{\text{co}}\,(E)$ is not remotal from 0.

In this connection, they asked in [MaRa10, Remark 6] whether the remotality of $\overline{\text{co}}\,(E)$ from a point $x \in X$, where E is a weakly closed and bounded subset of a Banach space X, implies the remotality of E from x. Example 2.0.1 below answers this question in the negative.

The second purpose of this note is to present an alternative proof of [MaRa10, Theorem 7]. To prove that in every non-reflexive Banach space there exists a closed convex bounded set which is not remotal, we use a simple construction using James' characterization of reflexivity. The case of reflexive spaces is covered by [Rao09, Theorem 2.3] or by [MaRa10, Remark 3].

It should be noted that the statement of [MaRa10, Theorem 7] has also been proved by L. Veselý in [Ves09, Remark 2.10].

Let us first summarize some notation. Let X be a Banach space. The topological dual of X is denoted by X^*. The weak closure of a subset E of X is denoted by $\overline{E}^w$, and the weak convergence in X is denoted by $\xrightarrow{w}$. The convex hull and the closed convex hull of a subset E of X are denoted by $\operatorname{co}(E)$ and $\overline{\operatorname{co}}(E)$ respectively. The symbol c_0 stands for the space of all real sequences vanishing at infinity, equipped with the supremum norm. If $x \in c_0$, we write x^k for the k-th coordinate of x.

Example 2.0.1. *There exists a weakly closed and bounded subset E of c_0 which is not remotal from 0, and such that $\overline{\operatorname{co}}(E)$ is remotal from 0.*

Proof. Define vectors $x_n \in c_0$, $n \in \mathbb{N}$, as

$$x_n := \left(2 - \tfrac{1}{n}, (-1)^n, (-1)^n, \dots, (-1)^n, 0, 0, \dots\right),$$

where the number of nonzero coordinates of x_n is $n+1$. Now, define $E := \{x_n : n \in \mathbb{N}\}$. Then E is a weakly closed and bounded subset of c_0 which is not remotal from 0, while $\overline{\operatorname{co}}(E)$ is remotal from 0.

Clearly the set E is bounded and not remotal from 0. Let us show that E is weakly closed. Assume for the contradiction that there exists $x \in \overline{E}^w \setminus E$. Let $k \in \mathbb{N}, k \geq 2$. We claim that $x^k \in \{-1, 1\}$. It is clear from the definition of the vectors x_n that there exists $m \in \mathbb{N}$ such that $x_n^k \in \{-1, 1\}$ for each $n > m$. And it is easy to see that $x \in \overline{E \setminus \{x_1, \dots, x_m\}}^w$. Then there exists a net $\{y_\alpha\}$ from $E \setminus \{x_1, \dots, x_m\}$ such that $y_\alpha \xrightarrow{w} x$. Applying a functional $\varphi \in (c_0)^*$ such that $\varphi(z) = z^k, z \in c_0$, we see that $y_\alpha^k \to x^k$. Since $y_\alpha^k \in \{-1, 1\}$ for all α, it follows that $x^k \in \{-1, 1\}$. But this is a contradiction with the fact that $x \in c_0$. Hence E is weakly closed.

Now, let us verify that $\overline{\operatorname{co}}(E)$ is remotal from 0. Clearly $D(0, \overline{\operatorname{co}}(E)) = D(0, E) = 2$ (for the first equality see [SaKh08, Lemma 2.1]). Let us show that $(2, 0, 0, \dots) \in \overline{\operatorname{co}}(E)$, which clearly implies the remotality of $\overline{\operatorname{co}}(E)$ from 0. To this end, we will show that if

$$a_n := \sum_{i=1}^{n} \frac{1}{n} x_i \in \operatorname{co}(E),$$

then $a_n \to (2, 0, 0, \dots)$.

First, it is easy to see that if $t_n, t \in \mathbb{R}$ and $t_n \to t$, then also

$$\sum_{i=1}^{n} \frac{1}{n} t_i \xrightarrow{n \to \infty} t.$$

Then

$$a_n^1 = \sum_{i=1}^{n} \frac{1}{n} x_i^1 \to 2,$$

since $x_n^1 = 2 - \frac{1}{n} \to 2$.

Further, let $k \in \mathbb{N}, k \geq 2$. It is clear from the definition of the vectors x_n that

$$\left(x_1^k, x_2^k, x_3^k, \dots\right) = \left(0, \dots, 0, (-1)^{m+1}, (-1)^m, (-1)^{m+1}, (-1)^m, \dots\right),$$

where the number $l \in \mathbb{N} \cup \{0\}$ of zero coordinates of the vector on the right hand side and the number $m \in \{0, 1\}$ depend on k (the precise values of l and m are not important for us). Then

$$\left|a_n^k\right| = \left|\sum_{i=1}^{n} \frac{1}{n} x_i^k\right| \leq \frac{1}{n}.$$

Hence

$$\|a_n - (2, 0, 0, \dots)\| \leq \max\left\{2 - a_n^1, \frac{1}{n}\right\} \to 0,$$

as desired. □

Let us now present the promised proof of [MaRa10, Theorem 7].

Theorem 2.0.2. *Let X be an infinite-dimensional Banach space. Then there exists a closed convex bounded subset of X which is not remotal.*

Proof. If X is in addition reflexive, then it fails the Schur property, and therefore we may apply the argument from [MaRa10, Remark 3] or follow [Rao09, Theorem 2.3].

Suppose that X is not reflexive. By James' theorem (see [Die75, p. 12]), there exists $\varphi \in X^*$ such that $\|\varphi\| = 1$ and φ is not norm-attaining, i.e. there exists no $x \in X$ such that $\|x\| \leq 1$ and $\varphi(x) = 1$. Define

$$K := \left\{x \in X : \|x\|^2 \leq \varphi(x)\right\}.$$

Then K is a closed convex bounded set which is not remotal from 0.

The set K is closed, because the functions $\|.\|^2$ and φ are continuous. To prove the convexity of K, let $x, y \in K$ and $\lambda \in [0, 1]$. Then (we use the fact that the function $t \mapsto t^2, t \in \mathbb{R}$, is convex and non-decreasing on $[0, \infty)$)

$$\begin{aligned} \|\lambda x + (1-\lambda)y\|^2 &\leq (\lambda\|x\| + (1-\lambda)\|y\|)^2 \leq \lambda\|x\|^2 + (1-\lambda)\|y\|^2 \\ &\leq \lambda\varphi(x) + (1-\lambda)\varphi(y) = \varphi(\lambda x + (1-\lambda)y). \end{aligned}$$

Hence K is convex.

Further, $\sup_{x \in K} \|x\| = 1$. Indeed, if $x \in K$, then $\|x\|^2 \leq \varphi(x) \leq \|\varphi\|\|x\| = \|x\|$, and therefore $\|x\| \leq 1$. On the other hand, if $\varepsilon > 0$, then, since $\|\varphi\| = 1$, there exists $y \in X$ such that $\|y\| = 1$ and $|\varphi(y)| > 1 - \varepsilon$. Let $x := \varphi(y)y$. Then $x \in K$, since $\|x\|^2 = \|\varphi(y)y\|^2 = \varphi(y)^2 = \varphi(\varphi(y)y) = \varphi(x)$, and $\|x\| = |\varphi(y)| > 1 - \varepsilon$.

Finally, let us show that there exists no $x \in K$ such that $\|x\| = 1$. Assume for the contradiction that there exists $x \in K$ such that $\|x\| = 1$. Then $1 = \|x\|^2 \leq \varphi(x) \leq \|\varphi\|\|x\| = 1$. Hence $\varphi(x) = 1$, a contradiction with the fact that φ is not norm-attaining. □

Acknowledgments. The author would like to thank Jaroslav Lukeš for suggesting the problems and Eva Pernecká for useful remarks on preliminary versions of this paper. The author also wishes to thank Aldric Brown and the referees for helpful comments.

This research was partially supported by the grant GAČR 201/07/0388.

Chapter 3

Polynomials and identities on real Banach spaces

(joint work with Petr Hájek)

3.1 Introduction

In our present paper we study linear identities via the duality theory for real polynomials and functions on Banach spaces, which allows for a unified treatment and generalization of some classical results in the area. The basic idea is to exploit point evaluations of polynomials, as e.g. in [Rez93]. As a by-product we also obtain a curious variant of the well-known Hilbert lemma on the representation of the even powers of the Hilbert norm as sums of powers of functionals (Corollary 3.3.5). In Theorems 3.3.8, 3.4.1 and 3.4.4 (generalizing [Wil18] and [Rez78]) we prove that under certain natural assumptions identities derived from point evaluations can be satisfied only by polynomials. We apply the Lagrange interpolation theory in order to create a machinery allowing the creation of linear identities which characterize spaces of polynomials of prescribed degrees (Theorem 3.5.2). We elucidate the special situation when all the evaluation points are collinear (Corollary 3.4.8 and Theorem 3.5.4). Our work is based on (and generalizes) the theory of functional equations in the complex plane due to Wilson [Wil18] and Reznick (in the homogeneous case) [Rez78], [Rez79], and the classical characterizations of polynomials due to Fréchet [Fré09a], [Fré09b], and Mazur and W. Orlicz, [MaOr34a], [MaOr34b], which can be summarized in the following theorem.

Theorem 3.1.1. *Let X, Y be real Banach spaces, $f : X \to Y$ be continuous, $n \in \mathbb{N} \cup \{0\}$. TFAE*

(i) $f \in \mathcal{P}^n(X;Y)$.

(ii) $\Delta^{n+1} f(x; h_1, \dots, h_{n+1}) = 0$ *for all* $x, h_i \in X$.

(iii) $f \restriction_E$ *is a polynomial of degree at most n for every affine one-dimensional subspace E of X.*

(iv)
$$\sum_{k=0}^{n+1}(-1)^{n+1-k}\binom{n+1}{k} f(x+kh)=0 \ \textit{for all}\ \ x,h\in X.$$

Here we use the higher order differences defined as follows.

$$\Delta^k f(x;h_1,\dots,h_k)=\sum_{j=0}^{k}\sum_{A\subset\{1,\dots,k\},|A|=j}(-1)^{k-j}f\left(x+\sum_{l\in A}h_l\right).$$

In particular,

$$\Delta^k f(x;h,\dots,h)=\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j} f(x+jh).$$

The theory of linear identities for Banach space norms was developed by many authors. Its first and well-known result is a theorem of Jordan and von Neumann.

Theorem 3.1.2 ([JoNe35]). *Let $(X,\|\cdot\|)$ be a Banach space such that*

$$\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2, x,y\in X.$$

Then X is isometric to a Hilbert space.

Note that a real Banach space X is isometric to a Hilbert space iff $\|.\|^2$ is a 2-homogeneous polynomial. Theorem 3.1.2 has been the basis of subsequent development with the aim of using similar identities in order to characterize the Hilbert spaces, or the classes of Banach spaces allowing the polynomial norms, e.g. Carlsson [Car64], Day [Day47], [Day59], Giles [Gil67], G.G. Johnson [Joh73], Koehler [Koe70], [Koe72], Lorch [Lor48], Reznick [Rez78], [Rez79] and Senechalle [Sen68]. This theory is closely related to the isometric Banach space theory, see e.g. Koldobsky and Konig [KoKö01] and references therein. In our paper, we develop an abstract approach to the theory of linear identities, generalizing Wilson's and Reznick's work. The novelty lies in giving a new functional-analytic meaning to these identities, finding the link to the Lagrange interpolation, and finding a general method for establishing new identitites with prescribed properties.

3.2 Basic facts and definitions

We begin developing our abstract framework. Let X,Y be real Banach spaces. We denote by $\mathcal{P}(^dX;Y)$ (resp. $\mathcal{P}^d(X;Y)$) the Banach space of continuous d-homogeneous polynomials from X to Y (resp. continuous polynomials of degree at most d).

Let $n\in\mathbb{N}, d\in\mathbb{N}\cup\{0\}$. We are going to use some notation and results in [Rez93]. We have a natural identification $(\mathbb{R}^n)^*=\mathbb{R}^n$, using the dot product. For simplicity of notation, we put $F_{n,d}=\mathcal{P}(^d\mathbb{R}^n;\mathbb{R})$. Denote the set of multi-indices by

$$\mathcal{I}(n,d)=\left\{\alpha:\{1,\dots,n\}\to\{0,\dots,d\}:|\alpha|=\sum_{i=1}^{n}\alpha(i)=d\right\}.$$

One gets $\dim F_{n,d} = |\mathcal{I}(n,d)| = \binom{n+d-1}{n-1}$. Further, we put $\Pi_{n,d} = \mathcal{P}^d(\mathbb{R}^n;\mathbb{R})$, $\mathcal{J}(n,d) = \bigcup_{l=0}^{d} \mathcal{I}(n,l)$ is the set of all multi-indices of degree at most d. Clearly, for every $P \in \Pi_{n,d}$ there exist a uniquely determined representation $P(x) = \sum_{\alpha\in\mathcal{J}(n,d)} a_\alpha x^\alpha$, where $x^\alpha = \prod_{i=1}^{n} x_i^{\alpha(i)}$ for $x = (x_1,\dots,x_n) \in \mathbb{R}^n$.

Fact 3.2.1.

$$\mathit{dim}\,\Pi_{n,d} = \sum_{l=0}^{d} \binom{n+l-1}{n-1} = \binom{n+d}{n} = \mathit{dim}\, F_{n+1,d}.$$

Moreover, there is a natural linear isomorphism $i : F_{n+1,d} \to \Pi_{n,d}$, given by the restriction $i(P) = P \restriction_E$, where $E = \{x \in \mathbb{R}^{n+1} : x_{n+1} = 1\}$ is an affine hyperplane. In other words, performing i on a d-homogeneous polynomial means replacing the $n+1$-st coordinate by the constant 1.

Let $C(\mathbb{R}^n)$ be the space of all continuous functions on $\mathbb{R}^n$. Point evaluations at $x \in \mathbb{R}^n$ belong to the linear dual of $C(\mathbb{R}^n)$. Point evaluations separate elements of $C(\mathbb{R}^n)$. For $z \in \mathbb{R}^n$ we are going to use the notation $\mathbf{z} = 1z \in C(\mathbb{R}^n)^*$ where $\mathbf{z}(f) = f(z), f \in C(\mathbb{R}^n)$, and we will call these evaluation functionals nodes. To simplify the language, we will occasionally identify $z \in \mathbb{R}^n$ with its corresponding node $\mathbf{z}$, calling the elements of $\mathbb{R}^n$ themselves nodes. We are going to introduce an abstract formalism suitable for working with nodes and their linear combinations. Consider the linear space $\mathcal{F}(\mathbb{R}^n)$ of all formal finite linear combinations of nodes. It is important to note that a linear multiple $\xi\mathbf{z}$ of the node $\mathbf{z}$ is not the same element as the node corresponding to the point $\xi z \in \mathbb{R}^n$. Informally, whenever we write ξz as an element of $\mathcal{F}(\mathbb{R}^n)$, it is understood that we are dealing with the element $\xi\mathbf{z}$. In order to distinguish the usual vector summation from the space $\mathbb{R}^n$ from the formal summation of the nodes we will introduce the new summation symbol $\boxplus$. So for every $\mathbf{x} \in \mathcal{F}(\mathbb{R}^n)$ there exist $a_i \in \mathbb{R}, x_i \in \mathbb{R}^n$ so that

$$\mathbf{x} = a_1x_1 \boxplus \cdots \boxplus a_kx_k = \boxplus - \sum_{i=1}^{k} a_ix_i$$

The previous expression is unique if x_i are assumed pairwise distinct and $a_i \neq 0, i = 1,\dots,k$.

