Finite Metric subsets of Banach Spaces

James Kilbane

Corpus Christi College

July 2018

This dissertation is submitted for the degree of

Doctor of Philosophy

DPMMS

University of Cambridge

As long as inequalities persist
in our world,
none of us can truly rest.

Dedication

Without my parents I would not be here.

Declaration

This dissertation is my own work and includes nothing which is the outcome of work done in collaboration, except as detailed below and where specifically noted.

Chapter 1 gives background for the thesis and is primarily classical, but Section 1.4 introduces a few new notions. Chapter 2 is primarily my own work. The main positive result of Chapter 3 was originally proven in [40] and here I give a different proof. The remainder of the chapter is original work. The first half of Chapter 4 (up to Section 4.5) is solo work presented in the paper [23], published in Proceedings of the American Mathematical Society. The second half of Chapter 4 was joint work with Mikhail Ostrovskii, published in the Houston Journal of Mathematics.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text

Acknowledgements

First and foremost, I would like to take the opportunity to thank my parents for their unwavering support over the past twenty something years. This has been a long road, and without their support I would not be here.

For the mathematical acknowledgements, I will begin with my supervisor András Zsák. Without him this thesis would either not exist, or (perhaps worse) be difficult to read. He has been a wonderful source of help, ideas and inspiration over the past few years. I also thank Mikhail Ostrovskii. He pointed out the question that led to the fourth section and then collaborated with me on the second part of this. He also answered my first mathoverflow question on this topic, where he pointed out a fact that did not make it into the logical flow of the thesis, but I'm happy to know is true.

For my friends, there are many, all of whom contributed in their own way. Without Samuel Edis this thesis (and my 'mathematical life') would never have slid into place. Without Alexandra Paivana, this thesis would exist in a state of 95% completion and have been submitted this way.[1] Out of my friends from home, I wish to thank Patrick Boss, Aaron Barber and James Doran for putting up with me in a way not many would. My friends in Cambridge included Nigel Burke, Benjamin Barrett, Tom Berrett, Lawrence Barrott, Jack Smith, Jo Evans and Nicolas Dupré, all of whom have influenced this thesis and my life in a positive way. As to my many friends online, I wish to take a moment to thank them: Chris, Wolf, Ferdia, Basti, Rob, Jeremy, Andy, Andres, Mike, Kieran and Ross/Jay. My online friends have given me support over the past few years in a way that has certainly kept me sane[2] through this process.

Here endeth the humorous footnotes.

[1] The examiners may believe that this is still the case.

[2] The examiners may believe that this is not the case.

Abstract

The central idea in this thesis is the introduction of a new isometric invariant of a Banach space. This is Property AI-I. A Banach space has Property AI-I if whenever a finite metric space almost-isometrically embeds into the space, it isometrically embeds. To study this property we introduce two further properties that can be thought of as finite metric variants of Dvoretzky's Theorem and Krivine's Theorem. We say that a Banach space satisfies the Finite Isometric Dvoretzky Property (FIDP) if it contains every finite subset of ℓ_2 isometrically. We say that a Banach space has the Finite Isometric Krivine Property (FIKP) if whenever p is such that ℓ_p is finitely representable in the space then it contains every subset of ℓ_p isometrically. We show that every infinite-dimensional Banach space *nearly* has FIDP and every Banach space *nearly* has FIKP. We then use convexity arguments to demonstrate that not every Banach space has FIKP, and thus we can exhibit classes of Banach spaces that fail to have Property AI-I. The methods used break down when one attempts to prove that there is a Banach space without the FIDP and we conjecture that every infinite-dimensional Banach space has the FIDP.

Contents

Introduction

One of the oldest themes in Banach space theory is to attempt to find structural properties that are true either for all Banach spaces, or all Banach spaces within a certain class. Over time, research interests have changed from finding *linear* structural properties, i.e., properties that are dependent upon *subspaces*, to finding *metric* structural properties, i.e., properties that are dependent upon *subsets*.

We will begin by introducing three metric properties of a Banach space. These metric properties are the key in understanding the motivation of the work undertaken here. The results we obtain are limited versions of 'compactness' that work for Banach spaces that exist in certain classes.

We will begin with an overview of these three properties - any notation and terminology will be defined later.

Property AI-I

The first metric property we consider is a very general one. We say that a Banach space X has *Property AI-I* (AI-I stands for "almost isometric to isometric") if whenever a finite metric space almost-isometrically embeds into X it isometrically embeds into X.

More specifically, we say that a Banach space X has Property AI-I if whenever M is a finite metric space such that for all $\epsilon > 0$ there is a $(1+\epsilon)$-distortion embedding of M into X, M isometrically embeds into X.

This attempts to capture a compactness principle - the way one would want to create such an isometric embedding is by considering $(1+1/n)$-distortion embeddings and then passing to a subsequence to construct an isometry. There is no naive way of doing this in infinite dimensions, so we have to use more complicated arguments.[3]

It should be relatively clear to the reader that the spaces ℓ_2 and ℓ_∞ have Property AI-I. It is trivial that ℓ_∞ has Property AI-I, indeed, ℓ_∞ contains every finite metric space isometrically. For ℓ_2, we can observe (by a Gram-Schmidt argument) that every n-point subset of ℓ_2 is isometric to an n-point subset of ℓ_2^n, and we can thus use the limiting procedure detailed above.

Our general question is whether a Banach space X has Property AI-I, ideally finding some specific 'nice' linear property that is equivalent to it (or even implies it.) The generality of Property AI-I makes this a hard task. We can restrict the generality of Property AI-I to ask easier questions, and we can restrict Property AI-I in two ways.

The first is that we could restrict to certain classes of Banach spaces. This results in questions of the type: Does every uniformly convex Banach space have Property AI-I? Does every Banach space with a 'very nice' basis have Property AI-I?

The second is that we could restrict the class of metric spaces under consideration. If $\mathcal{U}$ is a class of metric spaces, we could restrict the definition of Property AI-I to only metric spaces in this class, i.e., a Banach space X has Property AI-I($\mathcal{U}$) if whenever a finite metric space in $\mathcal{U}$ almost-isometrically embeds into X then it embeds isometrically into X. Evidently, for a space to have property AI-I, it must have Property AI-I($\mathcal{U}$) for any class of metric spaces $\mathcal{U}$.

We consider two classes $\mathcal{U}$ throughout our work. These classes (which we detail below) are chosen due to their interaction with the linear properties of a Banach space. A careful analysis of these two properties will give certain spaces that fail Property AI-I. Whenever our work on these properties yields information about Property AI-I we will point it out to the reader.

A Finite Isometric Dvoretzky Property

Suppose that X is an infinite-dimensional Banach space. By Dvoretzky's theorem, for any finite subset M of ℓ_2 there is a $(1+\epsilon)$-distortion embedding of M into X. We say that a Banach space X has the Finite Isometric Dvoretzky Property, or X has FIDP for short, if it contains every finite subset of ℓ_2 isometrically.

[3] In finite dimensions we iteratively pass to norm convergent sequences, and this gives us our isometry. In any dual Banach space we have weak-star convergence, however this does not help to construct isometries.

This is a restriction of Property AI-I that is of interest in all infinite-dimensional Banach spaces. We also have restrictions of Property AI-I that are of primary interest in Banach spaces lying in some fixed class.

A Finite Isometric Krivine Property

Suppose that X is an infinite-dimensional Banach space. There is some p such that ℓ_p is finitely represented in X (by Krivine's theorem). Then for any finite subset M of ℓ_p there is a $(1+\epsilon)$-distortion embedding of M into X. We say that a Banach space has the Finite Isometric Krivine Property, or X has $FIKP$ for short, if it contains every finite subset of ℓ_p isometrically

We note that the three properties above are all about *finite* subsets of Banach spaces. At various points throughout the text, we will have cause to mention infinite discrete subsets of a Banach space. This is not necessarily the main goal of our work, however, our methods will sometimes give us insights into infinite discrete subsets of a Banach space.

Structure of the Text

Chapter 1 of this work sets up the stage. We introduce the ideas necessary to understand the arguments given in the thesis, and we build up the necessary notation and terminology. We lean in the direction of giving more rather than fewer details.

Chapter 2 discusses the case of ℓ_∞, namely, it aims to discuss what is known about the question "If a Banach space X contains every finite dimensional Banach space isomorphically does it contain every finite metric space isometrically?" We show that such Banach spaces contain 'most' finite metric spaces isometrically, but they do not necessarily contain all finite metric spaces isometrically.

Chapter 3 discusses the case of ℓ_2, namely, it aims to discuss what is known about the question "If X is an infinite-dimensional Banach space, does it contain every subset of ℓ_2 isometrically?" We prove a positive result for affinely independent subsets of ℓ_2 then discuss what is known about affinely dependent subsets.

Chapter 4 discusses the case of ℓ_p, namely, it aims to discuss what is known about the question "If X is an infinite-dimensional Banach space that contains ℓ_p^n almost-isometrically for each n, then does it contain every subset of ℓ_p isometrically?"

We prove a positive result for 'almost all' finite subsets of ℓ_p and give a negative result in general.

Chapter 1

Definitions, Notation and Known Results

Our central objects of study are Banach spaces and metric spaces. This introduction is split into three parts: the first two parts are about terminology and known results for Banach spaces and metric spaces respectively. The third is an assorted collection of results that will be needed throughout.

Throughout this thesis we will be interested in embedding finite metric spaces into Banach spaces. Infinite-dimensional Banach spaces throughout will be labelled $X, Y, Z, \ldots$, with norms $\|.\|$ (or $\|.\|_X$ if the dependence on X is needed), finite-dimensional Banach spaces will be labelled $E, F, \ldots$ and metric spaces will be labelled $M, N, \ldots$ with distances d (or d_M if the dependence on M is needed.)

Points in X will be labelled $x_1, x_2, x_3, \ldots$, points in Y will be labelled $y_1, y_2, y_3, \ldots$ and moreover x and y *will only ever refer to elements of Banach spaces*. Points in M will be labelled as $m_1, m_2, m_3, \ldots$ and points in N will be labelled as $r_1, r_2, r_3 \ldots$ (generally the letter n will be used for dimension or sizes.) m and r *will only ever refer to elements of metric spaces.*

If we refer to a matrix, we will always refer to matrices with calligraphic font, i.e., matrices will be $\mathcal{M}$ or $\mathcal{P}$ (since M is a metric space!)

Functions will either have long names that are obviously functions (e.g.., see Section 1.2.4 for the function "CMDet"), or if we need to use a specific letter for a function we will use Greek letters, i.e., functions will thus be things like Θ or Φ. If we are considering continuous linear functions between Banach spaces, we will use $S, T, \ldots$.

1.1 Banach Spaces

Throughout this thesis, Banach spaces will normally be considered over $\mathbb{R}$: if we are considering a Banach space over $\mathbb{C}$ we will mention it. We will use various standard notation from the theory of Banach spaces including the following:

- $B_X = \{x \in X : \|x\| \leq 1\}$ is the closed unit ball of X
- $S_X = \{x \in X : \|x\| = 1 :\}$ is the unit sphere of X
- If $T : X \to Y$ is a bounded linear map between Banach spaces X and Y we define $\|T\| = \sup\{\|Tx\| : x \in B_X\}$. This norm turns $\mathcal{L}(X,Y)$, the space of all continuous linear maps between X and Y, into a Banach space.
- If X and Y are Banach spaces with dimension n, m respectively we identify $\mathcal{L}(X,Y)$ with $\text{Mat}_{n\times m}(\mathbb{R})$, the set of all $n \times m$ real matrices.
- If $(X, \|.\|)$ is a Banach space, and we equip it with a new norm $\|.\|'$, we say that $\|.\|$ and $\|.\|'$ are *equivalent* if there are strictly positive constants α, β such that $\alpha\|.\| \leq \|.\|' \leq \beta\|.\|$. If $\frac{\beta}{\alpha} \leq C$ we say that $\|.\|$ and $\|.\|'$ are *C-equivalent*.

A subset A of a Banach space X is *convex* if whenever $x, y \in A$ we have that $\lambda x + (1-\lambda)y \in X$ for all $\lambda \in (0,1)$. We note that B_X is convex. We say that A is an *$m-$dimensional convex set* if it is contained in some affine subspace of dimension m, and it is not contained in any affine subspace of strictly smaller dimension.

The space ℓ_p,[1] for $1 < p < \infty$ is the space of real sequences $(x_i)_{i\in\mathbb{N}}$ such that $\sum_{i=1}^{\infty} |x_i|^p < \infty$ and we equip ℓ_p with the norm $\|x\|_p = (\sum_{i=1}^{\infty} |x_i|^p)^{1/p}$. We will also be interested in the finite-dimensional variants of this, ℓ_p^n, which are vectors in $\mathbb{R}^n$ with norm $\|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p}$. The corresponding space ℓ_∞ is the space of bounded real sequences with norm $\|x\| = \sup_i |x_i|$, and ℓ_∞^n is defined analogously. We define c_0 as the space of real sequences tending to zero, and observe this is a closed subspace of ℓ_∞ in the $\|.\|_\infty$ norm, and thus $(c_0, \|.\|_\infty)$ is a Banach space. We define c_{00} to be those vectors of c_0 with finite support.

In ℓ_p we will let the vector e_i be the vector $(0, \ldots, 0, 1, 0, \ldots)$ where the i'th co-ordinate is 1 and every other co-ordinate is zero. We note that ℓ_p is the closure in the ℓ_p norm of c_{00} and c_0 is the closure in the ℓ_∞ norm of c_{00}.

[1]We note that we take ℓ_p with the subscript *at the bottom*. It is sometimes the case that people write ℓ^p instead. We prefer to keep the subscript at the bottom because we will be using a superscript n to denote dimensionality.

For $1 \leq p < \infty$ the dual of ℓ_p is canonically isometrically isomorphic to $\ell_{p'}$ where for $1 < p < \infty$ we have that $\frac{1}{p} + \frac{1}{p'} = 1$, and for $p = 1$ we have that $p' = \infty$. p' is called the *conjugate index* of p. The mapping from $\ell_{p'}$ to ℓ_p^* is given by the mapping $y \mapsto f_y$ where $f_y(x) = \sum_{n=1}^{\infty} y_n x_n$. Similarly we have c_0^* is canonically isometrically isomorphic to ℓ_1. However, ℓ_∞^* is *not* canonically isometrically isomorphic to ℓ_1.

ℓ_2 is a Hilbert space, and we shall use the language of basic Hilbert space theory throughout. The inner product is defined as $\langle x, y \rangle = \sum_{n=1}^{\infty} x_n y_n$.

If X and Y are two Banach spaces, we will write $X \oplus_p Y$ for the p-direct sum of X and Y. This is the space $\{(x,y) : x \in X, y \in Y\}$ with norm $\|(x,y)\| = (\|x\|^p + \|y\|^p)^{1/p}$, and we can show that it is a Banach space with this norm. If X_n is a sequence of Banach spaces we will write $\left(\oplus_{n=1}^{\infty} X_n\right)_p$ for the infinite p-direct sum of the X_n's which is the space

$$\{(x_n)_{n=1}^{\infty} : x_n \in X_n \text{ and } \| \left(\|x_n\|_{X_n}\right)_{n=1}^{\infty} \|_p \text{ is finite}\}$$

with norm $\|(x_n)_{n=1}^{\infty}\| = \| \left(\|x_n\|_{X_n}\right)_{n=1}^{\infty} \|_p$. One can show that this is a Banach space.

Correspondingly one can define the c_0 direct sum of Banach spaces X_n. One writes $\left(\oplus_{n=1}^{\infty} X_n\right)_{c_0}$ for the space $\{(x_n)_{n=1}^{\infty} : x_n \in X_n \text{ and } \|x_n\|_{X_n} \to 0\}$. Similar to the above this is a Banach space.

Basic results on the ℓ_p spaces, and basic results on the Hilbertian structure of ℓ_2, including things such as proofs that ℓ_p is a Banach space with the norm given, Hölder's inequality, Minkowski's inequality, the duality correspondence stated above, ... can be found in any introductory text on functional analysis, for example, see [5].

The Banach–Mazur distance will also be required throughout. If X and Y are Banach spaces, we let

$$d(X,Y) = \inf\{\|T\|\|T^{-1}\| : T \text{ is an isomorphism from } X \text{ to } Y\}.$$

We say that $d(X,Y) = \infty$ if no isomorphism exists between X and Y. It is clear that if two Banach spaces are isometrically isomorphic then $d(X,Y) = 1$. The converse is not true, i.e., $d(X,Y) = 1$ is *not* enough for us to say that X and Y are isometrically isomorphic. Indeed, all $d(X,Y) = 1$ means is that for any $\epsilon > 0$ there is an isomorphism T from X to Y such that $\|T\|\|T^{-1}\| < 1 + \epsilon$. We will see an example of this phenomenon as Theorem 1.1.19.

It is well known that whenever E and F are two Banach spaces of the same dimension the Banach–Mazur distance between E and F is less than infinity, i.e., any two Banach spaces of the same dimension are isomorphic. A slightly less well known fact is that whenever E and F are n-dimensional, we have that $d(E, F) \leq n$, see for example [5, Chapter 4].

If X is a Banach space and $(x_i)_{i=1}^{\infty}$ is a sequence in X, we say that $(x_i)_{i=1}^{\infty}$ is a *basic sequence* if for every element y in $\overline{span}\{x_i : i \in \mathbb{N}\}$ there are unique scalars $(\alpha_i)_{i=1}^{\infty}$ such that $y = \sum_{i=1}^{\infty} \alpha_i y_i$. If we choose an element y of $\overline{span}\{x_i : i \in \mathbb{N}\}$ we say that the *support of* y is the set $\{i \in \mathbb{N} : \alpha_i \neq 0\}$ where $y = \sum_{i=1}^{\infty} \alpha_i x_i$. We say $(y_i)_{i=1}^{n}$ is a *blocking* of the x_i's if whenever $i < j$, k is in the support of y_i and l is in the support of y_j then $k < l$.

If X and Y are Banach spaces, we say that $X \overset{C}{\hookrightarrow} Y$ if there is a mapping $T : X \to Y$ such that $\|x\| \leq \|Tx\| \leq C\|x\|$. We say that Y almost isometrically contains X if $X \overset{1+\epsilon}{\hookrightarrow} Y$ for each $\epsilon > 0$. We write $X \hookrightarrow Y$ if $X \overset{C}{\hookrightarrow} Y$ for some C.

If X_n is a collection of Banach spaces and Y is a fixed target space, we say that the *X_n's embed uniformly into Y*, or Y *uniformly contains* X_n, if there is some $C > 0$ such that $X_n \overset{C}{\hookrightarrow} Y$ for all n. We will say that Y *almost isometrically contains* X_n if the constant C in the above can be chosen to be $1 + \epsilon$ for any $\epsilon > 0$.

If X is a Banach space, then B_X is compact if and only if X is finite-dimensional.

1.1.1 Some Remarks on ℓ_p

Among the ℓ_p spaces, ℓ_∞ and ℓ_1 have certain universality properties. The first of these we will need to use later, and thus include a proof.

Lemma 1.1.1. *Every separable Banach space embeds isometrically into ℓ_∞.*

This result is the beginning of the theory of universal spaces - we say that the space ℓ_∞ is isometrically universal for the class of separable Banach spaces.

Proof. Since X is separable we can take some sequence $x_1, x_2, \ldots$ such that $\overline{\{x_n\}} = X$. Use the Hahn-Banach theorem to choose functionals x_i^* such that $x_i^*(x_i) = \|x_i\|$ and $\|x_i^*\| = 1$. Then the map $T :\to \ell_\infty$ given by $T(x) = (x_i^*(x))_{i=1}^{\infty}$ is the required isometric embedding.

Indeed, since the x_i^*'s have unit modulus it is clear that $\|T(x)\| \leq \|x\|$. By choosing a sequence n_j such that $x_{n_j} \to x$ and letting j tend to infinity in the inequality

$\|T(x)\| \geq |x^*_{n_j}(x)|$, we obtain that $\|T(x)\| = \|x\|$. $\square$

The second of these we include for interest, and will not include a proof of.

Lemma 1.1.2 (Unproved)**.** *Every separable Banach space is isomorphic to a quotient of* ℓ_1.

It is a classical fact that all of the ℓ_p spaces, for $1 \leq p \leq \infty$, together with c_0, are all mutually nonisomorphic. This can be seen either by the results of Section 1.1.6, or by Pitt's Theorem:

Theorem 1.1.3. *[[1, Chapter 2]] If $T : \ell_p \to \ell_q$ is a bounded linear map, with $1 \leq q < p < \infty$, then T is compact, i.e., $\overline{T(B_{\ell_p})}^{\ell_q}$ is compact in ℓ_q.*

This result (which we shall not prove) indicates that $d(\ell_p^n, \ell_q^n) \to \infty$ as $n \to \infty$. Later on, however, we will be interested in estimates on $d(\ell_p^n, \ell_q^n)$:

Lemma 1.1.4. *If $1 \leq p \leq q \leq \infty$ we have that $d(\ell_p^n, \ell_q^n) \leq n^{1/p-1/q}$ (where if $q = \infty$ we interpret $1/q$ as zero.)*

Proof. We consider the identity map $T : \ell_p^n \to \ell_q^n$, i.e., $T(x_1, \ldots, x_n) = (x_1, \ldots, x_n)$ and wish to estimate the norm of this map.

First we give an upper estimate. Note that

$$\sum_{i=1}^n |x_i|^p \leq \left(\sum_{i=1}^n 1\right)^{1-p/q} \left(\sum_{i=1}^n |x_i|^q\right)^{p/q}$$

by an application of Hölder's inequality with exponent q/p (since $q > p$ this exponent is larger than 1.) Raising both sides to the power of $1/p$ we see that $\|x\|_p \leq n^{1/p-1/q}\|x\|_q$.

For a lower estimate, suppose that $\sum_{i=1}^n |x_i|^p = 1$. Then

$$1 = \sum_{i=1}^n |x_i|^p \geq \sum_{i=1}^n |x_i|^q$$

(since $q > p$ and every $|x_i| \leq 1$). Raising both sides to the power of $1/q$ shows that $\|x\|_q \leq 1$. Homogeneity now shows that $\|x\|_p \geq \|x\|_q$.

These two arguments show that $\|T\| \cdot \|T^{-1}\| \leq n^{1/p-1/q}$. $\square$

Remark 1.1.5. A great deal more can be said about this bound. The interested reader can see Chapter 1 of the book [20] for more information. In the case $1 \leq p, q \leq 2$ or $2 \leq p, q \leq \infty$ the bound given *is* tight. However, if $1 \leq p \leq 2 \leq q \leq \infty$ we get that $d(\ell_p^n, \ell_q^n)$ is bounded above by $C \max(n^{1/p-1/2}, n^{1/2-1/q})$, where C is some absolute constant whose exact value is unknown. An example of this phenomenon is that the real Banach spaces ℓ_1^2 and ℓ_∞^2 are isometrically isomorphic, with the map $T : \ell_\infty^2 \to \ell_1^2$ given by

$$(x,y) \mapsto \left(\frac{x+y}{2}, \frac{x-y}{2}\right).$$

In ℓ_2 we have the property that any n-point subset of ℓ_2 embeds isometrically into ℓ_2^n. Indeed, if we have any $x_1, \ldots, x_n$ in ℓ_2 we may perform the Gram–Schmidt procedure on $x_1, \ldots, x_n$ to find orthonormal vectors $e_1, \ldots, e_j$ (with j possibly different from n) such that x_i lies in the span of $\{e_1, \ldots, e_j\}$. A different argument works in ℓ_∞ (see Section 1.2.3) and shows that any n-point subset of ℓ_∞ embeds into ℓ_∞^n.

One may hope that the same is true for ℓ_p, i.e., that for any n-point subset M of ℓ_p there is an isometric embedding of M into ℓ_p^n. However, this is not true, a result in the paper [3] shows that there is an n-point subset of ℓ_1 that embeds into ℓ_1^m just when $m \geq \binom{n-2}{2}$. However, we have the following theorem:

Theorem 1.1.6 ([3, Proposition 1])**.** *Suppose M is an n-point subset of ℓ_p. Then M is isometric to a subset of $\ell_p^{\binom{n}{2}}$.*

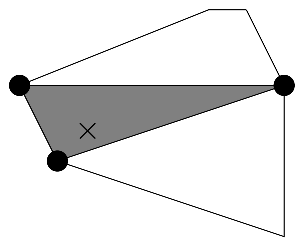
Remark 1.1.7. This result is true more generally, in fact, if M is an n-point subset of any $\mathcal{L}_p$ space (of which both ℓ_p and L_p form examples) then M is isometric to an n-point subset of $\ell_p^{\binom{n}{2}}$.

The result here is key and will inform certain definitions and choices made in the sequel. Theorem 1.1.6 gives us a mapping Θ that associates to any n-point subset of ℓ_p an n-point subset of $\ell_p^{\binom{n}{2}}$. In the sequel it is of relevance to know, say, whether the mapping Θ is continuous. This issue is discussed more completely in a remark following Lemma 4.3.1.

To prove Theorem 1.1.6 we will need to use Carathéodory's Theorem, which we recall here:

Theorem 1.1.8. *Suppose that V is a real vector space and $A \subset V$ is an $m-$dimensional convex set. If there is a collection $\{x_\gamma : \gamma \in \Gamma\}$ such that $A \subset conv(\{x_\gamma : \gamma \in \Gamma\})$ then for any point $x \in A$ there exist $\gamma_1, \ldots, \gamma_{m+1}$ and real numbers $\lambda_1, \ldots, \lambda_n \geq 0$ such that $\sum_{i=1}^m \lambda_i = 1$ and $\sum_{i=1}^m \lambda_i x_{\gamma_i} = x$.*

The proof of this statement is a relatively simple induction and can be found in, say, [12, Section 1].



Caratheodory's theorem in this example says that any point in the shape is contained in a triangle whose corners are three vertices.

The proof of Theorem 1.1.6 contains a lot of notation and ideas that are revisited throughout this section.

Proof of Theorem 1.1.6. We will prove this theorem in the case of ℓ_p - this is the only case we need. We consider the set $M_n = \{(\|x_i - x_j\|^p)_{i,j=1}^n : (x_i)_{i=1}^n$ are points in $\ell_p\}$. We will call an element of M_n *linear* if the associated set $x_1, \ldots, x_n$ can be isometrically embedded into $\mathbb{R}$. Let K denote the intersection of U_n with the compact convex set $\{\mathcal{A} \in \mathrm{Mat}_n(\mathbb{R}) : \sum_{i,j} \mathcal{A}_{ij} = 1\}$. Let $m = \binom{n}{2}$.

We claim that M_n is a convex set. We, in fact, show that M_n is a cone, i.e., a set closed under positive scaling and under addition.

If we take $\mathcal{A} \in M_n$ and $\mathcal{B} \in M_n$, then there are subsets of ℓ_p, $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ such that $\|x_i - x_j\|^p = \mathcal{A}_{ij}$ and $\|y_i - y_j\|^p = \mathcal{B}_{ij}$. We define elements z_j of ℓ_p as follows: the $2i$'th term of z_j is the i'th term of x_j and the $(2i+1)$'th term of z_j is the i'th term of y_j. Then $\|z_i - z_j\|^p = \|x_i - x_j\|^p + \|y_i - y_j\|^p$, i.e, we can find a subset $z_1, \ldots, z_n$ of ℓ_p such that if we let $\mathcal{C}_{ij} = \|z_i - z_j\|^p$ then $\mathcal{C}_{ij} = \mathcal{A}_{ij} + \mathcal{B}_{ij}$. It follows that K is convex, as it is the intersection of two convex sets.

Let L be the set of linear points points in K. Evidently the convex hull of L is contained within K. Our goal is now to show that K is contained within the convex hull of L. Since L is a compact set, the closure $\overline{\mathrm{conv}}(L)$ is equal to $\mathrm{conv}(L)$, so for the former all we need to do is show that every element is contained in the closure of the convex hull of L.

Consider a collection $x_1, \ldots, x_n \in \ell_p$. Then, for each r we have that the element $\Theta \in M_n$ given by $\Theta_{ij} = |x_i(r) - x_j(r)|^p$ is a linear point of M_n.[2] Hence every element of K is formed from a convex combination of elements from L.

Since K is an $(m-1)$-dimensional convex set, then (by Caratheodory's Theorem) every element of K can be written as a convex combination of at most m points of

[2]Simply, this is an embedding into the real line.

L. But this is the statement that every set of n points admits an isometric embedding into $\ell_p^{\binom{n}{2}}$. $\square$

1.1.2 The Quasi-Banach spaces ℓ_p with $0 < p < 1$

At times, it is interesting to consider the analogues of the ℓ_p spaces for which $0 < p < 1$. In this section (and only this section!) we shall fix $0 < p < 1$. The proofs from this section are primarily adapted from [14].

We can define ℓ_p to be the space of real sequences $(x_i)_{i\in\mathbb{N}}$ such that $\sum_{i=1}^{\infty} |x_i|^p < \infty$. We then consider a 'norm' on this space given by $\|x\|_p = (\sum_{i=1}^{\infty} |x_i|^p)^{1/p}$. We will prove three facts about the space ℓ_p. The first is that there is a sensible metric to take on it, the second is that it is a quasi-Banach space, and the third is that it is not a Banach space.

Lemma 1.1.9. *If $0 < p < 1$, the function $d(x, y) = \|x - y\|_p^p$ is a metric on ℓ_p.*

Proof. The main thing to prove here is the triangle inequality. Everything else follows straightforwardly. We wish to show that $d(x, y) + d(y, z) \geq d(x, z)$, i.e., $\sum |x_i - y_i|^p + |y_i - z_i|^p \geq \sum |x_i - z_i|^p$. It is evidently enough to prove that $(x + y)^p \leq x^p + y^p$ for x, y non-negative and $p \in (0, 1)$.

If $y = 0$ this is obvious. Otherwise divide by y^p and consider $f(x) = (1 + x)^p - 1 - x^p$. It is then sufficient to show that $f'(x)$ is negative, but $f'(x) = p(1 + x)^{p-1} - px^{p-1}$, and since $p - 1 < 0$ this is always negative. $\square$

This shows us that ℓ_p is a metric space when equipped with a certain metric. We will now introduce the notion of a quasi-Banach space, and show that the ℓ_p spaces, for $0 < p < 1$, form a natural class of examples. A quasi-Banach space $(X, \|.\|)$ is a vector space X equipped with a function $\|.\| : X \to \mathbb{R}$ such that:

- For every $x \in X$, $\|x\| \geq 0$
- There is a constant K such that for all x, y we have that $\|x + y\| \leq K(\|x\| + \|y\|)$
- For each $x \in X$ and $\lambda \in \mathbb{R}$, $\|\lambda x\| = |\lambda|\|x\|$.

Lemma 1.1.10. *ℓ_p for $0 < p < 1$ is a quasi-Banach space.*

Proof. Clearly $\|.\|_p$ satisfies the first and third of the conditions for a quasi-Banach space, so we just need to prove the modified triangle inequality.

We first observe that, if $\theta > 1$ then the function $\lambda \mapsto \lambda^\theta$ is convex as a positive real function. Thus, if x, y are positive reals,

$$(x+y)^{1/p} = \left(\frac{2x+2y}{2}\right)^{1/p} \leq \frac{(2x)^{1/p} + (2y)^{1/p}}{2} = 2^{1/p-1}(x^{1/p} + y^{1/p})$$

where the inequality is an application of Jensen's inequality.

Now $\|x+y\|_p = (\sum |x_n + y_n|^p)^{1/p}$ is less than $(\sum |x_n|^p + \sum |y_n|^p)^{1/p}$ by Lemma 1.1.9. This is less than $2^{1/p-1}\left(\sum |x_n|^p\right)^{1/p} + 2^{1/p-1}\left(\sum |y_n|^p\right)^{1/p}$, which is as required. $\square$

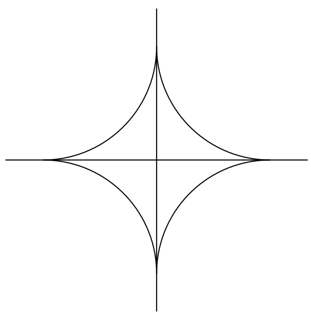
We finally show that the unit ball is not convex, in fact, we will show that whenever the coefficients of x and y have the same sign, then the triangle inequality reverses.