The operation $\boxplus$ formally acts on $\mathbf{x} = \boxplus - \sum_{i=1}^{k} a_ix_i$ and $\mathbf{y} = \boxplus - \sum_{i=1}^{l} b_iy_i$ as

$$\mathbf{x} \boxplus \mathbf{y} = \left(\boxplus - \sum_{i=1}^{k} a_ix_i\right) \boxplus \left(\boxplus - \sum_{i=1}^{l} b_iy_i\right).$$

Similarly, we define the scalar multiplication of $\xi \in \mathbb{R}$ and $\mathbf{x}$ as

$$\xi\mathbf{x} = \boxplus - \sum_{i=1}^{k} (\xi a_i)x_i.$$

With these operations $\mathcal{F}(\mathbb{R}^n)$ is a linear space. Then $\langle C(\mathbb{R}^n), \mathcal{F}(\mathbb{R}^n)\rangle$ form a dual pair ([FHHMZ11]) with the evaluation

$$\langle f, \mathbf{x}\rangle = \sum_{i=1}^{k} a_if(x_i).$$

Restricting this dual pairing to subspaces $F_{n,d}$ (resp. $\Pi_{n,d}$) of $C(\mathbb{R}^n)$ leads to a dual factorization of the action of $\boxplus$ on $\mathcal{F}(\mathbb{R}^n)$ so that $\mathbf{x}_d = \boxplus_d - \sum_{i=1}^k a_i x_i$ (resp. $\mathbf{x}^d = \boxplus^d - \sum_{i=1}^k a_i x_i$) and

$$\mathbf{x}_d = \boxplus_d - \sum_{i=1}^{k} a_i x_i = \mathbf{y}_d = \boxplus_d - \sum_{i=1}^{l} b_i y_i$$

iff

$$\langle f, \mathbf{x}_d \rangle = \langle f, \mathbf{y}_d \rangle \text{ holds for all } f \in F_{n,d}$$

(and the resp. case of $\Pi_{n,d}$).

Thus we have a (non-unique) representation of the elements of $F^*_{n,d}$ (resp. $\Pi^*_{n,d}$) as elements in $\mathcal{F}(\mathbb{R}^n)$, given by

$$\langle P, \mathbf{x} \rangle = \langle P, \boxplus - \sum_{i=1}^{k} a_i x_i \rangle = \sum_{i=1}^{k} a_i P(x_i).$$

$P \in F_{n,d}$ (resp. $\Pi_{n,d}$), $\mathbf{x} = \boxplus - \sum_{i=1}^k a_i x_i$. We let $\mathcal{K}_d \hookrightarrow \mathcal{F}(\mathbb{R}^n)$ be the subspace consisting of all elements for which

$$\langle P, \boxplus - \sum_{i=1}^{k} a_i x_i \rangle = 0 \text{ holds for all } P \in \Pi_{n,d}.$$

Then $\Pi^*_{n,d} = \mathcal{F}(\mathbb{R}^n)/\mathcal{K}_d$. Suppose $A = \{y_1, \dots, y_r\} \subset \mathbb{R}^n$. We say that the corresponding set of nodes $\mathbf{A} = \{\mathbf{y_1}, \dots, \mathbf{y_r}\}$ is $F_{n,d}$-independent if the nodes are linearly independent as elements of $F^*_{n,d}$. For simplicity, if the space $F_{n,d}$ is understood, we will often drop the boldface notation and say that A is a set of nodes, and that A is $F_{n,d}$-independent. It is clear from basic linear algebra that A is $F_{n,d}$-independent iff there exist dual elements $\{h_1, \dots, h_r\} \subset F_{n,d}$ so that $h_j(y_k) = \delta^k_j$, where δ is the Kronecker delta. If $\{y_1, ..., y_r\}$ are $F_{n,d}$-independent then $r \leq |\mathcal{I}(n,d)|$. In case of $r = |\mathcal{I}(n,d)|$, $F^*_{n,d} = \mathrm{span}(\{y_k\}^r_{k=1})$ and we call $\{y_k\}^r_{k=1}$ a basic set of nodes for $F_{n,d}$. A classical example of a basic set of nodes for $F_{n,d}$ is the set $\mathcal{I}(n,d)$ (Biermann, see [Rez93]). The following result is immediate.

Proposition 3.2.2. *Let $r = |\mathcal{I}(n,d)|$. If $\{y_k\}^r_{k=1}$ is a basic set of nodes for $F_{n,d}$ and $\{h_k\}^r_{k=1} \subset F_{n,d}$ is its dual basis, then for all $P \in F_{n,d}$*

$$P(x) = \sum_{k=1}^{r} P(y_k) h_k(x), x \in \mathbb{R}^n.$$

The following is a general characterization of basic sets of nodes [Lore92], [Rez93].

Theorem 3.2.3. *Let $r = |\mathcal{I}(n,d)|$, $\mathcal{I}(n,d) = \{\alpha_1, \dots, \alpha_r\}$. Let $\{y_k\}^r_{k=1} \subset \mathbb{R}^n$. Then $\{y_k\}^r_{k=1}$ is a basic set of nodes for $F_{n,d}$ iff it holds*

$$det \begin{pmatrix} y_1^{\alpha_1} y_1^{\alpha_2} \dots y_1^{\alpha_r} \\ y_2^{\alpha_1} y_2^{\alpha_2} \dots y_2^{\alpha_r} \\ \dots \\ y_r^{\alpha_1} y_r^{\alpha_2} \dots y_r^{\alpha_r} \end{pmatrix} \neq 0.$$

Moreover, if $\{y_k\}_{k=1}^r$ *is a basic set of nodes for* $F_{n,d}$*, then every* $P \in F_{n,d}$ *can be written uniquely as* $P(x) = \sum_{k=1}^r a_k \langle y_k, x\rangle^d$.

The same notation and terminology applies to the case of $\Pi_{n,d}$ spaces. Analogously, for $r = |\mathcal{J}(n,d)|$, we say that $\{y_k\}_{k=1}^r \subset \mathbb{R}^n$ is a basic set of nodes for $\Pi_{n,d}$ if these elements form a linear basis of $\Pi_{n,d}^*$. Observe that basic sets of nodes exist, as the pointwise evaluations form a separating set of functionals for $\Pi_{n,d}$. The following is a general characterization of basic sets of nodes for $\Pi_{n,d}$, analogous to Theorem 3.2.3, [Lore92].

Theorem 3.2.4. *Let* $r = |\mathcal{J}(n,d)|$, $\mathcal{J}(n,d) = \{\alpha_1, \ldots, \alpha_r\}$. *Let* $\{y_k\}_{k=1}^r \subset \mathbb{R}^n$. *Then* $\{y_k\}_{k=1}^r$ *is a basic set of nodes for* $\Pi_{n,d}$ *iff it holds*

$$det \begin{pmatrix} y_1^{\alpha_1} y_1^{\alpha_2} \ldots y_1^{\alpha_r} \\ y_2^{\alpha_1} y_2^{\alpha_2} \ldots y_2^{\alpha_r} \\ \ldots \\ y_r^{\alpha_1} y_r^{\alpha_2} \ldots y_r^{\alpha_r} \end{pmatrix} \neq 0.$$

Moreover, if $\{y_k\}_{k=1}^r$ *is a basic set of nodes then every node* $y \in \mathbb{R}^n \hookrightarrow \Pi_{n,d}^*$ *can be written uniquely as a linear combination of the elements in* $\{y_k\}_{k=1}^r$. *More precisely,* $\mathbf{y} = \boxplus^d - \sum_{k=1}^r a_k y_k$ *iff* $\{a_k\}_{k=1}^r$ *form a solution of the system of linear equations*

$$\sum_{k=1}^r a_k y_k^\alpha = y^\alpha, \ \alpha \in \mathcal{J}(n,d).$$

The Generalized Lagrange formula is an expression of linear dependence of nodes in the dual of $\Pi_{n,d}$.

Theorem 3.2.5 (Generalized Lagrange formula)**.** *Let* $r = |\mathcal{J}(n,d)|$, $\{y_k\}_{k=1}^r$ *be a basic set of nodes for* $\Pi_{n,d}$. *Then for every* $z \in \mathbb{R}^n$ *there exists a unique set of coefficients* $a_k(z) \in \mathbb{R}$ *such that* $\mathbf{z} = \boxplus^d - \sum_{k=1}^r a_k(z) y_k$. *The functions* $z \to a_k(z)$ *are polynomials of degree at most d, given by the formula*

$$a_k(z) = \frac{det \begin{pmatrix} y_1^{\alpha_1} y_1^{\alpha_2} \ldots y_1^{\alpha_r} \\ y_2^{\alpha_1} y_2^{\alpha_2} \ldots y_2^{\alpha_r} \\ \ldots \\ z^{\alpha_1} z^{\alpha_2} \ldots z^{\alpha_r} \\ \ldots \\ y_r^{\alpha_1} y_r^{\alpha_2} \ldots y_r^{\alpha_r} \end{pmatrix}}{det \begin{pmatrix} y_1^{\alpha_1} y_1^{\alpha_2} \ldots y_1^{\alpha_r} \\ y_2^{\alpha_1} y_2^{\alpha_2} \ldots y_2^{\alpha_r} \\ \ldots \\ y_k^{\alpha_1} y_k^{\alpha_2} \ldots y_k^{\alpha_r} \\ \ldots \\ y_r^{\alpha_1} y_r^{\alpha_2} \ldots y_r^{\alpha_r} \end{pmatrix}}.$$

Then $\{a_k, y_k\}_{k=1}^r$ *is a biorthogonal system in* $\Pi_{n,d} \times \Pi_{n,d}^*$ *and the formula*

$$P(z) = \sum_{k=1}^r a_k(z) P(y_k)$$

is valid for $P \in \Pi_{n,d}$.

We remark that the problem of characterizing geometrically basic sets of nodes for $\Pi_{n,d}$, when $n \geq 2$, is open, and it is important for approximation theory and its applications in numerical mathematics. We refer to [Lore92], [ChuYa77], [BoRo90] for more results and references. An interesting special case is due to Chung and Yao [ChuYa77], for certain implicitly described sets of nodes. Let us briefly describe this elegant result, although it is not central for our subsequent work.

Let $x_1, \ldots, x_k \in \mathbb{R}^n, k \geq n$, be such that every affine hyperplane in $\mathbb{R}^n$ contains at most n points of $0, x_1, \ldots, x_k$. Then for every $I \subset \{1, \ldots, k\}$ such that $\#I = n$ there exists a unique point $z_I \in \mathbb{R}^n$ such that $\langle z_I, x_i \rangle = -1$ for every $i \in I$ and $\langle z_I, x_i \rangle \neq -1$ for every $i \notin I$. Indeed, by the hypothesis the points $x_i, i \in I$, lie in an affine hyperplane H not containing 0, and $x_i \notin H$ for every $i \notin I$. Define

$$h_I(x) := \prod_{i=1, i \notin I}^{k} \frac{1 + \langle x, x_i \rangle}{1 + \langle z_I, x_i \rangle} \quad \text{for } x \in \mathbb{R}^n.$$

Then h_I is well-defined and $h_I \in \Pi_{n,k-n}$. Further, if $J \subset \{1, \ldots, k\}$ is such that $\#J = n$, then $h_I(z_J) = \delta_{I,J}$ (δ is the Kronecker delta). Hence the set $\{x_I : I \subset \{1, \ldots, k\}, \#I = n\}$ is a basic set of nodes for $\Pi_{n,k-n}$ (since the cardinality of this set is $\binom{k}{n} = \dim \Pi_{n,k-n}$).

Let $L \in \mathcal{L}(\mathbb{R}^N; \mathbb{R}^M)$. We let $\tilde{L} \in \mathcal{L}(\mathcal{F}(\mathbb{R}^N); \mathcal{F}(\mathbb{R}^M))$ be defined as

$$\tilde{L}\left(\boxplus - \sum_{i=1}^{k} a_i x_i\right) = \boxplus - \sum_{i=1}^{k} a_i L(x_i).$$

We introduce a partial ordering for elements of $\bigcup_{n=1}^{\infty} \mathcal{F}(\mathbb{R}^n)$ by setting for $\mathbf{x} = a_1 x_1 \boxplus \cdots \boxplus a_n x_n \in \mathcal{F}(\mathbb{R}^N)$ and $\mathbf{y} = b_1 y_1 \boxplus \cdots \boxplus b_m y_m \in \mathcal{F}(\mathbb{R}^M)$

$$\mathbf{x} \succ \mathbf{y} \text{ iff } \tilde{L}\mathbf{x} = \mathbf{y} \text{ for some } L \in \mathcal{L}(\mathbb{R}^N; \mathbb{R}^M).$$

Definition 3.2.6. *We say that a polynomial* $P \in \Pi_{n,d}$ *is compatible with* $\mathbf{x} \in \mathcal{F}(\mathbb{R}^m)$ *if*

$$\langle P \circ L, \mathbf{x} \rangle = \langle P, \tilde{L}\mathbf{x} \rangle = 0 \text{ for all } L \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n).$$

Let X, Y *be Banach spaces and* $f : X \to Y$ *be a continuous mapping. Then we say that* f *is compatible with* $\mathbf{x} \in \mathcal{F}(\mathbb{R}^m)$, $\mathbf{x} = \boxplus - \sum_{i=1}^{k} a_i x_i$, *if*

$$\langle f \circ L, \mathbf{x} \rangle = \sum_{i=1}^{k} a_i f(Lx_i) = 0 \text{ for all } L \in \mathcal{L}(\mathbb{R}^m; X).$$

Remark 3.2.7. Clearly, if X, Y are Banach spaces, then a continuous mapping $f : X \to Y$ is compatible with $\mathbf{x} = a_1 x_1 \boxplus \cdots \boxplus a_n x_n \in \mathcal{F}(\mathbb{R}^m)$, where $x_k = (x_k^1, \ldots, x_k^m)$, iff

$$\sum_{k=1}^{n} a_k f\left(\sum_{i=1}^{m} x_k^i z_i\right) = 0 \text{ for every } z_1, \ldots, z_m \in X. \tag{3.1}$$

The expression (3.1) is called a linear identity. In particular, Fréchet theorem 3.1.1 is equivalent to saying that f is a polynomial of degree at most n iff f is compatible with an element $\mathbf{x}_{\mathbf{M},n} \in \mathcal{F}(\mathbb{R}^{n+2})$ (resp. $\mathbf{x}_{\mathbf{F},n} \in \mathcal{F}(\mathbb{R}^2)$) where

$$\mathbf{x}_{\mathbf{M},n} = \boxplus - \sum_{j=0}^{n+1} \sum_{A \subset \{1,\ldots,n+1\}, |A|=j} (-1)^{n+1-j} \left(e_0 + \sum_{l \in A} e_l \right),$$

$$\mathbf{x}_{\mathbf{F},n} = \boxplus - \sum_{k=0}^{n+1} (-1)^{n+1-k} \binom{n+1}{k} (1,k). \tag{3.2}$$

Moreover, the linear operator $L : \mathbb{R}^{n+2} \to \mathbb{R}^2$ defined by

$$L(x_0, x_1, \ldots, x_{n+1}) = \left(x_0, \sum_{i=1}^{n+1} x_i \right)$$

satisfies $\tilde{L}(\mathbf{x}_{\mathbf{M},n}) = \mathbf{x}_{\mathbf{F},n}$, so in particular $\mathbf{x}_{\mathbf{M},n} \succ \mathbf{x}_{\mathbf{F},n}$. It is easy to see that $L : \mathbb{R}^N \to \mathbb{R}^M$ leads to a linear mapping $L^* : \Pi_{M,d} \to \Pi_{N,d}$ defined as $L^*(P) = P \circ L$. The adjoint linear operator $L^{**} : \Pi^*_{N,d} \to \Pi^*_{M,d}$ coincides with $\tilde{L}$ (if the duals are represented using the canonical evaluations). The following is a simple consequence of the definitions.