Lemma 1.1.11. *If x and y are non-negative sequences in ℓ_p then $\|x+y\|_p \geq \|x\|_p + \|y\|_p$.*

Proof. Let $q = 1/p$ and consider $w = (u, v) = (x^q, y^q)$ where x^q means $(x^q)_n = (x_n)^q$. We equip $\mathbb{R}^2$ with the $q-$norm. Let $I(w) = (\sum u_n, \sum v_n)$. Then $\|I(w)\|_q^q = (\sum u_n)^q + (\sum v_n)^q$. This is equal to $\|x\|_p + \|y\|_p$.

However, $(\sum \|w\|_q)^q = (\sum (u_n^q + v_n^q)^{1/q})^q = (\sum (x_n + y_n)^q)^{1/q}$. It is clear that $\|I(w)\| \leq \|w\|$, and so we are done. $\square$

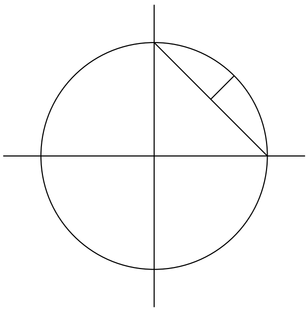
The "unit ball" of an ℓ_p space looks very strange - the fact that the triangle inequality reverses indicates that it is not convex. The following is a sketch of the unit ball of ℓ_p for $0 < p < 1$:



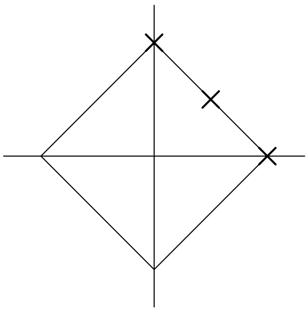
1.1.3 Convexity and Constants

Strict Convexity

The simplest notion of convexity is the notion of *strict convexity*. We say that a Banach space X is *strictly convex* if for any two points $x \neq y \in S_X$ we have that $\|x+y\| < 2$. We note that strict convexity is an isometric invariant of a space X: if we have two Banach spaces X and Y that are isometrically isomorphic, X is strictly convex if and only if Y is strictly convex. It is, however, not an isomorphic invariant. Indeed ℓ_1^2 and ℓ_2^2 are isomorphic, but we will see that ℓ_2^2 is strictly convex while ℓ_1^2 is not.



The space ℓ_2^2 is strictly convex because there is a 'gap' between the circle and the midpoint of the constructed line.



The space ℓ_1^2 given here is not strictly convex - the midpoint lies exactly on the line.

The simplest example of a strictly convex Banach space is Hilbert space. Suppose that we take $x \neq y \in S_{\ell_2}$. Then the parallelogram law states that $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$, ie, that $\|x+y\| = \sqrt{4 - \|x-y\|^2}$. This is evidently less than 2, so Hilbert space is strictly convex.

It is equally important to know that there are spaces that are not strictly convex. Two key examples of these are the spaces ℓ_1 and ℓ_∞. In ℓ_1, the two points $x = (1, 0, \dots)$ and $y = (0, 1, 0, \dots)$ have the property that $\|x\| = \|y\| = 1$ and $\|x+$

$y\| = 2$. In ℓ_∞ we can similarly take the points $x = (1,0,\dots)$ and $y = (1,1,0,\dots)$.

For $1 < p < \infty$ the spaces involved are strictly convex, however the proofs of this are slightly more complicated. In ℓ_2 we used the parallelogram law. However, this is not true in general Banach spaces. We can think of trying to generalise this to ℓ_p, attempting to relate $\|x+y\|_p^p$ and $\|x-y\|_p^p$ to $\|x\|_p^p$ and $\|y\|_p^p$. It turns out to be impossible to find an identity between these two things in general. However, we can find *inequalities*.

These inequalities are known as Clarkson's inequalities. They take different forms for $p < 2$ and $p > 2$. Clarkson's inequalities are very classical, dating back to Clarkson's seminal paper [8]. The proofs we present are from [6, Chapter 11] (who attributes them to Boas) with some of the details filled in.

Lemma 1.1.12 (Clarkson's inequality for $p > 2$)**.** *If $p > 2$ and x, y are in ℓ_p, then*

$$\left\|\frac{x+y}{2}\right\|_p^p + \left\|\frac{x-y}{2}\right\|_p^p \le \frac{1}{2}(\|x\|_p^p + \|y\|_p^p).$$

Proof. We first show for real numbers $a, b \in \mathbb{R}$ that

$$|a+b|^p + |a-b|^p \le 2^{p-1}(|a|^p + |b|^p).$$

Applying this inequality with $a = x_n$ and $b = y_n$, dividing by 2^p, then summing, gives the result.

Since the ℓ_p norms are decreasing, we have that $(|a+b|^p + |a-b|^p)^{1/p} \le (|a+b|^2 + |a-b|^2)^{1/2}$. By the parallelogram law, this is equal to $2^{1/2}(|a|^2 + |b|^2)^{1/2}$. Using the constant of equivalence (which is $2^{1/2-1/p}$, see Lemma 1.1.4) between ℓ_p^2 and ℓ_2^2, we get that this is at most $2^{1-1/p}(|a|^p + |b|^p)$, as required. □

For $p < 2$, Clarkson's inequality is more complicated:

Lemma 1.1.13 (Clarkson's inequality for $p < 2$)**.** *Suppose that $1 < p \le 2$, $x, y \in \ell_p$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then $\|x+y\|_p^q + \|x-y\|_p^q \le 2(\|x\|_p^p + \|y\|_p^p)^{q-1}$.*

We first establish the following pointwise inequality:

Lemma 1.1.14. *If $1 < p \le 2$, $\frac{1}{p} + \frac{1}{q} = 1$ and $0 \le x \le 1$ then*

$$(1+x)^q + (1-x)^q \le 2(1+x^p)^{q-1}.$$

Proof. Consider the function

$$f(\alpha) = (1+\alpha^{1-q}x)(1+\alpha x)^{q-1} + (1-\alpha^{1-q}x)(1-\alpha x)^{q-1}$$

for $0 \le \alpha \le 1$. Then $f(1)$ is the left hand side of the inequality, and $f(x^{p-1})$ is the right hand side of the inequality. Since $1 \ge x^{p-1}$ we show that $\partial_\alpha f \le 0$. But this is easy, we have that

$$\partial_\alpha f = (q-1)x(1-\alpha^{-q})\left((1+\alpha x)^{q-2} - (1-\alpha x)^{q-2}\right).$$

Now since $(1-\alpha^{-q}) \le 0$ (as $\alpha \le 1$) and the bracket

$$\left((1+\alpha x)^{q-2} - (1-\alpha x)^{q-2}\right) \ge 0$$

(as $q \ge 2$) we have that the derivative is negative. ☐

Proof of Lemma 1.1.13. We begin by noting that

$$\|x\|_p^q = (\sum |x_n|^p)^{1/(p-1)} = (\sum |x_n|^{q(p-1)})^{1/(p-1)} = \| |x|^q \|_{p-1},$$

where by $\|.\|_{p-1}$ we are referring to the quasi-Banach space ℓ_{p-1}, see Section 1.1.2. Here the triangle inequality reverses, so $\|x+y\|_p^q + \|x-y\|_p^q = \||x+y|^q\|_{p-1} + \||x-y|^q\|_{p-1} \le \||x+y|^q + |x-y|^q\|_{p-1}$. Applying the pointwise inequality proved above, we have that this is $\le 2\|(|x|^p + |y|^p)^{q-1}\|_{p-1}$. This is equal to $2(\|x\|_p^p + \|y\|_p^p)^{q-1}$ because $(q-1)(p-1) = 1$. ☐

We can now show very easily that the ℓ_p spaces are strictly convex:

Theorem 1.1.15. *The space ℓ_p is strictly convex for $1 < p < \infty$.*

Proof. First suppose that $p > 2$. If $x \neq y$ with $x, y \in S_{\ell_p}$, then by Lemma 1.1.12, we have that $\|x+y\| \le (2^p - \|x-y\|_p^p)^{1/p} < 2$. In the case $p < 2$, if $x \neq y$ with $x, y \in S_{\ell_p}$, then by Lemma 1.1.13, we have that $\|x+y\|_p \le (2^q - \|x-y\|_p^q)^{1/q}$, which is less than 2. ☐

The version of Clarkson's inequalities presented above are not the strongest forms possible. The proofs of the following stronger versions are more technical, generally depending on interpolation results that we do not want to be involved with. The reader can find a proof of these results in, for example, [14, Chapter 9].

Theorem 1.1.16. *Suppose that $x, y \in \ell_p$, where $1 < p < \infty$. We let p' be such that $p^{-1} + (p')^{-1} = 1$, and take $1 \le r \le \min(p, p')$. Then:*

- $\|x+y\|_p^{r'} + \|x-y\|_p^{r'} \le 2(\|x\|_p^r + \|y\|_p^r)^{r'-1}$
- $2(\|x\|_p^{r'} + \|y\|_p^{r'})^{r-1} \le \|x+y\|_p^r + \|x-y\|_p^r$
- $2(\|x\|_p^{r'} + \|y\|_p^{r'}) \le \|x+y\|_p^{r'} + \|x-y\|_p^{r'} \le 2^{r'-1}(\|x\|_p^{r'} + \|y\|_p^{r'})$

- $2^{r-1}(\|x\|_p^r + \|y\|_p^r) \leq \|x+y\|_p^r + \|x-y\|_p^r \leq 2(\|x\|_p^r + \|y\|_p^r)$.

Remark 1.1.17. The inequalities above are not the strongest inequalities relating to the convexity of the ℓ_p spaces. Hanner's inequality, as given below, is stronger.

Lemma 1.1.18. *Suppose that $1 \leq p \leq 2$ and $x, y \in \ell_p$. Then*

$$\|x+y\|_p^p + \|x-y\|_p^p \geq (\|x\|_p + \|y\|_p)^p + |\|x\|_p - \|y\|_p|^p.$$

If $2 \leq p < \infty$ the inequality reverses.

Since we have no need to use this strengthened inequality we will not give a proof.

Strict convexity is a property that is essentially only related to local geometric conditions on the space. We can, however, use it to deduce some theorems of a global nature:

Theorem 1.1.19. *There are two Banach spaces X and Y such that $d(X, Y) = 1$ but X and Y are not isometrically isomorphic.*

This result is interesting in and of itself. However, the construction given will be modified later to construct a counter-example. This is why we give such a careful proof of it.[3]

Proof. Take $(q_n)_{n=1}^{\infty}$ a sequence that is dense in $(1,2)$, and consider the space $X = \left(\oplus_{n=1}^{\infty} \ell_{q_n}^2\right)_2$ and $Y = \ell_1^2 \oplus_2 X$. We claim that, for any $\epsilon > 0$, there is a mapping T from X to Y such that $\|T\| \cdot \|T^{-1}\| < 1 + \epsilon$. We think of elements of X as sequences $(x_1, x_2, \dots)$ with $x_i \in \ell_{q_i}^2$ and elements of Y as sequences $(y_0, y_1, \dots)$ with $y_0 \in \ell_1^2$ and $y_i \in \ell_{q_i}^2$.

Fix some $\delta > 0$, that will be chosen later. The mapping T that we take will send the sequence $(x_1, x_2, \dots)$ to the sequence $(x_{\sigma(1)}, x_{\sigma(2)}, \dots)$ where σ is a permutation of $\mathbb{N}$ with the property that $q_i < q_{\sigma(i)}$ and either $|q_{\sigma(i)} - q_i| < \delta$ or $|q_i - q_{\sigma^{-1}(i)}| < \delta$.

We construct σ by a 'back-and-forth' construction. We take two sets $\mathbb{M}_1$ and $\mathbb{M}_2$ that we update at each stage, and set both of them equal to $\mathbb{N}$ to begin with.

Start by finding some $\sigma(1)$ such that $|q_{\sigma(1)} - 1| < \delta$, and some $\sigma^{-1}(1)$ to be chosen such that $|q_{\sigma^{-1}}(1) - q_1| < \delta$, and $q_1 < q_{\sigma(1)}$. We set $\mathbb{M}_1' = \mathbb{M}_1 \setminus \{1, \sigma^{-1}(1)\}$ and $\mathbb{M}_2' = \mathbb{M}_1 \setminus \{1, \sigma(1)\}$ and relabel to remove the $'$.

[3]The proof as written here was written by the author, however, there is no claim of originality to the result or the method of proof.

Assume we have performed n stages of this process. In the $n+1$'st stage, we choose the smallest element i of $\mathbb{M}_1$ and the smallest element j of $\mathbb{M}_2$. We then define $\sigma(i)$ to be any element of $\mathbb{M}_2$ for which $|q_i - q_{\sigma(i)}| < \delta$ and $q_i < q_{\sigma(i)}$ and $\sigma^{-1}(j)$ to be any element of $\mathbb{M}_1$ such that $|q_{\sigma^{-1}(j)} - q_j| < \delta$ and $q_{\sigma^{-1}}(j) < q_j$. Update $\mathbb{M}_1$ by removing i and $\sigma^{-1}(j)$ and $\mathbb{M}_2$ by removing $\sigma(i)$ and j.

We note that σ is, indeed, a permutation of the natural numbers. At each stage, we add in the smallest number in both the domain and range that have not been used.

We now check that, if δ is small enough, that $\|T\|\|T^{-1}\| < 1+\epsilon$. We first note, that since the ℓ_p norms are monotonically decreasing, we have that $\|T\| < 1$, by the condition that $q_i < q_{\sigma(i)}$. We now need to check that T^{-1} is bounded. $\|T^{-1}(x_1, x_2, \dots)\|^2 = \|(x_\sigma(1), x_\sigma(2), \dots)\|^2 = \sum \|x_{\sigma(i)}\|_i^2$. Now, by the computation in Lemma 1.1.4, $\|x_{\sigma(i)}\|_i \leq 2^{1/q_{\sigma(i)} - 1/q_i}\|x_i\|$. Since we have chosen $q_{\sigma(i)} - q_i$ to be small, the constant $2^{1/q_{\sigma(i)} - 1/q_i}$ can be made as arbitrarily close to $1+\epsilon$ (by shrinking δ at the beginning of the argument).

We now need to check that the spaces involved are not isometrically isomorphic. We will see later that the 2-direct sum of strictly convex spaces is strictly convex, and thus X is convex. It is, however, clear that the space Y is not as ℓ_1^2 isometrically embeds into Y. Thus there is no possible isometry from Y to X, if we take two distinct points $x, y \in S_Y$ such that $\|x+y\| = 2$, there is no possible image for them in X. $\square$

It is of interest to know whether a Banach space can be given an equivalent strictly convex norm. This is a topic of wide interest and much has been said on it (e.g., the interested reader can see the introduction to [32] for results and references.) In our thesis we will be interested in two such spaces, ℓ_1 and ℓ_∞. We will first show that ℓ_∞ has an equivalent strictly convex norm, and then use this to show that every separable Banach space has an equivalent strictly convex norm.

Lemma 1.1.20. *For any $\epsilon > 0$ there is a norm $\|.\|'$ on ℓ_∞ such that $\|.\|'$ is $(1+\epsilon)$-equivalent to the original norm and $\|.\|'$ is strictly convex.*

Proof. We first define a function $T : \ell_\infty \to \ell_2$ by $T(x_1, x_2, \dots) = (x_1, \frac{x_2}{2}, \frac{x_3}{3}, \dots)$. Note that this is bounded, $\|Tx\| = \sqrt{\sum_{i=1}^\infty |\frac{x_i}{i}|^2} \leq \|x\|_\infty \sqrt{\sum_{i=1}^\infty i^{-2}} \leq 2\|x\|_\infty$. For some $\delta > 0$, we define a new norm on ℓ_∞ by $\|x\|' = \|x\|_\infty + \delta\|Tx\|_2$. This is evidently a norm, and the distance from $(\ell_\infty, \|.\|')$ to ℓ_∞ is at most $1+2\delta$, so all we need to do is check strict convexity. Indeed, suppose $\|x\|' = \|y\|' = 1$ and

$\|x+y\|' = 2$. Then

$$\|x+y\|' = \sup|x_i+y_i| + \sqrt{\sum_{i=1}^{\infty} \frac{|x_i+y_i|^2}{n^2}}.$$

This is less than

$$\sup|x_i| + \sup|y_i| + \sqrt{\sum \frac{|x_i|^2}{n^2}} + \sqrt{\sum \frac{|y_i|^2}{n^2}}.$$

Equality in the triangle inequality for ℓ_2 occurs just when x and y are positive scalar multiples of each other. Since x, y both have prime-norm one, this means that both are the same, i.e. $x = y$. This is the condition of strict convexity. □

Theorem 1.1.21. *Suppose X is a separable Banach space. Then there is an equivalent norm on X such that X is strictly convex.*

Proof. We first recall, by our earlier results, that every separable Banach space has an isometric embedding T into ℓ_∞.

We define a new norm on X by $\|x\|_{sc} = \|T(x)\|'$, where the prime norm is as above. This embeds X into a subspace of a strictly convex space, and thus X is isomorphic to a strictly convex space. □

Uniform Convexity

Uniform convexity is a quantitative variant of the definition of strict convexity. We say that a Banach space is *uniformly convex* if for any $\epsilon > 0$ there is some $\delta > 0$ such that if we take $x, y \in S_X$ with $\|x-y\| \geq \epsilon$ then $\|x+y\| < 2-2\delta$. Evidently a uniformly convex space is strictly convex.

Uniform convexity is a strong condition. One may ask whether every Banach space can be given a uniformly convex norm, i.e., whether an analogue of Theorem 1.1.21 holds. This is, however, not true. Uniform convexity turns out to be a remarkably strong property, and we have the following result.

Theorem 1.1.22. *Suppose that X is a uniformly convex Banach space. Then X is reflexive.*

For a proof of this, see [35, Section 8].

This condition shows us that not every separable Banach space is uniformly convexifiable, i.e., there is no analogue of Theorem 1.1.21. Thus the condition of uniform convexifiability is a strong geometric condition on the Banach space.

The first thing to observe is that strict convexity implies uniform convexity in finite-dimensional Banach spaces. To show this, fix some $\epsilon > 0$ and assume that there is no $\delta > 0$ such that $\|x - y\| \geq \epsilon$ implies $\|x + y\| < 2 - 2\delta$. Set $\delta = 1/n$ and take points x_n, y_n such that $\|x_n - y_n\| \geq \epsilon$ and $\|x_n + y_n\| > 2 - 2/n$. Then, passing to subsequences, we get two points x and y such that $\|x - y\| \geq \epsilon$ but $\|x + y\| \geq 2$, a contradiction.

The second thing to observe is that there are infinite-dimensional Banach spaces that are strictly convex but not uniformly convex. We are going to prove a result later for spaces with this property, so we give three examples. The simplest example is $(\ell_\infty, \|.\|')$, which was defined above. Since ℓ_∞ is not reflexive we see that this space can not be uniformly convex.

The next example we give is $\left(\oplus_{i=1}^{\infty} \ell_{p_i}\right)_2$ where p_i is a sequence in $(1, \infty)$ that has a subsequence p_{n_j} converging to either 1 or ∞. In the case that $p_{n_i} \to 1$ simply note that $\|e_1 + e_2\|_{p_i} = 2^{1/p_i} \to 2$. In the case that $p_{n_i} \to \infty$ simply note that $(e_1 + e_2)/2^{1/p_i}$ and $(e_1 - e_2)^{2^{1/p_i}}$ are norm one vectors that, when added, give a vector whose norm tends to 2.

For a slightly more complicated example we give the construction of Talponen style sequence spaces (for more the interested reader can see [42]). To define this, given a sequence $p_n \in [1, \infty]$ we define a space which we think of as "a varying ℓ_p" sum. On $\mathbb{R}$ we define the norm $\|x\|_0 = |x|$. We inductively define the norm on $\mathbb{R}^n$ by $\|(x_1, \ldots, x_n)\|_n = \left(\|(x_1, \ldots, x_{n-1})\|_{n-1}^{p_{n-1}} + |x_n|^{p_{n-1}}\right)^{1/p_{n-1}}$. This evidently defines a norm on c_{00} and we define the space $\ell_{(p_i)}$ by the the completion of this norm.

This space has the property that $\ell^2_{p_n}$ isometrically embeds into it for each n. As in the previous case, one can see that if $1 < p_i < \infty$ we have that the constructed space is strictly convex, but the space is not uniformly convex if any subsequence of the p_i tends to 1 or any subsequence of the p_i tends to ∞.

We mention a third example without any details - this is the exponential Orlicz space. For the definition of this, the reader can see Section 4.5.4 where we take $r(t) = e^t - 1$. Proving that this space is strictly convex but not uniformly convex is relatively complicated, the reader can see the details in [31] (and simply check that $r(t)$ satisfies the conditions of the theorem in this paper.)

One may ask why we mention this space here if we give none of the details - this is simply because it isn't a "weird counterexample" space, the exponential Orlicz space is a relatively natural generalization of the L_p spaces and has received attention in physics. So, we can even say that not every "natural" or "obvious" Banach space has the property that strict convexity is equivalent to uniform convexity.

1.1.4 A way of defining Banach spaces

Suppose that E is a finite-dimensional Banach space. Then the closed unit ball B_E of E has the property that it is compact, convex, centrally symmetric and there is some Euclidean neighbourhood of zero contained in B_E. Conversely, if we have a subset U of $\mathbb{R}^n$ with these properties, then there is a Banach space E whose unit ball is U. We define the norm on E by the *Minkowski functional of* U, given by

$$p_U(x) = \inf\{\lambda \in \mathbb{R}^+ : x \in \lambda U\}.$$

Clearly $p(x)$ is well defined, and the set $\{x \in \mathbb{R}^n : p_U(x) \leq 1\}$ is E. The fact that p_U is a norm (i.e., obeys $p_U(\lambda x) = |\lambda| p_U(X)$ and the triangle inequality) is classical and can be found in many books. The presentation in [39] is understandable.

We will mainly be interested in the case $n = 2$ where the idea of the Minkowski functional gives us a convenient way of constructing counter-examples. To define a norm on $\mathbb{R}^2$ we can specify a certain type of curve Γ. Γ will play the role of the boundary of the set C above.

We begin by taking a continuous map $\gamma : [0,1] \to \mathbb{R}^2$ that is injective on $[0,1)$ and is such that $\gamma(0) = \gamma(1)$. We write Γ for the image of γ and call Γ a *simple closed curve*. The Jordan curve theorem says that Γ splits the plane into two connected components, one bounded and one unbounded. We call the bounded component the *interior* of the curve.

We say a *convex curve* is one whose interior is convex, and a symmetric curve is one for which if $x \in \Gamma$ then $-x \in \Gamma$. Since we will primarily be interested in specifying norms on 2-dimensional Banach spaces later on, we record the following lemma:

Lemma 1.1.23. *Suppose that Γ is a simple closed curve in $\mathbb{R}^2$ not containing 0 and let $\hat{\Gamma}$ be the curve together with its interior. Then the Minkowski functional $p_\Gamma = p_{\hat{\Gamma}}$ defines a norm on $\mathbb{R}^2$ with unit sphere Γ.*

We will use this later on where to specify a norm on $\mathbb{R}^2$ we will simply define a curve Γ.

1.1.5 Finite representability

Suppose that X and Y are Banach spaces. We say that X is *crudely finitely representable in* Y if there is a constant $C > 0$ such that for any finite-dimensional subspace E of X, there is a finite-dimensional subspace F of Y such that $d(E, F) \leq C$. We say that X is *finitely representable in* Y if for any finite-dimensional subspace E of X, and any $\epsilon > 0$ there is a finite-dimensional subspace F of Y such that $d(E, F) < 1 + \epsilon$.

We illustrate this concept with a lemma that will be of interest later:
Lemma 1.1.24. *For* $1 \leq p < \infty$, $L_p[0, 1]$ *is finitely representable in* ℓ_p.
Remark 1.1.25. The above statement is true for the case $p = \infty$, albeit for different reasons. We showed in Section 1.1.1 that every every finite-dimensional Banach space embeds linearly isometrically into ℓ_∞, and thus any Banach space is finitely representable in ℓ_∞.

The proof we give here does not use the structure of $[0, 1]$ at all, indeed, $L_p[0, 1]$ can be replaced by $L_p(\Omega, \mathcal{F}, \mu)$ for any measure space $(\Omega, \mathcal{F}, \mu)$.

Proof. Suppose that we are given an n-dimensional subspace E of $L_p[0, 1]$ and $\epsilon > 0$. Let $f_1, \ldots, f_n$ be a basis for E. Fix some $\delta > 0$. Since simple functions are dense in $L_p[0, 1]$ we can take $g_1, \ldots, g_n$ such that g_i are simple and $\|g_i - f_i\| < \delta$.

Since the g_i are simple we can find pairwise disjoint sets $E_1, \ldots, E_N$ such that $\cup E_i = [0, 1]$ and $g_i|_{E_j}$ is constant. It is clear that the span of the indicator functions of the E_i's is isometrically isomorphic to ℓ_p^N, and that the span of the $g_1, \ldots, g_n$ is contained in this space.

We now define an isomorphism T from $L_p[0, 1] \to L_p[0, 1]$ that has the property that $T(f_i) = g_i$. We can define functionals on the subspace E by the formula $\varphi_i(f_j) = \delta_{ij}$. These functionals can be extended by the Hahn-Banach theorem to all of $L_p[0, 1]$. Define a map $S : L_p[0, 1] \to L_p[0, 1]$ by $S(f) = \sum_{i=1}^n \varphi_i(f)(f_i - g_i)$. Then $\|S\| \leq \sum \|\varphi_i\| \|f_i - g_i\| \leq n\delta \max \|\varphi_i\|$.

We now define T by $(I - S)$. This clearly has the property that $T(f_i) = g_i$. Moreover we note that since, if δ is smaller than $\frac{1}{2n \max \|\varphi_i\|}$ that $(I - S)$ is invertible. Moreover, we note that $\|T\| \leq 1 + \|S\|$ and $\|T^{-1}\| \leq \frac{1}{1 - \|S\|}$, so

$$\|T\| \|T^{-1}\| \leq \frac{1 + \|S\|}{1 - \|S\|}.$$

Clearly $\|S\| \to 0$ as $\delta \to 0$, and thus $\|T\| \|T^{-1}\| \to 1$ as $\delta \to 0$.

Thus the map T from E to $T(E)$ is an into isomorphism from span$\{f_1, \ldots, f_n\}$ to span$\{1_{E_1}, \ldots, 1_{E_N}\}$, which is isometrically isomorphic to ℓ_p^N. $\square$

Remark 1.1.26. It is simple to show that ℓ_p isometrically embeds into $L_p[0,1]$, so for trivial reasons we also have that ℓ_p is finitely represented in $L_p[0,1]$.

Remark 1.1.27. We can combine Lemma 1.1.24 with Theorem 1.1.6 to deduce the following variant of Theorem 1.1.6: if we take n functions $f_1, \ldots, f_n$ in $L_p(\Omega)$ then we can isometrically embed the set $M = \{f_1, \ldots, f_n\}$ with the inherited ℓ_p-norm into $\ell_p^{\binom{n}{2}}$. Indeed, by Lemma 1.1.24 for any $\epsilon > 0$ we can find elements $x_1, \ldots, x_n$ of ℓ_p such that $|\|x_i - x_j\| - \|f_i - f_j\|| < \epsilon$. Applying Theorem 1.1.6 we can, without loss of generality, assume that $x_i \in \ell_p^{\binom{n}{2}}$.

Applying this process with $\epsilon = 1/m$ gives us points x_i^m such that $|\|x_i^m - x_j^m\| - \|f_i - f_j\|| < 1/n$. We now wish to apply compactness to find limits of the x_i^n's. By translation we may assume that $x_1^m = 0$ for all n and thus that the set $\{x_i^n : i \in \{1, \ldots, n\}, n \in \mathbb{N}\}$ is bounded. Applying compactness we can pass to a subsequence n_j such that $x_i^{n_j}$ is convergent. The limits x_i are the required isometric embedding of f_i into $\ell_p^{\binom{n}{2}}$.

We now define the concept of a *superreflexive Banach space*. A Banach space X is said to be *superreflexive* if whenever Y is finitely representable in X, then Y is reflexive. An example of a space that is reflexive but not superreflexive is $(\oplus_n \ell_\infty^n)_2$, c_0 is finitely representable in this space. The following remarkable theorem is due to Enflo:

Theorem 1.1.28 ([13]). *A Banach space X is superreflexive if and only if it is isomorphic to a uniformly convex space.*

Historically, it had been hoped that every infinite-dimensional Banach space contained an isomorphic copy of some ℓ_p space or c_0. Indeed, every classical Banach space has this property and it took until the introduction of Tsirelson's space in 1974 to give an example of a Banach space that contained no isomorphic copy of some ℓ_p space or c_0.

The local question is different. The following two theorems show that at a local level the ℓ_p spaces (together with c_0) form 'building blocks' of the infinite-dimensional Banach spaces.

The first of these theorems shows that "ℓ_2 is finitely representable in every infinite-dimensional Banach space". It was originally conjectured in [15] and first proved in [11]:[4]

Theorem 1.1.29 (Dvoretzky's Theorem)**.** *For all $n \in \mathbb{N}$ and $\epsilon > 0$ there is some $N \in \mathbb{N}$, depending on n and ϵ, such that if E is an N-dimensional Banach space then there is a subspace F of E such that $d(F, \ell_2^n) < 1 + \epsilon$.*

A combination of Dvoretzky's Theorem and Lemma 1.1.23 gives that if we choose any symmetric convex body with non-empty interior in $\mathbb{R}^N$, we can choose some n-dimensional cross-section of the body that resembles an ellipsoid.

The theorem immediately implies that whenever we have an infinite-dimensional Banach space X, any $n \in \mathbb{N}$ and any $\epsilon > 0$ there is an n-dimensional subspace E of X such that $d(E, \ell_2^n) < 1 + \epsilon$, i.e., ℓ_2 is finitely representable in X.

The second of these theorems very roughly says "some ℓ_p is finitely representable in any basic sequence". We state two different variations of it. The second is more useful for our purposes, but the first gives us an insight into why we would want to look at this in the first place.

The first version of this theorem was proved in [24][5]:

Theorem 1.1.30 (Krivine's Theorem (I))**.** *Suppose X is an infinite-dimensional Banach space and $x_1, x_2, \dots$ is a basic sequence in X. Then there is a $p \in [1, \infty]$ such that for any $\epsilon > 0$ and any $m \in \mathbb{N}$ we can choose vectors $y_1 = \sum_{i=1}^{n_1} \alpha_i x_i$, $y_2 = \sum_{i=n_1+1}^{n_2} \alpha_i x_i$, $\dots$, and $y_m = \sum_{n_{m-1}+1}^{n_m} \alpha_i x_i$ such that for any scalars $\alpha_1, \dots, \alpha_m$ we have that*

$$\|(\alpha_i)_{i=1}^m\|_p \leq \|\sum_{i=1}^n \alpha_i y_i\| \leq (1+\epsilon)\|(\alpha_i)_{i=1}^m\|_p$$

It is worth spelling out the difference between Dvoretzky's Theorem and Krivine's Theorem. Suppose that we have a basic sequence $x_1, x_2, \dots$ in an infinite-dimensional Banach space X. Then for any $n \in \mathbb{N}$, Dvoretzky's Theorem gives us a subspace of $\overline{\text{span}}(x_1, x_2, \dots)$ which is isomorphic to ℓ_2^n and Krivine's Theorem gives us a subspace of $\overline{\text{span}}(x_1, x_2, \dots)$ which is isomorphic to ℓ_p^n.

In Krivine's Theorem we can choose vectors that have *disjoint support* whereas it is impossible to guarantee this in Dvoretzky's Theorem.

Remark 1.1.31. As a concrete example of this, if $p \neq 2$ and we look at the unit vector basis e_i of ℓ_p, then any finite blocking of $e_1, e_2, \dots$ is a sequence that is 1-

[4] The reference given here is the original proof. However, it is a little hard to track down. The interested reader can see a proof in [1, Chapter 13].

[5] This paper is in French and a little hard to find. The interested reader can find an English version of the proof (that is slicker) in [29, Chapter 12]

equivalent to the unit basis of ℓ_p^n. Thus this blocking is not $(1+\epsilon)$-equivalent to the unit basis of ℓ_2. This means that no 'block version' of Dvoretzky is true.

The following finite quantitative variant was proved in [2],

Theorem 1.1.32 (Krivine's Theorem (II))**.** *For all $n \in \mathbb{N}$, $\epsilon > 0$, $C > 0$ and $1 < p < \infty$, there is some $N = N(n, \epsilon, C, p)$ such that if E is an N-dimensional Banach space with $d(E, \ell_p^N) < C$ there is an n-dimensional subspace F of E such that $d(F, \ell_p^n) < 1 + \epsilon$.*

The main way we will use this theorem is through the following corollary (that follows from Theorem 1.1.32 and Remark 1.1.31):

Corollary 1.1.33. *If $\ell_p \hookrightarrow X$, then for all $\epsilon > 0$ and $n \in \mathbb{N}$, $\ell_p^n \overset{1+\epsilon}{\hookrightarrow} X$*

1.1.6 Type and Cotype

The topic of finite representability of ℓ_p in a Banach space has natural links to the theory of type and cotype. We will provide a brief introduction to this theory here.

The idea of type and cotype comes from the desire to generalize a property of Hilbert space[6]. We begin by reminding the reader of the generalized parallelogram identity which states for any $x_1, \ldots, x_n$ that

$$\frac{1}{2^n} \sum_{(\epsilon_1,\ldots,\epsilon_n)\in\{-1,1\}^n} \left\| \sum_{i=1}^n \epsilon_i x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2.$$

Type and cotype form a natural generalisation of this equality.

We say that a Banach space has *type p* if there is some constant T such that for any $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in X$ we have that

$$\frac{1}{2^n} \sum_{(\epsilon_1,\ldots,\epsilon_n)\in\{-1,1\}^n} \left\| \sum_{i=1}^n \epsilon_i x_i \right\|^p \le T^p \sum_{i=1}^n \|x_i\|^p.$$

The least such constant is called the *type p constant of* X and is denoted T_p.

We say that a Banach space has *cotype q* if there is some constant C such that for any $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in X$ we have that

$$C^q \left(\frac{1}{2^n} \sum_{(\epsilon_1,\ldots,\epsilon_n)\in\{-1,1\}^n} \left\| \sum_{i=1}^n \epsilon_i x_i \right\|^q \right) \ge \sum_{i=1}^n \|x_i\|^q.$$

The least such constant is called the *type q constant of* X and is denoted C_q.

[6] A lot of good Banach space theory properties come from the same desire.

By taking every $x_i = x$ for some $\|x\| = 1$ we can see that no Banach space has type p for $p > 2$ and no Banach space has cotype q for $q < 2$. Every Banach space has type 1 and cotype infinity by an application of the triangle inequality. These two facts motivate the following definition, we set

$$p_X = \sup\{p : X \text{ has type } p\}$$

and

$$q_X = \inf\{q : X \text{ has cotype } q\}.$$

We see that $p_X \in [1, 2]$ and $q_X \in [2, \infty]$. We will say that a Banach space has *no non-trival type* if $p_X = 1$ and *no non-trivial cotype* if $q_X = \infty$.

It is clear that p_X and q_X are isomorphic invariants of a Banach space, thus we can show that two Banach spaces X and Y are not isomorphic by showing that $p_X \neq p_Y$ or $q_X \neq q_Y$. We have the following fact:

Lemma 1.1.34. *[[1, Chapter 6]] If $X = L_p$ for $1 \leq p \leq \infty$ then $p_X = \min\{p, 2\}$ and $q_X = \max\{q, 2\}$.*

A combination of the two previous lemmas shows that L^p is not isomorphic to L^q for any $p \neq q$ - a strengthening of the fact from the beginning of Section 1.1.1.

During our thesis we will prove theorems that are contingent on ℓ_p being finitely representable inside of a space. The following results demonstrate a fundamental link between ℓ_p being finitely representable in a space and linear properties of that space. The first of these results is called the Maurey-Pisier theorem.

Theorem 1.1.35. *[[30, Chapter 13] Suppose X is an infinite-dimensional Banach space. Then ℓ_{p_X} and ℓ_{q_X} are both finitely representable in X.*

We will primarily be interested in using the Maurey-Pisier Theorem in the following special case:

Theorem 1.1.36. *Suppose that X is a Banach space with no non-trivial cotype. Then for every $n \in \mathbb{N}$ and $\epsilon > 0$, $\ell_\infty^n \overset{1+\epsilon}{\hookrightarrow} X$. Moreover, every finite metric space embeds almost isometrically into X.*

Remark 1.1.37. The "moreover" part of the previous theorem easily follows from results on metric spaces presented in Section 1.2, specifically Lemma 1.2.7.

Remark 1.1.38. The converse of this statement is true, i.e., if we have a Banach space X and a constant C such that every finite metric embeds space into X with distortion C[7] then X contains ℓ_∞^n almost isometrically. The author does not know

[7]For the definition of distortion see Section 1.2, or the reader is free to ignore the comment if they do not want to read it.

a self-contained proof of this result[8].However, it follows from some developements of Mendel and Naor in [27] on the metric theory of cotype (combined with the Maurey-Pisier theorem.).

1.1.7 Some Results on Distortion

Much of our work is centred on the question of whether a certain finite metric space embeds into a given Banach space. A natural question is whether these results can be extended to the case of infinite metric spaces.

When we are interested in the case of finite metric spaces, and embeddings of finite subsets of ℓ_p, we will find the following corollary of Krivine's Theorem useful: if X is a Banach space such that for some $1 \leq p \leq \infty$ there is some $C > 0$ such that $\ell_p^n \overset{C}{\hookrightarrow} X$ then for any $\epsilon > 0$ we have that $\ell_p^n \overset{1+\epsilon}{\hookrightarrow} X$

It is natural to ask if there is an infinite analogue of this, i.e.,

Question 1.1.39. *Suppose that X is an infinite dimensional Banach space, $1 \leq p \leq \infty$, and that $\ell_p \hookrightarrow X$ (or $c_0 \hookrightarrow X$.) Then does $\ell_p \overset{1+\epsilon}{\hookrightarrow} X$ (or $c_0 \overset{1+\epsilon}{\hookrightarrow} X$) for all $\epsilon > 0$?*

This question was first studied in the cases of ℓ_1 and c_0 by James in [19], and was answered positively. The case of ℓ_∞ was answered positively by Partington in [36], which we state now (as we will need to refer to it later):

Theorem 1.1.40. *Suppose X is a Banach space containing a subspace isomorphic to ℓ_∞. Then, for any $\epsilon > 0$, X contains a subspace which is $(1+\epsilon)$-equivalent to ℓ_∞.*

It was a major unsolved problem whether the question has a positive answer for the case of ℓ_p with $1 < p < \infty$. The case of $p = 2$ was of most interest and was called "the distortion problem". It was answered negative by Odell and Schlumprecht in the paper [34].

This negative answer represents a barrier to us attempting to generalise our results to the infinite-dimensional case.

1.1.8 The Concept of a Null Set

One of the most powerful concepts in mathematics is that of a "small set". These can take various forms. However, there are two notions of smallness that are familiar to most mathematicians. These are *nowhere dense sets*, i.e., subsets N of a

[8]But would like to.

topological space for which $\overline{N}$ is empty, and *measure zero sets*, i.e., subsets N of a measure space which have zero measure.

If we denote the collection of measure zero sets or nowhere dense sets by $\mathcal{F}$, then:

- If $A_i \in \mathcal{F}$ for all $i \in \mathbb{N}$ then so is $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.
- If $A \in \mathcal{F}$ and $B \subset A$, then $B \in \mathcal{F}$.
- $\emptyset \in \mathcal{F}$.

We call a collection of sets satisfying the above three properties a *σ-ideal*.

Now let us specialise to X, a finite-dimensional Banach space, which we equip with Lebesgue measure.[9] In this case, the measure zero sets and the nowhere dense sets share another property, and that is translational invariance. Indeed, if A is measure zero or Lesbegue null, then so is $x + A$ for any $x \in X$.

When we wish to prove results about infinite-dimensional Banach spaces we would like to have a notion of "small" that has these two properties, i.e., we have an infinite-dimensional Banach space X and we are seeking a collection of sets $\mathcal{F}$ that is translational invariant and a σ-ideal.

A natural way of attempting to do this would be to introduce a translationally invariant Borel measure μ on X, i.e., a measure μ on the measurable space $(X, \mathcal{G})$ where $\mathcal{G}$ is the smallest σ-algebra that contains the open sets of X. We would then be able to set $\mathcal{F}$ as the measure zero sets of μ. However, the following theorem shows that this is impossible to do usefully in most cases:

Theorem 1.1.41. *Suppose that X is an infinite-dimensional separable Banach space. Then any translational-invariant Borel measure on X either assigns 0 or ∞ to every open subset of X.*

The author can not find the origin of this claim in the literature. However the result seems to date back until at least the 1930's. The proof as presented here is a modification of the proof in [18].

Proof. For contradiction, suppose that μ is a Borel measure on X and that $A \subset X$ is an open set with measure $0 < \mu(A) < \infty$. Since A is open, we can find some $\epsilon > 0$ and $x \in A$ such that $x + \epsilon B_X \subset A$. A standard Hahn-Banach argument shows that we can find a family of points $y_1, y_2, \ldots$ with $y_i \in \epsilon B_X$ and $\|y_i - y_j\| > \epsilon/3$.

[9] More specifically, since all n-dimensional Banach spaces are isomorphic to ℓ_2^n, if we take an n-dimensional Banach space X we can equip X with the pushforward of Lesbegue measure. Since we are only interested in measure zero sets, the choice of isomorphism does not matter.

Now we can look at the set $\cup_{i=1}^{\infty}(y_i + \epsilon/3B_x)$. This is a subset of A, and thus has finite measure. We now observe that $\mu(\cup_{i=1}^{\infty}(y_i + \epsilon/3B_X)) = \sum_{i=1}^{\infty} \mu(y_i + \epsilon/3B_X) = \sum_{i=1}^{\infty} \mu(\epsilon/3B_X)$, where the first inequality is due to the sets involved being disjoint, and the second inequality is due to the translational-invariance of the measure. We thus have that $\mu(\epsilon/3B_X) = 0$.

Since X is separable, there is some collection of points $z_1, z_2, \dots$ such that $A \subset \overline{\{z_1, z_2, \dots\}}$. It is now clear that $A \subset \cup_{i=1}^{\infty}(z_i + \epsilon/3B_X)$. However, we can bound the measure of $\cup_{i=1}^{\infty}(z_i + \epsilon/3B_X)$ by countable subadditivity, and we get that it is less than or equal to $\sum_{i=1}^{\infty} \mu(z_i + \epsilon/3B_X) = \sum_{i=1}^{\infty} \mu(\epsilon/3B_X) = 0$, a contradiction to A having non-zero measure. $\square$

Thus, in infinite-dimensional separable Banach spaces, it is hard to define a notion of measure that is sensible, i.e., one that isn't identically zero or infinity on all open sets. We can, however, introduce a different notion of null sets.

Let X be a Banach space. We say that Borel set A is *Haar null* if for every Borel probability measure on X, and every $x \in X$, $\mu(x + A) = 0$. The following results, of which the first three are taken from [4] and the last is obvious from the definition, give us that the Haar null sets are a σ-ideal of translationally-invariant sets that play an infinite-dimensional analogue of Lebesgue null sets in $\mathbb{R}^n$:

Theorem 1.1.42. 1. *Haar null sets form a σ-ideal.*

2. *In finite-dimensional Banach spaces the notions of Haar null and Lebesgue null coincide.*

3. *If A is a subset of X such that for some n-dimensional subspace Y of X and every $x \in X$ we have that the set $Y \cap (x + A)$ is Lebesgue null, then A is Haar null.*

4. *The translate of a Haar null set is Haar null.*

Remark 1.1.43. The notion of Haar null is not the only notion that satisfies properties 1,2 and 4 from the above theorem. There is a relative profusion of such notions, cube null, Gaussian null, ... The interested reader can read [4, Chapter 6] for explanations of these notions and the links between them.

1.2 Metric Spaces

Let (M, d_M) and (N, d_N) be metric spaces, and $f : M \to N$ be a map between them. We say that f is a *Lipschitz function* if there is a constant $\alpha > 0$ such that

$d_N(f(x), f(y)) \leq \alpha d_M(x, y)$. We say that f is *bilipschitz* if there exist constants $\alpha, \beta > 0$ such that

$$\alpha d_M(x,y) \leq d_N(f(x), f(y)) \leq \beta d_M(x,y).$$

We say that f has *distortion* $C > 0$, or f is a *C-distortion embedding*, or f is a *C-embedding*, if it is bilipschitz, and $\frac{\beta}{\alpha} \leq C$. We say f is *an isometry* if $\alpha = \beta = 1$. We say that M *almost-isometrically embeds into* N if there is a $(1+\epsilon)-$distortion embedding $f : M \to N$ for each $\epsilon > 0$.

Throughout this thesis, we will often need to view a metric space M as its matrix of distances, i.e., if we have a finite metric space $M = \{m_1, \ldots, m_n\}$, we can associate with it the symmetric matrix $\mathcal{M} = \left(d(m_i, m_j)\right)_{i,j=1}^{n}$, noting that $\mathcal{M}_{ii} = 0$ for each i. We say that $\mathcal{M}$ *represents the metric space* M or $\mathcal{M}$ *represents* M. Since the collection of $n \times n$ symmetric matrices with zero diagonal will be relevant later, we denote this by U_n. We consider U_n as a subspace of $\mathrm{Mat}_{n\times n}(\mathbb{R})$ and will consider the subspace topology on it.

Remark 1.2.1. We note that not all of the elements of U_n represent a metric space. The requirement that an element $\mathcal{M} \in U_n$ represents a metric space is that $\mathcal{M}_{ij} + \mathcal{M}_{jk} \geq \mathcal{M}_{ik}$ for every distinct triple i, j, k. If this is the case we can define a metric space on points $\{m_1, \ldots, m_n\}$ by setting $d(m_i, m_j) = \mathcal{M}_{ij}$ if $i < j$ and 0 if $i = j$. This then satisfies all the axioms for a metric space. In this case we will say that the matrix $\mathcal{M}$ *represents a metric space*.

We note that U_n has dimension $\binom{n}{2}$ as a vector space (or as a manifold). Throughout this thesis, we will consider U_n as $\mathbb{R}^J$ where $J = \{(i,j) : 1 \leq i < j \leq n\}$.

Remark 1.2.2. The above notation is the clearest for everything except, perhaps, the notion of taking a determinant. We will spell this out for the reader - if we have a linear map $T : \mathbb{R}^J \to \mathbb{R}^J$ we will use the definition

$$\det(T) = \sum_{\sigma \in S_J} \epsilon(\sigma) \prod_{(i,j)\in J} T_{(i,j),\sigma(i,j)}.$$

There are, in general, different matrices that can represent the metric space M, these correspond to picking some ordering of the points of M. When we say that $\mathcal{M}$ represents M, we really mean this with respect to some ordering. We will, in general, implicitly fix some ordering in advance.

1.2.1 Some remarks on embedding finite metric spaces into Banach spaces

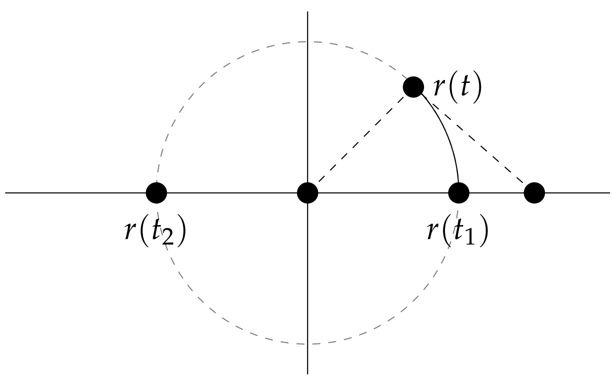
In this section we will give various simple remarks about embedding finite metric spaces into Banach spaces. These are things that are simple. However, without them the subject may seem a little pointless.

Lemma 1.2.3. *Every three-point metric space embeds isometrically into every Banach space of dimension ≥ 2.*

Proof. This is an intermediate value argument. Let M be a metric space on points m_1, m_2, m_3, and let X be a 2-dimensional Banach space with basis e_1, e_2, which we may assume have norm 1. We construct the isometry f as follows: first we set $f(m_1) = 0$ and $f(m_2) = d(m_1, m_2)e_1$.

We now have to find the image of m_3. Let us consider a parametrization of $d(m_3, m_1)S_X$, i.e., a function $r : [0,1] \to S_X$ such that r is a continuous function, bijective on $[0,1)$ and $r(0) = r(1)$. Such a function exists by, e.g., choosing $r(t)$ as the unique intersection of S_X with the ray forming an angle $2\pi t$ with the x axis. There are real numbers t_0 and t_1 such that $d(f(m_2), r(t_0)) = d(m_1, m_2) - d(m_1, m_3)$ and $d(f(m_2), r(t_1)) = d(m_1, m_2) + d(m_1, m_3)$, since $r[0,1]$ contains both of $\pm d(m_3, m_1)e_1$. We note that, by the triangle inequality, $d(m_2, m_3) \in [d(m_1, m_2) - d(m_1, m_3), d(m_1, m_2) + d(m_1, m_3)]$.

Finally, the function $t \mapsto \|r(t) - f(m_2)\|$ is continuous in t, and thus there is a point $t_2 \in [t_0, t_1]$ such that it this function attains the value $d(m_2, m_3)$, and $f(m_3) = r(t_2)$ is the required isometric extension.



This diagram shows the proof strategy.

$\square$

Lemma 1.2.4. *There is a four point metric space that does not isometrically embed into any Hilbert space.*

This example is taken from [35]. It is one of many examples in this book. However, this has the advantage of (perhaps) being the simplest.

Proof. Let $K_{1,3}$ be the metric space on points m_1, m_2, m_3, m_4 where $d(m_1, m_i) = 1$ for any $i \neq 1$ and $d(m_i, m_j) = 2$ for any other pairs i, j. The metric space $K_{1,3}$ is the complete bipartite graph with a vertex set of size 1 and a vertex set of size 3 equipped with the graph distance.

ℓ_2 has the following property: if x, y, z are in ℓ_2 such that $\|x - y\| + \|y - z\| = \|x - z\|$ then x, y, z lie on the same line. Indeed, we may suppose (wlog) that y is zero. Then $\|x - z\|^2 = \|x\|^2 + \|z\|^2 - 2\langle x, z\rangle$. If this is equal to $(\|x\| + \|z\|)^2$ that requires that $\langle x, z\rangle = -\|x|^2\|z\|^2$, which is the condition that x, z are co-linear and lie on opposite sides of 0.

This shows that m_1 is on the line joining m_2 to m_3, m_1 is on the line joining m_2 to m_4 and also on the line joining m_3 to m_4. Thus m_1, m_2, m_3 and m_4 are collinear, i.e., there exists an isometric embedding of $K_{1,3}$ into the real line. However this is impossible, there are not three real numbers a, b, c such that $|a - b| = |a - c| = |b - c|$. $\square$

We note that this argument requires the strict convexity of ℓ_2 and some idea of metric midpoints. We discuss these ideas more in the next section.

1.2.2 Convexity and Concavity of Metric Spaces

Just as there are a wealth of concepts of convexity and concavity for Banach spaces, there are a wealth of notions of convexity and concavity for metric spaces. The way that these concepts interact with the theory of Banach spaces will be of most interest to us.

We will call a metric space M *concave* if the triangle inequality is always an inequality, i.e., for distinct $x, y, z \in M$ we have that $d(x, y) + d(y, z) > d(x, z)$. This terminology was suggested by Nik Weaver on MathOverflow, see [16]. We will give some simple examples of concave metric spaces:

- If (M, d) is a metric space, then the *α-snowflake*, for $0 < \alpha < 1$, is defined to be the metric space (M, d^α), where $d^\alpha(m_1, m_2) = (d(m_1, m_2))^\alpha$. If we take m_1, m_2, m_3 a distinct triple, then $d(m_1, m_3) \leq d(m_1, m_2) + d(m_2, m_3)$. Taking the α power of both sides of this inequality, and using the fact that $(d(m_1, m_2) + d(m_2, m_3))^\alpha < d(m_1, m_2)^\alpha + d(m_2, m_3)^\alpha$ (if these are non-zero), shows that the α-snowflake of any metric space is concave.[10]

- If (M, d) is an infinite metric space such that $d(m_i, m_j) = 1$ for any distinct points $m_i, m_j \in M$, this space is concave. We call (M, d) the equilateral space.

- If (M, d) is an ultrametric space, i.e., a space where we have the strengthened triangle inequality $d(x, z) \leq \max(d(x, y), d(y, z))$ then the space is evidently concave.

A very useful notion is one of metric midpoints. If M is a metric space and $x, y, z \in M$ are distinct points, we say y is a *metric midpoint of x and z* if $d(x, y) = d(y, z) = \frac{1}{2}d(x, z)$. We say that a metric space has *unique metric midpoints* if there is at most one such y. The condition that a metric space has unique metric midpoints can be thought of as a convexity condition in light of the following result:

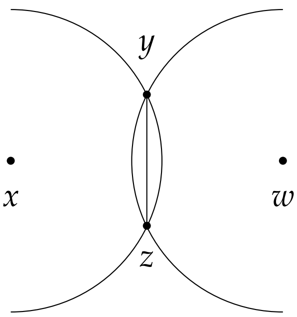
Lemma 1.2.5. *Suppose that X is a strictly convex Banach space. Then X has unique metric midpoints.*

Proof. Suppose that w, x, y, z are distinct points such that both y and z are metric midpoints of w and x. We define $2d$ to be the distance between x and w, i.e., $2d = \|x - w\|$. Then y and z both lie on $x + dS_X$ and $w + dS_X$. It is then clear to see that any point on the line segment joining y to z also lies in both of these sets, i.e., if we define $L = \{\alpha y + (1-\alpha)z : \alpha \in [0, 1]\}$ then $L \subset x + dS_X \cap w + dS_X$.

If $u \in L$ then $d(x, u) \leq d$ and $d(u, w) \leq d$ by the convexity of the norm. Since $2d = \|x - w\| \leq \|x - w\| + \|u - w\| \leq 2d$, we see that $\|x - u\| = \|u - w\| = d$. Setting $u = \frac{y+z}{2}$ we see that $\|x - u\| = d$, a contradiction to strict convexity.

The following picture explains what is occurring in this proof (and is simpler to follow than the words). The line between y and z has to lie on the interior of both the ball of radius d around x and the interior of the ball of radius d around w thus the points have to be closer than the distance from x to z. But this was chosen to be minimal.

[10]Compare to Section 1.1.2.



□

We obtain the following corollary that gives a huge class of finite metric spaces that do not embed into any strictly convex Banach space:

Corollary 1.2.6. *If M is a metric space that does not have unique metric midpoints, it does not isometrically embed into any strictly convex Banach space.*

1.2.3 The Fréchet–Kuratowski Embedding

The Fréchet–Kuratowski embedding is a foundational result about the theory of metric spaces, dating back to the very first paper on metric spaces. The first proof the author is aware of is in [25], a 1935 paper by Kuratowski which is in French. Our interest will be in both the result and its method of proof.

Lemma 1.2.7. *Suppose that M is a metric space consisting of n points $\{m_1, \dots, m_n\}$. Then there is an isometric embedding f_M of M into ℓ_∞^n.*

Proof. We define a map $f_M = f : M \to \ell_\infty^n$ by $f(m_j) = (d(m_j, m_i))_{i=1}^n$. We then claim that this map is an isometry. Indeed, $\|f(m_j) - f(m_k)\|_\infty = \sup_i |d(m_j, m_i) - d(m_k, m_i)|$. By the reverse triangle inequality this is less than $d(m_j, m_k)$.

To see that the upper bound is attained, set $i = k$. Then we get $d(m_j, m_k) - 0 = d(m_j, m_k)$. Thus the mapping f is an isometry. □

Remark 1.2.8. We will need the structure of this proof later on. Specifically we will need the fact that if we perturb the distances in the metric space M, then $f_M(M)$ changes continuously.

To make this precise, suppose that we have an open set $V \subset U_n$ such that every element of V represents a metric space. If $\mathcal{M}$ is an element of V, let the metric space it represents be denoted by $M = \{m_1, \dots, m_n\}$. Then the mapping

$M \mapsto f_M(\{m_1, \ldots, m_n\})$, as a mapping from $V \to \underbrace{\ell_\infty^n \times \cdots \times \ell_\infty^n}_{n \text{ times}}$, is evidently a continuous mapping; co-ordinate wise it is $d_M(m_i, m_j) = \mathcal{M}_{ij}$.

Remark 1.2.9. The statement that "every n-point metric space embeds into ℓ_∞^n" can be sharpened. It is easy enough to see that we can embed every n-point metric space into ℓ_∞^{n-1}, simply by mapping one of the points to zero. Recent work in [37] has shown that every n-point metric space embeds into $\ell_\infty^{\sigma(n)}$ where $n - \sigma(n) \to \infty$ as $n \to \infty$. The exact rate of growth of $n - \sigma(n)$ is unknown, however Ball proved in [3] that $\sigma(n) = o(n)$.

1.2.4 The Cayley–Menger Determinant

In Chapter 3 we will consider embeddings of finite metric spaces M into ℓ_2. There is a classical algorithm that checks whether a metric space M embeds as an affinely independent subset of ℓ_2 that we will detail here. As in the case of the Fréchet–Kuratowski embedding we will be interested in the exact details of the proof, and thus we will give the proof as found in [41, Section 4].

If we have n distinct points $m_1, m_2, \ldots, m_n$ in a metric space M, we consider the associated matrix $\mathcal{M}$ defined, as above, by $\mathcal{M}_{ij} = d(m_i, m_j)$. Define $\mathcal{P}$ to be the $n \times n$ matrix given by $\mathcal{P}_{ij} = \mathcal{M}_{ij}^2$. Then the *Cayley–Menger determinant* of $m_1, \ldots, m_n$, shortened CMDet, is defined as

$$\mathrm{CMDet}(m_1, \ldots, m_n) = \det \begin{pmatrix} \mathcal{P} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{pmatrix}$$

where $\mathbf{1}$ is the column vector of length n consisting entirely of 1's.

Remark 1.2.10. The reader should observe that permutations of $m_1, \ldots, m_n$ do not affect the value of CMDet, indeed, when we transpose m_i and m_j, the determinant is unaffected.

We have the following sharp criterion:

Theorem 1.2.11. *A metric space M consisting of $n+1$ distinct points $m_0, \ldots, m_n$ can be isometrically embedded into ℓ_2^n as an affinely independent set if and only if the sign of CMDet$(m_0, \ldots, m_r)$ is $(-1)^{r+1}$ for each $r = 1, \ldots, n$.*

Proof. First suppose that $\{m_0, \ldots, m_n\}$ is an affinely independent subset of ℓ_2^n. We will show that the sign of CMDet$(m_0, \ldots, m_n)$ is $(-1)^{n+1}$. Write the j'th co-ordinate of m_i as $m_i^{(j)}$. Let Δ denote the simplex with vertices at m_i, i.e., Δ is the convex hull of $\{m_0, \ldots, m_n\}$.

It is classical that the n-volume of a simplex with vertices $\{m_0, \ldots, m_n\}$ is given by the formula:

$$\mathrm{vol}(\Delta) = \frac{1}{n!}\left|\det\begin{pmatrix} m_0 & m_1 & m_2 & \ldots & m_n \\ 1 & 1 & 1 & \ldots & 1 \end{pmatrix}\right|$$

where we think of m_i as a column vector.

We note that this is equal to

$$\mathrm{vol}(\Delta) = \frac{1}{n!}\left|\det\begin{pmatrix} m_0 & m_1 & m_2 & \ldots & m_n & 0 \\ 1 & 1 & 1 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix}\right|.$$

Squaring this formula and using the fact that $\det \mathcal{M} = \det \mathcal{M}^T$, we get that

$$(n!\,\mathrm{vol}(\Delta)))^2 = \det\begin{pmatrix} m_0^T & 1 & 0 \\ m_1^T & 1 & 0 \\ \vdots & \vdots & \vdots \\ m_n^T & 1 & 0 \\ \mathbf{0}^T & 0 & 1 \end{pmatrix} \det\begin{pmatrix} m_0 & \ldots & m_n & \mathbf{0} \\ 1 & \ldots & 1 & 0 \\ 0 & \ldots & 0 & 1 \end{pmatrix}$$

where $\mathbf{0}$ is the column vector of length n consisting entirely of zeros. Interchanging the last two columns of the first matrix gives,

$$-(n!\,\mathrm{vol}(\Delta)))^2 = \det\begin{pmatrix} m_0^T & 0 & 1 \\ m_1^T & 0 & 1 \\ \vdots & \vdots & \vdots \\ m_n^T & 0 & 1 \\ \mathbf{0}^T & 1 & 0 \end{pmatrix} \det\begin{pmatrix} m_0 & \ldots & m_n & \mathbf{0} \\ 1 & \ldots & 1 & 0 \\ 0 & \ldots & 0 & 1 \end{pmatrix}$$

We can then multiply these two matrices together, using the multiplicativity of the the determinant to get that

$$-(n!\,\mathrm{vol}(\Delta))^2 = \det\begin{pmatrix} \langle m_0, m_0\rangle & \langle m_0, m_1\rangle & \ldots & \langle m_0, m_n\rangle & 1 \\ \langle m_1, m_0\rangle & \langle m_1, m_1\rangle & \ldots & \langle m_1, m_n\rangle & 1 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ \langle m_n, m_0\rangle & \langle m_n, m_1\rangle & \ldots & \langle m_n, m_n\rangle & 1 \\ 1 & 1 & \ldots & 1 & 0 \end{pmatrix}$$

We can now expand $2\langle m_i, m_j\rangle = \langle m_i, m_i\rangle + \langle m_j, m_j\rangle - d(m_i, m_j)^2$, and subtract multiples of the last row/column to get that this is equal to

$$-(n!\,\mathrm{vol}(\Delta))^2 = \frac{(-1)^n}{2^{n+1}}\mathrm{CMDet}(m_0, \ldots, m_n).$$

This shows that if a metric space on $n+1$ points embeds into ℓ_2^n as an affinely independent set, then its Cayley Menger determinant has sign $(-1)^{n+1}$.

Remark 1.2.12. The above proof shows that if $m_0, \ldots, m_n$ embeds as an affinely dependent set into ℓ_2^n, then $\mathrm{CMDet}(m_0, \ldots, m_n) = 0$.

We now need to show the converse, i.e., if $\mathrm{CMDet}(m_0, \ldots, m_k)$ has sign $(-1)^{k+1}$ for each k then $m_0, \ldots, m_n$ embeds isometrically into ℓ_2^n as an affinely independent set. We prove this inductively. For $n = 0$ this theorem is obvious. We may assume inductively that the theorem is true for $n-1$, i.e., we have embedded $m_0, \ldots, m_{n-1}$ into $\overline{\mathrm{span}}\{e_1, \ldots, e_{n-1}\}$, where e_i is the standard orthonormal basis of ℓ_2^n. Suppose $m_0 \mapsto 0$, and $m_i \mapsto x_i$ with the span of x_i being all of ℓ_2^{n-1}.

We want to find a point $x_n \in \ell_2^n$, which we denote by x, such that $\|x\|_2 = \mathcal{M}_{0n}$ and $\|x - x_i\|_2 = \mathcal{M}_{in}$. Write $x = v + \lambda e_n$ where $v \in \ell_2^{n-1}$, i.e., $\langle v, e_n \rangle = 0$.