Fact 3.2.8. *Let $\mathbf{x} \in \mathcal{F}(\mathbb{R}^m), \mathbf{y} \in \mathcal{F}(\mathbb{R}^n)$, X, Y be Banach spaces and $f : X \to Y$ be continuous. Suppose that $\mathbf{x} \succ \mathbf{y}$. Then the compatibility of f with $\mathbf{x}$ implies the compatibility of f with $\mathbf{y}$. Consequently, if $\tilde{L}\mathbf{x} = \mathbf{y}$ for some bijection $L \in \mathcal{L}(\mathbb{R}^m)$, then f is compatible with $\mathbf{x}$ iff f is compatible with $\mathbf{y}$.*

The implication in Fact 3.2.8 cannot be reversed. For example, let $n \in \mathbb{N}$ and let $\mathbf{x}, \mathbf{y} \in \mathcal{F}(\mathbb{R}^3)$ be defined by

$$\mathbf{x} = (-1)^{n+1}(1,0,1) \boxplus \left(\boxplus - \sum_{k=1}^{n+1} (-1)^{n+1-k} \binom{n+1}{k} (1,k,0) \right),$$

$$\mathbf{y} = \boxplus - \sum_{k=0}^{n+1} (-1)^{n+1-k} \binom{n+1}{k} (1,k,0)$$

($\mathbf{x}$ and $\mathbf{y}$ differ only in the third coordinate of the first node). Then clearly $\mathbf{x} \succ \mathbf{y}$. It is also clear that the compatibility of a continuous $f : \mathbb{R}^3 \to \mathbb{R}$ with $\mathbf{y}$ is equivalent to the compatibility of f with $\mathbf{x}_{\mathbf{F},n}$ from (3.2), and therefore the space of those continuous $f : \mathbb{R}^3 \to \mathbb{R}$ compatible with $\mathbf{y}$ is $\Pi_{3,n}$. On the other hand, if $P \in \Pi_{3,1}$ is defined as $P : (x,y,z) \mapsto z$, then $\langle P, \mathbf{x} \rangle = (-1)^{n+1} \neq 0$, and therefore P is not compatible with $\mathbf{x}$. In fact, it will follow from Theorem 3.4.4 that the only continuous functions on $\mathbb{R}^3$ compatible with $\mathbf{x}$ are the constant functions.

3.3 Fundamental properties of compatibility

In this section, we establish basic results concerning compatibility and show that, under some natural assumptions, polynomials are the only continuous mappings satisfying linear identities.

Lemma 3.3.1. *Let* $\mathbf{x} \in \mathcal{F}(\mathbb{R}^m)$*. TFAE*

(i) *For every Banach spaces* X, Y *every* $P \in \mathcal{P}(^d X; Y)$ *is compatible with* $\mathbf{x}$*.*

(ii) *Every* $P \in F_{m,d}$ *is compatible with* $\mathbf{x}$*.*

(iii) $\langle P, \mathbf{x}\rangle = 0$ *for every* $P \in F_{m,d}$*.*

Proof. The implications (i)⇒(ii)⇒(iii) are clear.

(iii)⇒(ii): Suppose that (iii) holds, and let $P \in F_{m,d}$. If $L \in \mathcal{L}(\mathbb{R}^m)$, then $P \circ L \in F_{m,d}$, hence $\langle P \circ L, \mathbf{x}\rangle = 0$, and therefore P is compatible with $\mathbf{x}$.

(ii)⇒(i): Suppose that every $P \in F_{m,d}$ is compatible with $\mathbf{x}$. Let X, Y be Banach spaces and $P \in \mathcal{P}(^d X; Y)$. Let $L \in \mathcal{L}(\mathbb{R}^m; X)$ and choose $\varphi \in Y^*$ arbitrary. Then $\varphi \circ P \circ L \in F_{m,d}$, and therefore $0 = \langle \varphi \circ P \circ L, \mathbf{x}\rangle = \varphi(\langle P \circ L, \mathbf{x}\rangle)$. Since φ was arbitrary, we conclude that $\langle P \circ L, \mathbf{x}\rangle = 0$. ☐

Lemma 3.3.2. *Let* X, Y *be Banach spaces and let* $P = \sum_{k=0}^{d} P_k \in \mathcal{P}^d(X; Y)$*, where* $P_k \in \mathcal{P}(^k X; Y)$ *are* k*-homogeneous summands. If* P *is compatible with* $\mathbf{x} \in \mathcal{F}(\mathbb{R}^m)$*, then each nonzero summand* P_k *is compatible with* $\mathbf{x}$*.*

Proof. By assumption,

$$\langle P \circ L, \mathbf{x}\rangle = \sum_{k=0}^{d} \langle P_k \circ L, \mathbf{x}\rangle = 0 \text{ for all } L \in \mathcal{L}(\mathbb{R}^m; X).$$

In particular, fixing L, composing $L \circ (tId_{\mathbb{R}^m})$, and using the homogenity of P_k we obtain

$$0 = \langle P \circ (L \circ (tId_{\mathbb{R}^m})), \mathbf{x}\rangle = \sum_{k=0}^{d} t^k \langle P_k \circ L, \mathbf{x}\rangle \text{ for all } L \in \mathcal{L}(\mathbb{R}^m; X).$$

The right hand side, for a fixed L, is an Y-valued polynomial in t. Thus each $\langle P_k \circ L, \mathbf{x}\rangle = 0$, otherwise for some t the total value could not be zero. ☐

The following result was proved by Reznick. We give a proof using our formalism.

Lemma 3.3.3. *Let* X, Y *be Banach spaces and let* $0 \neq P \in \mathcal{P}(^d X; Y)$*,* $\mathbf{x} \in \mathcal{F}(\mathbb{R}^m)$*. Then* P *is compatible with* $\mathbf{x}$ *iff the polynomial* $t \to t^d$ *from* $F_{1,d}$ *is compatible with* $\mathbf{x}$*.*

Proof. On one hand, there exists a one dimensional subspace $E \hookrightarrow X$ such that $P \restriction_E = at^d, a \neq 0$. So for every $L : \mathbb{R}^m \to E$ we have that $\langle P \circ L, \mathbf{x}\rangle = 0$. Consequently, t^d is compatible with $\mathbf{x}$. On the other hand, if t^d is compatible with $\mathbf{x}$, then so is every $\phi^d(y)$, where $\phi \in (\mathbb{R}^m)^*$. Indeed, $\phi^d(y)$ is a composition of a linear projection of $\mathbb{R}^m$ onto a one dimensional subspace $F \hookrightarrow \mathbb{R}^m$, and the polynomial t^d defined on $F = \mathbb{R}$. If $Q \in F_{m,d}$, then by Theorem 3.2.3 $Q(y) = \sum a_k \phi_k^d(y)$, so Q is compatible with $\mathbf{x}$, being a sum of finitely many polynomials compatible with $\mathbf{x}$. Lemma 3.3.1 then finishes the proof. ☐

Corollary 3.3.4. *An element* $\mathbf{x} = a_1x_1 \boxplus \cdots \boxplus a_nx_n \in \mathcal{F}(\mathbb{R}^m)$ *is compatible with* $t \to t^d$ *(or any other nonzero d-homogeneous polynomial) iff* $a_1x_1 \boxplus_d \cdots \boxplus_d a_nx_n = 0$ *in* $F^*_{m,d}$.

Corollary 3.3.5. *Let* $0 \neq P \in F_{n,d}$. *Then for any* $Q \in F_{n,d}$ *there exist a finite collection of linear* $L_k \in \mathcal{L}(\mathbb{R}^n)$ *and* $a_k \in \{\pm 1\}, k = 1, \dots, r = |\mathcal{I}(n,d)|$, *such that* $Q = \sum_{k=1}^r a_k P \circ L_k$.

Proof. Suppose, by contradiction, that the linear span $H = \operatorname{span}\{P \circ L : L \in \mathcal{L}(\mathbb{R}^n)\}$ in the space $F_{n,d}$ is a proper subspace, i.e. there exists some $Q \in F_{n,d} \setminus H$ and a linear functional $\mathbf{x}$ which is zero on H and nonzero on Q. Thus P is compatible with $\mathbf{x}$, but Q is not. This contradicts Lemma 3.3.3. Hence $H = F_{n,d}$ and Q is a finite linear combination of elements of the form $P \circ L_k$. By Carathéodory lemma [Rez93], we infer that the number of summands can be chosen not to exceed the dimension of the space $F_{n,d}$. $\square$

The above corollary is analogous to the celebrated Hilbert lemma, which claims that for given $l, n \in \mathbb{N}$ there exists a finite collection $\{\phi_1, \dots, \phi_N\} \subset (\mathbb{R}^n)^*$ such that

$$\|x\|_{\ell_2}^{2l} = \sum_{i=1}^N \phi_i^{2l}(x), x \in \mathbb{R}^n.$$

The difference lies in the value of coefficients a_k, which in the Hilbert case can be chosen to be positive. Such conclusion is false in our setting, by easy examples when Q is non-positive or non-convex and $P(x) = \phi^n(x)$. Much subtler counterexamples follow from the work of Neyman [Ney84], who proved that there exists a finite dimensional Banach space whose norm taken to n-th power is an n-homogeneous polynomial Q but the space is not isometric to a subspace of ℓ_n space. It follows that the polynomial $Q \in F_{n,d}$ may be convex and non-negative and yet it admits no formula with all $a_k \geq 0$.

Next, we investigate under which conditions on $\mathbf{x} \in \mathcal{F}(\mathbb{R}^m)$ the only continuous mappings compatible with $\mathbf{x}$ are polynomials. Under the assumption that there is k such that for every $i \neq k$ the vector x_k is not a multiple of x_i, we will prove in Theorem 3.3.8 that every continuous mapping which is compatible with $\mathbf{x} = a_0x_0 \boxplus \cdots \boxplus a_{n+1}x_{n+1} \in \mathcal{F}(\mathbb{R}^m)$ is necessarily a polynomial of degree at most n. In particular, the Jordan-von Neumann Theorem 3.1.2 follows immediately from this statement. A similar result was proved by Reznick for homogeneous functions.

The assumption that there is k such that for every $i \neq k$ the vector x_k is not a multiple of x_i is in some sense optimal. Indeed, it is easy to see that if the vectors x_i fail this condition and if $d \in \mathbb{N}$, then there are nonzero a_i such that every d-homogeneous continuous mapping is compatible with $\mathbf{x} = a_0x_0 \boxplus \cdots \boxplus a_{n+1}x_{n+1}$.

An interesting example in this direction is derived from the polarization formula below. The second part (3.3) is an easy observation of the present authors, which follows by inspection of the classical proof (e.g. [Din99, p.8]).

Proposition 3.3.6 ([BoHi31], [MaOr34a], Polarization formula)**.** *For every* $P \in \mathcal{P}(^nX; Y)$, *where* X, Y *are Banach spaces, there exists a unique symmetric n-linear form* $\check{P} \in \mathcal{L}^s(^nX; Y)$ *such that* $P(x) = \check{P}(x, \dots, x)$. *The following formula* *holds.*

$$\check{P}(x_1,\dots,x_n) = \frac{1}{2^n n!}\sum_{\varepsilon_i=\pm 1}\varepsilon_1\dots\varepsilon_n P\left(\sum_{i=1}^n \varepsilon_i x_i\right).$$

On the other hand, for every $0 \neq P \in \mathcal{P}({}^k X;Y), k < n$, *or* $k-n$ *odd and positive the following formula holds.*

$$\sum_{\varepsilon_i=\pm 1}\varepsilon_1\dots\varepsilon_n P\left(\sum_{i=1}^n \varepsilon_i x_i\right) = 0, x_i \in X. \tag{3.3}$$

In the remaining case when $k > n$ *and* $k-n$ *is even, there exists* $x \in X$ *such that the left hand side in* (3.3) *for* $x_i = x, i = 1,\dots,n$, *is nonzero.*

Translated into our language, we see that

$$\mathbf{x_B} = \boxplus - \sum_{\varepsilon_i=\pm 1}\varepsilon_1\dots\varepsilon_n\left(\sum_{i=1}^n \varepsilon_i e_i\right) \in \mathcal{F}(\mathbb{R}^n)$$

is compatible with k-homogeneous polynomials iff either $k < n$ or $k-n$ is a positive odd number.

In the proof of Theorem 3.3.8 we will use the following result due to Wilson [Wil18]. Since Wilson's paper may be difficult to acces, we will also give a proof. The original statement in [Wil18] is for functions on $\mathbb{R}^2$, but the proof works with no change for arbitrary mappings between Banach spaces. By a direction determined by a nonzero vector $x \in \mathbb{R}^m$ we mean a one dimensional subspace $\{tx : t \in \mathbb{R}\}$ of $\mathbb{R}^m$.

Theorem 3.3.7 ([Wil18]). *Let* X, Y *be Banach spaces,* $f : X \to Y$ *be a continuous mapping and let* $\mathbf{x} = a_0x_0 \boxplus \dots \boxplus a_{n+1}x_{n+1} \in \mathcal{F}(\mathbb{R}^2), n \in \mathbb{N}\cup\{0\}$. *Suppose that for every* $k \neq n+1$ *the vectors* x_k *and* x_{n+1} *are linearly independent, and that* $a_{n+1} \neq 0$. *Let* $p+2$ *be the number of distinct directions determined by the vectors* $x_0,\dots,x_{n+1}$. *If* f *is compatible with* $\mathbf{x}$, *then* f *is compatible with* $\mathbf{x}_{\mathbf{F},p}$ *from* (3.2).

Proof. Let $x_i = (r_i, s_i), i = 0,\dots,n+1$. By Fact 3.2.8 we may suppose WLOG that $r_{n+1} = 0$ and $s_{n+1} = 1$. Then $r_i \neq 0$ for every $i \neq n+1$. By Remark 3.2.7 the mapping f for every $x, y \in X$ satisfies

$$\sum_{i=0}^n a_i f(r_i x + s_i y) + a_{n+1} f(y) = 0. \tag{3.4}$$

First, let us suppose that $x_0,\dots,x_{n+1}$ are pairwise linearly independent, i.e. $p = n$. Put $\Delta_{j,i} = s_i - r_i\frac{s_j}{r_j}$ for $i,j \in \{0,\dots,n\}$. Then $\Delta_{j,i} = 0$ iff $j = i$.

In the first step, we subtract (3.4) from the equation derived from (3.4) by replacing x by $x - \frac{s_0}{r_0}x$ and y by $y + x$. We obtain

$$\sum_{i=1}^n a_i\Big(f(r_i x + s_i y + \Delta_{0,i}x) - f(r_i x + s_i y)\Big) + a_{n+1}(f(y+x) - f(y)) = 0. \tag{3.5}$$

Note that since $\Delta_{0,0} = 0$, we have eliminated the terms with $i = 0$.

In the second step, we subtract (3.5) from the equation derived from (3.5) by replacing x by $x - \frac{s_1}{r_1}x$ and y by $y + x$. We obtain

$$\sum_{i=2}^{n} a_i \Big(f(r_i x + s_i y + (\Delta_{1,i} + \Delta_{0,i})x) - f(r_i x + s_i y + \Delta_{1,i} x) - f(r_i x + s_i y + \Delta_{0,i} x) + f(r_i x + s_i y) \Big) + a_{n+1} \Big(f(y + 2x) - 2f(y + x) + f(y) \Big) = 0.$$

In this step we have eliminated the terms with $i = 1$.

We continue in this manner. In the k-th step we subtract the last equation from the equation derived from the last one by replacing x by $x - \frac{s_{k-1}}{r_{k-1}}x$ and y by $y + x$. Since the substitutions replace $r_i x + s_i y$ by $r_i x + s_i y + \Delta_{k-1,i} x$ and since $\Delta_{k-1,k-1} = 0$, the subtraction eliminates the terms with $i = k - 1$. After $n + 1$ steps we arrive at

$$a_{n+1} \sum_{k=0}^{n+1} (-1)^{n+1-k} \binom{n+1}{k} f(y + kx) = 0,$$

and since $a_{n+1} \neq 0$, we see that f is compatible with $\mathbf{x}_{\mathbf{F},n}$.

Now consider the case when some pairs of the vectors $x_0, \ldots, x_n$ are linearly dependent. Then in some steps we eliminate terms corresponding to more than one value of i. It is easy to see that after $p + 1$ steps we arrive at

$$a_{n+1} \sum_{k=0}^{p+1} (-1)^{p+1-k} \binom{p+1}{k} f(y + kx) = 0,$$

and therefore f is compatible with $\mathbf{x}_{\mathbf{F},p}$. □

Theorem 3.3.8. *Let X, Y be Banach spaces, $f : X \to Y$ be a continuous mapping and $\mathbf{x} = a_0 x_0 \boxplus \cdots \boxplus a_{n+1} x_{n+1} \in \mathcal{F}(\mathbb{R}^m), m \geq 2, n \in \mathbb{N} \cup \{0\}$. Suppose that for every $k \neq n + 1$ the vector x_{n+1} is not a multiple of x_k, and that $a_{n+1} \neq 0$. Let q be the number of distinct directions determined by the vectors $x_0, \ldots, x_{n+1}$ (0 does not determine a direction), and let $p = \max\{q - 2, 0\}$ (hence $p \leq n$). If f is compatible with $\mathbf{x}$, then f is a polynomial of degree at most p.*

Proof. We may suppose WLOG that $x_0, \ldots, x_{n+1}$ are distinct. We will distinguish between two cases.

Case 1: $x_k \neq 0$ for every $k \neq n + 1$. Then for every $k \neq n + 1$ the vectors x_k and x_{n+1} are linearly independent.

First let $m = 2$. Since f is compatible with $\mathbf{x}$, it is compatible with $\mathbf{x}_{\mathbf{F},p}$ by Theorem 3.3.7. By Theorem 3.1.1 the mapping f is a polynomial of degree at most p.

Now let $m > 2$. By Fact 3.2.8 it is enough to find a $\mathbf{y} \in \mathcal{F}(\mathbb{R}^2)$, such that $\mathbf{x} \succ \mathbf{y}$, and $\mathbf{y}$ satisfies the assumptions of the case $m = 2$. So we claim that there exists a linear operator $T : \mathbb{R}^m \to \mathbb{R}^2$ such that the couple $T(x_k)$ and $T(x_{n+1})$ is linearly independent for all $k \neq n + 1$. (The number of distinct directions

determined by $T(x_0), \dots, T(x_{n+1})$ is clearly less than or equal to q.) This is easily seen as follows. Let $E_k = \operatorname{span}\{x_k, x_{n+1}\} \hookrightarrow \mathbb{R}^m$, $k \in \{0, \dots, n\}$, be a system of 2-dimensional subspaces of $\mathbb{R}^m$. There exists an $(m-2)$-dimensional subspace $F \hookrightarrow \mathbb{R}^m$ such that $F \cap E_k = \{0\}, k \in \{0, \dots, n\}$. (Equivalently, $F + E_k = \mathbb{R}^m$). Then the orthogonal projection T in $\mathbb{R}^m$, with kernel F and two dimensional range $F^\perp \hookrightarrow \mathbb{R}^m$, clearly satisfies the condition.
Case 2: $x_k = 0$ for some $k \neq n+1$. We may suppose that $x_0 = 0$. Let us first show that if f is compatible with $\mathbf{x}$, then $f - f(0)$ is compatible with $\mathbf{y} = a_1x_1 \boxplus \cdots \boxplus a_{n+1}x_{n+1}$. Note first that the compatibility of f with $\mathbf{x}$ yields $\sum_{i=0}^{n+1} a_i f(0) = 0$. Let $L \in \mathcal{L}(\mathbb{R}^m; X)$. Then

$$\begin{aligned}\langle (f - f(0)) \circ L, \mathbf{y} \rangle &= \sum_{i=1}^{n+1} a_i (f - f(0))(Lx_i) = \sum_{i=1}^{n+1} a_i f(Lx_i) - \sum_{i=1}^{n+1} a_i f(0) \\ &= \sum_{i=0}^{n+1} a_i f(Lx_i) - \sum_{i=0}^{n+1} a_i f(0) = \langle f \circ L, \mathbf{x} \rangle = 0.\end{aligned}$$

Hence $f - f(0)$ is compatible with $\mathbf{y}$.

Now if $n \geq 1$, then $\mathbf{y}$ satisfies the hypotheses of Case 1, and therefore $f - f(0)$ is a polynomial of degree at most p. If $n = 0$, then $f - f(0) = 0$, hence it is a polynomial of degree $0 = p$. In both cases, f is a polynomial of degree at most p. □

3.4 The space of compatible mappings

If $\mathbf{x} \in \mathcal{F}(\mathbb{R}^m)$ and X, Y are Banach spaces, then the set of all continuous mappings from X to Y which are compatible with $\mathbf{x}$ is clearly a linear space. We are now ready to describe this space more precisely.

Theorem 3.4.1. *Let $\mathbf{x} = a_0x_0 \boxplus \cdots \boxplus a_{n+1}x_{n+1} \in \mathcal{F}(\mathbb{R}^m), m \geq 2, n \in \mathbb{N} \cup \{0\}$. Suppose that for every $k \neq n+1$ the vector x_{n+1} is not a multiple of x_k, and that $a_{n+1} \neq 0$. Let q be the number of distinct directions determined by the vectors $x_0, \dots, x_{n+1}$ (0 does not determine a direction), and let $p = \max\{q-2, 0\}$. Then there exists $A \subset \{0, \dots, p\}$ such that if X, Y are Banach spaces and $f : X \to Y$ is a continuous mapping, then f is compatible with $\mathbf{x}$ iff $f = \sum_{k \in A} P_k$ for some $P_k \in \mathcal{P}(^kX; Y)$ (if A is empty, the sum is understood to be equal to 0).*

Proof. Let A be the set of all $k \in \{0, \dots, p\}$ for which there exist Banach spaces X, Y and a nonzero polynomial from $\mathcal{P}(^kX; Y)$ which is compatible with $\mathbf{x}$. By Lemma 3.3.3, if $k \in A$, then for every Banach spaces X, Y every polynomial from $\mathcal{P}(^kX; Y)$ is compatible with $\mathbf{x}$, and the same holds also for their linear combinations. Let now X, Y be Banach spaces and $f : X \to Y$ be a continuous mapping compatible with $\mathbf{x}$. By Theorem 3.3.8 the mapping f is a polynomial of degree at most p. Say $f = \sum_{k=0}^{n} P_k$, where $P_k \in \mathcal{P}(^kX; Y)$. If $P_k \neq 0$ for some $k \in \{0, \dots, p\}$, then it follows from Lemma 3.3.2 that $k \in A$. Hence $f = \sum_{k \in A} P_k$. □

It may happen that the set A from the above theorem contains some gaps. In fact, we have even the following.

Theorem 3.4.2. *Let $0 \leq d_1 < d_2 < \cdots < d_k \leq d$ be given integers and let $m \geq 2$. Then there exists $\mathbf{x} = a_1x_1 \boxplus \cdots \boxplus a_nx_n \in \mathcal{F}(\mathbb{R}^m), n \geq 2$, where $x_1, \ldots, x_n$ are pairwise linearly independent vectors and $a_i \neq 0$ for $i = 1, \ldots, n$, such that $\mathbf{x}$ is compatible with $t \to t^l, l \leq d$, iff $l \in \{d_1, d_2, \ldots, d_k\}$.*

Proof. Consider the linear subspace E of $\Pi_{m,d}$ generated by $\bigcup_{i=1}^k F_{m,d_i}$. Let $M = \{0, \ldots, d\} \setminus \{d_1, d_2, \ldots, d_k\}$. Choose for every $l \in M$ some nonzero l-homogeneous polynomial $P_l \in F_{m,l}$. Then $P_l \notin E$ for every $l \in M$.

Now, let $x_1, \ldots, x_n \in \mathbb{R}^m$ be pairwise linearly independent vectors such that the restriction map $\Phi : \Pi_{m,d} \to C(\{x_1, \ldots, x_n\}) = \mathbb{R}^n$ defined by

$$\Phi(P) = P \restriction_{\{x_1,\ldots,x_n\}}, P \in \Pi_{m,d},$$

is one-to-one and not surjective (for example, take a pairwise linearly independent basic set of nodes for $\Pi_{m,d}$ and add one point which is not a multiple of any of the nodes). Then $\Phi(P_l) \notin \Phi(E)$ for every $l \in M$ and $\Phi(\Pi_{m,d})$ is a proper subspace of $\mathbb{R}^n$. It is easy to see that there exists $f = (a_1, \ldots, a_n) \in (\mathbb{R}^n)^* \setminus \{0\}$ such that $f(\Phi(E)) = 0$ and $f(\Phi(P_l)) \neq 0$ for every $l \in M$. It is clear that if $\mathbf{x} = a_1x_1 \boxplus \cdots \boxplus a_nx_n \in \mathcal{F}(\mathbb{R}^m)$, then $\mathbf{x}$ is not compatible with P_l for every $l \in M$, but it is compatible with members of E by Lemma 3.3.1. We may of course suppose that $a_i \neq 0$ for $i = 1, \ldots, n$. Lemma 3.3.3 then concludes the proof. $\square$

More can be said if the points $x_0, \ldots, x_{n+1}$ lie in an affine hyperplane not containing 0.

Lemma 3.4.3. *Let $\mathbf{x} = a_0x_0 \boxplus \cdots \boxplus a_{n+1}x_{n+1} \in \mathcal{F}(\mathbb{R}^m), m \geq 2, n \in \mathbb{N} \cup \{0\}$, where $x_0, \ldots, x_{n+1}$ are distinct and lie in an affine hyperplane not containing 0. If every polynomial from $F_{m,d}$ is compatible with $\mathbf{x}$, then the same holds for every polynomial from $\Pi_{m,d}$.*

Proof. Let H be an affine hyperplane in $\mathbb{R}^m$ which contains $x_0, \ldots, x_{n+1}$ and does not contain 0. Suppose that every polynomial from $F_{m,d}$ is compatible with $\mathbf{x}$. If $P \in \Pi_{m,d}$, then it is clear from Fact 3.2.1 that there exists $Q \in F_{m,d}$ such that $Q \restriction_H = P \restriction_H$. Since Q is compatible with $\mathbf{x}$, we see that $\langle P, \mathbf{x} \rangle = \langle Q, \mathbf{x} \rangle = 0$. By Lemma 3.3.1 every polynomial from $\Pi_{m,d}$ is compatible with $\mathbf{x}$. $\square$

Theorem 3.4.4. *Let $\mathbf{x} = a_0x_0 \boxplus \cdots \boxplus a_{n+1}x_{n+1} \in \mathcal{F}(\mathbb{R}^m), m \geq 2, n \in \mathbb{N} \cup \{0\}$, where $x_0, \ldots, x_{n+1}$ are distinct and lie in an affine hyperplane not containing 0, and $a_k \neq 0$ for $k = 0, \ldots, n+1$.*

If $\sum_{k=0}^{n+1} a_k = 0$, then there exists $l \in \{0, \ldots, n\}$ such that if X, Y are Banach spaces and $f : X \to Y$ is a continuous mapping, then f is compatible with $\mathbf{x}$ iff f is a polynomial of degree at most l.

If $\sum_{k=0}^{n+1} a_k \neq 0$, then there is no nonzero mapping compatible with $\mathbf{x}$.

Proof. Since $x_0, \ldots, x_{n+1}$ are pairwise linearly independent, Theorem 3.4.1 applies. Let $A \subset \{0, \ldots, n\}$ be a set whose existence is ensured by Theorem 3.4.1. If $\sum_{k=0}^{n+1} a_k = 0$, then $t \mapsto 1, t \in \mathbb{R}$, is compatible with $\mathbf{x}$ and therefore A is nonempty. Let $l \in A$ be maximal. Since every polynomial from $F_{m,l}$ is compatible with $\mathbf{x}$, by Lemma 3.4.3 every polynomial from $\Pi_{m,l}$ is also. Hence $A = \{0, \dots, l\}$. This argument also shows that if A is nonempty, then $t \mapsto 1$ is compatible with $\mathbf{x}$, and consequently $\sum_{k=0}^{n+1} a_k = 0$. Hence if $\sum_{k=0}^{n+1} a_k \neq 0$, then there is no nonzero mapping compatible with $\mathbf{x}$. □

Some information on the exact value of l can be derived from the geometrical properties of the set $\{x_0, \dots, x_{n+1}\}$. Clearly there is no lower bound on l, since to each $x_0, \dots, x_{n+1}$ we may take $a_0, \dots, a_{n+1}$ such that $\sum_{k=1}^{n+1} a_k \neq 0$, and then there is no nonzero mapping compatible with $\mathbf{x}$. Even if we demand that $\sum_{k=0}^{n+1} a_k = 0$, it is easy to find such $a_0, \dots, a_{n+1}$ so that some $P \in F_{m,1} = (\mathbb{R}^m)^*$ is not compatible with $\mathbf{x}$. Indeed, take $P \in F_{m,1}$ which is not constant on $x_0, \dots, x_{n+1}$ and then find $a_0, \dots, a_{n+1}$ such that $\sum_{k=0}^{n+1} a_k = 0$ and $\sum_{k=0}^{n+1} a_k P(x_k) \neq 0$. However, there is a simple upper bound in terms of the dimension of the affine hull of the points $x_0, \dots, x_{n+1}$. It will be given in Corollary 3.4.8. In the proof of Lemma 3.4.6 we will use the following simple fact.

Fact 3.4.5. *If $M \subset \mathbb{R}^2$ is a union of n distinct lines containing 0, then M is a nullspace of an n-homogeneous polynomial $P : \mathbb{R}^2 \to \mathbb{R}$. Indeed, let $P(x) = \Pi_{i=1}^n \phi_i(x)$, where $\phi \in (\mathbb{R}^2)^*$ are chosen so that their kernels coincide with the given lines.*

If $M \subset \mathbb{R}^m$, we denote by $\operatorname{aff}(M)$ the affine hull of M.

Lemma 3.4.6. *Let $x_0, \dots, x_{n+1} \in \mathbb{R}^m, n \in \mathbb{N} \cup \{0\}$, be distinct and denote by d the dimension of $\operatorname{aff}(\{x_0, \dots, x_{n+1}\})$. Then there exists $k_0 \in \{0, \dots, n+1\}$ and a polynomial $P : \mathbb{R}^m \to \mathbb{R}$ of degree at most $n+2-d$ such that $P(x_{k_0}) \neq 0$ and $P(x_k) = 0$ for every $k \in \{0, \dots, n+1\} \setminus \{k_0\}$.*

Proof. We may WLOG suppose that $m = d$. The case $d = 1$ is trivial. Let $d \geq 2$. We may further suppose WLOG that $x_0, \dots, x_{d-1}$ are affinely independent, that $M = \operatorname{aff}(\{x_0, \dots, x_{d-1}\})$ is a hyperplane in $\mathbb{R}^d$ (i.e. it is a subspace), and that $x_{n+1} \notin M$. Using a similar argument as in the proof of Theorem 3.3.8 we construct a linear mapping $L : \mathbb{R}^d \to \mathbb{R}^2$ such that $L(x_0), \dots L(x_{d-1})$ lie on a line $p \subset \mathbb{R}^2$, $L(x_{n+1}) \notin p$ and $L(x_{n+1}) \neq L(x_k)$ for all $k \neq n+1$.