Squaring the equation $\|x - x_i\|_2 = \mathcal{M}_{in}$ we get $2\langle x_i, v \rangle = \|x\|_2^2 + \mathcal{M}_{0i}^2 - \mathcal{M}_{in}^2$, which is equal to $2\langle x_i, v \rangle = \mathcal{M}_{0n}^2 + \mathcal{M}_{0i}^2 - \mathcal{M}_{in}^2$. Since the x_i's are linearly independent, this is solvable, i.e., there is a unique solution v to this system of simultaneous equations.

Remark 1.2.13. We observe that this unique solution v varies continuously on the values of $\mathcal{M}_{i,j}$, e.g., by using the cofactor formula for the determinant.

The equation $\|x\|_2 = \mathcal{M}_{0n}$ can be squared and simplified to $\lambda^2 = \mathcal{M}_{0n}^2 - \|v\|_2^2$. It remains to show that $\mathcal{M}_{0n}^2 - \|v\|_2^2$ is positive.

Remark 1.2.14. We observe that the value of λ varies continuously on the values $\mathcal{M}_{i,j}$ (if $\mathcal{M}_{0,n}^2 - \|v\|_2^2$ is positive). It is just a square root.

By Remark 1.2.12, we have that the Cayley–Menger determinant of the points $0, x_1, \ldots, x_{n-1}, v$ is equal to zero, i.e.,

$$\det \begin{pmatrix} 0 & \mathcal{M}_{0,1}^2 & \ldots & \mathcal{M}_{0,n-1}^2 & \|v\|^2 & 1 \\ \mathcal{M}_{1,0}^2 & 0 & \ldots & \mathcal{M}_{1.n-1}^2 & \|v - x_1\|_2^2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathcal{M}_{n-1,0}^2 & \mathcal{M}_{n-1,1}^2 & \ldots & 0 & \|v - x_{n-1}\|^2 & 1 \\ \|v\|_2^2 & \|v - x_1\|_2^2 & \ldots & \|v - x_{n-1}\|_2^2 & 0 & 1 \\ 1 & 1 & \ldots & 1 & 1 & 0 \end{pmatrix} = 0$$

We have that $\|v - x_i\|_2^2 = \|v\|_2^2 - 2\langle v, x_i \rangle + \|x_i\|_2^2$, which is equal to $\|v\|_2^2 - \mathcal{M}_{0,n}^2 + \mathcal{M}_{in}^2$. Since the determinant is invariant under row and column operations we can subtract the final row multiplied by $(\|v\|_2^2 - \mathcal{M}_{0,n}^2)$ from the penultimate row and not affect the determinant. We also can subtract the final column multiplied by $(\|v\|_2^2 - \mathcal{M}_{0,n}^2)$ from the penultimate column and not affect the determinant. This gives us that

$$\det\begin{pmatrix} 0 & \mathcal{M}_{0,1}^2 & \dots & \mathcal{M}_{0,n-1}^2 & \mathcal{M}_{0,n}^2 & 1 \\ \mathcal{M}_{1,0}^2 & 0 & \dots & \mathcal{M}_{1.n-1}^2 & \mathcal{M}_{1,n}^2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathcal{M}_{n-1,0}^2 & \mathcal{M}_{n-1,1}^2 & \dots & 0 & \mathcal{M}_{n-1,n}^2 & 1 \\ \mathcal{M}_{0,n}^2 & \mathcal{M}_{1,n}^2 & \dots & \mathcal{M}_{n-1,n}^2 & -2(\|v\|_2^2 - \mathcal{M}_{0,n}^2) & 1 \\ 1 & 1 & \dots & 1 & 1 & 0 \end{pmatrix} = 0.$$

This determinant is the same as the determinant $\mathrm{CMDet}(m_0, \dots, m_n)$ except in the $(n+1, n+1)$ entry. We can write the penultimate column as

$$\begin{pmatrix} \mathcal{M}_{0,n}^2 \\ \mathcal{M}_{1,n}^2 \\ \vdots \\ \mathcal{M}_{n-1,n}^2 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -2(\|v\|_2^2 - \mathcal{M}_{0,n}^2) \\ 0 \end{pmatrix}.$$

We now use the multilinearity of the determinant in the $(n+1)$'st column and the fact that to obtain that

$$\mathrm{CMDet}(m_0, \dots, m_n) + \det\begin{pmatrix} 0 & \mathcal{M}_{0,1}^2 & \dots & \mathcal{M}_{0,n-1}^2 & 0 & 1 \\ \mathcal{M}_{1,0}^2 & 0 & \dots & \mathcal{M}_{1.n-1}^2 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathcal{M}_{n-1,0}^2 & \mathcal{M}_{n-1,1}^2 & \dots & 0 & 0 & 1 \\ \mathcal{M}_{0,n}^2 & \mathcal{M}_{1,n}^2 & \dots & \mathcal{M}_{n-1,n}^2 & -2(\|v\|_2^2 - \mathcal{M}_{0,n}^2) & 1 \\ 1 & 1 & \dots & 1 & 0 & 0 \end{pmatrix} = 0$$

and expanding the second determinant along the penultimate column gives us that

$$\mathrm{CMDet}(m_0, \dots, m_n) - 2(\|v\|_2^2 - \mathcal{M}_{0,n}^2)\mathrm{CMDet}(m_0, \dots, m_{n-1}) = 0.$$

Since the signs of $\mathrm{CMDet}(m_0, \dots, m_n)$ and $\mathrm{CMDet}(m_0, \dots, m_{n-1})$ are different by assumption, we necessarily have that $\|v\|_2^2 - \mathcal{M}_{0,n}^2 < 0$, which was as required. □

If M is a metric space satisfying the assumptions of Theorem 1.2.11, let g_M denote the embedding of M into ℓ_2^n that is constructed by following the proof of the previous theorem.

Remark 1.2.15. We can observe, similar to Remark 1.2.8 in the case of the Fréchet–Kuratowski embedding, that the embedding $g_M(M)$ is a continuous function of the distances involved in M. This is not quite as obvious as in the case of the Frechet–Kuratowski embedding but can be read out of the above proof by combining Remarks 1.2.13 and 1.2.14.

1.3 Assorted Extras

Throughout this thesis we will need some theorems from analysis and geometry that did not naturally fit into any of the earlier sections. We collect these results in this section.

1.3.1 Brouwer's Fixed Point Theorem

At key points in our arguments later on we will need Brouwer's fixed point theorem.

Theorem 1.3.1. *Let K be a non-empty compact convex subset of $\mathbb{R}^n$. Then every continuous function $f : K \to K$ has a fixed point.*

This forms a natural generalization of the intermediate value theorem. Indeed, in the case where $n = 1$ it can be recovered from this theorem.

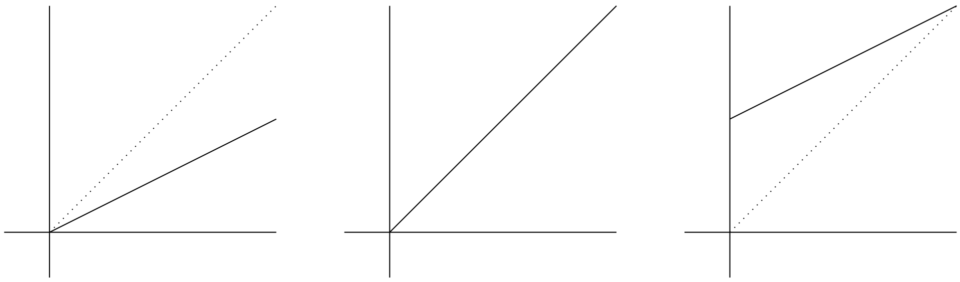
We will not give a proof of this theorem. Indeed, there are many proofs available and the interested reader can check [5, Chapter 15] for a self contained proof.

Even though the statement of the theorem is straightforward, there is an aspect of the theorem we would like to bring the reader's attention to: the map sending a function to its fixed point is not well behaved.

Remark 1.3.2. Suppose that K is a compact convex subset of $\mathbb{R}^n$, and that, for $r \in [0,1]$, f_r is a family of continuous functions $f_r : K \to K$. Moreover, suppose that f_r depends continuously on r, i.e., the function $f : [0,1] \times K \to K$ given by $f(r,x) = f_r(x)$ is a continuous function. Then there is a set-valued function $\Sigma : [0,1] \to 2^K$ that sends a function f_r to the set of its fixed points. We may ask a very natural question: is there a continuous map $\sigma : [0,1] \to K$ such that

$\sigma(r) \in \Sigma(r)$?

The answer in general is no. We give an example for $n = 1$ and $K = [0,1]$. Then consider maps defined by: $f_r(x) = (2r)x$ for $r \leq 1/2$ and $f_r(x) = (2r-1) + (2-2r)x$ for $r \geq 1/2$.



$r < 1/2$ $\qquad$ $r = 1/2$ $\qquad$ $r > 1/2$

It is now easy to see that the map σ associated to this family of functions must be discontinuous. Indeed, $\sigma(r) = 0$ for $r < 1/2$ and $\sigma(r) = 1$ for $r > 1/2$.

There are more pathological examples that show that the function sending a function to its set of fixed point must be badly behaved. There are set theoretic results saying (very roughly) that if a function is of a certain 'complexity' its fixed points can be of higher 'complexity'. There are examples showing, for example, that if we have a function $f : [0,1]^2 \to [0,1]^2$ that is *computable* its fixed points can fail to be computable.

The fact that this map is badly behaved will illustrate why some of our methods of proof later can not be extended.

The poor behaviour of Brouwer's Fixed Point Theorem should be contrasted with the good behaviour of the Contraction Mapping Theorem.

Theorem 1.3.3. *Let (X,d) be a non-empty complete metric space, $0 \leq \lambda < 1$, $f : X \to X$ and suppose f satisfies $d(f(x), f(y)) \leq \lambda d(x,y)$. Then there is a unique fixed point x^* of f. Moreover, $x^* = \lim_n f^{(n)}(x)$ for any $x \in X$, where $f^{(n)}$ is the n-fold iterate of f with itself.*

We call a function with the property that $d(f(x), f(y)) \leq \lambda d(x,y)$ a *λ-contraction*.

The uniqueness of the fixed point in this case, coupled with the property of λ-contractiveness allows us to show that the map sending a λ-contraction to its unique fixed point is well behaved. Indeed, if f and g are two $\lambda-$contractions, with fixed points x_f, x_g, respectively, then

$$d(x_f, x_g) \leq d(f(x_f), f(x_g)) + d(f(x_g), g(x_g)).$$

This shows that $d(x_f, x_g) \leq \frac{1}{1-\lambda} \sup_x(d(f(x), g(x)))$. For compact metric spaces this shows that the map sending a contraction to its fixed point is continuous, with respect to the supremum distance.

Remark 1.3.4. There is a remarkable counter-point to Remark 1.3.2 whose existence I learned of on MathOverflow, see [17].

Theorem 1.3.5. *Let K be a non-empty, compact, convex subset of $\mathbb{R}^n$ and suppose that $\{f_i : i \in I\}$ is a family of continuous functions on K indexed by some compact subset of I of $\mathbb{R}^m$ (for m and n possibly different.) Suppose that f_i depends continuously on I, i.e. that the function $G : I \times K \to K$ given by $G(i, x)$ is a continuous function. Moreover, suppose that the fixed points of such functions are unique. Then the map sending a function to its fixed point is continuous in i.*

This theorem will follow easily from the Closed Graph Theorem for topological spaces which we now recall.

Theorem 1.3.6. *Suppose that $f : S \to T$ is a map from a topological space S to a compact Hausdorff space T. Then f is continuous if and only if the graph $\Gamma(f) = \{(s, f(s)) : s \in S\}$ is closed.*

Proof. Let E be the set $\{(i, x) : G(i, x) - x = 0\}$. It is easy to see that E is closed, and that E is the graph of the function that sends a function f_i to its fixed point. Applying Theorem 1.3.6 we see that the function that sends a function to its fixed point is continuous. □

1.3.2 The Inverse Function Theorem and The Submersion Theorem

At various points in the text we will need to use differentiability arguments. To fix notation and terminology, we say that a function $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ is a C^1-map if Φ is continuous, differentiable, and the map that sends $x \mapsto (D\Phi)(x)$ is continuous.

We begin by recalling the Inverse Function Theorem.

Theorem 1.3.7. *Let $x \in \mathbb{R}^n$ and W be an open neighbourhood of x. Let $\Phi : W \to \mathbb{R}^n$ be a C^1-map such that $(D\Phi)(x)$ is an invertible linear map.*

Then there is an open subset U of x, an open subset V of $\Phi(x)$ and a map $\Psi : V \to U$ such that $\Phi|_U$ is a bijection, $\Phi(\Psi(z)) = z$ for all $z \in V$ and $\Psi(\Phi(z)) = z$ for all $z \in U$.

The Submersion Theorem and the Immersion Theorem are close relatives of the Inverse Function Theorem. Both relate the behaviour of a function at a point to the rank of the derivative there. A function $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ is a *submersion at a point* x if $(D\Phi)(x)$ is surjective, i.e., has rank m. This necessarily means that $n \geq m$. This is dual to the concept of an *immersion at a point* x where $(D\Phi)(x)$ is injective, i.e., has rank n. This necessarily means that $m \geq n$.

Theorem 1.3.8. *Suppose* $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ *is a* $C^1 - map$.

- *If* Φ *is a submersion at a point* x*, there is an open set* U *containing* x*, and an open set* V *containing* $\Phi(x)$ *such that that* $\Phi(U) = V$*. Moreover, there is a continuous map* $\Psi : V \to U$ *such that* $\Phi \circ \Psi$ *is the identity and* $\Psi(\Phi(x)) = x$*.*

- *If* Φ *is an immersion at a point* x*, there is an open set* U *containing* x*, and an open set* V *containing* $\Phi(x)$ *such that* $\Phi(U) \subset V$*. Moreover, there is a continuous map* $\Psi : V \to U$ *such that* $\Psi(V) = U$ *and* $\Psi \circ \Phi$ *is the identity on* U*.*

Remark 1.3.9. The following consequence of Theorem 1.3.8 is important for our purposes: if a map $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ is a submersion at a point x then it is locally surjective, i.e., there is an open set V around $\Phi(x)$ such that $\Phi(\mathbb{R}^n)$ contains V.

1.3.3 Green's Theorem and a Corollary

At some point in the text, we will need to use an argument based on integration. A key technical ingredient in these arguments will be Green's Theorem. We will mainly be interested in these arguments in the context of Lemma 1.1.23, i.e., we will have a closed convex curve that we will be the boundary of the unit sphere of a 2-dimensional Banach space.

Theorem 1.3.10. *Let* C *be a simple closed smooth curve in* $\mathbb{R}^2$*, bounding a region* D*. Moreover, suppose that we take* C *being traversed counter clockwise. Suppose that* Θ, Ψ *are real-valued differentiable functions of* x, y *in an open region containing* D*. Then*

$$\oint_C \Theta\, dx + \Psi\, dy = \iint_D \partial_x \Psi - \partial_y \Theta \; dxdy.$$

An incredibly simple corollary of this is:

Corollary 1.3.11. *Suppose that* D *is a convex subset of the plane bounded by a piecewise smooth curve* Γ*. Then*

$$-\oint_\Gamma y\, dx = \oint_\Gamma x\, dy = Area(D).$$

1.4 Some Preliminary Observations about the Property AI-I

This section contains our first observations about Property AI-I. We first recall the definition of Property AI-I, which the reader is now equipped to understand. We say that a Banach space X has *Property AI-I* if whenever M is a finite metric space such that M almost isometrically embeds into X then M isometrically embeds into X. We say that a Banach space X has the *Finite Isometric Dvoretzky Property* if every finite subset of ℓ_2 isometrically embeds into X. We say that a Banach space X has the Finite Isometric Krivine Property if, whenever ℓ_p is finitely representable in X, every finite subset of ℓ_p embeds isometrically into X.[11]

This section will be a little vaguer in places than the rest of the introduction, however the reader should be satisfied that every question we pose will be made precise. We will also discuss a possible generalization of Property AI-I.

Theorem 1.4.1. *The following spaces have Property AI-I:*

1. *Any space that contains every finite metric space isometrically;*
2. ℓ_2*;*
3. ℓ_p *for* $1 \leq p \leq \infty$*;*
4. *any finite dimensional space.*

The only hard one of these is ℓ_p for $1 < p < \infty$ which follows from Theorem 1.1.6 and compactness.

1.4.1 Some Questions

We will now list some questions that we will answer throughout this thesis.

Question 1.4.2.
1. *Does every Banach space have Property AI-I?*
2. *What classes of Banach spaces have Property AI-I?*
3. *Does every reflexive Banach space have Property AI-I? Every superreflexive space?*
4. *How does Property AI-I relate to the properties of bases that Banach spaces have?*

[11] These definitions make sense in light of Dvoretzky's Theorem (Theorem 1.1.29) and Krivine's Theorem (Theorem 1.1.32.)

5. *Is Property AI-I 'open'? If two Banach spaces are close, then does one having Property AI-I imply the other does?*

6. *Does every Banach space have the Finite Isometric Krivine Property?*

During this thesis we will answer Questions 1, 3, 5 and 6 in the negative. Due to the counterexamples constructed here, the following conjecture is natural:
Conjecture 1.4.3. *Suppose that X is an infinite-dimensional Banach space with Property AI-I and that $X \neq \ell_2$. Then for every $\epsilon > 0$ there is a Banach space Y with $d(X, Y) < 1 + \epsilon$ which lacks Property AI-I.*

The reader should bear this conjecture in mind, especially within Section 4, where we show that if $X = \ell_p$, then it holds.

In Section 3 we will discuss the Dvoretzky analogue of the sixth question above, i.e., we will conjecture:
Conjecture 1.4.4. *Every Banach space has the Finite Isometric Dvoretzky Property.*

1.4.2 A Generalisation

The reader should notice that Property AI-I only makes reference to *finite* metric spaces. We can ask the following analogue of part 2 of Question 1.4.2:
Question 1.4.5. *Suppose that X is a Banach space. What classes of Banach spaces X have the property that whenever M is a countable metric space that almost isometrically embeds into X, M isometrically embeds into X?*

We say such a Banach space has Property AI-I(countable). We can provide some insights into this question here that the reader may find interesting.

We first note that if X is a Banach space that is isometrically universal for all countable metric spaces then X has Property AI-I(countable). This means implies that ℓ_∞ has this property.

Second we note that ℓ_2 has this property. Indeed, suppose we have a metric space $M = \{m_1, m_2, \dots\}$ such that for every $\epsilon > 0$ we have a map $f_\epsilon : M \to \ell_2$ such that $d(m_i, m_j) \leq \|f_\epsilon(m_i) - f_\epsilon(m_j)\|_2 \leq (1+\epsilon)d(m_i, m_j)$. Fix some orthonormal basis $e_1, e_2, \dots$ of ℓ_2. Then by translation and the Gram-Schmidt process we can assume that $f_\epsilon(m_1) = 0$ and $f_\epsilon(m_i) \in \operatorname{span}\{e_1, \dots, e_{i-1}\}$.

Now focus on $f_{1/n}$. We can pass to subsequences, and use a diagonal argument, to construct some subsequence f_{1/n_j} such that $(f_{1/n_j}(m_i))_j$ is convergent for each

i. Setting $f(m_i) = \lim_j f_{1/n_j}(m_i)$ yields an isometric embedding.

However, have the following two theorems:

Theorem 1.4.6. *c_0 does not have Property AI-I(countable).*

Proof. This follows from work of Kalton and Lancien in [21]. Specifically Theorem 2.9 of this paper exhibits a countable metric space M that embeds into c_0 with distortion $1+\epsilon$ for every ϵ positive, but does not isometrically embed into c_0. $\square$

Theorem 1.4.7. *Any strictly convex Banach space that does not contain any subspace isometrically isomorphic to ℓ_2^2 does not have Property AI-I(countable), in particular ℓ_p for $1 < p < \infty$ and $p \notin 2\mathbb{N}$ does not have Property AI-I(countable).*

We will give the proof of the first part of this theorem later, in Theorem 3.3.5. To deduce the second part we need to observe that ℓ_2^2 does not isometrically isomorphically embed into ℓ_p which was proven in [10].

These two results seem to indicate that the class of spaces that have Property AI-I(countable) is small. Indeed, the author can not think of a space that is not a Hilbert space[12] or universal for all countable metric spaces that has Property AI-I(countable). We give this as a conjecture, to which we have neither dedicated particularly large amounts of time, nor have any idea how to approach:

Conjecture 1.4.8. *The only spaces with Property AI-I countable are those that are either isometrically isomorphic to a Hilbert space or those that are isometrically universal for countable metric spaces.*

Although this is not the main focus of our thesis we feel that this is an interesting open problem. One can ask similar questions for metric spaces of higher and higher cardinality where we would expect the classes of spaces would get narrower and narrower. This is an interesting problem which the author has not given much attention to.

[12]'A' Hilbert space here means 'a Hilbert space of arbitrary cardinality'.

Chapter 2

The ℓ_∞ Case

We begin by talking about isometric embeddings of finite metric spaces into a Banach space with no non-trivial cotype. We begin with this case because it is simpler than the other cases that we will consider later. This is primarily due to the Fréchet-Kuratowski embedding which makes all of our considerations very explicit. This allows us to give a more in-depth exposition and explain the method of proof that we generalise later.

2.1 The Setup

We begin by considering an infinite-dimensional Banach space X that has no non-trivial cotype, or equivalently (by the Maurey-Pisier Theorem, Theorem 1.1.36) a Banach space X for which ℓ_∞ is finitely representable in X. We observe that for any n-point metric space M, and any $\epsilon > 0$ there is a $(1+\epsilon)-$distortion embedding of M into X. Indeed, the Fréchet-Kuratowski embedding (see Lemma 1.2.7) shows that there is an isometric embedding f of M into ℓ_∞^n. Then, taking a subspace E of X for which $d(E, \ell_\infty^n) < 1+\epsilon$ gives us a map $T : \ell_\infty^n \to E$ for which $\|T\|\|T^{-1}\| < 1+\epsilon$. Scaling, we can assume that $\|T\| = 1$ and $\|T^{-1}\| < 1+\epsilon$.

The composite map $T \circ f : M \to E \leq X$ is the required $(1+\epsilon)-$distortion embedding of M into X.

We are thus in the position where any finite metric space almost-isometrically embeds into X. The question of whether X has Property AI-I is thus the question of whether X is isometrically universal for all finite metric spaces. This brings us to the question that drives our study for the rest of the chapter:

Question 2.1.1. *Suppose X is an infinite-dimensional Banach space that lacks non-trivial cotype. Then is X isometrically universal for all finite metric spaces?*

We structure the rest of the chapter as follows: we begin by proving a positive result for finite concave metric spaces. After this we give a counterexample showing that the positive result does not extend to all finite metric spaces. To end the chapter we discuss an extension of our positive result to the case of a certain subclass of infinite concave metric spaces.

2.2 A Positive Result

We will begin by talking about *concave* metric spaces. We recall our definition of concave metric spaces from Section 1.2.2: a metric M space is *concave* if for every distinct triple x, y, z we have that $d(x, y) + d(y, z) > d(x, y)$.

Theorem 2.2.1. *Let $n \in \mathbb{N}$. Suppose that M is an n-point concave metric space, and X is an infinite-dimensional Banach space for which ℓ_∞^n almost isometrically embeds into X. Then M isometrically embeds into X.*

The hypothesis of this theorem is implied by any of the following assumptions in increasing[1] order of strength: X has no non-trivial cotype, X uniformly contains the ℓ_∞^n's, c_0 embeds into X or ℓ_∞ embeds into X. We reduce Theorem 2.2.1 to the following quantitative finite dimensional result:

Theorem 2.2.2. *Let $n \in \mathbb{N}$. If M is an n-point concave metric space then there is some $\delta > 0$ such that if E is an n-dimensional Banach space with $d(\ell_\infty^n, E) < 1 + \delta$, then M isometrically embeds into E.*

We will give three proofs of this result - even though the proofs are substantively similar we wish to give all three to illustrate different aspects of the proof.

The first proof will be a proof similar to the author's work in [23]. It is less clear conceptually to extend to the case of infinite metric spaces, but has the advantage of being the best indicator of a proof strategy for the ℓ_p case.

The second proof is the cleanest proof, taken from the author's work in [22]. It is the shortest and gives us the best insight as to how to extend to the case of infinite metric spaces.

[1] Increasing, but not strictly increasing. The first two are equivalent by the Maurey-Pisier theorem.

The above two proofs use Brouwer's fixed point theorem to find an isometric embedding of the metric space M. In both of these proofs we simultaneously find the image points, i.e., we apply Brouwer's fixed point theorem once and the resulting fixed point is the image of M. In fact, we can do this iteratively, i.e., we have an isometric embedding of $n-1$ points that we extend to an isometric embedding of all n points (again by Brouwer's fixed point theorem). In the third proof we give the details of this. We also attempt to explain why the choice of perturbation (which is present in all three proofs) is the most natural choice of perturbation.

Proof One of Theorem 2.2.2

Let M consist of points $\{m_1, \ldots, m_n\}$ and set $\mathcal{M}$ to be matrix that represents the space M, i.e., $\mathcal{M}_{ij} = d(m_i, m_j)$. We remind the reader that U_n is the space of $n \times n$ symmetric matrices with zeroes on the diagonal. Let $A \subset U_n$ be the collection of matrices that represent concave n-point metric spaces. In the following we will take the ℓ_∞ norm on $\mathrm{Mat}_{n\times n}(\mathbb{R})$ as it is the most convenient computationally.

We will first show that A is an open subset of U_n. Indeed, a metric space M is convex if and only if $\mathcal{M}_{ij} + \mathcal{M}_{jk} - \mathcal{M}_{ik} > 0$ for each distinct triple of numbers i, j, k. Thus $A = \cap_{i,j,k}\{\mathcal{N} \in U_n : \mathcal{N}_{ij} + \mathcal{N}_{jk} - \mathcal{N}_{ik} > 0\}$. This is an intersection of open sets and thus A is an open subset of $\mathrm{Mat}_{n\times n}(\mathbb{R})$.

Since A is an open subset of U_n, there is some $\eta > 0$ such that the set $B = \{\mathcal{N} \in U_n : \|\mathcal{N} - \mathcal{M}\|_\infty < \eta\} \subset A$.

For $\mathcal{N} \in B$, with $\mathcal{N} = \{r_1, \ldots, r_n\}$ we consider the points $x_i(\mathcal{N}) = (d(r_j, r_i))_{j=1}^n \in \ell_\infty^n$. We recall from Section 1.2.3 that this is the Fréchet-Kuratowski embedding of N into ℓ_∞^n. This mapping satisfies $\|x_i(\mathcal{N}) - x_j(\mathcal{N})\|_\infty = d(x_i, x_j)$. Moreover, the mapping $\mathcal{N} \mapsto (x_i(\mathcal{N}))_{i=1}^n$ is continuous when considered as a mapping from B to $(\ell_\infty^n)^n$ (see Remark 1.2.8).

We have now shown that if M is a concave metric space, then there is some $\eta > 0$ such that 'η-small perturbations' of M remain finite concave metric spaces, and thus still embed into ℓ_∞^n. We now wish to prove a dual statement to this, namely, that if we consider a finite-dimensional Banach space E that is a small perturbation of ℓ_∞^n, then we can find an isometric copy of M inside E.

We fix $\delta > 0$ (to be determined later) and consider a Banach space E with the property that $d(E, \ell_\infty^n) < 1 + \delta$. Without loss of generality we may assume that

$E = (\mathbb{R}^n, \|.\|_E)$ with $\|.\|_E \leq \|.\|_\infty \leq (1+\delta)\|.\|_E$. Define $\iota : \ell_\infty^n \to E$ to be the formal identity map. We now define a map $\Theta : [0,\eta]^{\binom{n}{2}} \to E^n$ by

$$\Theta(\mathcal{N}) = (\iota(x_i(\mathcal{N} + \mathcal{M})))_{i=1}^n$$

where an element of $[0,\eta]^{\binom{n}{2}}$ is indexed by $\{(i,j) : 1 \leq i < j \leq n\}$ which we identify as an element of U_n in the obvious way.

We then consider the map $\varphi : [0,\eta]^{\binom{n}{2}} \to \mathbb{R}^{\binom{n}{2}}$ given by

$$\varphi(\mathcal{N})_{ij} = \mathcal{N}_{ij} + \mathcal{M}_{ij} - \|\Theta(\mathcal{N})_i - \Theta(\mathcal{N})_j\|_E.$$

We claim that if δ is sufficiently small (and the size of δ will depend only on $\mathcal{M}$) then $\varphi(\mathcal{N}) \in [0,\eta]^{\binom{n}{2}}$.

To establish the lower bound, note that

$$\|\Theta(\mathcal{N})_i - \Theta(\mathcal{N})_j\|_E \leq \|\Theta(\mathcal{N})_i - \Theta(\mathcal{N})_j\|_\infty = \mathcal{N}_{ij} + \mathcal{M}_{ij}$$

and thus $\varphi(\mathcal{N})_{ij} \geq 0$. For the upper bound, we have that

$$\|\Theta(\mathcal{N})_i - \Theta(\mathcal{N})_j\|_E \geq \frac{1}{1+\delta}\|\Theta(\mathcal{N})_i - \Theta(\mathcal{N})_j\|_\infty = \frac{1}{1+\delta}(\mathcal{N}_{ij} + \mathcal{M}_{ij}).$$

We thus have that

$$\varphi(\mathcal{N})_{ij} \leq \mathcal{N}_{ij} + \mathcal{M}_{ij} - \frac{1}{1+\delta}(\mathcal{N}_{ij} + \mathcal{M}_{ij}) = \frac{\delta}{\delta+1}(\mathcal{N}_{ij} + \mathcal{M}_{ij})$$

and thus if δ is sufficiently small, then this is $< \eta$.

By construction, φ is a continuous mapping from $[0,\eta]^{\binom{n}{2}}$ to itself. The set $[0,\eta]^{\binom{n}{2}}$ is a compact convex subset of $\mathbb{R}^{\binom{n}{2}}$ and thus by the Brouwer fixed point theorem, there is some $\mathcal{N}$ such that $\varphi(\mathcal{N}) = \mathcal{N}$. For this choice of $\mathcal{N}$ we have that $\|\Theta(\mathcal{N})_i - \Theta(\mathcal{N})_j\|_E = \mathcal{M}_{ij}$, and thus $m_i \mapsto \Theta(\mathcal{N})_i$ is the required isometric embedding of M into E. □

Remark 2.2.3. This proof really demonstrates the role of concavity in the argument - the assumption of concavity shows that there is an open ball *in* U_n such that every element of this ball represents a concave metric space. If a metric space is not concave, then for any $\epsilon > 0$, there is a matrix $\mathcal{N} \in U_n$ with $\|\mathcal{N} - \mathcal{M}\|_\infty < \epsilon$ and $\mathcal{N}$ not representing a metric space.

We reiterate this - it is not just the case that $\mathcal{N}$ does not represent a concave metric space, it does not have to represent a metric space at all.

Indeed, consider a metric space on three points, x, y, z with $d(x,y) = d(y,z) = 1$ and $d(x,z) = 2$. This has associated matrix

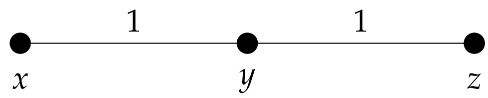
$$\mathcal{M} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

but then, for any $\epsilon > 0$, we can consider the matrix

$$\mathcal{M}_\epsilon = \begin{pmatrix} 0 & 1 & 2+\epsilon \\ 1 & 0 & 1 \\ 2+\epsilon & 1 & 0 \end{pmatrix}.$$

This is a bona fide element of U_n, but does not represent a metric space.

Pictorially M is represented by:



Since the points of M are collinear in this fashion, we can not move x and z further apart without affecting the distances from x to y and x to z.

Proof Two of Theorem 2.2.2

Let M consist of the points $\{m_1, \ldots, m_n\}$. Since M is concave and finite we can fix some $\eta > 0$ such that for any distinct $x, y, z \in M$, we have that $d(x,y) + d(y,z) - d(x,z) > 2\eta$.