Now, there exists $z \in p$ such that the line $q \subset \mathbb{R}^2$ which contains z and $L(x_{n+1})$ does not contain $L(x_k)$ for all $k \neq n+1$. Let $p_1, \dots, p_r$ be distinct lines which contain z and some $L(x_k), k \neq n+1$. Then $r \leq n+2-d$. By Fact 3.4.5 (since a translation of a polynomial of degree r is again a polynomial of degree r) there exists a polynomial $Q : \mathbb{R}^2 \to \mathbb{R}$ of degree $r \leq n+2-d$ such that the nullspace of Q is $\bigcup_{i=1}^r p_i$. Then $P = Q \circ L \in \Pi_{d,n+2-d}$ is the desired polynomial. □

Proposition 3.4.7. *Let $\mathbf{x} = a_0 x_0 \boxplus \cdots \boxplus a_{n+1} x_{n+1} \in \mathcal{F}(\mathbb{R}^m), n \in \mathbb{N} \cup \{0\}$, where $x_0, \dots, x_{n+1}$ are distinct and $a_k \neq 0$ for $k = 0, \dots, n+1$, and denote by d the dimension of $\operatorname{aff}(\{x_0, \dots, x_{n+1}\})$. If every $P \in \Pi_{m,k}$ is compatible with $\mathbf{x}$, then $k \leq n+1-d$.*

Proof. By Lemma 3.4.6 there exists $k_0 \in \{0, \dots, n+1\}$ and a polynomial $P : \mathbb{R}^m \to \mathbb{R}$ of degree at most $n+2-d$ such that $P(x_{k_0}) \neq 0$ and $P(x_k) = 0$ for every

$k \in \{0, \ldots, n+1\} \setminus \{k_0\}$. Then P cannot be compatible with $\mathbf{x}$, since otherwise we would have

$$0 = \langle P, \mathbf{x} \rangle = \sum_{k=0}^{n+1} a_k P(x_k) = a_{k_0} P(x_{k_0}),$$

and therefore $a_{k_0} = 0$, a contradiction. Hence $k \leq n + 1 - d$. $\square$

Corollary 3.4.8. *Let* $\mathbf{x} = a_0 x_0 \boxplus \cdots \boxplus a_{n+1} x_{n+1} \in \mathcal{F}(\mathbb{R}^m), m \geq 2, n \in \mathbb{N} \cup \{0\}$, *where* $x_0, \ldots, x_{n+1}$ *are distinct and lie in an affine hyperplane not containing* 0, $a_k \neq 0$ *for* $k = 0, \ldots, n+1$ *and* $\sum_{k=0}^{n+1} a_k = 0$. *Let* l *be as in Theorem 3.4.4 and denote by* d *the dimension of* $\operatorname{aff}(\{x_0, \ldots, x_{n+1}\})$. *Then* $l \leq n + 1 - d$.

For example, if in Corollary 3.4.8 the points $x_0, \ldots, x_{n+1}$ are affinely independent, then $d = n + 1$ and therefore $l = 0$. Corollary 3.4.8 also shows that in order to achieve the maximal possible value of l in Theorem 3.4.4 (i.e. $l = n$), it is necessary that $x_0, \ldots, x_{n+1}$ be collinear; see Theorem 3.5.4 for more general result.

3.5 Generating linear identities

In order to generate linear identities, we can use Theorem 3.2.5 on the generalized Lagrange formula. In fact, the Lagrange formula is an expression of linear dependence of functionals in the dual of $\Pi_{m,d}$. Let $\{x_k\}_{k=1}^r \subset \mathbb{R}^m$ be a basic set of nodes for $\Pi_{m,d}$ and let $\{h_k\}_{k=1}^r \subset \Pi_{m,d}$ be its dual basis. Given $z \in \mathbb{R}^m \setminus \{x_k\}_{k=1}^r$, there exists a unique set of coefficients $a_k = a_k(z) \in \mathbb{R}$ such that

$$P(z) = \sum_{k=1}^{r} a_k(z) P(x_k) \text{ for every } P \in \Pi_{m,d},$$

and $a_k(z) = h_k(z), k = 1, \ldots, r$. Then every $P \in \Pi_{m,d}$ is compatible with $a_1(z)x_1 \boxplus \cdots \boxplus a_r(z)x_r \boxplus (-1)z$.

Lemma 3.5.1. *Let* $\{x_k\}_{k=1}^r \subset \mathbb{R}^m$ *be a basic set of nodes for* $\Pi_{m,d}$, $z \in \mathbb{R}^m \setminus \{x_k\}_{k=1}^r$, *and let* $\mathbf{x} = a_1(z)x_1 \boxplus \cdots \boxplus a_r(z)x_r \boxplus (-1)z$. *If every* $P \in \Pi_{m,l}$ *is compatible with* $\mathbf{x}$, *then* $l \leq d$.

Proof. Assume WLOG that $a_1(z) \neq 0$. Considering the dual basis of $\{x_k\}_{k=1}^r$ we see that there exists $Q \in \Pi_{m,d}$ such that $Q(x_1) \neq 0$ and $Q(x_k) = 0$ for $k = 2, \ldots, r$. Further, it is clear that there exists $R \in \Pi_{m,1}$ (these are the affine functions on $\mathbb{R}^m$) such that $R(x_1) \neq 0$ and $R(z) = 0$. Then clearly $P = QR \in \Pi_{m,d+1}$, $P(x_1) \neq 0$, $P(x_k) = 0$ for $k = 2, \ldots, r$ and $P(z) = 0$. But then $\langle P, \mathbf{x} \rangle = a_1(z)P(x_1) \neq 0$, hence P is not compatible with $\mathbf{x}$, and therefore $l \leq d$. $\square$

The following theorem describes a method of generating linear identities which, for prescribed d, characterize polynomials of degree at most d.

Theorem 3.5.2. *Let* $\{x_k\}_{k=1}^r \subset \mathbb{R}^m$ *be a basic set of nodes for* $\Pi_{m,d}$, $z \in \mathbb{R}^m \setminus \{x_k\}_{k=1}^r$, *and let* $\mathbf{x} = a_1(z)x_1 \boxplus \cdots \boxplus a_r(z)x_r \boxplus (-1)z$. *Let* $T : \mathbb{R}^m \to \mathbb{R}^n, n > m$, *be an affine one-to-one mapping such that* $0 \notin T(\mathbb{R}^m)$. *Then* $a_1 = a_1(z), \dots, a_r = a_r(z)$ *are the unique coefficients with the following property. Let* $\mathbf{y} = a_1T(x_1) \boxplus \cdots \boxplus a_rT(x_r) \boxplus (-1)T(z)$. *If* X, Y *are Banach spaces and* $f : X \to Y$ *is continuous, then* f *is compatible with* $\mathbf{y}$ *iff* f *is a polynomial of degree at most* d.

Proof. Since $T(x_1), \dots, T(x_r), T(z)$ lie in an affine hyperplane not containing 0, Theorem 3.4.4 applies. It follows that it suffices to prove the theorem for $X = \mathbb{R}^n$ and $Y = \mathbb{R}$, and it also follows that the space of those continuous $f : \mathbb{R}^n \to \mathbb{R}$ which are compatible with $\mathbf{y}$ is $\Pi_{n,l}$ for some $l \in \mathbb{N} \cup \{0\}$ or a trivial space. If $P \in \Pi_{n,d}$, then $P \circ T \in \Pi_{m,d}$, so $P \circ T$ is compatible with $\mathbf{x}$, and therefore $\langle P, \mathbf{y} \rangle = 0$. By Lemma 3.3.1 every member of $\Pi_{n,d}$ is compatible with $\mathbf{y}$. Hence the space of compatible functions is nontrivial and $l \geq d$. On the other hand, if $P \in \Pi_{m,l}$, then $P \circ T^{-1} : T(\mathbb{R}^m) \to \mathbb{R}$ can be extended to a member of $\Pi_{n,l}$, which is compatible with $\mathbf{y}$ by the definition of l. It follows from Lemma 3.3.1 that every polynomial from $\Pi_{m,l}$ is compatible with $\mathbf{x}$. By Lemma 3.5.1 we conclude that $l \leq d$. Theorem 3.2.5 then yields the uniqueness part. □

A special case of Theorem 3.5.2 in dimension one corresponds to the classical Lagrange interpolation polynomial.

Theorem 3.5.3 (Classical Lagrange interpolation)**.** *Let* $x_0, \dots, x_{n+1} \in \mathbb{R}, n \in \mathbb{N} \cup \{0\}$, *be distinct. Then there exist a unique set of coefficients* $a_0, \dots, a_n \in \mathbb{R} \setminus \{0\}$, *such that every* $P \in \Pi_{1,n}$ *is compatible with* $a_0x_0 \boxplus \cdots \boxplus a_nx_n \boxplus (-1)x_{n+1}$. *Moreover,*

$$a_k = \prod_{i=0, i\neq k}^{n} \frac{x_{n+1} - x_i}{x_k - x_i}, \ \ k = 0, \dots, n.$$

The following theorem characterizes those $a_0x_0 \boxplus \cdots \boxplus a_{n+1}x_{n+1} \in \mathcal{F}(\mathbb{R}^m)$ which can be used to characterize polynomials of degree at most n, the highest possible degree. It is a generalization of the equivalence of the conditions (i) and (iv) in Theorem 3.1.1.

Theorem 3.5.4. *Let* $x_0, \dots, x_{n+1} \in \mathbb{R}^m, m \geq 2, n \in \mathbb{N}$, *be distinct points. TFAE*

(i) *The points* $x_0, \dots, x_{n+1}$ *lie on a line not containing* 0.

(ii) *There exist* $a_0, \dots a_n \in \mathbb{R} \setminus \{0\}$ *such that if* X, Y *are Banach spaces and* $f : X \to Y$ *is a continuous mapping, then* f *is compatible with* $\mathbf{x} = a_0x_0 \boxplus \cdots \boxplus a_nx_n \boxplus (-1)x_{n+1}$ *iff* f *is a polynomial of degree at most* n.

Moreover, the coefficients $a_0, \dots, a_n$ *from (ii) are uniquely determined, and if* $T : \mathbb{R} \to \mathbb{R}^m$ *is an affine one-to-one map and* $y_k \in \mathbb{R}, k = 0, \dots, n+1$, *are such that* $T(y_k) = x_k$, *then*

$$a_k = \prod_{i=0, i\neq k}^{n} \frac{y_{n+1} - y_i}{y_k - y_i}, \ \ k = 0, \dots, n.$$

Proof. (i)⇒(ii): Suppose that (i) holds. Since $x_0, \ldots, x_{n+1}$ lie on a line not containing 0, there exists an affine one-to-one map $T : \mathbb{R} \to \mathbb{R}^m$ and $y_k \in \mathbb{R}, k = 0, \ldots, n+1$, such that $T(y_k) = x_k$ and $0 \notin T(\mathbb{R})$. Combining Theorem 3.5.2 with Theorem 3.5.3 gives (ii) and also the moreover part.

(ii)⇒(i): Denote by d the dimension of aff $(\{x_0, \ldots, x_{n+1}\})$. If (ii) holds, then it follows from Proposition 3.4.7 that $n \leq n + 1 - d$, and therefore $x_0, \ldots, x_{n+1}$ are collinear.

Suppose by contradiction that $x_0, \ldots, x_{n+1}$ lie on a line containing 0. It is easy to construct a continuous function $f : \mathbb{R}^m \to \mathbb{R}$ which is not a polynomial but it is linear on every one dimensional subspace of $\mathbb{R}^m$. Let $L \in \mathcal{L}(\mathbb{R}^m)$. As $x_0, \ldots, x_{n+1}$ lie in a one dimensional subspace, the same holds for $L(x_0), \ldots, L(x_{n+1})$. Hence there exists $P \in \Pi_{m,1}$ such that $P(L(x_k)) = f(L(x_k))$ for all k. Since P is compatible with $\mathbf{x}$, we obtain $0 = \langle P \circ L, \mathbf{x} \rangle = \langle f \circ L, \mathbf{x} \rangle$. Hence f is compatible with $\mathbf{x}$. But this is a contradiction, since f is not a polynomial. □

Acknowledgments. P. Hájek was supported in part by Institutional Research Plan AV0Z10190503 and GAČR P201/11/0345. M. Kraus was supported in part by GAČR 201/07/0388 and SVV-2010-261316.

Chapter 4

Coarse and uniform embeddings between Orlicz sequence spaces

4.1 Introduction

Let $(M, d_M), (N, d_N)$ be metric spaces and let $f : M \to N$ be a mapping. Then f is called a *coarse embedding* if there exist nondecreasing functions $\rho_1, \rho_2 : [0, \infty) \to [0, \infty)$ such that $\lim_{t\to\infty} \rho_1(t) = \infty$ and

$$\rho_1(d_M(x, y)) \leq d_N(f(x), f(y)) \leq \rho_2(d_M(x, y)) \quad \text{for all } x, y \in M.$$

We say that f is a *uniform embedding* if f is injective and both f and $f^{-1} : f(M) \to M$ are uniformly continuous. If f is both a coarse embedding and a uniform embedding, then f is called a *strong uniform embedding*. Naturally we say that M *coarsely embeds* into N if there exists a coarse embedding of M into N, and similarly for other types of embeddings. Let us mention that what we call a coarse embedding is called a uniform embedding by some authors. We use the term coarse embedding because in the nonlinear geometry of Banach spaces the term uniform embedding has a well established meaning as above.

The study of conditions under which a Banach space coarsely (or uniformly) embeds into another Banach space has been a very active area of the nonlinear geometry of Banach spaces. Coarse embeddability has received much attention in recent years mainly because of its connection with geometric group theory, whereas the study of uniform embeddability may be regarded as classical. See [Kal08] for a recent survey on the nonlinear geometry of Banach spaces.

Not much is known in general, but there are some partial results. The coarse and uniform embeddability between ℓ_p-spaces is now completely characterized. Let us recall the results. Nowak proved that ℓ_p coarsely embeds into ℓ_2 if $1 \leq p < 2$ [Now05, Proposition 4.1] and that ℓ_2 coarsely embeds into ℓ_p for any $1 \leq p < \infty$ [Now06, Corollary 4]. A construction due to Albiac in [Alb08, proof of Proposition 4.1(ii)], originally used to show that ℓ_p Lipschitz embeds into ℓ_q if $0 < p < q \leq 1$, can be used to show that ℓ_p strongly uniformly embeds into ℓ_q if $1 \leq p < q$ (see also [AlBa12], where this construction is performed for all $0 < p < q$). This fact also follows from Proposition 4.4.1 below, whose proof is based on Albiac's construction. On the other hand, Johnson and Randrianarivony proved that ℓ_p does not coarsely embed into ℓ_2 if $p > 2$ [JoRa06, Theorem 1]. Later, results of Mendel and Naor [MeNa08, Theorems 1.9 and 1.11] showed that ℓ_p actually does not coarsely or uniformly embed into ℓ_q if $p > 2$ and $q < p$. Furthermore, ℓ_2 uniformly embeds into ℓ_p if $1 \leq p < \infty$. Indeed, by [BeLi00, Corollary 8.11], ℓ_2 uniformly embeds into S_{ℓ_2}, which is uniformly homeomorphic to S_{ℓ_p} by [BeLi00, Theorem 9.1]. In fact, ℓ_2 even strongly uniformly embeds into ℓ_p if $1 \leq p < 2$. This will be proved in Theorem 4.3.1 below. We can summarize the results as follows.