We fix $\delta > 0$ (to be determined later) and consider a Banach space E with the property that $d(E, \ell_\infty^n) < 1 + \delta$. Without loss of generality we may assume that $E = (\mathbb{R}^n, \|.\|_E)$ with $\|.\|_E \leq \|.\|_\infty \leq (1+\delta)\|.\|_E$.

We define a map $\Theta : [0,\eta]^{\binom{n}{2}} \to \underbrace{E \times \cdots \times E}_{n \text{ times}}$ as follows:

$$\Theta(\boldsymbol{\epsilon})_1 = (d(m_1,m_1), d(m_2,m_1), d(m_3,m_1) \ldots, d(m_n,m_1))$$
$$\Theta(\boldsymbol{\epsilon})_2 = (d(m_1,m_2) + \epsilon_{1,2}, d(m_2,m_2), d(m_3,m_2), \ldots, d(m_n,m_2))$$
$$\Theta(\boldsymbol{\epsilon})_3 = (d(m_1,m_3) + \epsilon_{1,3}, d(m_2,m_3) + \epsilon_{2,3}, d(m_3,m_3), \ldots, d(m_3,m_n))$$
$$\vdots$$
$$\Theta(\boldsymbol{\epsilon})_n = (d(m_1,m_n) + \epsilon_{1,n}, d(m_2,m_n) + \epsilon_{2,n}, \ldots, d(m_{n-1},m_n) + \epsilon_{n,n-1}, d(m_n,m_n))$$

We note that this is the Fréchet embedding of M into ℓ_∞^n in the case that $\boldsymbol{\epsilon} = 0$. For $i \neq j$ we have that

$$\|\Theta(\boldsymbol{\epsilon})_i - \Theta(\boldsymbol{\epsilon})_j\|_\infty = \sup_k |d(m_k, m_i) - d(m_k, m_j) + \epsilon_{i,k} - \epsilon_{j,k}|$$

where we set $\epsilon_{s,t} = 0$ for $s \geq t$. This is equal to $d(m_i, m_j) + \epsilon_{ij}$ as if $k \neq i, j$ we have that

$$d(m_i, m_k) - d(m_j, m_k) + \epsilon_{ik} - \epsilon_{kj} \leq d(m_i, m_j) - 2\eta + \epsilon_{ik} - \epsilon_{jk} \leq d(m_i, m_j)$$

where the first inequality is where the hypothesis of concavity is needed.

We now continue similarly to the previous proof, with different notation. Define a map $\varphi : [0, \eta]^{\binom{n}{2}} \to \mathbb{R}^{\binom{n}{2}}$ by

$$\varphi(\boldsymbol{\epsilon})_{i,j} = d(m_i, m_j) + \epsilon_{ij} - \|\Theta(\boldsymbol{\epsilon})_i - \Theta(\boldsymbol{\epsilon})_j\|_E.$$

We claim that, if δ is sufficiently small φ sends $[0, \eta]^{\binom{n}{2}}$ to itself.

To show that $\varphi(\boldsymbol{\epsilon})_{i,j} > 0$ we note that

$$\|\Theta(\boldsymbol{\epsilon})_i - \Theta(\boldsymbol{\epsilon})_j\|_E \leq \|\Theta(\boldsymbol{\epsilon})_i - \Theta(\boldsymbol{\epsilon})_j\|_\infty = d(m_i, m_j) + \epsilon_{i,j}$$

where the last equality is due to the above computation.

To show that $\varphi(\boldsymbol{\epsilon})_{ij} < \eta$ for sufficiently small δ, we note that

$$\|\Theta(\boldsymbol{\epsilon})_i - \Theta(\boldsymbol{\epsilon})_j\|_E \geq \frac{1}{1+\delta}\|\Theta(\boldsymbol{\epsilon})_i - \Theta(\boldsymbol{\epsilon})_j\|_\infty = \frac{1}{1+\delta}(d(m_i, m_j) + \epsilon_{i,j})$$

and thus

$$\varphi(\boldsymbol{\epsilon})_{i,j} \leq \frac{\delta}{1+\delta}(d(m_i, m_j) + \epsilon_{i,j})$$

and thus whenever δ is sufficiently small $\varphi(\boldsymbol{\epsilon})_{i,j}$ is less than η.

By construction, φ is a continuous mapping from $[0, \eta]^{\binom{n}{2}}$ to itself. The set $[0, \eta]^{\binom{n}{2}}$ is a compact convex subset of $\mathbb{R}^{\binom{n}{2}}$ and thus by the Brouwer fixed point theorem, there is some $\boldsymbol{\epsilon}$ such that $\varphi(\boldsymbol{\epsilon}) = \boldsymbol{\epsilon}$. For this choice of $\boldsymbol{\epsilon}$ we have that $\|\Theta(\boldsymbol{\epsilon})_i - \Theta(\boldsymbol{\epsilon})_j\|_E = d(m_i, m_j)$, and thus $m_i \mapsto \Theta(\boldsymbol{\epsilon})_i$ is the required isometric embedding of M into E. $\square$

Proof Three of Theorem 2.2.2

In this proof, our goal is to more clearly understand the constructions given in Proof Two above so that we can generalise them to give a (slightly different) result in the case of infinite metric spaces (see Section 2.5 for these details.) Our proof here is essentially identical to Proof Two above, however we choose the $\epsilon_{i,j}$ inductively. This small tweak will allow us to generalise to an infinite result.

Let M consist of the points $\{m_1, \ldots, m_n\}$ and identify M with its image in ℓ_∞^n under the Frechét-Kolmogorov emebdding, i.e., $m_i = (d(m_1, m_i), \ldots, d(m_n, m_i))$. Suppose that $E = \mathbb{R}^n$ with a norm $\|.\|_E$ that is 'close' to the ℓ_∞^n norm in the Banach-Mazur sense, i.e., we can (without loss of generality) assume that there is some constant $\delta > 0$ such that

$$\|.\|_E \leq \|.\|_\infty \leq (1+\delta)\|.\|_E. \tag{2.1}$$

This assumption is a normalisation of the norm $\|.\|_E$. We aim to show that if δ is sufficiently small (depending on M) then M isometrically embeds into E.

In what follows, we are going to attempt a very "bare hands" inductive proof - at stage i we will find a possible image $m_i' \in E$ of m_i. At each stage we will try and enforce that, for $j < i$, $\|m_i' - m_j'\|_E = d(m_i, m_j)$.

For m_1, the isometry condition is vacuous, we may as well set

$$m_1' = (d(m_1, m_1), \ldots, d(m_n, m_1)) = m_1.$$

What do we do with m_2? We have n co-ordinates to play with, so we could send them anywhere. The simplest thing we could do is change only one co-ordinate. This is a degrees of freedom argument: we have one condition (namely $\|m_2' - m_1'\| = d(m_1, m_2)$) and thus only need to change one thing to force this condition.

So, which co-ordinate do we change? First let's see what happens if we change the first co-ordinate.

Changing the first co-ordinate

Fix $R > 0$ (to be determined later) and for ϵ in $[0, R]$ define

$$f(\epsilon) = (d(m_1, m_2) + \epsilon, d(m_2, m_2), d(m_3, m_2), \ldots)$$

which is an ϵ-perturbation of m_2. We note that f is a continuous function from $[-R, R]$ to E.

We are looking for some ϵ such that $\|f(\epsilon) - m_1'\|_E = d(m_2, m_1)$. To find this we are going to estimate $\|f(0) - m_1'\|_E$ and $\|f(R) - m_1'\|_E$ and show that one of these is greater than $d(m_2, m_1)$ and one of these is less than $d(m_2, m_1)$.

So, what is $\|f(0) - m_1'\|_E$? Since

$$\|f(0) - m_1'\|_\infty = \|m_2 - m_1\|_\infty = d(m_1, m_2)$$

we can use our choice of normalisation (2.1) to see that

$$\frac{1}{1+\delta} d(m_1, m_2) \leq \|f(0) - m_1'\|_E \leq d(m_1, m_2).$$

The main thing to observe is that $\|f(0) - m_1'\|_E \leq d(m_1, m_2)$.

Now we look at $\|f(R) - m_1'\|_E$. Similar to the above, we have to work out what $\|f(R) - m_1'\|_\infty$ is. We can write this out explicitly as

$$\begin{aligned}\|(d(m_1, m_2) + R - d(m_1, m_1), d(m_2, m_2) - d(m_2, m_1),\\ d(m_3, m_2) - d(m_3, m_1), \dots, d(m_n, m_2) - d(m_n, m_1)\|_\infty.\end{aligned}$$

This supremum is easy to compute since the first term is $d(m_1, m_2) + R$, and every other term can be bounded above using the triangle inequality by $d(m_1, m_2)$. We thus have that $\|f(R) - m_1'\|_\infty = d(m_1, m_2) + R$.

Using the choice of normalisation (2.1) we see that

$$\frac{1}{1+\delta}(d(m_1, m_2) + R) \leq \|f(R) - m_1'\| \leq d(m_1, m_2) + R.$$

If we choose R and δ such that

$$\frac{d(m_1, m_2) + R}{1+\delta} > d(m_1, m_2)$$

then $\|f(R) - m_1'\|_E$ is larger than $d(m_1, m_2)$.

We wish to apply the intermediate value theorem one some interval to find a point ϵ such that $\|f(\epsilon) - m_1'\|_E = d(m_1, m_2)$. So, define a function $\varphi : [0, R] \to \mathbb{R}$ by

$$\varphi(\epsilon) = d(m_1, m_2) - \|f(\epsilon) - m_1'\|_E.$$

It is clear that φ is a continuous function. Our estimates above show that $\varphi(0) \geq 0$ and that (if δ is sufficiently small) $\varphi(R) \leq 0$. So there is some ϵ for which $\varphi(\epsilon) = 0$, and we set

$$m_2' = (d(m_1, m_2) + \epsilon, d(m_2, m_2), d(m_3, m_2), \dots).$$

Changing a different co-ordinate

We could have, instead, opted to alter a different co-ordinate. What would have happened then?

Changing the second co-ordinate of m_2' is the same thing as changing the first co-ordinate after interchanging the roles of m_1 and m_2 in the proof above. We thus do not consider changes in the second co-ordinate.

We will begin by considering the case that we alter some co-ordinate that is not the second - let us imagine changing the third (the same argument works for any co-ordinate that isn't the first or second.) Then we would set

$$m_2' = (d(m_1, m_2), d(m_2, m_2), d(m_3, m_2) + K, d(m_4, m_2), \dots)$$

for some K, and we want that $\|m_1' - m_2'\|_E = d(m_1, m_2)$. Then

$$\begin{aligned}\|m_1' - m_2'\|_\infty &= \max(\sup_{i\neq 3} |d(m_i, m_2) - d(m_i, m_1)|, |d(m_3, m_2) + K - d(m_3, m_1)| \\ &= \max(d(m_1, m_2), |d(m_3, m_2) + K - d(m_3, m_1)|).\end{aligned}$$

Since the metric space is concave, we may assume that there is some $\eta > 0$ such that for every triple i, j, k we have that $d(m_i, m_j) + d(m_j, m_k) > d(m_i, m_k) + 2\eta$[2]. Now if $|K| < \eta$ we have that $\|m_1' - m_2'\|_\infty = d(m_1, m_2)$. So, unless we happen to have $d(m_1, m_2) = \|m_1 - m_2\|_E$ estimates of the kind used above can not help us to find a point m_2' with $\|m_2' - m_1'\| = d(m_1, m_2)$.

We thus have that $|K|$ would have to be larger than η. Large perturbations in this co-ordinate change all of the future distances, and violate this 'inductive' process we are trying to accomplish for a possible infinite generalisation.

We note that it *is* possible that a 'non-inductive' process works when an inductive one fails. This statement is a little vague (intentionally) and we make this precise in Section 3.4.

To recap m_2'

So to recap our process up to the point, we are trying to embed the metric space M, which we are treating as a subset of ℓ_∞^n, into E, which we are treating as $\mathbb{R}^n$ with a norm $\|.\|_E$ that satisfies (2.1). To do this we iteratively construct a map f, by first setting $f(m_1) = m_1$. To construct $f(m_2)$, we looked at $m_2 + \theta e_1$ where θ is some small parameter, and showed (if $d(E, \ell_\infty^n)$ was small) that there is a small θ such that $\|m_2 + \theta e_1 - m_1\|_E = \|m_2 - m_1\|_\infty$. We have also shown that if we had

[2]The factor 2 here will become relevant later, any concern the reader has may be assuaged that $d(m_i, m_j) + d(m_j, m_k) > d(m_i, m_k) + 2\eta$ certainly implies $d(m_i, m_j) + d(m_j, m_k) > d(m_i, m_k) + \eta$.

considered $m_2 + \theta e_i$, with $i \neq 1, 2$ to ensure that $\|m_2 + \theta e_i - m_1\|_E = \|m_2 - m_1\|_\infty$ then θ would have to be large. This 'largeness' is not what we want - we want small perturbations to lead to small perturbations and the behaviour to not be 'chaotic' in some sense.

Finding m_3'

Similar to our thought process with m_2', we have n co-ordinates to play with, and we could send them anywhere. The simplest thing we can do is change two co-ordinates. We have two conditions (that $\|m_1' - m_3'\|_E = d(m_1, m_3)$ and that $\|m_2' - m_3'\|_E = d(m_2, m_3)$,) and thus we have to change at least two co-ordinates for there to be enough degrees of freedom in the system to solve it.

Which co-ordinates do we change? First let's see what happens when we change the first two co-ordinates.

Changing the first two co-ordinates

In the case of m_2' we (eventually) considered a map $f : [0, R] \to E$ where $f(r)$ perturbed the first co-ordinate by r. In this case we consider a map $f : [0, R]^2 \to E$ by

$$f(\epsilon_1, \epsilon_2) = (d(m_1, m_3) + \epsilon_1, d(m_2, m_3) + \epsilon_2, d(m_3, m_3), \dots)$$

We wish to make an argument that looks like the argument in the previous section: we wish to find a point (ϵ_1, ϵ_2) such that $\|f(\epsilon_1, \epsilon_2) - m_1'\|_E = d(m_1, m_3)$ and $\|f(\epsilon_1, \epsilon_2) - m_2'\|_E = d(m_2, m_3)$. In the previous section we used an intermediate value argument. The general way of extending an intermediate value argument to more than one dimension is the Brouwer fixed point theorem, and here it makes its appearance.

If we want $\|f(\epsilon_1, \epsilon_2) - m_1'\|_E = d(m_1, m_3)$ and $\|f(\epsilon_1, \epsilon_2) - m_2'\|_E = d(m_2, m_3)$ to correspond to fixed points of a function on $[0, R]^2$, the way we do this is by setting

$$\epsilon_1 = d(m_1, m_3) + \epsilon_1 - \|f(\epsilon_1, \epsilon_2) - m_1'\|_E$$

and

$$\epsilon_2 = d(m_2, m_3) + \epsilon_2 - \|f(\epsilon_1, \epsilon_2) - m_2'\|_E.$$

This informs our choice of function $\varphi(\epsilon_1, \epsilon_2)$, i.e., we set

$$\varphi(\epsilon_1, \epsilon_2)_i = d(m_i, m_3) + \epsilon_i - \|f(\epsilon_1, \epsilon_2) - m_i'\|_E.$$

The reader should note that this is the same as the choice of φ in the previous proof (just in a restricted number of dimensions.) The same bounding as in the previous proof gives that this continuous function does have a fixed point.

The reader should observe that the concavity of M is needed here. In Proof Two we needed concavity to deduce that the function φ had a fixed point - and here we need the concavity of M.

Perturbing other co-ordinates

A similar symmetry argument to the case of m_2' shows us that we can perturb any two of the first three co-ordinates and find a point m_3' as required. Since this is an identical argument we do not consider it here.

However, if we perturb at least one of the co-ordinates above 3, the same issue arises as for m_2': the change in the co-ordinate required to find m_3' such that $\|m_3' - m_1'\| = d(m_3, m_1)$ is *large*. Identical reasoning to m_2' indicates that we do not want to use large perturbations.

We will omit the details.

Finding further m_k'

Finding m_k' for $k \geq 4$ is then a simple inductive procedure where at stage k we consider the function $\varphi_k : [0, R]^{k-1} \to [0, R]^{k-1}$ with

$$\varphi_k(\epsilon_1, \ldots, \epsilon_{k-1})_i = d(m_i, m_k) + \epsilon_i - \|f(\epsilon_1, \ldots, \epsilon_{k-1}) - m_i'\|_E.$$

A conclusion of this method

In this method, the reader should find it clear why the previous result worked. At each stage, we embed the i'th point and use the Brouwer's fixed point theorem in i dimensions to find a point m_i' such that $\|m_i' - m_j'\|_E = d(m_i, m_j)$ for $j < i$. This inductive process allows us to find the embedding of each point in turn.

The proof we have given here emphasises two aspects of the method. The first is that the method is truly an *inductive* one. We send points in turn to a new one, and at each stage we force the distances to be correct. It should not be a surprise that there is a generalization to the infinite case, see Section 2.5, where we essentially just iterate this infinitely many times. The reader is cautioned that the infinite generalization requires some slightly stronger conditions - the passage to the infinite requires us that the bounds throughout this proof (which we have been slightly sloppy with) do not converge to zero.

The second aspect that this proof emphasises is that it is a relatively natural thing to do. At each stage we, in some sense, do the obvious thing and it works. The author feels this is the key to the previous two proofs (even though they are much simpler to read and understand.)

2.2.1 Some concluding thoughts

Before we continue to the negative results, the author wants to point out two things from the previous theorem and its proofs.

The 'key' idea

Here we want to extract a key idea from the arguments in the previous section. We will make bountiful use of this lemma later. The reader can think of this as a duality principle - given a pair (A, E) where E is a Banach space and A is an n-point subset of E, if there is an $\epsilon > 0$ such that every ϵ-perturbation of A isometrically embeds into A then there is a $\delta > 0$ such that if F is a Banach space that is a δ-perturbation of E then A isometrically embeds into F. We make this precise in the following,

Lemma 2.2.4. *Suppose that E is a Banach space and A is an n-point concave subset of E on points $\{a_1, \dots, a_n\}$. Suppose that there is some $\epsilon > 0$ such that there is a continuous function $\Theta : [0, \eta]^{\binom{n}{2}} \to E^n$ such that $\|\Theta(\epsilon)_i - \Theta(\epsilon)_j\|_E = d(a_i, a_j) + \epsilon_{ij}$. Then there is some $\delta > 0$ such that if $d(E, F) < 1 + \delta$ then A embeds isometrically into F.*

Note that the assumption on Θ implies that all ϵ-perturbations of the distance matrix $(d(a_i, a_j))_{i,j=1}^n$ are distance matrices of a metric space. This implies that (viewing A as a metric space) A is concave. Moreover, note that the dimension of E is not assumed to be equal to n.

Proof. This proof is essentially contained in the above work, but we give the details explicitly so we may make use of the result in the sequel.

Fix $\delta > 0$ to be determined later, and assume that F is a Banach space with $d(E, F) \leq 1 + \delta$. Moreover, we may assume that $\|x\|_E \leq \|x\|_F \leq (1 + \delta)\|x\|_E$.

The idea is that we define the 'helper' function $\varphi : [0,\eta]^{\binom{n}{2}} \to [0,\eta]^{\binom{n}{2}}$ by setting

$$\varphi(\boldsymbol{\epsilon})_{ij} = d(a_i, a_j) + \epsilon_{ij} - \|\Theta(\boldsymbol{\epsilon})_i - \Theta(\boldsymbol{\epsilon})_j\|_F.$$

We have to show that if δ is sufficiently small (depending on A and η) then φ is well defined, i.e., that $0 \leq \varphi(\boldsymbol{\epsilon})_{ij} \leq \eta$. An application of the Brouwer fixed point theorem provides us a point $\boldsymbol{t}$ such that $\varphi(\boldsymbol{\epsilon}) = \boldsymbol{\epsilon}$, and by the definition of φ, the image of $\Theta(\boldsymbol{\epsilon})$ is the required isometric embedding of A into F.

First we note that

$$\varphi(\boldsymbol{\epsilon})_{ij} \geq d(a_i, a_j) + \epsilon_{ij} - \|\Theta(\boldsymbol{\epsilon})_i - \Theta(\boldsymbol{\epsilon})_j\|_E = 0$$

where we have used $\|.\|_F \geq \|.\|_E$. This gives us that the lower bound is satisfied.

For the upper bound, we have that

$$\varphi(\boldsymbol{\epsilon})_{ij} \leq d(a_i, a_j) + \epsilon_{ij} - \frac{1}{1+\delta}\|\Theta(\boldsymbol{\epsilon})_i - \Theta(\boldsymbol{\epsilon})_j\|_E = \frac{\delta}{\delta+1}(d(a_i, a_j) + \epsilon_{ij}).$$

This is less than $\frac{\delta}{\delta+1}(d(a_i, a_j) + \eta)$, and if $\delta \leq \frac{\eta}{d(a_i,a_j)}$ for all i, j this quantity is less than η. Thus the function φ is well defined, and we are done. $\square$

This proof of Lemma 2.2.4 reveals that one can take the following generalisation which demonstrates that the fact we are in normed spaces is not particularly important. We can, in fact, take suitable generalisations to metric vector spaces with a notion of 'closeness' that is similar to the Banach-Mazur notion.

Lemma 2.2.5. *Suppose that M is a metric space and that we have a continuous function $G : M \to \mathbb{R}$, a collection $\{R_{ij} : 1 \leq i < j \leq n\}$ an $\epsilon > 0$ and a continuous function $F : [0,\epsilon]^{\binom{n}{2}} \to M^n$ such that $G(F(\boldsymbol{\epsilon})_i - F(\boldsymbol{\epsilon})_j) = R_{ij} + \epsilon_{ij}$. Then there is some $\delta > 0$ such that if we have a continuous $H : \mathbb{R}^n \to \mathbb{R}$ with the property that $H(x) \leq F(x) \leq (1+\delta)H(x)$ then there exist points $x_1, \ldots, x_n$ in $\mathbb{R}^n$ such that $H(x_i - x_j) = R_{ij}$.*

The reader may (rightly) think that this gives a generalisation of every result in this thesis. Every time we apply Lemma 2.2.4 we could instead apply this lemma to obtain a generalisation to the case of metric spaces with this unusual notion of closeness.

Our main reason for not presenting the ideas in this more general form is that our interest is primarily in the case of Banach spaces, where we can apply results such as Krivine's Theorem, Dvoretzky's Theorem or the Maury-Pisier Theorem to find the function G in this lemma. Analysing metric variants of these properties would take us into an active research area where we would run into many open problems. Presenting the ideas in this general metric setting would clutter the presentation.

The proof is identical to that of Lemma 2.2.4.

How large can δ be?

It is perhaps of interest to look at how large δ can be in the above calculation.

To state the next result as clearly as possible, we need to introduce the notion of an $\eta-$concave metric space. A metric space M is *η-concave* if $d(x,y) + d(y,z) - d(x,z) > 2\eta$ for each distinct triple $x, y, z \in M$

Corollary 2.2.6. *Let M be an η-concave metric space consisting of m points, and let $K = \max d(m_i, m_j)$. If $d(\ell_\infty^m, E) < 1 + \eta/K$ then M embeds isometrically into E.*

The key here is that this is dependent only on two things - the η-concavity and the diameter of M. There is *no* dependence on $|M|$.

Proof. All we need to do is analyse the exact bounds in the proof of Lemma 2.2.4. If one looks at the last paragraph, we require that $\frac{\delta}{\delta+1}(d(m_i, m_j) + \epsilon_{ij}) \leq \eta$. Since ϵ_{ij} is at most η, this follows from $\delta \leq \frac{\eta}{d(m_i,m_j)}$. Since $\min_{i,j} \frac{\eta}{d(m_i,m_j)} = \frac{\eta}{K}$ the result follows. $\square$

2.3 A Negative Result

In this section we will show that there is a metric space containing ℓ_∞^n uniformly that does not isometrically contain every finite metric space. We will, in fact, show the following theorem:

Theorem 2.3.1. *There is a space that is isomorphic to ℓ_∞ that is not universal for all finite metric spaces.*

The assumption in this theorem is much stronger than that of containing ℓ_∞^n uniformly.

In the previous section, to show that an n-point metric space embedded into spaces close to ℓ_∞^n we had to use concavity of the metric space. This assumption seemed to be a technical one. However, the construction below will show that it is a genuine obstruction, and that the hypotheses of Theorem 2.2.1 can not be weakened.

The proof of Theorem 2.3.1 is a generalization of the following, slightly easier to understand, counter-example.

Lemma 2.3.2. *There is a space X that almost isometrically contains the family ℓ_∞^n, $n \in \mathbb{N}$, that is not universal for all finite metric spaces.*

Proof. Consider the space $X = (\oplus_{n=2}^{\infty} \ell_n^n)_2$. We first claim that ℓ_∞^n is finitely representable in X. This follows from our bound on the distance between the spaces ℓ_p^n: Lemma 1.1.4 says that $d(\ell_p^n, \ell_\infty^n) \le n^{1/p}$. So, for fixed n and $\epsilon > 0$, if m is sufficiently large $d(\ell_m^n, \ell_\infty^n) \le 1 + \epsilon$. Since ℓ_m^n isometrically embeds into ℓ_m^m, we have that ℓ_∞^n almost isometrically embeds into X.

We claim that if $x, y \in S_X$ with $x \neq y$ then $\|x + y\| < 2$. Indeed, take $x_n \in \ell_n^n$ and $y_n \in \ell_n^n$, and suppose that $\sum \|x_n\|^2 = \sum \|y_n\|^2 = 1$. Moreover, suppose that we have some m such that $x_m \neq y_m$.

First, we look at the case that $\|x_n\| = \|y_n\|$ for each n. In this case, by the strict convexity of ℓ_m^m, we have that $\|x_m + y_m\| < 2\|x_m\|$. For every other n we have that $\|x_n + y_n\| \le 2\|x_n\|$. Thus $\sum \|x_n + y_n\|^2 < \sum 4\|x_n\|^2 = 4$, i.e., $\|x + y\| < 2$.

Second, we look at the case that $\|x_m\|$ and $\|y_m\|$ are different for some m. In this case we can bound differently, $\sum \|x_n + y_n\|^2 \le \sum(\|x_n\| + \|y_n\|)^2 = \sum \|x_n\|^2 + \sum \|y_n\|^2 + 2\sum \|x_n\|\|y_n\|$. Applying Cauchy-Schwarz, we see that this is smaller than the quantity $2 + 2\sqrt{\sum \|x_n\|^2 \sum \|y_n\|^2}$, which is equal to 4.

Equality in the above statement holds just when the vector $(\|x_n\|)_n$ is a scalar multiple of the vector $(\|y_n\|)$. Since both vectors have ℓ_2 norm 1, this scalar multiple has to be 1. We assumed that there was some m such that $\|x_m\| \neq \|y_m\|$, and thus $\sum \|x_n + y_n\|^2$ is strictly less than 4.

However, it is easy to see that in ℓ_∞^2 there are two points such that $\|x\| = \|y\| = 1$, $x \neq y$ and $\|x + y\| = 2$, take (for example) $(1, 0)$ and $(1, 1)$. So the metric space $\{(0,0), (1,0), (1,1), (1,1/2)\}$ with the inherited ℓ_∞^2 norm does not isometrically embed into X. This is the content of Theorem 1.2.5. $\square$

Remark 2.3.3. The following classical result can be extracted from the above proof.

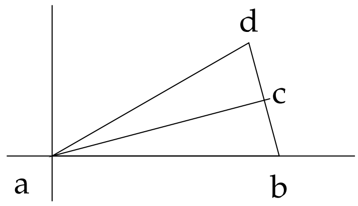
Lemma 2.3.4. *If X_n are strictly convex spaces then $(\oplus X_n)_2$ is a strictly convex space.*

Later on we will want to prove variants of this lemma which will be much less straightforward: the use of 2-direct sums allowed us to use Cauchy-Schwarz.

In the above proof we used a finite metric space M that lacked the unique metric midpoint property. If a strictly convex Banach space contains $(1+\epsilon)$-distortion copies of every finite metric space the above lemma shows that it is not finitely universal as it does not contain M. The class of Banach spaces containing $(1+\epsilon)-$distortion copies of every finite metric space is, by Remark 1.1.38, the class of Banach spaces containing ℓ_∞^n almost isometrically. This is a strict class of Banach spaces. The following theorem shows a larger class of spaces that are guaranteed to almost isometrically contain such a space M, but not isometrically contain M.

Theorem 2.3.5. *Suppose that X is a Banach space that is strictly convex but not uniformly convex. Then there is a four point metric space that $(1+\nu)$-embeds into X for any $\nu > 0$ but that does not isometrically embed into X.*

Proof. Since X is not uniformly convex, there is some $\epsilon > 0$ such that for any $\delta > 0$ we can find points x_δ and y_δ such that $\|x_\delta - y_\delta\| = \epsilon$ but $\|x_\delta + y_\delta\| > 2 - 2\delta$. We construct our metric space M as follows: take points a, b, c, d with $d(a,b) = d(a,c) = d(a,d) = 1$, $d(b,c) = d(c,d) = \epsilon/2$ and $d(b,d) = \epsilon$.



A pictoral representation of the space M

Assume that there is an isometric embedding $f : M \to X$. Assume, without loss of generality, that $f(a) = 0$. Then, since c is a metric midpoint of b, d and the space X is strictly convex, $f(c) = (f(b) + f(d))/2$ and $\|f(c)\| = 1$. This is a contradiction of strict convexity, as $f(b)$ and $f(d)$ lie in S_X (and are distinct.)

However, for any $\nu > 0$ there is a $(1+\nu)$ embedding of M into X. Indeed, defining f_n by $f_n(a) = 0$, $f_n(b) = x_{1/n}$, $f_n(d) = y_{1/n}$ and $f_n(c) = 1/2(x_{1/n} + y_{1/n})$ we see that the map is isometric on all pairs of points except (a, c) where

$$1 > \|f_n(c) - f_n(a)\| > 1 - \frac{1}{n}.$$

Taking the limit as n tends to infinity we see that the distortion of the embedding f_n tends to 1. $\square$

We can end this section with the proof of Theorem 2.3.1, i.e., a space isomorphic to ℓ_∞ that is not finitely universal.

Proof of Theorem 2.3.1. Combine Theorem 2.3.5 with Lemma 1.1.20 (which is where we constructed an explicit norm on ℓ_∞ that was strictly convex.[3]) □

2.3.1 Some Remarks on Property AI-I

The work in the above section allows us to make some non-trivial remarks about Property AI-I. Theorem 2.3.5 shows us that any space that is strictly convex but not uniformly convex fails Property AI-I. We listed several of these in Section 1.1.3. For example, we have the following:

Theorem 2.3.6. *There are reflexive spaces with unconditional bases that fail Property AI-I.*

Proof. Simply note that the space $(\oplus_{n=2}^{\infty} \ell_n^n)_2$ is reflexive (as it is the 2 direct sum of finite dimensional spaces), strictly convex, not uniformly convex and that space has an unconditional basis. □

We observe that the above space is not superreflexive, indeed c_0 is finitely representable in the space. This method of constructing spaces that fail Property AI-I is not particularly helpful in finding a superreflexive example due to Theorem 1.1.28.[4] We will see an example of a superreflexive space without Property AI-I later.

The work we have established so far also allows us to show that Property AI-I does not interact well with the notion of Banach-Mazur distance. One may hope that the property is 'open' in some sense, i.e., if a Banach space X has Property AI-I there is some $\epsilon > 0$ such that if $d(X, Y) \leq 1 + \epsilon$ then Y has Property AI-I. The next theorem shows that this is false in a very strong sense, you can even fail with $\epsilon = 0$.

Theorem 2.3.7. *There is a Banach space X with Property AI-I, and a Banach space without Property AI-I such that $d(X, Y) = 1$.*

[3]To observe that the norm constructed on ℓ_∞ is not uniformly convex recall that Theorem 1.1.22 showed that a uniformly convex Banach space was reflexive. ℓ_∞ is not reflexive.

[4]It is possible that such a method will work due to Theorem 1.1.28 being an isomorphic result instead of an isometric result. This does not seem particularly fruitful avenue to the author.