Theorem 4.1.1. *Let $p, q \in [1, \infty)$. Then the following assertions are equivalent:*

(i) *ℓ_p coarsely embeds into ℓ_q.*

(ii) *ℓ_p uniformly embeds into ℓ_q.*

(iii) *ℓ_p strongly uniformly embeds into ℓ_q.*

(iv) *$p \leq q$ or $q < p \leq 2$.*

Our aim is to generalize this classification to a wider class of Banach spaces, namely to Orlicz sequence spaces. Let h_M and h_N be Orlicz sequence spaces associated with Orlicz functions M and N, and let β_M and β_N be the upper Matuszewska-Orlicz indices of the functions M and N. We will show that the coarse (uniform) embeddability of h_M into h_N is in most cases determined only by the values of β_M and β_N. The dependence of the embeddability of h_M into h_N on the values of β_M and β_N is very similar to the dependence of the embeddability of ℓ_p into ℓ_q on the values of p and q from Theorem 4.1.1 (note that the upper Matuszewska-Orlicz index of ℓ_p is p). In some cases, however, the embeddability of h_M into h_N is not determined by the values of β_M and β_N. A brief summary of our results is given at the end of the paper.

It is worth mentioning that Borel-Mathurin proved in [Bor10a] the following result concerning uniform homeomorphisms (i.e. bijections which are uniformly continuous and their inverses are also uniformly continuous) between Orlicz sequence spaces. Let M and N be Orlicz functions and let α_M and α_N be their lower Matuszewska-Orlicz indices. If h_M and h_N are uniformly homeomorphic, then $\alpha_M = \alpha_N$ and $\beta_M = \beta_N$. The fact that $\alpha_M = \alpha_N$ was published also in [Bor10b], the fact that $\beta_M = \beta_N$ is a consequence of results of Kalton [Kal12].

This paper is organized as follows. In Section 4.2 we summarize the notation and terminology, and recall basic facts concerning Orlicz sequence spaces. In Section 4.3 we give the proof of the fact that ℓ_2 strongly uniformly embeds into ℓ_p if $1 \leq p < 2$. Section 4.4 then contains the results concerning the coarse and uniform embeddability between Orlicz sequence spaces.

4.2 Preliminaries

Our notation and terminology for Banach spaces is standard, as may be found for example in [LiTz77] and [LiTz79]. All Banach spaces throughout the paper are supposed to be real. The unit sphere of a Banach space X is denoted by S_X. If

$(X_n)_{n=1}^\infty$ is a sequence of Banach spaces and $1 \le p < \infty$, then $(\sum_{n=1}^\infty X_n)_{\ell_p}$ stands for the ℓ_p-sum of these spaces, i.e. the space of all sequences $x = (x_n)_{n=1}^\infty$ such that $x_n \in X_n$ for every n, and $\|x\| = (\sum_{n=1}^\infty \|x_n\|^p)^{\frac{1}{p}} < \infty$. If a Banach space X is isomorphic to a subspace of a Banach space Y, we will sometimes say that X *linearly embeds into* Y.

Let us give the necessary background concerning Orlicz sequence spaces. Details may be found in [LiTz77] and [LiTz79].

A function $M : [0, \infty) \to [0, \infty)$ is called an *Orlicz function* if it is continuous, nondecreasing and convex, and satisfies $M(0) = 0$ and $\lim_{t\to\infty} M(t) = \infty$.

Let M be an Orlicz function. We denote by ℓ_M the Banach space of all real sequences $(x_n)_{n=1}^\infty$ satisfying $\sum_{n=1}^\infty M\left(\frac{|x_n|}{\rho}\right) < \infty$ for some $\rho > 0$, equipped with the norm defined for $x = (x_n)_{n=1}^\infty \in \ell_M$ by

$$\|x\| = \inf\left\{\rho > 0 : \sum_{n=1}^\infty M\left(\frac{|x_n|}{\rho}\right) \le 1\right\}.$$

Let h_M denote the closed subspace of ℓ_M consisting of all $(x_n)_{n=1}^\infty \in \ell_M$ such that $\sum_{n=1}^\infty M\left(\frac{|x_n|}{\rho}\right) < \infty$ for every $\rho > 0$. The sequence $(e_n)_{n=1}^\infty$ of canonical vectors then forms a symmetric basis of h_M. Clearly if $M(t) = t^p$ for some $1 \le p < \infty$, then h_M is just the space ℓ_p with its usual norm.

If $M(t) = 0$ for some $t > 0$, then M is said to be *degenerate*. In this case, h_M is isomorphic to c_0 and ℓ_M is isomorphic to ℓ_∞. In the sequel, *Orlicz functions are always supposed to be nondegenerate.*

We will be interested in the spaces h_M. Note that $\ell_M = h_M$ if and only if ℓ_M is separable if and only if $\beta_M < \infty$, where β_M is defined below.

An important observation is that if two Orlicz functions M_1 and M_2 coincide on some neighbourhood of 0, then h_{M_1} and h_{M_2} consist of the same sequences and the norms induced by M_1 and M_2 are equivalent.

The lower and upper Matuszewska-Orlicz indices of M are defined by

$$\alpha_M = \sup\left\{q \in \mathbb{R} : \sup_{\lambda,t\in(0,1]} \frac{M(\lambda t)}{M(\lambda)t^q} < \infty\right\},$$

$$\beta_M = \inf\left\{q \in \mathbb{R} : \inf_{\lambda,t\in(0,1]} \frac{M(\lambda t)}{M(\lambda)t^q} > 0\right\},$$

respectively. Then $1 \le \alpha_M \le \beta_M \le \infty$. Note also that if $M(t) = t^p$ for some $1 \le p < \infty$, then $\alpha_M = \beta_M = p$. We will need the following theorem due to Lindenstrauss and Tzafriri (see [LiTz77, Theorem 4.a.9]).

Theorem 4.2.1. *Let M be an Orlicz function and let $1 \le p \le \infty$. Then ℓ_p if $p < \infty$, or c_0 if $p = \infty$, is isomorphic to a subspace of h_M if and only if $p \in [\alpha_M, \beta_M]$.*

Let M be an Orlicz function and $x = (x_n)_{n=1}^\infty \in h_M$. Using the Lebesgue's dominated convergence theorem we see that the function

$$\rho \mapsto \sum_{n=1}^\infty M\left(\frac{|x_n|}{\rho}\right),\ \rho > 0,$$

is continuous. In particular,

$$\sum_{n=1}^{\infty} M\left(\frac{|x_n|}{\|x\|}\right) = 1. \tag{4.1}$$

The following lemma is a simple consequence of the convexity of M combined with the fact that $M(0) = 0$, and (4.1).

Lemma 4.2.2. *Let M be an Orlicz function and let $x = (x_n)_{n=1}^{\infty} \in h_M$.*

(a) *If $\|x\| \le 1$, then $\sum_{n=1}^{\infty} M(|x_n|) \le \|x\|$.*

(b) *If $\|x\| \ge 1$, then $\sum_{n=1}^{\infty} M(|x_n|) \ge \|x\|$.*

If X is a Banach space, define $q_X = \inf\{q \ge 2 : X \text{ has cotype } q\}$. Then if M is an Orlicz function, we have

$$q_{h_M} = \max(2, \beta_M). \tag{4.2}$$

This can be proved as follows. Suppose first that $\beta_M < \infty$. Note that h_M, equipped with the natural order, is a Banach lattice. By Remark 2 after Proposition 2.b.5 in [LiTz79], we have

$$\beta_M = \inf\{1 < q < \infty : h_M \text{ satisfies a lower } q\text{-estimate}\}.$$

By [LiTz79, Theorem 1.f.7], if a Banach lattice satisfies a lower r-estimate for some $1 < r < \infty$, then it is q-concave for every $r < q < \infty$. And by [LiTz79, Proposition 1.f.3(i)], if a Banach lattice is q-concave for some $q \ge 2$, then it is of cotype q. Hence $q_{h_M} \le \max(2, \beta_M)$. The opposite inequality follows from the fact that ℓ_{β_M} is isomorphic to a subspace of h_M by Theorem 4.2.1, and $q_{\ell_{\beta_M}} = \max(2, \beta_M)$. If $\beta_M = \infty$, then, by Theorem 4.2.1, h_M contains c_0, and the result follows.

4.3 Embeddings of ℓ_2

In this section we give the promised proof of the fact that ℓ_2 strongly uniformly embeds into ℓ_p if $1 \le p < 2$. The proof is inspired by Nowak's construction of coarse embeddings between these spaces in [Now06, proof of Corollary 4].

Recall that a kernel K on a set X (i.e. a function $K : X \times X \to \mathbb{C}$ such that $K(y, x) = \overline{K(x, y)}$ for every $x, y \in X$) is called

(a) *positive definite* if $\sum_{i,j=1}^{n} K(x_i, x_j) c_i \overline{c_j} \ge 0$ for every $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$ and $c_1, \dots, c_n \in \mathbb{C}$,

(b) *negative definite* if $\sum_{i,j=1}^{n} K(x_i, x_j) c_i \overline{c_j} \le 0$ for every $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$ and $c_1, \dots, c_n \in \mathbb{C}$ satisfying $\sum_{i=1}^{n} c_i = 0$.

Note that if the kernel K is real-valued, then in order to check the positive or negative definiteness of K it suffices to use only the real scalars.

Recall also that for $p, q \in [1, \infty)$, the Mazur map $M_{p,q} : S_{\ell_p} \to S_{\ell_q}$, defined for $x = (x_n)_{n=1}^{\infty}$ by

$$M_{p,q}(x) = \left(|x_n|^{\frac{p}{q}} \operatorname{sign} x_n\right)_{n=1}^{\infty},$$

is a uniform homeomorphism between these unit spheres. If $p > q$, then it satisfies for all $x, y \in S_{\ell_p}$ and for some $C > 0$ the inequalities

$$C\|x - y\|^{\frac{p}{q}} \le \|M_{p,q}(x) - M_{p,q}(y)\| \le \frac{p}{q}\|x - y\|, \tag{4.3}$$

and the opposite inequalities if $p < q$ (with different C) because clearly $M_{q,p} = M_{p,q}^{-1}$. See [BeLi00, Theorem 9.1] for a proof.

Theorem 4.3.1. *Let $1 \le p < 2$. Then ℓ_2 strongly uniformly embeds into ℓ_p.*

Proof. First, for every $t > 0$ there exists a mapping $\varphi_t : \ell_2 \to S_{\ell_2}$ such that for all $x, y \in \ell_2$ we have

$$\|\varphi_t(x) - \varphi_t(y)\|^2 = 2\left(1 - e^{-t\|x-y\|^2}\right). \tag{4.4}$$

To prove this statement, fix $t > 0$. By a simple computation, the function $(x, y) \to \|x - y\|^2$, $(x, y) \in \ell_2 \times \ell_2$, is a negative definite kernel on ℓ_2, and therefore, by [BeLi00, Proposition 8.4], the function $(x, y) \to \mathrm{e}^{-t\|x-y\|^2}$, $(x, y) \in \ell_2 \times \ell_2$, is a positive definite kernel on ℓ_2. By [BeLi00, Proposition 8.5(i)], there exists a Hilbert space H and a mapping $T : \ell_2 \to H$ such that $\mathrm{e}^{-t\|x-y\|^2} = \langle T(x), T(y)\rangle$ for every $x, y \in \ell_2$. The rest is clear.

Let $t_n > 0$, $n \in \mathbb{N}$, be such that $\sum_{n=1}^{\infty} \sqrt{t_n} < \infty$. For each $n \in \mathbb{N}$, define $f_n = M_{2,p} \circ \varphi_{t_n}$. Let $x_0 \in \ell_2$ be arbitrary and define $f : \ell_2 \to (\sum_{n=1}^{\infty} \ell_p)_{\ell_p}$ by $f(x) = (f_n(x) - f_n(x_0))_{n=1}^{\infty}$ (that $f(x) \in (\sum_{n=1}^{\infty} \ell_p)_{\ell_p}$ for every $x \in \ell_2$ will follow from the estimate (4.5) below). Let us show that f is a strong uniform embedding. Since the spaces $(\sum_{n=1}^{\infty} \ell_p)_{\ell_p}$ and ℓ_p are isometric, the proof will be then complete.

Let $x, y \in \ell_2$. Then

$$\begin{aligned}\|f(x) - f(y)\| &\le \sum_{n=1}^{\infty} \|f_n(x) - f_n(y)\| = \sum_{n=1}^{\infty} \|M_{2,p}(\varphi_{t_n}(x)) - M_{2,p}(\varphi_{t_n}(y))\| \\ &\le \frac{2}{p}\sum_{n=1}^{\infty} \|\varphi_{t_n}(x) - \varphi_{t_n}(y)\| = \frac{2\sqrt{2}}{p}\sum_{n=1}^{\infty}\left(1 - e^{-t_n\|x-y\|^2}\right)^{\frac{1}{2}} \\ &\le \frac{2\sqrt{2}}{p}\sum_{n=1}^{\infty}\left(t_n\|x - y\|^2\right)^{\frac{1}{2}} = \|x - y\|\frac{2\sqrt{2}}{p}\sum_{n=1}^{\infty}\sqrt{t_n},\end{aligned} \tag{4.5}$$

where the first inequality follows from the triangle inequality, the second inequality from (4.3), the second equality from (4.4), and the third inequality from the fact that $1 - \mathrm{e}^{-t} \le t$ for all $t \in \mathbb{R}$. By our assumption, $\sum_{n=1}^{\infty} \sqrt{t_n} < \infty$.

On the other hand,

$$
\begin{aligned}
\|f(x)-f(y)\|^p &= \sum_{n=1}^{\infty} \|f_n(x)-f_n(y)\|^p \\
&= \sum_{n=1}^{\infty} \|M_{2,p}(\varphi_{t_n}(x)) - M_{2,p}(\varphi_{t_n}(y))\|^p \\
&\geq C^p \sum_{n=1}^{\infty} \|\varphi_{t_n}(x)-\varphi_{t_n}(y)\|^2 \\
&= 2C^p \sum_{n=1}^{\infty} \left(1-e^{-t_n\|x-y\|^2}\right), \qquad (4.6)
\end{aligned}
$$

where the inequality follows from (4.3).

Define functions ρ_1, ρ_2 on $[0,\infty)$ by

$$\rho_1(s) = 2^{\frac{1}{p}} C \left(\sum_{n=1}^{\infty} \left(1-e^{-t_n s^2}\right)\right)^{\frac{1}{p}}$$

and

$$\rho_2(s) = s\,\frac{2\sqrt{2}}{p} \sum_{n=1}^{\infty} \sqrt{t_n}.$$

Then, by (4.5) and (4.6), for every $x, y \in \ell_2$ we have

$$\rho_1(\|x-y\|) \leq \|f(x)-f(y)\| \leq \rho_2(\|x-y\|).$$

Clearly both ρ_1, ρ_2 are nondecreasing. Let us show that $\rho_1(s) \to \infty$ as $s \to \infty$. Let $N \in \mathbb{N}$. Then there exists $K > 0$ such that for each $1 \leq n \leq N$ and $s > K$ we have $1 - e^{-t_n s^2} \geq \frac{1}{2}$. For such s we then obtain

$$\rho_1(s) \geq 2^{\frac{1}{p}} C \left(\sum_{n=1}^{N} \left(1-e^{-t_n s^2}\right)\right)^{\frac{1}{p}} \geq C N^{\frac{1}{p}}.$$

Hence $\rho_1(s) \to \infty$ as $s \to \infty$, and therefore f is a coarse embedding.

Since $\rho_2(s) \to 0$ as $s \to 0_+$, and

$$\rho_1(s) \geq 2^{\frac{1}{p}} C \left(1-e^{-t_1 s^2}\right)^{\frac{1}{p}} > 0$$

for every $s > 0$, we see that f is also a uniform embedding. □

4.4 Main Results

Let us start with a sufficient condition for the strong uniform embeddability of Orlicz sequence spaces into ℓ_p-spaces. The proof of the following proposition is based on a construction due to Albiac [Alb08, proof of Proposition 4.1(ii)].