Proof. This is a modification of the proof of Theorem 1.1.19 (the proof that there are two Banach spaces of distance 1 away from each other that are not isometrically isomorphic.) We take the space $X = \left(\oplus_2 \ell_{p_n}^{q_n}\right)$ and $Y = (\oplus_2 \ell_\infty^n) \oplus_2 \left(\oplus_2 \ell_{p_n}^{q_n}\right)$ where p_n is a dense subset of $[2, \infty)$ and q_n a sequence of natural numbers that takes every $m \in \mathbb{N}$ infinitely often. The reader can observe that the proof of Theorem 1.1.19 can be repeated with the spaces X and Y to show that they are isomorphic. The space X is strictly convex, and the space Y evidently has Property AI-I (simply, it is universal for finite metric spaces.)

□

Remark 2.3.8. The reader may be interested in the fact that there are many examples of non-isomorphic spaces that satisfy the assumptions in Theorem 2.3.5. For an example, we can pick any two non-isomorphic separable spaces that are not uniformly convexifiable (i.e., not superreflexive), apply the procedure in Theorem 1.1.21 and then observe that the resulting spaces remain non-isomorphic. A concrete example of this would be $\ell_1 \oplus_2 \ell_p$ for $1 \leq p \leq \infty$ - an uncountable family of pairwise non-isomorphic spaces. A less trivial example would be the family of generalized Tsirelson spaces of order α.

2.4 Two Open Problems

Over this section we have developed techniques to handling questions about Property AI-I in the case that the spaces involved contain, for every $\epsilon > 0$, $(1+\epsilon)$-distortion copies of every finite metric space. Over the previous two subsections we have shown that:

- If M is a finite concave metric space and X contains ℓ_∞^n almost-isometrically, then M isometrically embeds into X.
- There exists a non-concave metric space M and a space isomorphic to ℓ_∞ such that M does not isometrically embed into X.

The author sees two clear directions of further study here. The first is to ask:

Question 2.4.1. *Let $\mathcal{R}$ be the collection of finite metric spaces that isometrically embed into every Banach space with no non-trivial cotype. Which metric spaces are in $\mathcal{R}$?*

We have shown that $\{$finite concave spaces$\} \subset \mathcal{R}$. It is also clear, from the definition of a Banach space, that $\{$subsets of $\mathbb{R}\} \subset \mathcal{R}$. In some sense subsets of the reals are the 'least concave' spaces: if we pick any three points then one is the convex combination of the other two.

The author believes that it is possible that these are the only counter-examples, i.e., he is of the opinion that 'two dimensional concavity is a barrier to embeddability.' We formulate this here as a conjecture:
Conjecture 2.4.2. $\mathcal{R} = \{$*finite concave spaces*$\} \cup \{$*subsets of* $\mathbb{R}\}$.

The second direction of study that the author sees is to ask the following:
Question 2.4.3. *Is there a characterization of Banach spaces that contain every finite metric space isometrically?*

Of particular note to the author is:
Question 2.4.4. *Is Tsirelson's original space finitely isometrically universal?*

2.5 An Infinite Result

The main result of this section is a generalisation of Theorem 2.2.2, originally proven in [28]. We include it here because it is a simple extension of our idea for the finite case (specifically we use Corollary 2.2.6.) The presentation of the proof we have given above allows for a very easy generalization to the infinite case (which is slightly different from the proof in [28], simply we keep the proof in line with our presentation to this point.[5].)
Theorem 2.5.1 ([28]). *Suppose that M is an η-concave countably infinite bounded metric space. Then there is some $\delta > 0$ such that if X contains an $(1+\delta)$-isomorphic copy of ℓ_∞ then M embeds isometrically into X.*

Proof. Fix $\delta = 1 + \eta/2\,\mathrm{diam}(M)$ and consider a Banach space X with $d(X, \ell_\infty) < 1 + \delta$. Without loss of generality in what follows we may assume that $X = (\ell_\infty, \|.\|_E)$ such that $\|.\|_X \leq \|.\|_\infty \leq (1+\delta)\|.\|_E$. We will inductively embed M into X, following precisely the argument of Proof 3 of 2.2.2.

So, assume we have isometrically embedded $\{m_1, \dots, m_n\}$ into X by a map f_n where

$$f_n(m_i) = (d(m_i, m_1) + \epsilon_{i,1},\ d(m_i, m_2) + \epsilon_{i,2}, \dots,\\ d(m_i, m_{i-1}) + \epsilon_{i,i-1}, d(m_i, m_i),\ d(m_i, m_{i+1}), \dots)$$

[5]Our proof also does not rely on the Schauder fixed point theorem: the iterative approach we opt for allows us to only ever use Brouwer's fixed point theorem.

where $0 \leq \epsilon_{i,j} \leq \eta$. We now construct an embedding f_{n+1} as follows: we set $f_{n+1}(m_i) = f_n(m_i)$ if $i \leq n$ and we will construct $f_{n+1}(m_{n+1})$.

Set

$$\Theta(\epsilon_1, \dots, \epsilon_n) = (d(m_{n+1}, m_1) + \epsilon_1, d(m_{n+1}, m_2) + \epsilon_2, \dots \\ , d(m_{n+1}, m_n) + \epsilon_n, \ d(m_{n+1}, m_{n+1}), \dots)$$

and our goal is to show that for some choice of $(\epsilon_1, \dots, \epsilon_n) \in [0, \eta]^n$ we have that

$$\|f_n(m_i) - \Theta(\epsilon_1, \dots, \epsilon_n)\|_X = d(m_i, m_{n+1}).$$

So, as per before, we take $\varphi : [0, \eta]^n \to [0, \eta]^n$ and set

$$\varphi(\epsilon_1, \dots, \epsilon_n)_i = d(m_i, m_{n+1}) + \epsilon_i - \|f_n(m_i) - \Theta(\epsilon_1, \dots, \epsilon_n)\|_X.$$

Then, the same computation as in Lemma 2.2.4 and Corollary 2.2.6 shows that for our choice of δ there is a fixed point of φ_i, and this choice of $\epsilon_1, \dots, \epsilon_n$ gives us our choice, i.e., $f_{n+1}(m_{n+1}) = \Theta(\epsilon_1, \dots, \epsilon_n)$. □

We can give an interesting application of this:

Theorem 2.5.2. *Suppose that X is a Banach space that contains an isomorphic copy of ℓ_∞. Then X isometrically contains an infinite subset $M = \{m_1, m_2, \dots, \}$ such that $d(m_i, m_j) = 1$ for all $i \neq j$.*

Proof. This is a combination of the previous result (with the trivial observation that M satsifies the assumptions of the theorem) and Partington's theorem, Theorem 1.1.40. □

Chapter 3

The ℓ_2 Case

In the previous section we looked at Banach spaces with no non-trivial cotype. Using the Maurey-Pisier theorem we deduced that these were almost-isometrically universal for finite metric spaces. We then extended this to isometric universality for certain finite metric spaces.

In this section we look at arbitrary infinite-dimensional Banach spaces. By Dvoretzky's theorem we will observe that these are almost-isometrically universal for all finite subsets of ℓ_2. This is a more complicated case than that of ℓ_∞ and will require us to use different classical techniques to get a handle.

In the case of ℓ_∞ we found a restriction on finite metric spaces that guaranteed that they embedded isometrically into any Banach space with no non-trivial cotype. In this section we similarly find a condition on finite subsets of ℓ_2 that guarantee they embed into every infinite-dimensional Banach space. However, we do not establish the equivalent of Theorem 2.3.1. Indeed, it is still an open problem to establish whether every infinite-dimensional Banach space is isometrically universal for all finite subsets of ℓ_2. We discuss this in Section 3.4, where we talk about issues with extending the results proven here.

3.1 The setup

We begin by considering an arbitrary infinite-dimensional Banach space X. By Dvoretzky's Theorem, Theorem 1.1.29, ℓ_2 is finitely representable in X, i.e., for every n and $\epsilon > 0$, $\ell_2^n \overset{1+\epsilon}{\hookrightarrow} X$. Let M be a finite subset of ℓ_2.

By first applying the Gram-Schmidt process to M we may assume that M is a subset of ℓ_2^n. Let E be an n-dimensional subspace of X such that $d(E, \ell_2^n) < 1+\epsilon$, and suppose that $T : \ell_2^n \to E$ is a mapping for which $\|T\|\|T^{-1}\| < 1+\epsilon$. Scaling, we may assume that $\|T\| = 1$ and $\|T^{-1}\| < 1+\epsilon$. By taking $\iota : M \to \ell_2^n$ to be the formal inclusion mapping, the map $T \circ \iota : M \to E \leq X$ is a $(1+\epsilon)-$distortion embedding of M into X.

We are thus in the position where any finite subset of ℓ_2 almost isometrically embeds into X. We are in the position to investigate the Finitely Isometric Dvoretzky Property for any space X, i.e.,

Question 3.1.1. *Suppose X is an infinite dimensional Banach space. Does every finite subset of ℓ_2 embed isometrically into X?*

3.2 A positive result for affinely independent subsets

A subset M of ℓ_2 is called *affinely independent* if for some $x \in \ell_2$, the set $x + M$ is linearly independent. The following Theorem was proven as Lemma 3 in [40].

Theorem 3.2.1. *[Lemma 3 of [40]] Suppose that X is an infinite-dimensional Banach space and that M is a finite, affinely independent subset of ℓ_2. Then M isometrically embeds into X.*

We note that affinely independent subsets of ℓ_2 are concave. If our Banach space X has no non-trivial cotype this theorem will not give us a larger class of metric spaces that embed into X. The converse is not true.

As in the ℓ_∞ case, we establish this result by proving the following more specific theorem and then appealing to Dvoretzky's theorem:

Theorem 3.2.2. *If M is a finite, affinely independent subset of ℓ_2 then there is some $\delta > 0$ such that if E is an n-dimensional Banach space with $d(\ell_2^n, E) < 1+\delta$ then M isometrically embeds into E.*

To very briefly spell out how Theorem 3.2.1 follows from Theorem 3.2.2, suppose we have some M, an affinely independent subset of ℓ_2, and an infinite-dimensional Banach space X. By Theorem 3.2.2, there is some $\delta > 0$ such that if E is an n-dimensional Banach space with $d(\ell_2^n, E) < 1+\delta$ then M isometrically embeds into E. By Dvoretzky's theorem there is some subspace F of X such that $d(F, \ell_2^n) < 1+\delta$, and thus M embeds isometrically into F (and hence X.)

Similar to the previous case, the proof will be an application of Brouwer's fixed point theorem. We are going to give two proofs - one inductive and one not inductive. The proof that is not inductive will be very similar to the first proof given in the ℓ_∞ case, essentially we will replace all of our observations about the Fréchet-Kolmogorov embedding with observations about the Cayley-Menger determinant. This is the clearest way for the reader to understand what happens in the ℓ_p case.

The proofs we will give of Theorem 3.2.2 are different than those in [40]. The proof in this paper focuses on an argument via homotopy, whereas we will proceed using Lemma 2.2.4.

Proof One of Theorem 3.2.2

First we shall recall the definition of the Cayley-Menger determinant. If we label the $(n+1)$-points of M by $m_0, \ldots, m_n$, we consider the associated distance matrix $\mathcal{M} = \left(d(m_i, m_j)\right)_{i,j=0}^{n}$. Setting $\mathcal{P}_r = \left(\mathcal{M}_{ij}^2\right)_{i,j=0}^{r}$, we define

$$\mathrm{CMDet}(m_0, \ldots, m_r) = \det \begin{vmatrix} \mathcal{P}_r & \mathbf{1} \\ \mathbf{1}^T & 0 \end{vmatrix}$$

where $\mathbf{1}$ is the length n vector of 1's. Theorem 1.2.11 says that M isometrically embeds into ℓ_2^n just when the sign of $\mathrm{CMDet}(m_0, \ldots, m_r)$ is $(-1)^{r+1}$ for all $r = 1, \ldots, n$.

We remind the reader that U_n is the space of symmetric $n \times n$-matrices with 0 on the diagonal.

In the case of the Fréchet-Kuratowski embedding we noted that if we looked at the matrix $\mathcal{M}$ associated to the metric space M, the mapping $M \mapsto f_M(M)$ (where f_M is the Fréchet-Kuratowski embedding) was a continuous mapping from U_n to $\underbrace{\ell_\infty^n \times \cdots \times \ell_\infty^n}_{n \text{ times}}$. We can say something similar here.

First note that if $\mathcal{M}$ is a matrix associated to a metric space M such that the $\mathrm{CMDet}(m_0, \ldots, m_i)$ has sign $(-1)^{i+1}$, then for some $\epsilon > 0$, if $\mathcal{N}$ is another element of U such that $\|\mathcal{N} - \mathcal{M}\|_\infty < \epsilon$ then $\mathcal{N}$ represents a metric space $N = \{r_0, \ldots, r_n\}$ such that $\mathrm{CMDet}(r_0, \ldots, r_i)$ has sign $(-1)^{i+1}$ for all $i = 1, \ldots, n$. Let $C = \{\mathcal{N} : \|\mathcal{N} - \mathcal{M}\|_\infty < \epsilon\}$. We observe that the fact that M isometrically embeds into ℓ_2 in an affinely independent way shows that M is concave. Thus, by possibly shrinking ϵ we may assume that all elements of C are matrices representing concave metric spaces.

For any element $\mathcal{M}$ of C we construct a metric space M such that $\mathcal{M}$ represents M: simply define M as a metric space on points $\{m_1, \ldots, m_n\}$ with distance function d defined by $d(m_i, m_j) = \mathcal{M}_{ij}$. We recall the function $g_M : M \to \ell_2^n$ defined after the proof of Theorem 1.2.11. The function g_M was constructed to have the property that

$$\|g_M(m_i) - g_M(m_j)\|_2 = d(m_i, m_j).$$

We now define a function $g : C \to \underbrace{\ell_2^n \times \ldots \ell_2^n}_{n \text{ times}}$ by $g(\mathcal{M}) = (g_M(m_1), \ldots, g_M(m_n))$.

We now, essentially, replicate the proof of Theorem 2.2.2. We have shown that if M is a metric space that embeds into ℓ_2 in an affinely independent way, then there is some ball C around $\mathcal{M}$ and a continuous function from C to $\underbrace{\ell_2^n \times \cdots \times \ell_2^n}_{n \text{ times}}$ such that if $\mathcal{N}$ is in C then $\mathcal{N}$ represents a metric space N and the image of $\mathcal{N}$ is an isometric copy of N.

Fix $\delta > 0$ and consider a Banach space E with $d(E, \ell_2^n) < 1 + \delta$. Without loss of generality we may assume that $E = (\mathbb{R}^n, \|.\|_E)$ with $\|.\|_E \leq \|.\|_2 \leq (1+\delta)\|.\|_E$. Set $\iota : \ell_2^n \to E$ to be the formal identity map and define a map $\Theta : [0, \epsilon]^{\binom{n}{2}} \to \underbrace{E \times \cdots \times E}_{n \text{ times}}$ by $\Theta(\mathcal{N}) = \iota \circ g_{N'}(N')$ where N' is the metric space associated to $\mathcal{M} + \mathcal{N}$.

We are now in the position where an application of Lemma 2.2.4 completes the proof.

Proof Two of Theorem 3.2.2

The following argument is going to use the Cayley-Menger determinant as little as possible. It amounts to the same proof as the previous section. However, we will emphasise the inductive nature of the argument.

Let $M = \{m_0, m_1, \ldots, m_n\}$ be an affinely independent $(n+1)$-point subset of ℓ_2^n. Without loss of generality we may assume that $m_0 = 0$. Since M is affinely independent, it is concave. By Theorem 1.2.11 we have that $(-1)^{k+1}$ is the sign of $\text{CMDet}(m_0, \ldots, m_k)$ for each $k = 0, \ldots, n$.

Suppose that $N = \{r_0, \ldots, r_n\}$ is a set of $(n+1)$ points. Then there is some $\epsilon > 0$ such that if $|d_N(r_i, r_j) - d_M(m_i, m_j)| < \epsilon$, then (N, d_N) is a metric space and $\text{CMDet}(r_0, \ldots, r_k) = (-1)^{k+1}$ for each $k = 0, \ldots, n$.

Fix $\delta > 0$ and consider a Banach space E with $d(E, \ell_2^n) < 1 + \delta$. Without loss of generality we may assume that $E = (\mathbb{R}^n, \|.\|_E)$ with $\|.\|_E \leq \|.\|_2 \leq (1+\delta)\|.\|_E$. Let $\{e_1, \ldots, e_n\}$ be the standard basis for $\mathbb{R}^n$.

We are going to construct an isometric embedding $f : M \to E$ inductively. Begin by setting $f(m_0) = 0$, and it is clear we can find a point $f(m_1)$ such that $\|f(m_0) - f(m_1)\|_2 = d(m_0, m_1)$. Suppose that for some $1 \leq k < n$ we have defined $f(m_0), \ldots, f(m_k)$ such that f is an isometry and $f(m_i) \in \text{span}\{e_1, \ldots, e_i\}$.

Fix $\epsilon_0, \ldots, \epsilon_k$ and set $\boldsymbol{\epsilon} = (\epsilon_0, \ldots, \epsilon_k)$. We de fine a metric space $N(\boldsymbol{\epsilon})$ consisting of $k+2$ distinct points $r_0, \ldots, r_{k+1}$ with metric $d = d_{\boldsymbol{\epsilon}}$ as follows:

- $d(r_i, r_j) = \|f(m_i) - f(m_j)\|_2$ for each $i, j \in \{0, \ldots, k\}$
- $d(r_i, r_{k+1}) = d(m_i, m_{k+1}) + \epsilon_i$ for each $i \in \{0, \ldots, k\}$.

Provided that δ is sufficiently small, we have that $|d_M(m_i, m_j) - d(r_i, r_j)| < \epsilon$ for all pairs i, j. Hence, by the choice of ϵ, d is a metric on $N(\boldsymbol{\epsilon})$. By Theorem 1.2.11 there is an isometric embedding G of $N(\boldsymbol{\epsilon})$ into ℓ_2^{k+1} whose image is an affinely independent set. After a series of suitable reflections we may assume that $G(r_i) = f(m_i)$. Set $g(\boldsymbol{\epsilon}) = G(r_{k+1})$. Note that, by affine independence, there are two possible choices of $g(\epsilon)$ depending on the sign of the coefficient of e_{k+1}. By fixing a sign we obtain a continuous function $g : [0, \epsilon]^{k+1} \to \ell_2^{k+1}$.

We are now in a position to apply Lemma 2.2.4 to obtain the theorem.

3.3 On extensions of the proof of Theorem 3.2.2

In this section we will detail some results that are related to the proof of Theorem 3.2.2. These will primarily be centred around how possible extensions to the proof would have to work.

The reader may note that we only study these possible results in the ℓ_2 case - not the ℓ_p or ℓ_∞ cases. This is because we have definitive counter-examples in these cases, so (of course) there can be no general positive result. We will mainly be interested in the notion that any possible extension to Theorem 3.2.2 must be *global*. Up until this point, the two positive results we have proven are *local*, i.e., the proofs proceed by looking at images of points, fixing them, and then embedding others.

The results here are mostly based on the following simple lemma:

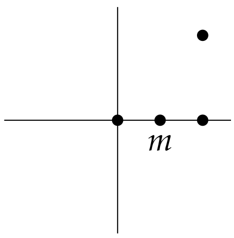
Lemma 3.3.1. *Suppose that M is a metric space, X is a strictly convex Banach space and $\alpha \in (0,1)$. Suppose that there are three distinct points m_1, m_2, m_3 in M such that $d(m_1, m_2) = \alpha d(m_1, m_3)$ and $d(m_2, m_3) = (1-\alpha)d(m_1, m_3)$. If $f : M \to X$ is an isometry, then $f(m_2) = \alpha f(m_1) + (1-\alpha) f(m_3)$.*

Proof. This is a simple modification of the proof of Lemma 1.2.5. □

Theorem 3.3.2. *There is a 3-point subset M of ℓ_2^2 and a point m of ℓ_2^2 such that for all $\epsilon > 0$ there is a Banach space E with $d(E, \ell_2^2) \leq 1 + \epsilon$, and an isometric embedding $f : M \to X$ such that there is no extension $\tilde{f} : M \cup \{m\} \to X$ that remains isometric.*

This is saying that there is no possible method of extending the method of the proof of Theorem 3.2.2 to the case of affine dependence - we must break free of local considerations and move to global ones.

Proof. Let M be the subset of ℓ_2^2 given by $M = \{(0,0), (1,1), (0,1)\}$ and let m be the point $(1/2, 0)$ The following is a diagram of M.



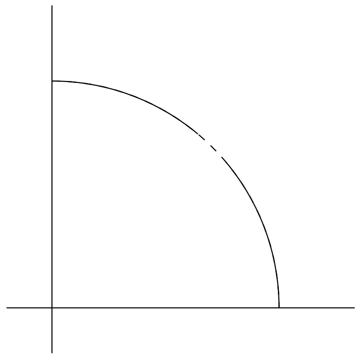
A diagrammatic representation of M and the point m.

We now construct a Banach space E such that $d(E, \ell_2^2)$ is small. Fix some $\epsilon > 0$. We will construct the Banach space E by specifying a curve Γ and using Lemma 1.1.23: a Banach space in $\mathbb{R}^2$ can be defined by specifying its unit sphere. Let $(x_1, x_2) \in \mathbb{R}^2$. In the following, when we say "argument of x", we mean the argument of $x_1 + ix_2$ where argument takes values in $[0, 2\pi)$.

- If the argument of x is not in $[\pi/4 - \epsilon, \pi/4 + \epsilon]$, $x \in \Gamma$ if and only if $x \in S_{\ell_2^2}$.
- If the argument of x is in $[\pi/4 - \epsilon, \pi/4 + \epsilon]$, we say $x \in \Gamma$ if and only if $C(\epsilon)(x_1^{3/2} + x_2^{3/2}) = 1$ where $C(\epsilon)$ is chosen such that Γ is continuous.[1] We think of this sector of the unit disc as the "bad sector".
- We say that Γ has rotational symmetry of order 4, i.e., there are 'bad' sectors around $\pi/4, 3\pi/4, 5\pi/4$ and $7\pi/4$.

[1]We can do this because both the ℓ_2^2 norm and the $\ell_{3/2}^2$ norm are symmetric when we interchange x and y.

The following is a diagram of the unit sphere of E, where the dashed line is "not quite the ℓ_2^2 norm":



The norm $\|.\|_\epsilon$ is evidently strictly convex, and moreover $d(E, \ell_2^2) \to 1$ as $\epsilon \to 0$.

By Lemma 1.2.3 (and its proof) we know that we can take an isometry $f : M \to E$ that sends $(0,0)$ to $(0,0)$, $(0,1)$ to e_1 and $(1,1)$ to some point x.

Now, assume that we can extend f to the point $(1/2, 0)$. Then, by the strict convexity of E we have that $f(1/2, 0) = (1/2, 0)$. Now, we can use the isometry condition, noting that $d_E((1/2,0),(1,1)) = d_{\ell_2}((1/2,0),(1,1))$ and $d_E((1,0),(1,1)) = d_{\ell_2}((1,0),(1,1))$, simply all these directions are in the 'good' sector. Thus, for f to be an isometry we have that $x \in ((1/2,0) + \sqrt{5/4}S_{\ell_2^2}) \cap ((1,1) + S_{\ell_2^2})$, which is exactly $(1,1)$ and $(-1,1)$. Both of these do not have norm $\sqrt{2}$, by construction, so f fails to be an isometry. $\square$

Remark 3.3.3. The above proof also shows that there exists an embedding of $((0,0),(1/2,0),(1,0)) \subset \ell_2^2$ into a Banach space E with no possible extension to $(1,1)$. This is an example where the new point added is affinely independent of the previous points, yet there is no possible example.

Theorem 3.3.4. *There is a four-point subset M of ℓ_2^2 such that for any strictly convex Banach space X that contains no subspace isometric to ℓ_2^2, and for any isometric embedding $f : M \to X$ there is a point $m \in \ell_2^2$ such that no extension $\tilde{f} : M \cup \{m\} \to X$ remains isometric.*

The reader should observe that here M is a four point metric space. It is possible that we could reduce 4 to 3, however it does not seem particularly necessary to optimise this result to best possible.

Proof. This is a modification of the previous result. We choose a slightly different subset M of ℓ_2^2, $M = \{(0,0),(1,0),(-1,0),(0,1)\}$. Let X be any strictly convex

Banach space with no subspaces isometric to ℓ_2^2, and let $f : M \to X$ be an isometric embedding with $f(0,0) = 0$. Now suppose that for any point $m \in \ell_2^2$ but not in M we could extend f isometrically to $M \cup \{m\}$.

First observe that if we take $m = \alpha(1,0) + (1-\alpha)(0,1)$ by Lemma 3.3.1 the extension $\tilde{f}$ must map m to $\alpha f(1,0) + (1-\alpha) f(0,1)$. We thus have that $\|\alpha f(1,0) + (1-\alpha) f(0,1)\| = \sqrt{\alpha^2 + (1-\alpha)^2}$ for $\alpha \in (0,1)$. By homogeneity we get that $\|\alpha f(0,1) + \beta f(1,0)\| = \sqrt{\alpha^2 + \beta^2}$ whenever α, β have the same sign.

Second, if we take $m = \alpha(-1,0) + (1-\alpha)(0,1)$, repeating the above shows that $\|\alpha f(-1,0) + (1-\alpha) f(0,1)\| = \sqrt{\alpha^2 + (1-\alpha)^2}$. By homogeneity we get that $\|\alpha f(0,1) + \beta f(1,0)\| = \sqrt{\alpha^2 + \beta^2}$ whenever α, β have different signs. □

The final theorem is that we can not hope to extend any embedding indefinitely: if X is strictly convex, finite-dimensional, and doesn't contain any two dimensional subspace isometric to ℓ_2^2 then we can find some M such that an M-point subset of ℓ_2^2 does not embed into X. This is an isometric invariant of the space. It is, perhaps, of interest to compute this number for some spaces, however, the author has given little thought to this problem.

This theorem also tells us something else (which may be of interest when we deal with the ℓ_p case.) By Theorem 3.2.2, for every $n \in \mathbb{N}$ there is an $N \in \mathbb{N}$ such that every affinely independent subset of ℓ_2^n embeds into every N-dimensional Banach space sufficiently close to ℓ_2^N (indeed, $N = n + 1$ in this case). This shows that without the hypotheses of affinely independent subsets, this can not be true.

Theorem 3.3.5. *If E is a strictly convex, finite-dimensional Banach space with no subspace isometric to ℓ_2^2, there is a subset M of ℓ_2^2 that does not embed into E.*

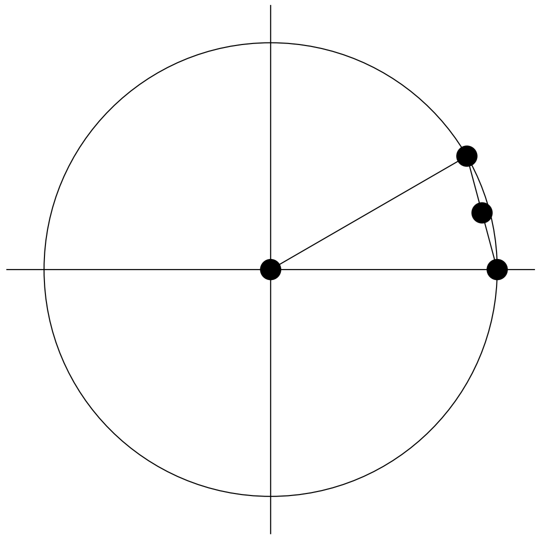
Proof. Let $m_1, m_2, \ldots$ be a dense subset of ℓ_2^2 with $m_1 = 0$ and consider embeddings $f_n : \{m_1, \ldots, m_n\} \to E$ with $f_n(0) = 0$. By a diagonal argument we can pass to a subsequence such that $f_n(m_i)$ is a convergent sequence (defining $f_n(m_i) = 0$ if $i > n$.) We denote the limit of $f_n(m_i)$ as $f(m_i)$. By the continuity of the norm this extends to an isometric embedding $\hat{f} : \ell_2^2 \to X$, which is linear due to the strict convexity of X and Lemma 3.3.1. □

3.4 A positive result for some affinely dependent subsets

In this section we shall modify a proof of Nordlander from [33], and we shall show that the simplest, non-trivial (i.e., not points on a line), case of an affinely dependent subset of ℓ_2 *can* be embedded into any Banach space.

Theorem 3.4.1. *Suppose that we take the subset M of ℓ_2^2 given by $\{0, m, n, \alpha m + (1-\alpha)n\}$ where $m, n \in S_{\ell_2^2}$ are distinct and $0 < \alpha < 1$. Then M isometrically embeds into every Banach space X of dimension 2 or greater.*

The following is a diagram of M:



We include this argument here not because of its generality: it is in fact not very general at all. However, it does (perhaps) include an insight into how one would wish to proceed. The author cannot find a way of extending this argument to more general subsets of ℓ_2 (or even ℓ_2^2). However the method of proof here provides some evidence that such an extension may be possible in this case.

This argument is also useful because all of the counter-examples we have in the other cases use subsets of this type.

After the proof we will mention the aspects of it that make it hard to extend to the case of more than 4 points (or a more complicated affine dependence.)

Proof. Without loss of generality we may assume that X has dimension two. Let $\epsilon = \|m-n\|$. It is easy to work out the size of $\|\alpha m + (1-\alpha)n\|_2$, indeed, since $\langle m-n, m-n\rangle = \epsilon^2$ we have that $2 - 2\langle m, n\rangle = \epsilon^2$, i.e., $\langle m, n\rangle = 1 - \epsilon^2/2$. We thus have that $\|\alpha m + (1-\alpha)\|_2^2 = 1 - \alpha(1-\alpha)\epsilon^2$.

Some technicalities are going to crop up during this proof. The first is that we are primarily going to be interested in the case of differentiable parametrization of the unit sphere - (evidently) this is only possible when the sphere is differentiable. After we show that the space M isometrically embeds into every 2-dimensional Banach space whose unit sphere has a differentiable parametrization we use the following Lemma:

Lemma 3.4.2. *If a finite metric space embeds into every n-dimensional Banach space with differentiable unit sphere then it embeds into every n-dimensional Banach space.*

This theorem follows from classical results about approximation of convex bodies, specifically there are a wealth of results on showing that we have a convex body C there is a sequence of smooth convex bodies C_n such that C_n $'' \to ''$ C in some sense. An example of this was formalised by Weil in [43].

We now take two functions $r, s : [0, 1] \to S_X$ that have the following properties:

- $\|r(t) - s(t)\| = \epsilon$.
- r, s are continuous bijections of $[0, 1)$ and S_X, with $r(0) = r(1)$ and $s(0) = s(1)$.
- r, s are smooth, with the right hand derivatives at 1 agreeing with the left hand derivatives at 0.
- $r - s$ is a smooth parametrization of ϵS_X.
- r and s are taken so that they traverse S_X *clockwise*.
- $r - s$ traverses ϵS_X clockwise.

Remark 3.4.3. We construct the functions r and s as follows. We first fix $\theta(t)$ to be the standard clockwise parameterisation of the circle, i.e., we set $\theta(t) = (\cos(2\pi t), -\sin(2\pi t))$. Then $r(t) = \frac{\theta(t)}{\|\theta(t)\|}$. This gives a bijection from $[0, 1)$ to S_X. We choose $s(t)$ as follows. By the Intermediate Value Theorem there is some point t_0 such that $\|r(t + t_0) - r(t)\| = \epsilon$, where we interpret the sum $t + t_0$ modulo 1. Indeed, $\|r(t + \frac{1}{2}) - r(t)\| = 2$ and $\|r(t) - r(t)\| = 0$, and we have that the value of t_0 changes continuously. We set $s(t) = r(t + t_0)$. The fact that $f(s) = \|r(t + s) - r(t)\|$ is injective on $[0, \frac{1}{2}]$ is due to the smoothness of S_X.