Proposition 4.4.1. *Let M be an Orlicz function with $\beta_M < \infty$ and let $p > \beta_M$. Then h_M strongly uniformly embeds into ℓ_p.*

Proof. We may clearly suppose that $M(1) = 1$. Fix arbitrary $\beta_M < q < p$. Then there is $C > 0$ such that

$$\frac{M(\lambda t)}{M(\lambda)t^q} \geq C \quad \text{for every } \lambda, t \in (0,1]. \tag{4.7}$$

We may suppose without loss of generality that

$$\frac{M(\lambda t)}{M(\lambda)t^q} \geq C \quad \text{for every } \lambda > 0 \text{ and } t \in (0,1]. \tag{4.8}$$

Indeed, if (4.7) holds, then in particular $M(t) \geq Ct^q$ for every $0 < t \leq 1$. We may clearly suppose that $Ct^q \leq M(t) \leq Dt^q$ for some $D \geq 1$ and for every $t > 1$. Then if $\lambda > 1$ and $t \in (0,1]$, we have $M(\lambda t) \geq C(\lambda t)^q = C\lambda^q t^q \geq \frac{C}{D}M(\lambda)t^q$. Since $\frac{C}{D} \leq C$, we may take as C in (4.8) the number $\frac{C}{D}$.

We will proceed in two steps.

Step 1: We will construct functions $f_{n,k} : \mathbb{R} \to [0,\infty)$, $n, k \in \mathbb{Z}$, such that for certain constant $A \geq 1$ and for every $s, t \in \mathbb{R}$ we have

$$M(|s-t|) \leq \sum_{n,k=-\infty}^{\infty} |f_{n,k}(s) - f_{n,k}(t)|^p \leq A\, M(|s-t|). \tag{4.9}$$

Suppose that $n \in \mathbb{Z}$. Let $a_n = 2^{n+2} M\left(\frac{1}{2^{n+1}}\right)^{\frac{1}{p}}$ and define

$$f_n(t) = \begin{cases} a_n t & \text{if } t \in \left[0, \frac{1}{2^n}\right], \\ -a_n\left(t - \frac{1}{2^{n-1}}\right) & \text{if } t \in \left(\frac{1}{2^n}, \frac{1}{2^{n-1}}\right], \\ 0 & \text{otherwise.} \end{cases}$$

For $k \in \mathbb{Z}$ then define the translation of f_n by

$$f_{n,k}(t) = f_n\left(t - \frac{k-1}{2^{n+1}}\right), \ t \in \mathbb{R}.$$

Note that for every $n, k \in \mathbb{Z}$ the estimate $0 \leq f_{n,k} \leq a_n \frac{1}{2^n}$ holds, the Lipschitz constant of $f_{n,k}$ is a_n, and the support of $f_{n,k}$ is $\left[\frac{k-1}{2^{n+1}}, \frac{k-1}{2^{n+1}} + \frac{1}{2^{n-1}}\right]$.

For the upper estimate in (4.9), let $s, t \in \mathbb{R}$, $s \neq t$, and let $N \in \mathbb{Z}$ be such that $\frac{1}{2^{N+1}} < |s-t| \leq \frac{1}{2^N}$.

If $n > N$ and $k \in \mathbb{Z}$, then

$$\begin{aligned} |f_{n,k}(s) - f_{n,k}(t)|^p &\leq a_n^p \frac{1}{2^{np}} = 4^p M\left(\frac{1}{2^{n+1}}\right) = 4^p M\left(\frac{1}{2^{n-N}}\frac{1}{2^{N+1}}\right) \\ &\leq 4^p \frac{1}{2^{n-N}} M\left(\frac{1}{2^{N+1}}\right) \leq 4^p \frac{1}{2^{n-N}} M(|s-t|) \end{aligned} \tag{4.10}$$

(the first inequality follows from the fact that $0 \leq f_{n,k} \leq a_n \frac{1}{2^n}$, while the second one from the convexity of M and the fact that $M(0) = 0$).

If $n \le N$ and $k \in \mathbb{Z}$, then

$$
\begin{aligned}
|f_{n,k}(s) - f_{n,k}(t)|^p &\le a_n^p |s-t|^p \le 2^{p(n+2)} M\left(\frac{1}{2^{n+1}}\right)\frac{1}{2^{pN}} \\
&= 4^p \frac{1}{2^{p(N-n)}} M\left(\frac{1}{2^{n+1}}\right) = 4^p \left(\frac{1}{2^{\frac{p}{q}(N-n)}}\right)^q M\left(\frac{1}{2^{n+1}}\right) \\
&\le \frac{4^p}{C} M\left(\frac{1}{2^{\frac{p}{q}(N-n)}}\frac{1}{2^{n+1}}\right) = \frac{4^p}{C} M\left(\frac{1}{2^{\left(\frac{p}{q}-1\right)(N-n)}}\frac{1}{2^{N+1}}\right) \\
&\le \frac{4^p}{C}\frac{1}{2^{\left(\frac{p}{q}-1\right)(N-n)}} M\left(\frac{1}{2^{N+1}}\right) \\
&\le \frac{4^p}{C}\frac{1}{2^{\left(\frac{p}{q}-1\right)(N-n)}} M(|s-t|) \qquad (4.11)
\end{aligned}
$$

(the first inequality follows from the fact that the Lipschitz constant of $f_{n,k}$ is a_n, the third one from (4.8), and the fourth one from the convexity of M and the fact that $M(0) = 0$).

Note that the estimates (4.10) and (4.11) do not depend on k. For $n \in \mathbb{Z}$, denote $S_n = \{k \in \mathbb{Z} : f_{n,k}(s) > 0 \text{ or } f_{n,k}(t) > 0\}$. Clearly the cardinality of S_n is at most 8. Hence, using (4.10) and (4.11),

$$
\begin{aligned}
&\sum_{n,k=-\infty}^{\infty} |f_{n,k}(s) - f_{n,k}(t)|^p \\
&\quad = \sum_{n>N}\sum_{k\in S_n} |f_{n,k}(s) - f_{n,k}(t)|^p + \sum_{n\le N}\sum_{k\in S_n} |f_{n,k}(s) - f_{n,k}(t)|^p \\
&\quad \le 8\cdot 4^p \left(\sum_{n>N}\frac{1}{2^{n-N}} + \frac{1}{C}\sum_{n\le N}\frac{1}{2^{\left(\frac{p}{q}-1\right)(N-n)}}\right) M(|s-t|) \\
&\quad = 8\cdot 4^p \left(1 + \frac{1}{C}\frac{1}{1-2^{1-\frac{p}{q}}}\right) M(|s-t|).
\end{aligned}
$$

So we may take

$$
A = 8\cdot 4^p \left(1 + \frac{1}{C}\frac{1}{1-2^{1-\frac{p}{q}}}\right).
$$

For the lower estimate in (4.9), suppose that $s, t \in \mathbb{R}$, $s < t$, and let now $N \in \mathbb{Z}$ be such that $\frac{1}{2^{N+2}} < |s-t| \le \frac{1}{2^{N+1}}$. Let K be the largest $k \in \mathbb{Z}$ such that s belongs to the support of $f_{N,k}$. Then $s \in \left[\frac{K-1}{2^{N+1}}, \frac{K-1}{2^{N+1}} + \frac{1}{2^{N+1}}\right)$ and $t \in \left[\frac{K-1}{2^{N+1}}, \frac{K-1}{2^{N+1}} + \frac{1}{2^{N}}\right)$. Hence

$$
\begin{aligned}
|f_{N,K}(s) - f_{N,K}(t)|^p = a_N^p |s-t|^p &\ge 2^{p(N+2)} M\left(\frac{1}{2^{N+1}}\right)\frac{1}{2^{p(N+2)}} \\
&= M\left(\frac{1}{2^{N+1}}\right) \ge M(|s-t|),
\end{aligned}
$$

and therefore

$$
\sum_{n,k=-\infty}^{\infty} |f_{n,k}(s) - f_{n,k}(t)|^p \ge |f_{N,K}(s) - f_{N,K}(t)|^p \ge M(|s-t|).
$$

Step 2: Define $f : h_M \to \ell_p(\mathbb{N} \times \mathbb{Z} \times \mathbb{Z})$ by

$$f(x) = (f_{n,k}(x_i) - f_{n,k}(0))_{(i,n,k)\in\mathbb{N}\times\mathbb{Z}\times\mathbb{Z}},$$

where $x = (x_i)_{i=1}^{\infty}$ (the fact that $f(x) \in \ell_p(\mathbb{N}\times\mathbb{Z}\times\mathbb{Z})$ for every $x \in h_M$ will follow from the estimates below). Let us show that f is a strong uniform embedding.

Let $x = (x_i)_{i=1}^{\infty}, y = (y_i)_{i=1}^{\infty} \in h_M$. By (4.9), for each $i \in \mathbb{N}$ we have

$$M(|x_i - y_i|) \le \sum_{n,k=-\infty}^{\infty} |f_{n,k}(x_i) - f_{n,k}(y_i)|^p \le A\, M(|x_i - y_i|),$$

and therefore

$$\sum_{i=1}^{\infty} M(|x_i - y_i|) \le \|f(x) - f(y)\|^p \le A \sum_{i=1}^{\infty} M(|x_i - y_i|).$$

By Lemma 4.2.2, if $\|x - y\| \le 1$, then

$$\|f(x) - f(y)\|^p \le A \sum_{i=1}^{\infty} M(|x_i - y_i|) \le A\|x - y\|,$$

and if $\|x - y\| \ge 1$, then

$$\|f(x) - f(y)\|^p \ge \sum_{i=1}^{\infty} M(|x_i - y_i|) \ge \|x - y\|.$$

If $\|x - y\| > 1$, then, by (4.8), for every $i \in \mathbb{N}$ we have

$$M\left(\frac{|x_i - y_i|}{\|x - y\|}\right) \ge CM(|x_i - y_i|)\frac{1}{\|x - y\|^q},$$

and therefore, using also (4.1), we obtain

$$\begin{aligned}\|f(x) - f(y)\|^p &\le A \sum_{i=1}^{\infty} M(|x_i - y_i|) \le \frac{A}{C} \sum_{i=1}^{\infty} M\left(\frac{|x_i - y_i|}{\|x - y\|}\right) \|x - y\|^q \\ &= \frac{A}{C}\|x - y\|^q.\end{aligned}$$

If $\|x - y\| < 1$, then similarly

$$\begin{aligned}\|f(x) - f(y)\|^p &\ge \sum_{i=1}^{\infty} M(|x_i - y_i|) \ge C \sum_{i=1}^{\infty} M\left(\frac{|x_i - y_i|}{\|x - y\|}\right) \|x - y\|^q \\ &= C\|x - y\|^q.\end{aligned}$$

Now define

$$\rho_1(t) = \begin{cases} C^{\frac{1}{p}} t^{\frac{q}{p}} & \text{if } t \in [0, 1), \\ t^{\frac{1}{p}} & \text{if } t \ge 1, \end{cases}$$

and

$$\rho_2(t) = \begin{cases} A^{\frac{1}{p}} t^{\frac{1}{p}} & \text{if } t \in [0,1], \\ \left(\frac{A}{C}\right)^{\frac{1}{p}} t^{\frac{q}{p}} & \text{if } t > 1. \end{cases}$$

Then ρ_1, ρ_2 are nondecreasing (since $C \leq 1$), $\lim_{t\to\infty} \rho_1(t) = \infty$, and for every $x, y \in h_M$ we have

$$\rho_1(\|x-y\|) \leq \|f(x) - f(y)\| \leq \rho_2(\|x-y\|).$$

Hence f is a coarse embedding, and clearly it is also a uniform embedding. Since $\ell_p(\mathbb{N} \times \mathbb{Z} \times \mathbb{Z})$ is isometric to ℓ_p, we have obtained a strong uniform embedding of h_M into ℓ_p. □

We are now ready to give a sufficient condition for the strong uniform embeddability between Orlicz sequence spaces. Recall that if M, N are metric spaces and $f : M \to N$ is a mapping, then f is called a *Lipschitz embedding* provided f is injective and both f and $f^{-1} : f(M) \to M$ are Lipschitz mappings. Clearly if f is a Lipschitz embedding, then f is a strong uniform embedding.

Theorem 4.4.2. *Let M, N be Orlicz functions. If $\beta_M < \beta_N$ or $\beta_N \leq \beta_M < 2$ or $\beta_M = \beta_N = \infty$, then h_M strongly uniformly embeds into h_N.*

Proof. If $\beta_N = \infty$, then c_0 linearly embeds into h_N by Theorem 4.2.1, and since every separable metric space Lipschitz embeds into c_0 by [Aha74], we conclude that any h_M even Lipschitz embeds into h_N. So suppose that $\beta_N < \infty$.

If $\beta_M < \beta_N$, then h_M strongly uniformly embeds into ℓ_{β_N} by Proposition 4.4.1, and ℓ_{β_N} linearly embeds into h_N by Theorem 4.2.1. Hence h_M strongly uniformly embeds into h_N.

If $\beta_N \leq \beta_M < 2$, then h_M strongly uniformly embeds into ℓ_2 by Proposition 4.4.1. By Theorem 4.3.1, ℓ_2 strongly uniformly embeds into ℓ_{β_N}, which in turn linearly embeds into h_N by Theorem 4.2.1, and therefore h_M strongly uniformly embeds into h_N. □

To give a condition ensuring the nonexistence of a coarse or uniform embedding between two Orlicz sequence spaces, we will use the following result due to Mendel and Naor. Recall that if X is a Banach space, then we define $q_X = \inf\{q \geq 2 : X \text{ has cotype } q\}$.

Theorem 4.4.3 ([MeNa08, Theorems 1.9 and 1.11])**.** *Let Y be a Banach space with nontrivial type and let X be a Banach space which coarsely or uniformly embeds into Y. Then $q_X \leq q_Y$.*

Theorem 4.4.4. *Let M, N be Orlicz functions. If $\beta_M > 2$ and $\beta_N < \beta_M$, then h_M does not coarsely or uniformly embed into h_N.*

Proof. Assume first that h_N has nontrivial type. Since, by (4.2),

$$q_{h_M} = \max(2, \beta_M) > \max(2, \beta_N) = q_{h_N},$$

it follows from Theorem 4.4.3 that h_M does not coarsely or uniformly embed into h_N.

Now suppose that h_N does not have nontrivial type and suppose for the contradiction that h_M coarsely or uniformly embeds into h_N. Pick any $p \in (\beta_N, \beta_M)$. Then h_N strongly uniformly embeds into ℓ_p by Proposition 4.4.1, and therefore h_M coarsely or uniformly embeds into ℓ_p. But ℓ_p has nontrivial type (since $p > 1$) and its upper Matuszewska-Orlicz index is equal to $p < \beta_M$, which is in contradiction with the first part of the proof. □

Theorems 4.4.2 and 4.4.4 give almost complete classification of the coarse (uniform) embeddability between Orlicz sequence spaces. In the remaining cases, when $\beta_N \leq \beta_M = 2$ or $2 < \beta_M = \beta_N < \infty$, the situation is more complicated.

Let us now investigate the case when $\beta_N \leq \beta_M = 2$. We will show that in this case the coarse (uniform) embeddability of h_M into h_N is not determined by the values of β_M and β_N. More precisely, for any $1 \leq p \leq 2$ we can find Orlicz functions M_1, N_1, M_2, N_2 such that $\beta_{M_1} = \beta_{M_2} = 2$ and $\beta_{N_1} = \beta_{N_2} = p$, and such that h_{M_1} coarsely (uniformly) embeds into h_{N_1} and h_{M_2} does not coarsely (uniformly) embed into h_{N_2}. Of course, by Theorem 4.3.1, ℓ_2 strongly uniformly embeds into ℓ_p, providing thus examples of M_1 and N_1. Let us give examples of M_2 and N_2.