We write $r(t) = (x_1(t), y_1(t)), s(t) = (x_2(t), y_2(t))$. Since r and s are both parametrization of S_X we know that the area contained by r and s is the same, which we denote by R. Moreover, since $r - s$ is a smooth parametrization of ϵS_X, we know that the area contained by $r - s$ is equal to $\epsilon^2 R$. By Corollary 1.3.11 (the area formula implied by Green's Theorem) we have that

$$R = \int_0^1 y_1\, dx_1 = \int_0^1 y_2\, dx_2.$$

We now consider the area contained in the curve $\alpha r + (1 - \alpha s)$, and denote this by $\tilde{R}$. We apply Corollary 1.3.11 again to see that

$$\tilde{R} = \int_0^1 (\alpha y_1 + (1-\alpha) y_2)\ d(\alpha x_1 + (1-\alpha)x_2),$$

which we can expand to get that

$$\tilde{R} = \alpha^2 \int_0^1 y_1\, dx_1 + (1-\alpha)^2 \int_0^1 y_2\, dx_2 + \alpha(1-\alpha)\left(\int_0^1 y_1\, dx_2 + \int_0^1 y_2\, dx_1\right).$$

We can also note that

$$R = \alpha \int_0^1 y_1\, dx_1 + (1-\alpha) \int_0^1 y_2\, dx_2$$

and subtracting the previous two equalities gives that

$$R - \tilde{R} = \alpha(1-\alpha)\left[\int_0^1 y_1\, dx_1 + \int_0^1 y_2\, dx_2 - \int_0^1 y_1\, dx_2 - \int_0^1 y_2\, dx_1\right].$$

We thus get that

$$R - \tilde{R} = \alpha(1-\alpha) \int_0^1 (y_1 - y_2)\, d(x_1 - x_2).$$

Since $y_1 - y_2$ forms a smooth parametrization of ϵS_X we get that the left hand side of this equality is equal to $\alpha(1-\alpha)\epsilon^2 R$, e.g., $\hat{R} = (1 - \alpha(1-\alpha)\epsilon^2)R$.Thus the curve $\alpha r + (1-\alpha)s$ is neither strictly contained within, nor strictly contains, the curve $\sqrt{1 - \alpha(1-\alpha)\epsilon^2} S_X$. So, there is some point in the intersection, i.e, there are points $v, w \in S_X$ with $u = \alpha v + (1-\alpha)w \in \sqrt{1 - \alpha(1-\alpha)\epsilon^2} S_X$. This is as required, and constructs the isometry. □

3.4.1 Difficulty with extending this proof

There are several difficulties with extending this proof. It is perhaps worth explaining a high level view of the proof and then detailing whether or not it is possible to extend each of these steps into higher dimensions.

We have shown that any affinely independent subset of ℓ_2 embeds isometrically into every infinite-dimensional Banach space, and we have shown that the simplest possible affinely dependent set (namely, the set M from Theorem 3.4.1) embeds into every infinite-dimensional Banach space. We wish to establish a similar result for other affinely dependent sets. There are two 'obvious' next choices:

1. A collection $M = \{x_1, \ldots, x_n\}$ of affinely independent points in ℓ_2^n and a point x_{n+1} which is an affine combination of points in M.
2. The collection $M = \{0, x_1, x_2\}$ (with $0, x_1$ and x_2 not being collinear) and points y_1, y_2 that are affine combinations of $0, x_1$ and x_2.

Below we will talk about the first of these cases, i.e., we will fix an $M \subset \ell_2^n$ of affinely independent points and a point x_{n+1} that is an affine combination of points in M. Discussing this case will show us issues in extending the proof of Theorem 3.4.1 to more complicated cases. Similar ideas and issues exist in the second case.

After some preamble in the proof (and some 'trivial' approximation arguments), we start with:

Difficulty 1 - Parametrizing the unit sphere

This is the first non trivial step in the proof. We choose points in the unit sphere that have some special properties - namely we choose points r, s such that $0, r$ and s form isometric copies of the isosceles triangle with base ϵ. This is the contents of Remark 3.4.3.

Suppose that $x_{n+1} = \sum_{i=1}^{n} \alpha_i x_i$ with $\sum_{i=1}^{n} \alpha_i = 1$. For ease of stating what is going on we will be attempting to embed M into a Banach space E, with $\dim(E) = n+1$ and $d(E, \ell_2^n) = 1 + \delta$, where δ is the δ that appears in the statement of Theorem 3.2.2.[2]

If we want to replicate the proof of Theorem 3.4.1, we need to do the following iterative process:

- Take a parametrization of $\|x_1\| S_E$, we can think of this as a smooth function $\varphi_1 : [0,1]^n \to \|x_1\| S_E$.
- Take a second smooth parametrization, this time of $\|x_2\| S_E$, given by $\varphi_2 : [0,1]^n \to \|x_2\| S_E$ such that $\|\varphi_1(\boldsymbol{t}) - \varphi_2(\boldsymbol{t})\| = \|x_1 - x_2\|$ for all t. Moreover, we want $\varphi_1(\boldsymbol{t}) - \varphi_2(\boldsymbol{t})$ to form a parametrization of $\|x_1 - x_2\| S_E$.
- Take a third parametrization , this time of $\|x_3\| S_E$, given by $\varphi_3 : [0,1]^n \to \|x_3\| S_E$ such that $\|\varphi_i(\boldsymbol{t}) - \varphi_3(\boldsymbol{t})\| = \|x_i - x_3\|$ for each $i \in \{1, 2\}$. Moreover,

[2]The author is not claiming that $n+1$ is correct - merely that it gets even more confusing to state without doing this.

we want $\varphi_i(\boldsymbol{t}) - \varphi_3(\boldsymbol{t})$ to form a parametrization of $\|x_i - x_3\| S_E$ for $i \in \{1, 2\}$.

- $\vdots$
- Once we have all of our parametrization, we let $\varphi_{n+1}(\boldsymbol{t}) = \sum_{i=1}^{n} \alpha_i \varphi_i(\boldsymbol{t})$.

It is unclear how to achieve these goals - indeed the process described for the simple case above is certainly not enough. We note that in the simple case, we had to use the intermediate value theorem to do it. So, to have any chance at this, we will have to use the Brouwer's fixed point theorem. This is where we run into...

Difficulty 2 - Brouwer's Fixed Point Theorem

We are now going to use Brouwer's fixed point theorem, as in Theorem 3.2.2 to construct the embeddings φ. Since we have E as an $(n+1)$-dimensional space, we can look at an n-dimensional slice of it, say, projected onto its first n co-ordinates. We then have to 'glue' them together to attempt to do something similar to the volume arguments in Theorem 3.4.1.

So, let us go down this route. Fix some n-dimensional subspace U of $\mathbb{R}^{n+1}$. We are interested in 'how many ways this can be embedded into $\mathbb{R}^{n+1}$'. The main idea here is that the linear embedding of $\mathbb{R}^n$ into $\mathbb{R}^{n+1}$ has n degrees of freedom - if we fix some parametrization $\varphi(t)$ of S^n, we can embed the n-plane into $\mathbb{R}^{n+1}$ in such a way that it is perpendicular to the point $\varphi(t)$. Then S^n is an n-manifold, so we have n degrees of freedom - a good sign that we may be getting somewhere.

Now what happens? Since we assumed that $d(E, \ell_2^{n+1}) < 1 + \delta$, we can embed M into any of the n-dimensional slices of $\mathbb{R}^{n+1}$ using Theorem 3.2.2. This gives us the embeddings $\varphi_1, \ldots, \varphi_{n+1}$ as described above - indeed, suppose we fix some parametrization of the sphere of ℓ_2^{n+1}, this is a map $F : [0,1]^n \to S_E$ that is bijective and smooth. Then we set $\varphi_i(\boldsymbol{t})$ to be the image of x_i under the isometric embedding constructed in Theorem 3.2.2 where we take the target Banach space to be $\{x : x \cdot F(\boldsymbol{t}) = 0\}$.

However, the problem is that the construction in Theorem 3.2.2 uses Brouwer's fixed point theorem and as commented in Remark 1.3.2 it is certainly possible that these embeddings are not continuous.

Remark 3.4.4. This is where Theorem 1.3.5 can come into play. If we, somehow, had engineered the Banach spaces so that the slices had *unique* places for M to sit, we would be able to engineer that these embeddings were continuous. This seems counterintuitive - to find nice properties we look where very not nice properties exist.

This is, however, very hard. Even if we did manage this we run into...

Difficulty 3 - Volumes

Let us look at the final part of the proof of Theorem 3.4.1. Now we are trying to compare the volume contained in $\varphi_{n+1}(\boldsymbol{t})$ with the volume of the Euclidean n-sphere, and check whether it is right or not. In the case of two points we used Green's theorem in the plane. A natural generalization of this is the divergence theorem that states $\int_\Omega \nabla \cdot F dx = \int_{\partial\Omega} F \cdot n ds$, where n is a normal to $\partial\Omega$ and Ω is a domain. Let us write it out in co-ordinate form where $F = (p_1, \ldots, p_n)$, in which case the divergence theorem says that

$$\int_{\partial\Omega} \sum_{i=1}^{n} p_i \prod_{j \neq i} dx_j = \int_\Omega \left(\sum \partial_{x_i} p_i\right) dx_1 \ldots dx_n.$$

This means that we can look at any choice of functions p_i that make the right hand side constant. Taking cues from the proof of Theorem 3.4.1 we can, e.g., choose $F = x_i e_i + x_j e_j$ to find statements such as

$$\mathrm{Vol}(\Omega) = \int_\Omega x_i \, (dx_1 \ldots dx_n)^i + x_j \, (dx_1 \ldots dx_n)^j,$$

where the superscript j indicates we omit dx_j.

Let us write $\varphi_i(t) = (x_i^1, \ldots, x_i^n)$. The volume R contained in $\varphi_{n+1}(\boldsymbol{t})$ is thus equal to $\mathrm{Vol}(R) = \int_\Omega (\sum \alpha_i x_i^n) d(\sum_i \alpha_i \prod_{j=1}^n dx_j^i)$. We can expand this and we get a sum of terms of the form $\int_\Omega x_i^n \prod_j dx_j^i$ and $\int_\Omega x_i^n \prod_j dx_j^k$, with $k \neq i$. Terms of the first type we know the value of, and terms of the second type we can deduce the value of (since we know that $\varphi_i - \varphi_k$ parametrizes some scaled copy of the n-sphere).

So difficulty 3 is not a difficulty! The linear relation will be *independent* of the norm on the space. This means that $\mathrm{Vol}(R)$ will be the same as the volume as if we were in the Euclidean case. Thus, if we can solve difficulties 1 and 2 we will be done.

Summary

We thus have a proof strategy that is nearly complete, all we have to do is either:

- Choose our norm such that embeddings are unique, so Theorem 1.3.5 applies.
- Choose our norm in a way with some 'nice' properties so that the non-uniqueness does not matter. One possible idea here would be to choose an 'algebraic' norm, i.e., a norm for which $\|x\| = 1$ is the solution to some polynomial equation. We would then impose *algebraic* conditions to deduce existence of embeddings, however, the author can not see a way of completing this.[3]

[3]There is an issue where the solutions of polynomial equations can fail to be connected, and thus a similar problem happens. However if the generated set is an irreducible variety then it *is* connected. The reader may find it interesting (the author certainly does) that the condition of irreducibility on the variety seems remarkably close to the condition of uniqueness of solutions for the norm.

Chapter 4

The ℓ_p Case

We now arrive at the case of a general $1 < p < \infty$. Since this is the most complicated case, we will assume the reader's familiarity with the general strategy taken in the last two sections. Namely, the proof strategy is to show that "most" n-point subsets of ℓ_p^n have the property that small perturbations of them embed isometrically into ℓ_p^n. This 'openness' is then leveraged to show that "most" n-point subsets of ℓ_p embed isometrically into spaces almost isometrically containing ℓ_p^m for each $m \in \mathbb{N}$.

In each of these two steps we have complicating factors because of the fact that we are in ℓ_p and not in ℓ_2 or ℓ_∞. We shall list the major complicating factors now.

First complicating factor: In the case of ℓ_2 and ℓ_∞ we have a 'dimension reduction' result, i.e., if we take an n-point subset of ℓ_2 (or ℓ_∞) it embeds isometrically into ℓ_2^n or ℓ_∞^n.

This is false in general - a fact that we mentioned before the statement of Theorem 1.1.6. Indeed, to briefly remind the reader, the paper [3] has an example of an n-point subset of ℓ_1 that embeds into ℓ_1^m just when $m \geq \binom{n-2}{2}$.

Second complicating factor: In the case of ℓ_2 (and ℓ_∞) we have the following statement: Every n-point subset $\{x_1, \ldots, x_n\}$ of ℓ_2^n (or ℓ_∞^n) is isometric to an n-point subset $\{y_1, \ldots, y_n\}$ of ℓ_2^n (or ℓ_∞^n) such that the $y_i \in \operatorname{span}\{e_1, \ldots, e_i\}$.

In the case of ℓ_2 this follows from the Gram-Schmidt process. In the case of ℓ_∞ this follows from a minor modification of Lemma 1.2.7 - indeed if we have any n-point metric space $M = \{m_1, \ldots, m_n\}$ then the mapping $\tilde{f}(m_j) = (d(m_j, m_i))_{i=1}^{j}$ is an isometric embedding of M into ℓ_∞^n.

This fact was used implicitly in the proofs of Theorem 2.2.2 and 3.2.2.

In the case of ℓ_p^n it is not true that an n-point subset $\{x_1, \ldots, x_n\}$ of ℓ_p^n is isometric to a subset $\{y_1, \ldots, y_n\}$ of ℓ_p^n such that $y_i \in \operatorname{span}\{e_1, \ldots, e_i\}$. This is a major complicating factor in establishing analogues of Theorems 3.2.2 and 2.2.2.

Third Complicating Factor: We have a classical method to establish whether an n-point metric space isometrically embeds into ℓ_2, namely, the Cayley-Menger determinant. Moreover, the conditions arising in the Cayley-Menger determinant lead to the construction of an isometric embedding of this metric space into ℓ_2. The isometric embedding constructed this way is continuous in the distances in the metric space, as observed in Remark 1.2.15.

In ℓ_∞ we have the stronger result that *every* n-point metric space isometrically embeds into ℓ_∞, through the Frechét-Kolmogorov embedding. We observed in Remark 1.2.8 that this isometric embedding is continuous in the distances involved.

To the best of the author's knowledge there is no known generalisation of these ideas to the case of ℓ_p, i.e., there is no known condition on the distances in a metric space that imply that it is embeddable into ℓ_p. Moreover, we would require that such an idea would lead to an explicit embedding of the metric space into ℓ_p that gave rise to an embedding that was continuous in the distances, similar to the continuity discussed in Remark 1.2.8 and 1.2.15.

This lack of global condition is the main obstacle in proving an equivalent result for ℓ_p.

4.1 The setup

Suppose that X is a Banach space that uniformly contains ℓ_p^n, i.e., there is some universal constant C such that $\ell_p^n \overset{C}{\hookrightarrow} X$ for each $n \in \mathbb{N}$. Fix some n-point subset $M = \{m_1, \ldots, m_n\}$ of ℓ_p.

By Krivine's Theorem, Theorem 1.1.32, for each $m \in \mathbb{N}$ and $\epsilon > 0$ we can find a subspace $E_{m,\epsilon}$ of X and a map $T_{m,\epsilon} : \ell_p^m \to E_{m,\epsilon}$ such that $\|T_{m,\epsilon}\| \|T_{m,\epsilon}^{-1}\| < 1 + \epsilon$. Without loss of generality we may suppose that $\|x\| \leq \|T_{m,\epsilon} x\| \leq (1 + \epsilon)\|x\|$.

It easily follows from Ball's result, Theorem 1.1.6, that M almost isometrically embeds into X. We give a direct proof of this result[1] since we will use a similar

[1] As direct as a proof can be that invokes Krivine's Theorem.

strategy later.

The way we show that M almost isometrically embeds into X is to show that for each $\epsilon > 0$ we can choose m large enough and δ small enough such that there exists $f_\epsilon : M \to E_{m,\delta} \leq X$ such that f_ϵ has distortion $(1+\epsilon)$.

For this aim, fix some $N > 0$ such that for each i we have $\sum_{j=N+1}^{\infty} |m_i^{(j)}|^p < \delta^p$, and let $\tilde{m}_i = (m_i^{(1)}, \ldots, m_i^{(N)})$. We now define a mapping f_ϵ from M to X by sending m_i to $T_{m,\delta}(\tilde{m}_i)$.

We can now see that

$$\|m_i - m_j\| - 2\delta \leq \|f_\epsilon(\tilde{m}_i) - f_\epsilon(\tilde{m}_j)\| \leq (1+\delta)(\|m_i - m_j\| + 2\delta),$$

so evidently the distortion of this mapping tends to one as $\delta \to 0$.

We can now ask the following question, which is us spelling out the sixth part of Question 1.4.2.

Question 4.1.1. *Suppose that M is a finite subset of ℓ_p and X is a Banach space uniformly containing ℓ_p^n. Then does M isometrically embed into X?*

In this chapter we will show that the answer to this question is, in general, no. However, we will show that if we are willing to restrict to a 'large collection' of finite subsets of ℓ_p, the answer becomes true:

Theorem 4.1.2. *Suppose that X is a Banach space such that ℓ_p^n almost isometrically embeds into X for each n. Then, for each $n \in \mathbb{N}$, the set of n-point subsets of ℓ_p that do not embed into X is nowhere dense and Haar null.*

Since the positive partial answer to Question 4.1.1 is more involved than the positive answers we have obtained in the previous sections, we will split it into sections in an attempt to make it easier to follow.

We begin with a definition: if we have M, a metric space on $\{m_1, \ldots, m_n\}$ and N, a metric space on $\{r_1, \ldots, r_n\}$, we say that N is an *ϵ-perturbation of M* if

$$|d_M(m_i, m_j) - d_N(r_i, r_j)| < \epsilon$$

for all i, j.

4.2 A result on ϵ-perturbations of subsets of ℓ_p

Throughout this section, where working in $\mathbb{R}^n$, *almost all* will be refering to "with respect to Lesbegue measure". We will fix some $1 < p < \infty$ and some $n \in$

$\mathbb{N}$. We will denote the p-norm on ℓ_p^n by $\|.\|$. The following theorem may be of independent interest, but for us it provides the 'openness' criterion for subsets of ℓ_p^n.

In what follows there is a slight technical difference between 'unordered n-point subsets of $\mathbb{R}^n$' and 'n-tuples of elements of $\mathbb{R}^n$'. We will primarily prove all results in the case of n-tuples of elements, since if our results are true for n-tuples of elements then they are clearly true for n-point subsets.

Theorem 4.2.1. *For almost all n-point subsets M of ℓ_p^n, there is an $\epsilon > 0$ such that if N is an $\epsilon-$perturbation of M then N isometrically embeds into ℓ_p^n.*

Before this moment, we have been using U_n to refer to the set of symmetric matrices with zero on the diagonal. We will take a brief departure here and use $U = U_n$ to refer to upper triangular matrices.[2] We hope that the reader can forgive this slight abuse of notation given the oncoming notational storm.

Let us (very briefly) motivate the proof strategy, since this is a departure from everything we have done so far. Let

$$\Xi : \{\text{ordered } n\text{-point metric spaces}\} \to \mathrm{Mat}_{n\times n}(\mathbb{R})$$

be given by $\Xi(M) = \mathcal{N}$ where $\mathcal{N}_{ij} = d_M(m_i, m_j)_{i=1}^{j}$. Then the statement that a metric space embed isometrically into ℓ_p^n is the statement that $\Xi(M) \in \Xi(\{n\text{-point subsets of } \ell_p\})$.

Let Ξ' denote Ξ restricted to the set of ordered n-point subsets of ℓ_p. The statement in Theorem 4.2.1 is equivalent to the following: for almost all n-point subsets M of ℓ_p^n there is an $\epsilon > 0$ such that if $\mathcal{N}$ represents a metric space and $\|\Xi'(M) - \mathcal{N}\|_\infty < \epsilon$ then there is some $N \subset \ell_p^n$ of size n, such that $\Xi'(N) = \mathcal{N}$. This is an *equivalent* rephrasing of Theorem 4.2.1 and (hopefully) will inform the following definitions and notation.

Let $\mathrm{Mat}_n(\mathbb{R}) = \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ times}}$, which throughout this proof we are considering as the set of ordered n-point subsets of ℓ_p^n. Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$ and e_i^j be the element of $\mathrm{Mat}_n(\mathbb{R})$ with e_j in the i'th co-ordinate and zero everywhere else. Given $m = (m_1, \ldots, m_n)$ we denote the j'th co-ordinate of m_i (with respect to the standard basis) as m_i^j so that $m = \sum_{i,j} m_i^j e_i^j$. Let $\mathcal{M}^{ij}$ be the $n \times n$ matrix with 1 in the (i, j)-entry and 0 elsewhere. We have that $\mathcal{M}^{ij}$, $1 \le i < j \le n$ forms a basis for U.

[2]This simplifies notation in the proof a little - one can do the entire proof with U_n being the space of symmetric matrices, the modifications are slight but increase the amount of notation.

We define the map $\Theta = \Theta_n : \mathrm{Mat}_n(\mathbb{R}) \to U$ by

$$\Theta(m_1, \ldots, m_n) = \left(\|m_i - m_j\|^p\right)_{1 \le i < j \le n}.$$

The definition of Θ is chosen as a more convenient definition to work with than Ξ'.

We observe that Θ is a C^1-map[3], indeed, by computing the partial derivative in the direction e_l^k we get that

$$\frac{\partial \Theta}{\partial e_l^k}(r_1, \ldots, r_n) = \left(p|r_i^k - r_j^k|^{p-1}\mathrm{sgn}(r_i^k - r_j^k)(\delta_{il} - \delta_{jl})\right)_{1 \le i < j \le n} \tag{4.1}$$

Theorem 4.2.1 *follows* from the statement that Θ is locally open at almost all n-tuples in $\mathbb{R}^n$. It is this rephrasing we work with, i.e., we attempt (and succeed) to establish a stronger result than Theorem 4.2.1, which is Theorem 4.2.2 below. Namely, we will need that an ϵ-perturbation of M embeds into ℓ_p^n in a way that continuously depends on the perturbation (in a way we make precise below):

Theorem 4.2.2. *Let $\Theta : Mat_n(\mathbb{R}) \to U$ be defined as above. Set $C = C_n = \{m \in Mat_n(\mathbb{R}) : D\Theta|_m$ has rank $\binom{n}{2}\}$. Then C is an open subset of $Mat_n(\mathbb{R})$ whose complement is measure zero (and is thus nowhere dense). Moreover, given $m \in Mat_n(\mathbb{R})$, there is an open subset A of $Mat_n(\mathbb{R})$ containing m, a open subset B of U containing $\Theta(m)$ and a C^1-map $\Phi : B \to A$ such that $\Theta \circ \Phi = Id_B$ and $\Phi(\Theta(m)) = m$.*

Let us now spell out (in considerable detail) how the proof of Theorem 4.2.1 follows from the proof of Theorem 4.2.2. Suppose $M = \{m_1, \ldots, m_n\}$ is an n-point subset of ℓ_p^n such that $m = (m_1, \ldots, m_n) \in C$, and let $\mathcal{P}_{ij} = \|m_i - m_j\|^p$. Then, since $m \in C$, by Theorem 4.2.2 there are open subsets A of $\mathrm{Mat}_n(\mathbb{R})$ and B of U such that $m \in A$, $\Theta(m) \in B$ and $\Theta(A) = B$. Thus there is some $\epsilon > 0$ such that if $\mathcal{Q} \in U$ and $|\mathcal{Q}_{ij} - \mathcal{P}_{ij}| < \epsilon$ for all i, j then $\mathcal{Q}$ is an element of B and thus is the image under Θ of some $\{r_1, \ldots, r_n\} \in A$. Hence $\mathcal{N}_{ij} = \mathcal{Q}_{ij}^{1/p}$ defines a metric on an n-point set and the resulting metric space embeds isometrically into ℓ_p^n. This recovers the statement that ϵ-perturbations of the metric space $\{m_1, \ldots, m_n\}$ with the inherited metric embed isometrically into ℓ_p^n.

Proof of Theorem 4.2.2. Let us first show that C is an open set. Indeed, if $m \in C$, there is a linear map $B : U \to \mathrm{Mat}_n(\mathbb{R})$ such that $(D\Theta)|_m \circ B = \mathrm{Id}_U$. Since $D\Theta$ is continuous, there is some $\epsilon > 0$ such that whenever y is such that $\|m - y\| < \epsilon$, $\|D\Theta|_y \circ B - \mathrm{Id}_U\| < 1$. Thus $D\Theta|_y \circ B$ is invertible, and $D\Theta|_y$ has full rank.

[3] This is where we are using the fact that $1 < p < \infty$.

Our goal is to now show that $\mathrm{Mat}_n(\mathbb{R}) \setminus C$ has measure zero. Once we have done this, the rest of Theorem 4.2.2 will follow from the submersion theorem. Our first goal is to establish a certain subset of C.

Lemma 4.2.3. *Let $D = \{(m_1, \ldots, m_n) \in Mat_n(\mathbb{R}) : m_i = e_i + \sum_{j=1}^{i-1} m_i^j e_j$ for each $i = 1, \ldots, n\}$. Then, if $m \in D$, the partial derivatives $\frac{\partial \Theta}{\partial e_l^k}(m)$, $1 \leq k < l \leq n$, are linearly independent. In particular $D \subset C$.*

Remark 4.2.4. The reader should bear in mind "what is going on if we specialised to $p = 2$ or $p = \infty$?" In either of these cases, one can ask if the function Θ is locally open at the points of D.

If $p = 2$ then it is clear a point of D corresponds to an affinely independent subset of ℓ_2, and thus (by the first part of the proof of Theorem 3.2.2) we see that Θ is locally open here.

If $p = \infty$ then it is clear that a point of D corresponds to a concave metric space, and thus (by the first part of the proof of Theorem 2.2.2) we see that Θ is locally open here.

Remark 4.2.5. In what follows, we only ever look at the partial derivatives $\frac{\partial \Theta}{\partial e_l^k}(m)$ with $1 \leq k < l \leq n$ and show that these are linearly independent. If one carefully checks the proof of the Submersion Theorem we, in fact, obtain the following theorem:

Theorem 4.2.6. *Almost all n-tuples $(m_1, \ldots, m_n)$ of ℓ_p^n have the following property: there is some $\epsilon > 0$ such that if M is an n-point metric space $\{y_1, \ldots, y_n\}$ with the property that $|d(y_i, y_j) - \|m_i - m_j\|| < \epsilon$, then there are points $\{m_1', \ldots, m_n'\}$ such that $m_i' = m_i + \sum_{j=1}^{i-1} \alpha_{i,j} e_j$ and $\|m_i' - m_j'\| = d(y_i, y_j)$.*

This stronger form will not be used in what follows. However, it can be compared to the case of $p = 2$ and $p = \infty$. A careful analysis of these proofs show that the same is true for these.

Proof of 4.2.3. Fix $m = (m_1, \ldots, m_n) \in D$. By (4.1) we see that the (i, j) entry of $\frac{\partial \Theta}{\partial e_l^k}(m) = 0$ unless $j = l$ and $i \leq k$. We can hence expand $\frac{\partial \Theta}{\partial e_l^k}(m)$ in terms of the matrices $\mathcal{M}^{kl}$ as follows,

$$\frac{\partial \Theta}{\partial e_l^k}(m) = -p\mathcal{M}^{kl} + \sum_{i=1}^{k-1} \alpha_i^k \mathcal{M}^{il}$$

where α_i^k are constants depending on m. It follows by induction on k that $\mathcal{M}^{kl}$ is in the span of $\frac{\partial\Theta}{\partial e_i^j}(m)$ for all $1 \leq k < l \leq n$. This completes the proof of the lemma. $\square$

We now set $V = \{m = (m_1, \ldots, m_n) \in \mathrm{Mat}_n(\mathbb{R})$: there are $i, j, k \in \{1, \ldots, n\}$ such that $i \neq j$ and $m_i^k = m_j^k\}$. $\mathrm{Mat}_n(\mathbb{R}) \setminus V$ has finitely many components; these components are open and convex. Since $\mu(V) = 0$ it suffices to show that for any connected component W of $\mathrm{Mat}_n(\mathbb{R}) \setminus V$ we have that $\mu(W \setminus C) = 0$. The following lemma will be vital to this aim:

Lemma 4.2.7. *Suppose that $m = (m_1, \ldots, m_n)$ and $r = (r_1, \ldots, r_n)$ are two points in the same connected component of $Mat_n(\mathbb{R}) \setminus V$, and suppose that $\frac{\partial\Theta}{\partial e_j^i}(m)$, $1 \leq i < j \leq n$, are linearly independent. Then, for all but finitely many values of $t \in [0,1]$, the partial derivatives $\frac{\partial\Theta}{\partial e_j^i}((1-t)m + tr)$, $1 \leq i < j \leq n$ are linearly independent. In particular, for all but finitely many values of $t \in [0,1]$, we have that $(1-t)m + tr \in C$.*

It is the author's opinion that this lemma is the technical heart of this proof.

Proof. Define J to be the set $\{(k,l) : 1 \leq k < l \leq n\}$. For $\sigma = (i,j) \in J$ we will write $e_\sigma = e_j^i$, and for $\mathcal{M} \in U$ we will write $\mathcal{M}_\sigma$ for the (i,j)-entry of $\mathcal{M}$. By assumption the $J \times J$ matrix given by $\left(\left(\dfrac{\partial\Theta}{\partial e_\sigma}(m)\right)_\rho\right)$ has non-zero determinant. We now define a function $g : [0,1] \to \mathbb{R}$ by setting

$$g(t) = \det\left(\left(\frac{\partial\Theta}{\partial e_\sigma}((1-t)m + tr)\right)_\rho\right) = \det(\mathcal{M}(t)).$$

Using Equation (4.1) and the fact that m and r are in the same component of $\mathrm{Mat}_n(\mathbb{R}) \setminus V$, for each $\rho, \sigma \in J$, the matrix $\mathcal{M}(t)$ has (σ, ρ)-entry given by

$$p(a_{\sigma,\rho}t + b_{\sigma,\rho})^{p-1}\epsilon_{\sigma,\rho}$$

where $a_{\sigma,\rho}$ and $b_{\sigma,\rho}$ are non-zero constants with $a_{\sigma,\rho}t + b_{\sigma,\rho} > 0$ for all $t \in [0,1]$ and $\epsilon_{\sigma,\rho} \in \{-1, 0, 1\}$.

By compactness there is an open connected subset Ω of $\mathbb{C}$ containing $[0,1]$ such that the real part of $a_{\sigma,\rho}t + b_{\sigma,\rho}$ is positive for each $t \in \Omega$. It follows that the function g extends analytically to all of Ω, since the function $z \mapsto z^{p-1}$ is an analytic function on $\mathbb{C} \setminus \{\Re z \leq 0\}$. Since g is the restriction of an analytic function, the identity principles shows that it has at most finitely many zeroes in $[0,1]$.

Observe that $g(t)$ being non-zero implies that $\mathcal{M}(t)$ has full rank. This means that, at the point $(1-t)m + tr$, $D\Theta$ is surjective. $\square$

Remark 4.2.8. We have now shown that if there is some 'good' point in some connected component W of $\mathrm{Mat}_n(\mathbb{R}) \setminus V$ then if we choose another point of this connected component, only finitely many points on the line joining them are 'bad'. To complete the proof, we will show that there is a 'good' point in every connected component - essentially we show this by looking at the points in Lemma 4.2.3 and permuting the co-ordinates. The following argument is technical but the 'natural' thing to do.

We consider the subset R of $\mathrm{Mat}_n(\mathbb{R})$ defined by

$$R = \{(m_1, \ldots, m_n) \in \mathrm{Mat}_n(\mathbb{R}) : m_i^i > m_j^i \text{ for each } 1 \leq i < j \leq n\}.$$

Note that for each component W of $\mathrm{Mat}_n(\mathbb{R}) \setminus V$ either $W \subset R$ or $W \cap R = \emptyset$. We next show that in order to prove that $\mu(W \setminus C) = 0$ for all connected components W it suffices to consider components W that are contained within R.