We will use the following theorem due to Johnson and Randrianarivony.

Theorem 4.4.5 ([JoRa06, Theorem 1])**.** *Let X be a Banach space with a normalized symmetric basis $(e_n)_{n=1}^{\infty}$ such that*

$$\liminf_{n\to\infty} \frac{1}{n^{\frac{1}{2}}} \left\| \sum_{i=1}^{n} e_i \right\| = 0.$$

Then X does not coarsely or uniformly embed into a Hilbert space.

This theorem was originally stated only for coarse embeddability; the statement about uniform embeddability follows from a result of Randrianarivony [Ran06, a paragraph before Theorem 1], who proved that a Banach space coarsely embeds into a Hilbert space if and only if it uniformly embeds into a Hilbert space.

Proposition 4.4.6. *Let M be an Orlicz function such that*

$$\lim_{t\to 0_+} \frac{M(t)}{t^2} = 0.$$

Then h_M does not coarsely or uniformly embed into ℓ_2.

Proof. We may suppose without loss of generality that $M(1) = 1$. Then the sequence of canonical vectors $(e_n)_{n=1}^{\infty}$ forms a normalized symmetric basis of h_M. Furthermore,

$$\begin{aligned} \left\| \sum_{i=1}^{n} e_i \right\| &= \inf\left\{ \rho > 0 : \sum_{i=1}^{n} M\left(\frac{1}{\rho}\right) \leq 1 \right\} = \inf\left\{ \rho > 0 : M\left(\frac{1}{\rho}\right) \leq \frac{1}{n} \right\} \\ &= \inf\left\{ \rho > 0 : \frac{1}{\rho} \leq M^{-1}\left(\frac{1}{n}\right) \right\} = \frac{1}{M^{-1}\left(\frac{1}{n}\right)}, \end{aligned}$$

and therefore

$$\frac{1}{n^{\frac{1}{2}}}\left\|\sum_{i=1}^{n} e_i\right\| = \frac{1}{n^{\frac{1}{2}}M^{-1}\left(\frac{1}{n}\right)}.$$

Let $t_n = M^{-1}\left(\frac{1}{n}\right)$. Then $t_n \to 0$ and $M(t_n) = \frac{1}{n}$, and therefore

$$\frac{1}{n^{\frac{1}{2}}M^{-1}\left(\frac{1}{n}\right)} = \frac{M(t_n)^{\frac{1}{2}}}{t_n} \xrightarrow{n\to\infty} 0,$$

since $\lim_{t\to 0_+} \frac{M(t)}{t^2} = 0$.

Hence

$$\liminf_{n\to\infty} \frac{1}{n^{\frac{1}{2}}}\left\|\sum_{i=1}^{n} e_i\right\| = 0,$$

and therefore, by Theorem 4.4.5, the space h_M does not coarsely or uniformly embed into ℓ_2. $\square$

Example 4.4.7. *There exists an Orlicz function M such that $\alpha_M = \beta_M = 2$ and h_M does not coarsely or uniformly embed into ℓ_p for any $1 \le p \le 2$.*

Proof. Let

$$f(t) = \frac{t^2}{1-\log t},\ t \in (0, \mathrm{e}).$$

Then using simple calculus we see that f is a continuous convex function, $f(t) > 0$ for each $t \in (0, \mathrm{e})$ and $\lim_{t\to 0_+} f(t) = 0$. Clearly there exists an Orlicz function M such that $M(t) = f(t)$ for every $t \in (0, 1]$.

Let us show that $\alpha_M = \beta_M = 2$. Let $q \le 2$ and $\lambda, t \in (0, 1]$. Then

$$\frac{M(\lambda t)}{M(\lambda)t^q} = \frac{\frac{(\lambda t)^2}{1-\log(\lambda t)}}{\frac{\lambda^2}{1-\log\lambda}t^q} = t^{2-q}\frac{1-\log\lambda}{1-\log(\lambda t)} \le t^{2-q} \le 1$$

(the first inequality follows from the fact that $s \mapsto 1 - \log s$ is decreasing), and therefore $\alpha_M \ge 2$.

Let $q > 2$. If $\lambda, t \in (0, 1]$, then

$$\frac{M(\lambda t)}{M(\lambda)t^q} = t^{2-q}\frac{1-\log\lambda}{1-\log(\lambda t)} = t^{2-q}\frac{1-\log\lambda}{1-\log\lambda-\log t} = \frac{t^{2-q}}{1+\frac{-\log t}{1-\log\lambda}} \ge \frac{t^{2-q}}{1-\log t},$$

where the inequality holds since $-\log t \ge 0$ and $1 - \log\lambda \ge 1$. Now if we define $g(s) = \frac{s^{2-q}}{1-\log s}$, $s \in (0, 1]$, then $\lim_{s\to 0_+} g(s) = \infty$, $g(1) = 1$ and $g(s) > 0$ for each $s \in (0, 1]$. It follows that there is $C > 0$ such that $g(s) \ge C$ for each $s \in (0, 1]$. Hence

$$\frac{M(\lambda t)}{M(\lambda)t^q} \ge C$$

for every $\lambda, t \in (0, 1]$. This implies that $\beta_M \le 2$.

Finally, if $t \in (0, 1]$, then

$$\frac{M(t)}{t^2} = \frac{\frac{t^2}{1-\log t}}{t^2} = \frac{1}{1-\log t} \xrightarrow{t\to 0_+} 0.$$

Hence, by Proposition 4.4.6, h_M does not coarsely or uniformly embed into ℓ_2. Let $1 \leq p < 2$. Since ℓ_p strongly uniformly embeds into ℓ_2 by Theorem 4.1.1, it follows that h_M does not coarsely or uniformly embed into ℓ_p. □

The last remaining case is when $2 < \beta_M = \beta_N < \infty$. In this case, we can of course always have the coarse (uniform) embeddability (since any Banach space strongly uniformly embeds into itself). However, we do not know whether there exist Orlicz functions M, N satisfying $2 < \beta_M = \beta_N < \infty$, such that h_M does not coarsely (uniformly) embed into h_N.

Let us conclude with a brief summary of the results. Let M, N be Orlicz functions.

(1) If $\beta_M < \beta_N$ or $\beta_N \leq \beta_M < 2$ or $\beta_M = \beta_N = \infty$, then h_M strongly uniformly embeds into h_N.

(2) If $\beta_M > 2$ and $\beta_N < \beta_M$, then h_M does not coarsely or uniformly embed into h_N.

(3) If $\beta_N \leq \beta_M = 2$, then the coarse (uniform) embeddability of h_M into h_N is not determined by the values of β_M and β_N.

(4) If $2 < \beta_M = \beta_N < \infty$, then the question of the coarse (uniform) embeddability of h_M into h_N is open.

Acknowledgments. The author would like to thank Gilles Lancien for many inspiring discussions and helpful comments on this paper. He would also like to thank Tony Procházka for several valuable conversations. The author is grateful to F. Baudier for pointing out to him Albiac's construction of embeddings between ℓ_p-spaces, and to both F. Albiac and F. Baudier for providing him with their manuscript [AlBa12].

This work was completed during the author's stay at Université de Franche-Comté in Besançon. The author wishes to thank the staff for their hospitality.

This research was partially supported by the grant PHC Barrande 2012-26516YG.

Bibliography

[Aha74] I. Aharoni, *Every separable metric space is Lipschitz equivalent to a subset of* c_0, Israel J. Math. 19 (1974), 284–291.

[Alb08] F. Albiac, *Nonlinear structure of some classical quasi-Banach spaces and F-spaces*, J. Math. Anal. Appl. 340 (2008), 1312–1325.

[AlBa12] F. Albiac and F. Baudier, *Embeddability of snowflaked metrics with applications to the Lipschitz geometry of the spaces* L_p *and* ℓ_p *for* $0 < p < \infty$, preprint.

[Bač09] M. Bačák, *Point simpliciality in Choquet representation theory*, Illinois J. Math. 53 (2009), 289–302.

[BeLi00] Y. Benyamini and J. Lindenstrauss, *Geometric Nonlinear Functional Analysis*, Coll. Pub. 48, Amer. Math. Soc., Providence, RI, 2000.

[BiLe59] E. Bishop and K. de Leeuw, *The representations of linear functionals by measures on sets of extreme points*, Ann. Inst. Fourier (Grenoble) 9 (1959), 305–331.

[BoHi31] H.F. Bohnenblust and E. Hille, *On the absolute convergence of Dirichlet series*, Ann. of Math. 32 (1931), 600–622.

[BoRo90] C. de Boor and A. Ron, *On multivariate polynomial interpolation*, Constr. Approx. 6 (1990), 287–302.

[Bor10a] L. Borel-Mathurin, *Isomorphismes non linéaires entre espaces de Banach*, Ph.D. thesis, 2010.

[Bor10b] L. Borel-Mathurin, *The Szlenk index of Orlicz sequence spaces*, Proc. Amer. Math. Soc. 138 (2010), 2043–2050.

[Car64] S.O. Carlsson, *A characteristic property of Euclidean spaces*, Ark. Mat. 5 (1964), 327–330.

[ChuYa77] K.C. Chung and T.H. Yao, *On lattices admitting unique Lagrange interpolations*, SIAM J. Numer. Anal. 14 (1977), 735–741.

[Day47] M.M. Day, *Some characterizations of inner product spaces*, Trans. Amer. Math. Soc. 62 (1947), 320–337.

[Day59] M.M. Day, *On criteria of Kasahara and Blumenthal for inner product spaces*, Proc. Amer. Math. Soc. 10 (1959), 92–100.

[Die75] J. Diestel, *Geometry of Banach Spaces — Selected Topics*, Lecture Notes in Math., vol. 485, Springer-Verlag, Berlin, Heidelberg, 1975.

[Din99] S. Dineen, *Complex Analysis on Infinite Dimensional Spaces*, Springer Monographs in Mathematics, London, Springer Verlag, 1999.

[DLS06] P. Dostál, J. Lukeš and J. Spurný, *Measure convex and measure extremal sets*, Canad. Math. Bull. 49 (2006), 536–548.

[FHHMZ11] M. Fabian, P. Habala, P. Hájek, V. Montesinos, and V. Zizler, *Banach Space Theory. The Basis for Linear and Nonlinear Analysis*, Canadian Math. Soc. Books in Mathematics, Springer-Verlag, New York (2011).

[Fré09a] M. Fréchet, *Toute fonctionelle continue est développable en série de fonctionelles d'order entier*, C. R. Acad. Sci. Paris 148 (1909), 155–156.

[Fré09b] M. Fréchet, *Une definition fonctionelle des polynomes*, Nouv. Ann. Math. 9 (1909), 145–162.

[Gil67] J.R. Giles, *Classes of semi-inner-product spaces*, Trans. Amer. Math. Soc. 129 (1967), 436–446.

[Joh73] G.G. Johnson, *Inner products characterized by difference equations*, Proc. Amer. Math. Soc. 37 (1973), 535–536.

[JoRa06] W.B. Johnson and N.L. Randrianarivony, ℓ_p $(p > 2)$ *does not coarsely embed into a Hilbert space*, Proc. Amer. Math. Soc. 134 (2006), 1045–1050.

[JoNe35] P. Jordan and J. von Neumann, *On inner products in linear metric spaces*, Ann. of Math. 36 (1935), 719–723.

[Kal08] N.J. Kalton, *The nonlinear geometry of Banach spaces*, Rev. Mat. Complut. 21 (2008), no. 1, 7–60.

[Kal12] N.J. Kalton, *Uniform homeomorphisms of Banach spaces and asymptotic structure*, to appear in Trans. Amer. Math. Soc.

[Koe70] D.O. Koehler, G_{2n} *spaces*, Trans. Amer. Math. Soc. 150 (1970), 507–518.

[Koe72] D.O. Koehler, *The analytic properties of* G_{2n} *spaces*, Proc. Amer. Math. Soc. 35 (1972), 201–206.

[Köh70] J. Köhn, *Barycenters of unique maximal measures*, J. Funct. Anal. 6 (1970), 76–82.

[KoKö01] A. Koldobsky and H. König, *Aspects of Isometric Theory of Banach Spaces*, Chap. 21, Handbook of the Geometry of Banach Spaces, vol. I, Elsevier (2001).

[LiTz77] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces I*, Springer-Verlag, 1977.

[LiTz79] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces II*, Springer-Verlag, 1979.

[Lor48] E.R. Lorch, *On certain implications which characterize Hilbert spaces*, Ann. of Math. 49 (1948), 523–532.

[Lore92] R.A. Lorentz, *Multivariate Birkhoff Interpolation*, Lecture Notes in Math. 1516, (1992).

[LMNS10] J. Lukeš, J. Malý, I. Netuka and J. Spurný, *Integral Representation Theory. Applications to Convexity, Banach Spaces and Potential Theory*, Walter de Gruyter, Berlin, New York, 2010.

[MaRa10] M. Martín and T.S.S.R.K. Rao, *On remotality for convex sets in Banach spaces*, J. Approx. Theory 162 (2010), 392–396.

[MaOr34a] S. Mazur and W. Orlicz, *Grundlegende Eigenschaften der polynomischen Operationen I*, Studia Math. 5 (1934), 50–68.

[MaOr34b] S. Mazur and W. Orlicz, *Grundlegende Eigenschaften der polynomischen Operationen II*, Studia Math. 5 (1934), 179–189.

[MeNa08] M. Mendel and A. Naor, *Metric cotype*, Ann. of Math. (2) 168 (2008), 247–298.

[Ney84] A. Neyman, *Representation of L_p-norms and isometric embedding in L_p-spaces*, Israel J. Math. 48 (1984), 129–138.

[Now05] P.W. Nowak, *Coarse embeddings of metric spaces into Banach spaces*, Proc. Amer. Math. Soc. 133 (2005), 2589–2596.

[Now06] P.W. Nowak, *On coarse embeddability into ℓ_p-spaces and a conjecture of Dranishnikov*, Fund. Math. 189 (2006), 111–116.

[Phe01] R.R. Phelps, *Lectures on Choquet's Theorem, 2nd ed.*, Springer, Berlin, Heidelberg, 2001.

[Ran06] N.L. Randrianarivony, *Characterization of quasi-Banach spaces which coarsely embed into a Hilbert space*, Proc. Amer. Math. Soc. 134 (2006), 1315–1317.

[Rao09] T.S.S.R.K. Rao, *Remark on a paper of Sababheh and Khalil*, Numer. Funct. Anal. Optim. 30 (2009), 822–824.

[Rez78] B. Reznick, *Banach spaces which satisfy linear identities*, Pacific J. Math. 74 (1978), 221–233.

[Rez79] B. Reznick, *Banach spaces with polynomial norms*, Pacific J. Math. 82 (1979), 223–235.

[Rez93] B. Reznick, *Sums of even powers of real linear forms*, Mem. Amer. Math. Soc. 463 (1993).

[SaKh08] M. Sababheh and R. Khalil, *Remotality of closed bounded convex sets in reflexive spaces*, Numer. Funct. Anal. Optim. 29 (2008), 1166–1170.

[Sen68] D.A. Senechalle, *A characterization of inner product spaces*, Proc. Amer. Math. Soc. 19 (1968), 1306–1312.

[Sim70] S. Simons, *Minimal sublinear functionals*, Studia Math. 37 (1970), 37–56.

[Ves09] L. Veselý, *Convex sets without diametral pairs*, Extracta Math. 24(3) (2009), 271–280.

[Wil18] W.H. Wilson, *On a certain general class of functional equations*, Amer. J. Math. 40 (1918), 263–282.