Fix $(m_1, \ldots, m_n) \in \mathrm{Mat}_n(\mathbb{R}) \setminus V$. We define a permutation $\pi \in S_n$ recursively as follows: for $j = 1, \ldots, n$ let $\pi(j)$ be the unique $i \in \{1, \ldots, n\} \setminus \{\pi(1), \ldots, \pi(j-1)\}$ such that

$$m_i^j > m_k^j \text{ for all } k \in \{1, \ldots, n\} \setminus \{\pi(1), \ldots, \pi(j-1), i\}.$$

It then follows that $m^j_{\pi(j)} > m^j_{\pi(k)}$ for all $1 \leq j < k \leq n$, and hence we have that $(m_{\pi(1)}, \ldots, m_{\pi(n)}) \in R$.

Define a map $A_\pi : \mathrm{Mat}_n(\mathbb{R}) \to \mathrm{Mat}_n(\mathbb{R})$ by $A_\pi(y_1, \ldots, y_n) = (y_{\pi(1)}, \ldots, y_{\pi(n)})$ and a map $B_\pi : U \to U$ by $B_\pi(\mathcal{M}_{ij}) = \mathcal{N}_{ij}$ where

$$\mathcal{N}_{ij} = \begin{cases} \mathcal{M}_{\pi(i),\pi(j)} & \text{if } \pi(i) < \pi(j) \\ \mathcal{M}_{\pi(j),\pi(i)} & \text{if } \pi(i) > \pi(j) \end{cases}$$

It is apparent that A and B are diffeomorphisms of $\mathbb{R}^{n^2}$. We note that $B_\pi^{-1}\Theta A_\pi = \Theta$, and thus $B_\pi^{-1} D\Theta|_{A_\pi(m)} A_\pi = D\Theta|_m$, so to verify that Θ has full rank at x it is sufficient to verify that F has full rank at $A_\pi(m)$, which lies in R. This completes the proof that it is sufficient to show that $\mu(W \setminus C) = 0$ whenever W is a connected component that is contained in R.

Remark 4.2.9. We now have some sets with some strange properties and we need to show that they are Lebesgue null. If we take a connected component W of R, we can find some point $r \in C$ such that every line through r that intersects W has, at worst, countably many points on it that are not in C. This property we can exploit to show that the set of points not in C has measure zero.

Fix some component W of $\mathrm{Mat}_n(\mathbb{R}) \setminus V$ with $W \subset R$. If $\mu(W \setminus C) > 0$ we can apply Lebesgue's density theorem to find a point $r \in W \setminus C$ on which the set $W \setminus C$ is 'dense', i.e.,

$$\lim_{\epsilon \to 0} \frac{\mu(B_\epsilon(r) \cap (W \setminus C))}{\mu(B_\epsilon(r))} \to 1.$$

For $i, j \in \{1, \ldots, n\}$ we define

$$m_i^j = \begin{cases} 1 & \text{if } i = j \\ 1 & \text{if } r_i^j > r_j^j \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to verify that if $r_i^k < r_j^k$ then $m_i^k \leq m_j^k$ and thus $(1-t)m + tr \in W$ for each $t \in (0, 1]$. Moreover, since $r \in R$ we have that $m \in D$. It follows, by Lemma 4.2.3, that the partial derivatives $\frac{\partial \Theta}{\partial e_j^i}(m)$, $1 \leq i < j \leq n$, are linearly independent. Hence there is some $\epsilon > 0$ such that at each $s \in B_\epsilon(m)$ the same holds. Choose $t \in (0, 1)$ such that $u = (1-t)r + tm \in B_\epsilon(m)$. Then $u \in B_\epsilon(m) \cap W$, so there is some $\delta > 0$ such that $B_\delta(u) \subset B_\epsilon(m) \cap W$.

The Lebesgue density of $W \setminus C$ is equal to 1 at r, so by making δ smaller we may assume that $B_\delta(u) \subset W$ and $\mu(B_\delta(r) \setminus C) > 0$. By Lemma 4.2.7, each line in the direction $r - n$ through a point in $B_\delta(r)$ intersects $B_\delta(n) \setminus G$ in finitely many points. The lines in the direction $r - n$ through $B_\delta(r)$ can be parametrised by where they first intersect the hyperplane through u whose normal is in the direction $r - n$, an $\left(\binom{n}{2} - 1\right)$ $-$dimensional hyperplane. The measure of $B_\delta(r) \setminus C$ can be given by Fubini's Theorem as

$$\mu(B_\delta(r) \setminus C) = \int_{\mathbb{R}^{\binom{n}{2}-1}} \int_{[a_s, b_s]} 1_{L(s) \cap B_{\delta(r)} \setminus C} \, d\mu' ds$$

where $L(s)$ is the line through the point s in the previously mentioned hyperplane, $[a_s, b_s]$ is the interval for which $L(s)$ intersects the sphere $B_\delta(r)$ and μ' is one-dimensional Lebesgue measure. This integral is equal to zero as $L(s) \cap B_\delta(n)$ is finite. This is a contradiction on r being a point of Lebesgue density and thus of $W \setminus C$ having non-zero measure. Thus $\mu(W \setminus C) = 0$, and the proof is complete.

□

Remark 4.2.10. The above proof uses the definition of the ℓ_p norm explicitly. At multiple times throughout the proof we make explicit use of the *definition* of the ℓ_p norm instead of merely the *properties* of the ℓ_p norm has.

Remark 4.2.11. Theorem 4.2.1 may be surprising to the reader. However, we can justify it by looking at the proof of Theorem 1.1.6. In the proof of this theorem we demonstrated that the set $M_n = \{\|x_i - x_j\|^p : \{x_1, \ldots, x_n\}$ is a subset of $\ell_p\}$ was a convex cone. It can be shown that the cone M_n is 'full-dimensional', i.e., the interior of the cone contains a ball. This indicates that the following is true:

Theorem 4.2.12. *If $\{x_1, \ldots, x_n\}$ is a subset of ℓ_p such that $\Theta(x_1, \ldots, x_n)$ lies inside of the interior of the cone M_n, then there is an $\epsilon > 0$ such that whenever A_{ij}, $1 \le i < j \le n$, are numbers such that $|A_{ij} - \Theta(x_1, \ldots, x_n)_{ij}| < \epsilon$, there is some subset $\{y_1, \ldots, y_n\}$ of ℓ_p such that $\Theta(y_1, \ldots, y_n)_{ij} = A_{ij}$.*

Our choice of presentation, i.e., deducing Theorem 4.2.1 from Theorem 4.2.2, may seem strange in light of this. However, by using Theorem 4.2.2 we can deduce that for almost all $\{x_1, \ldots, x_n\}$ the set $\{y_1, \ldots, y_n\}$ is such that $\|x_i - y_i\|$ is *small*, and can be chosen in a continuous way. These two properties are key in what follows.

To attempt to use Theorem 4.2.12 in what follows we would have to show that the mapping Θ is open for almost all n-tuples of elements from ℓ_p^n. This is, essentially, the contents of Theorem 4.2.2 and we preferred to avoid reintroducing the set M_n into our presentation.

4.3 The Proof of Theorem 4.1.2

In this section we first take the result of the previous section, i.e., "small perturbation of most subsets of size n of ℓ_p^n embed isometrically into ℓ_p^n, and generalize them to ℓ_p. This takes two parts, the first is a reduction from ℓ_p to ℓ_p^N, and then we prove a technical result about ℓ_p^N.

After this we will use (essentially) the argument from the proof of Lemma 2.2.4 to finish the proof of Theorem 4.1.2.

Given a subset $\mathbb{S} = \{s_1, \ldots, s_n\}$ of $\mathbb{N}$ with $s_1 < s_2 < \cdots < s_n$, if $x = (x_i)_{i=1}^{\infty} \in \ell_p$ or $x = (x_i)_{i=1}^{N} \in \ell_p^N$ with $N > s_n$, we define the projection operator $P_{\mathbb{S}}$ by $P_{\mathbb{S}}(x) = (x_{s_1}, \ldots, x_{s_n})$. If $N \in \mathbb{N}$ we write P_N instead of $P_{\{1,\ldots,N\}}$.

We say that an n-tuple $(x_1, \ldots, x_n)$ in ℓ_p (or ℓ_p^N with $N \ge n$) has *Property K* if there is an $\mathbb{S} \subset \mathbb{N}$ of size n (or $\mathbb{S} \subset \{1, \ldots, N\}$ respectively) such that the n-tuple $(P_{\mathbb{S}}(x_1), \ldots, P_{\mathbb{S}}(x_n)) \in C_n$, where C_n is the set in Theorem 4.2.2. We will motivate the choice of Property K later: there are a few technical possibilities here that attempt to leverage the openness of Θ at a point. Note that the set of n-tuples with Property K is open since the set C_n is open, and we are taking an infinite union.

We prove Theorem 4.1.2 by showing that the closed set of n-tuples without Property K is Haar null (and thus nowhere dense), and that an n-tuple with Property K embeds isometrically into a Banach space that satisfies the assumptions of Theorem 4.1.2.

Lemma 4.3.1. *Suppose that $x = (x_1, \dots, x_n)$ is an n-tuple in ℓ_p with Property K. Then there is some $N \in \mathbb{N}$, and vectors $y_1, \dots, y_n$ in ℓ_p^N such that $\|y_i - y_j\| = \|x_i - x_j\|$ and the n-tuple $(y_1, \dots, y_n)$ has Property K.*

Remark 4.3.2. This lemma is a variant of Theorem 1.1.6. We will, briefly, discuss why we use this more technical lemma than Theorem 1.1.6. The issue with the proof of Theorem 1.1.6 is that the proof is very non-constructive - the embedding goes via a convexity argument and it is unclear whether the mapping involved is continuous. Indeed, a careful analysis of Caratheodory's theorem shows that the mapping involved isn't necessarily continuous. It is unclear *how* the mapping depends on the points, indeed, it is still an open problem whether the $\binom{n}{2}$ in the statement of Theorem 1.1.6 is the best bound. We could be in a situation where the image of every n-point subset of ℓ_p lands on some smaller dimensional manifold and we would have no hope for openness. Thus, we prefer to use the lemma here, which is very much a weaker result than Theorem 1.1.6 in terms of the bound it gives (the bound we have depends on the subset as well as n), but allows us to use Property K to apply the local openness of the mapping Θ in the previous section.

Proof. Let $\mathbb{S} \subset \mathbb{N}$ be such that $|\mathbb{S}| = n$ and $(P_{\mathbb{S}}(x_1), \dots, P_{\mathbb{S}}(x_n)) \in C_n$. After an isometry (permuting the indices) we may assume, without loss of generality, that $\mathbb{S} = \{1, \dots, n\}$. Then since $(P_n x_1, \dots, P_n x_n) \in C_n$ and C_n is open, there is some $\epsilon > 0$ such that if $z_i \in \ell_p^n$ and $\|z_i - P_n x_i\| < \epsilon$ then $(z_1, \dots, z_n) \in C_n$.

Since $(P_n x_1, \dots, P_n x_n) \in C_n$, by Theorem 4.2.2, there are open sets A containing $(P_n x_1, \dots, P_n x_n)$, $B \ni \Theta(P_n x_1, \dots, P_n x_n)$ and a C^1-map $\Phi : B \to A$ such that $\Theta \circ \Phi = \mathrm{Id}_B$ and $\Phi(\Theta(x_1, \dots, x_n)) = (x_1, \dots, x_n)$.

Fix $N \geq n$ and define $\rho_{ij} = \rho_{ij}(N)$ by $\|x_i - x_j\|^p = \|P_N x_i - P_N x_j\|^p + \rho_{ij}$. Since $\rho_{ij} \to 0$ as $N \to \infty$ there is an $N > n$ such that the element $\mathcal{N} = \mathcal{N}(N) = (\|P_n x_i - P_n x_j\|^p + \rho_{ij})_{1 \leq i < j \leq n}$ of U is in the set B. Set $z = z(N) = (z_1, \dots, z_n) = \Phi(\mathcal{N})$. By the continuity of Φ at the point $\Theta(P_n x_1, \dots, P_n x_n)$, if N is sufficiently large then

$\|z_i - P_n x_i\| < \epsilon$ and hence $(z_1, \ldots, z_n) \in C_n$.

Now define the points $y_1, \ldots, y_n \in \ell_p^N$ by:

- $P_n y_i = z_i$
- $(P_N - P_n) y_i = (P_N - P_n) x_i$.

By construction, $(P_n y_1, \ldots, P_n y_n)$ is in C_n, and thus $(y_1, \ldots, y_n)$ has Property K. We need to verify that $\|y_i - y_j\| = \|x_i - x_j\|$. Note that

$$\|y_i - y_j\|^p = \|P_n y_i - P_n y_j\|^p + \|(P_N - P_n) y_i - (P_N - P_n) y_j\|^p$$

which is equal to

$$\|z_i - z_j\|^p + \|(P_N - P_n) x_i - (P_N - P_n) x_j\|^p.$$

By the definition of z_i we have that $\|z_i - z_j\|^p = \|P_n x_i - P_n x_j\|^p + \rho_{ij}$.

□

We have shown that an n-tuple in ℓ_p with Property K is isometric to an n-tuple in ℓ_p^N with Property K. We now show an analogue of Theorem 4.2.2, i.e., that the continuity of the local inverse of Θ in ℓ_p^n can be extended to ℓ_p^N.

Lemma 4.3.3. *Suppose $(x_1, \ldots, x_n)$ is an n-tuple in ℓ_p^N with Property K. Then there is some $\epsilon > 0$ such that any ϵ-perturbation of the metric space consisting of points $(x_1, \ldots, x_n)$ with distance given by the ℓ_p norm embeds into ℓ_p^N with the embedding depending continuously on the perturbation.*

At the beginning of the proof we will make clear what "continuous dependence on the perturbation" means - it is the exact analogue of the statement of Theorem 4.2.2.

Proof. Define $\tilde{\Theta} : \underbrace{\mathbb{R}^N \times \cdots \times \mathbb{R}^N}_{n \text{ times}} \to U_n$ by

$$\tilde{\Theta}(y_1, \ldots, y_n) = (\|y_i - y_j\|)_{1 \le i < j \le n},$$

where the lack of p'th power is intentional. Our goal is to show that there is an open subset $\tilde{B}$ of U_n and a continuous map $\Psi : \tilde{B} \to \underbrace{\mathbb{R}^N \times \cdots \times \mathbb{R}^N}_{n \text{ times}}$ such that:

- $\tilde{\Theta}(x) \in \tilde{B}$
- $\Psi(\tilde{\Theta}(x)) = x$
- $\tilde{\Theta} \circ \Psi = \mathrm{Id}_{\tilde{B}}$.

Let $\mathbb{S} \subset \{1, \dots, N\}$ be such that $|\mathbb{S}| = n$ and $(P_{\mathbb{S}}(x_1), \dots, P_{\mathbb{S}}(x_n)) \in C_n$. Again, without loss of generality, we may assume that $\mathbb{S} = \{1, \dots, n\}$.

By Theorem 4.2.2, there exists an open set $A \ni (P_n x_1, \dots, P_n x_n)$, an open set $B \ni \Theta(P_n x_1, \dots, P_n x_n)$ and a C^1-mapping $\Phi : B \to A$ such that $\Phi(\Theta(P_n x_1, \dots, P_n x_n)) = (P_n x_1, \dots, P_n x_n)$ and $\Theta \circ \Phi = \mathrm{Id}_B$. Fix $\epsilon > 0$ such that if $\mathcal{N} \in U_n$ is specified such that

$$|\mathcal{N}_{ij} - \|P_n x_i - P_n x_j\|^p| < \epsilon$$

then $\mathcal{N} \in B$.

Choose $\delta = \delta(\epsilon)$ to be specified later. We set

$$\tilde{B} = \{\mathcal{N} \in U_n : |\mathcal{N}_{ij} - \|x_i - x_j\|| < \delta \text{ for all pairs } i, j\}.$$

Fix $\mathcal{N} \in \tilde{B}$. We define $\Psi(\mathcal{N})$ similarly to the way we defined the points $(y_1, \dots, y_n)$ in the proof of Lemma 4.3.1. We define $\rho_{ij} = \mathcal{N}_{ij} - \|x_i - x_j\|$ and $\epsilon_{ij} = \epsilon_{ij}(\mathcal{N})$ by $(\|x_i - x_j\| + \rho_{ij})^p = \|x_i - x_j\|^p + \epsilon_{ij}$. If $|\rho_{ij}|$ is sufficiently small (i.e., our choice of δ is sufficiently small) then $(\|P_n x_i - P_n x_j\|^p + \epsilon_{ij})_{1 \le i < j \le n}$ is in B. Define $z = \Phi((\|P_n x_i - P_n x_j\|^p + \epsilon_{ij})_{1 \le i < j \le n})$. We then set $\Psi(\mathcal{N})$ to be the n-tuple $(y_1, \dots, y_n)$ where :

- $P_n y_i = z_i$
- $(P_N - P_n) y_i = (P_N - P_n) x_i$.

We verify that $\|y_i - y_j\| = \|x_i - x_j\| + \rho_{ij} = \mathcal{N}_{ij}$, i.e. that $\tilde{F}(\Psi(\mathcal{N})) = \mathcal{N}$, as this is the only one of the three properties of Ψ and $\tilde{B}$ listed above that is non-trivial.

Indeed,

$$\|y_i - y_j\|^p = \|P_n y_i - P_n y_j\|^p + \|(P_N - P_n) y_i - (P_N - P_n) y_j\|^p,$$

which by the definition of y_i is equal to

$$\|z_i - z_j\|^p + \|(P_N - P_n) x_i - (P_N - P_n) x_j\|^p$$

which is equal to, by the definition of z_i,

$$\|x_i - x_j\|^p + \epsilon_{ij}.$$

By our definition of ϵ_{ij} this is equal to $(\|x_i - x_j\| + \rho_{ij})^p$, which is as required. $\square$

Our next, and final, lemma in this section shows that an n-point subset of ℓ_p^N with Property K embeds isometrically into any Banach space satisfying the assumptions of Theorem 4.1.2. This result is in some sense dual of Theorem 4.1.2, where

Theorem 4.1.2 says "small perturbations of metric spaces embed into the Banach space", this is saying that "the metric space embeds into small perturbations of the Banach space".

Lemma 4.3.4. *Let $x = (x_1, \dots, x_n)$ be an n-tuple in ℓ_p^N, $N \geq n$, with Property K. Then there is some $\delta > 0$ such that if $d(E, \ell_p^N) < 1 + \delta$ then $\{x_1, \dots, x_n\}$ with the metric inherited from ℓ_p embeds isometrically into E.*

From our quite laborious set up, it will turn out that the majority of this proof is similar to the cases 2 and ∞.

Proof. Let $\tilde{\Theta}$, $\tilde{B}$ and Ψ be as in the proof of Lemma 4.3.3. We now simply apply Lemma 2.2.4.[4] □

We (finally) give our proof of Theorem 4.1.2.

Proof of Theorem 4.1.2. By a combination of Lemmas 4.3.1 and 4.3.4, we see that if an n-tuple $(x_1, \dots, x_n)$ in ℓ_p has Property K, then there is some $N \in \mathbb{N}$ and $\delta > 0$ such that if E is a Banach space with $d(E, \ell_p^N) < 1 + \delta$ then $\{x_1, \dots, x_n\}$ with the metric inherited from ℓ_p embeds isometrically into E. By Krivine's Theorem, Theorem 1.1.32, any Banach space X satisfying the assumptions of the theorem (i.e., containing the spaces ℓ_p^n, $n \in \mathbb{N}$, uniformly), contains a subspace E with $d(E, \ell_P^N) < 1 + \delta$. Thus $\{x_1, \dots, x_n\}$ with the metric inherited from ℓ_p embeds isometrically into X.

To conclude we just need to see that the set A of all n-tuples of elements of ℓ_p that do not have Property K is Haar null. Indeed, the intersection of A with the finite-dimensional subspace $\ell_p^n \times \cdots \times \ell_p^N$ is contained in the complement of G_n, which by Theorem 4.2.2 has measure zero. Note also that A is translation invariant: thus by the characterization of Haar null stated in Theorem 1.1.42, A is Haar null. Since A is closed, it follows that A is nowhere dense. □

4.4 Some Remarks on the Proof of Theorem 4.1.2

We have some remarks that we would like to give about the proof of Theorem 4.1.2.

[4]This was the 'key idea' lemma for ℓ_∞ that was our use of Brouwer's fixed point theorem.

The first is our notion of Property K. If the reader analyses the above proof, there is a much more "obvious"[5] choice of Property K. We could say that an n-tuple $(x_1, \ldots, x_n)$ of ℓ_p has Property K' if the map $\Theta : \underbrace{\ell_p \times \cdots \times \ell_p}_{n \text{ times}} \to U_n$ has a continuous local right inverse at $\Theta(x_1, \ldots, x_n)$.

This choice of Property K works, in fact. All of the result of the previous section go through, and the following theorem can be deduced:

Theorem 4.4.1. *Suppose that X is a Banach space that contains an isomorphic copy of ℓ_p. Then, up to Haar null complement, every finite subset of ℓ_p embeds isometrically into X.*[6]

The weaker results of this theorem compared to the results of Theorem 4.1.2 explains why we chose the seemingly weird definition of Property K.

We could also have chosen the definition "a subset has Property K if the projection onto the first n co-ordinates are in C_n." This definition would work, but has several drawbacks that make it unsatisfying. The clearest of these is related to zero padding: suppose that $(x_1, \ldots, x_n)$ is an n-tuple that satisfies the suggested Property K and that we consider the n-tuple $(\tilde{x}_1, \ldots, \tilde{x}_n)$ where we set $\tilde{x}_i^j = x_i^{j+1}$ and $\tilde{x}_i^0 = 0$. Then the n-tuple $(\tilde{x}_1, \ldots, \tilde{x}_n)$ does not have the suggested Property K, yet is clearly isometric to the original n-tuple.

The results presented here are generalizations of our results in the previous chapters.

Theorem 4.4.2. *Every finite affinely independent subset of ℓ_2 isometrically embeds into every infinite-dimensional Banach space X.*

Proof. First note that every affinely independent set in ℓ_2 has a linearly independent translate, so without loss of generality, we may reduce to the case of linearly independent sets. Let $e_1, e_2, \ldots$ be an orthonormal basis of ℓ_2. If $\{x_1, \ldots, x_n\}$ is a linearly independent subset of ℓ_2 then there is some isometry Ψ such that $\Psi x_i \in \text{span}(e_1, \ldots, e_i)$ with the coefficient in the direction e_i of Ψx_i is non zero. Such a Ψ is constructed using the Gram-Schmidt process.

Then a minor variant of Lemma 4.2.3 (in which the coefficients of e_i in x_i is non-zero but not necessarily one) shows that the n-tuple $(\Psi x_1, \ldots, \Psi x_n)$ belongs to C_n. Thus $(\Psi x_1, \ldots, \Psi x_n)$ has Property K.

[5] Obvious to the mind of the author, not necessarily the reader.

[6] This theorem is easier to see if ℓ_p almost-isometrically embeds into X, but with a little more work one can deduce the case for ℓ_p isomorphically embedding into X.

Applying Lemma 4.3.4 to $(\Psi x_1, \ldots, \Psi x_n)$ we see that there is some $\delta > 0$ such that whenever E is an n-dimensional Banach space with $d(E, \ell_2^n) < 1 + \delta$ then $(\Theta x_1, \ldots, \Theta x_n)$ embeds isometrically into E. By Dvoretzky's Theorem there is a subspace Z of X such that $d(Z, \ell_2^n) < 1 + \delta$, and thus Z contains an isometric copy of $(\Psi x_1, \ldots, \Psi x_n)$ (which is isometric to $(x_1, \ldots, x_n)$). $\square$

In the case of ℓ_∞, our results in Section 2 essentially proceed by showing that, if $(m_1, \ldots, m_n)$ is a convex space then the mapping $\tilde{\Theta}$ is locally open at $(m_1, \ldots, m_n)$. This argument does not use differentiation - it is simpler in this case.

4.4.1 What about ℓ_1?

Our results are much less clear in the case of ℓ_1. So let us think about what happens in this case. Throughout this section we will use notation from the previous two sections.

Suppose we have an n-tuple $(x_1, \ldots, x_n)$ that has Property K. This is defined exactly as in the case of ℓ_p, i.e., there is a projection onto n co-ordinates such that the map Θ has a continuous local right inverse at this n-tuple.

A careful reading of Section 4.3 reveals that at no point did we use the assumption $1 < p < \infty$ and indeed the result holds for $p = 1$, i.e., if X contains ℓ_1^n uniformly then X contains an isometric copy of $(x_1, \ldots, x_n)$. Indeed, the main place that $1 < p < \infty$ was used is in the proof that the mapping Θ was a C^1-mapping.

In the case of $1 < p < \infty$ we used differentiation (and the submersion theorem) to show the existence of n-tuples with Property K. Here we attempt the same approach, to see how it lends itself to the case $p = 1$.

Looking at Equation (4.1) in the case $p = 1$ gives that:

$$\frac{\partial \Theta}{\partial e_l^k}(r_1, \ldots, r_n) = \Big(\mathrm{sgn}(r_i^k - r_j^k)(\delta_{il} - \delta_{jl}) \Big)_{1 \le i < j \le n}, \tag{4.2}$$

and Θ is not differentiable if $r_i^k = r_j^k$ for any i, j, k.

Note that this is only dependent on the *order* of $\{r_i^k : i = 1, \ldots, n\}$, and not on the values. Moreover, we have that $D\Theta$ is a matrix consisting of only 1's and -1's. We can observe that $D\Theta$ is invertible just when this collection of matrices has $\binom{n}{2}$ linearly independent elements.

This allows us to form the following theorem:

Theorem 4.4.3. *Let $n \in \mathbb{N}$ and $\pi_1, \ldots, \pi_n \in S_n$. We define $R = R_{\pi_1,\ldots,\pi_n}$ to be the cone of elements in $\underbrace{\ell_1^n \times \cdots \times \ell_1^n}_{n \text{ times}}$ given by the condition that $(x_1, \ldots, x_n) \in R$ just when $x^i_{\pi_i(1)} < x^i_{\pi_i(2)} < \cdots < x^i_{\pi_i(n)}$ for each i. Then the following are equivalent:*

- $R \subset C_n$
- $R \cap C_n \neq \emptyset$.

Using this theorem, and the observation that all of the theorems in Section 4.3, we can form the following:

Theorem 4.4.4. *Suppose that $(x_1, \ldots, x_n) \in \underbrace{\ell_1 \times \cdots \times \ell_1}_{n \text{ times}}$. Suppose that there is some subset $i_1, \ldots, i_n$ of $\mathbb{N}$ such that $P_{\{i_1,\ldots,i_n\}}x_1, \ldots, P_{\{i_1,\ldots,i_n\}}x_n \in R_{\pi_1,\ldots,\pi_n}$ with $R_{\pi_1,\ldots,\pi_n} \subset C_n$.[7] Then if ℓ_1 is finitely representable in a Banach space X, X isometrically contains a copy of $(x_1, \ldots, x_n)$.*

This result is not particularly satisfying - however we can not do better. One can show that certain choices of $\pi_1, \ldots, \pi_n$ lead to $R_{\pi_1,\ldots,\pi_n}$ not being in C_n - a good example is in the case of 3 points x_1, x_2 and x_3 where $x^i_1 < x^i_2 < x^i_3$ for all i, j. This corresponds to a metric space where $d(x_1, x_2) + d(x_2, x_3) = d(x_1, x_3)$ - a space where equality is attained in the triangle inequality. Identical reasoning to Remark 2.2.3 allows us to see that the function Θ is not open at this point - and similar reasoning to Theorem 2.3.1 shows us that some minor variants of such a space will not necessarily embed into a space that is a small perturbation of ℓ_1.

Remark 4.4.5. Note that there are quite a lot of degrees of freedom in different possible choices of R. The natural correspondence (through taking the derivative of Θ at a point in some $R_{\pi_1,\ldots,\pi_n}$) between $\{\pi_1, \ldots, \pi_n\}$ and $\binom{n}{2} \times \binom{n}{2}$ matrices consisting of only ± 1's is surjective: any possible choice of ± 1's can be attained.

4.5 Negative Results

In this section we will present two negative results. These results show that the we cannot drop the phrase "almost all" from Theorem 4.1.2. We will first present an example to show that the we cannot strengthen Theorem 4.1.2 to all n-point subsets of ℓ_p, i.e., the following question has a negative answer,

[7] We write $\in R$ to accentuate that this is a condition on the *ordering* of the co-ordinates and not on the *values* of the co-ordinates.

Question 4.5.1. *If X is a Banach space such that $\ell_p^n \hookrightarrow X$ for every $\epsilon > 0$ and $n \in \mathbb{N}$, then does every finite subset of ℓ_p embed isometrically into X?*

We are also interested in the following (harder) variant:
Question 4.5.2. *If X is a Banach space isomorphic to ℓ_p, then does every n-point subset of ℓ_p embed isometrically into X?*

We note that a negative answer to Question 4.5.2 forms a counter example to a possible strengthening of Theorem 4.1.2. However, the counter example for Question 4.5.2 is harder, and will require us to develop more theory. We will first take a detour into some classical theory.

4.5.1 James constants and Jordan-von Neumann constants

James constants and Jordan-von Neumann constants are isometric quantities that characterize how close a Banach space is to being Hilbert. We will first give the definition and give some examples.

Suppose X is a Banach space. We define the *Jordan-von Neumann constant of X* by

$$JvN(X) = \sup\left\{\frac{\|x+y\|^2 + \|x-y\|^2}{2\|x\|^2 + 2\|y\|^2} : x, y \neq 0\right\}$$

and the *James constant of X* by

$$J(X) = \sup\left\{\min\left(\|x+y\|, \|x-y\|\right), \|x\| = \|y\| = 1\right\}.$$

Let us start by computing the values of these constants for the ℓ_p spaces.
Lemma 4.5.3. *If $1 < p \leq 2$ then $J(\ell_p) = 2^{1/p}$ and $JvN(\ell_p) = 2^{2/p-1}$, and if $2 \leq p < \infty$ then $J(\ell_p) = 2^{1-1/p}$ and $JvN(\ell_p) = 2^{1-2/p}$.*

These results are classical. The proof we present for the Jordan-von Neumann constant is due to Clarkson in [7]. The proof that we write for the James constant is the author's, but similar proofs may well appear in the literature. We will need to generalise this proof to answer Question 4.5.2.

Proof. In both of these cases, it is easier to write down the answer, and then prove that this is the correct value. By Ball's result, Theorem 1.1.6, we have that $J(\ell_p) = J(\ell_p^2)$ (and similar for the Jordan von-Neumann constant).

The proofs split up naturally into the case $p = 2$, $1 < p < 2$ and $2 < p < \infty$. As before, we will let p' be the conjugate index of p, i.e., $p^{-1} + (p')^{-1} = 1$.

$p = 2$

To compute $JvN(\ell_2)$ we note that the parallelogram identity says that $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$ and the Jordan-von Neumann constant of ℓ_2 is 1.

To compute the James constant, note that $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle$ and $\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x, y\rangle$, i.e., the minimum of these two quantities is maximized when $\langle x, y\rangle = 0$. In this case we have that $\|x+y\|^2 = \|x-y\|^2 = 2$, and thus the James constant of ℓ_2 is equal to $\sqrt{2}$.

$1 < p < 2$

In the case $1 < p < 2$ we will show that the suprema in the definition of the Jordan-von Neumann constant and the James constant are both attained with the points $x = e_1$ and $y = e_2$. In this case we have that $\|x+y\| = \|x-y\| = 2^{1/p}$, i.e., the James constant is $2^{1/p}$ and the Jordan-von Neumann constant is $2^{2/p-1}$.

We now need to show that these are maximal. First let us show that the Jordan-von Neumann constant is equal to $2^{2/p-1}$. Note that, by an application of Hölder's inequality, $\|x+y\|^2 + \|x-y\|^2 \le 2^{(p'-2)/p'}(\|x+y\|^{p'} + \|x-y\|^{p'})^{2/p'}$. We now use the inequalities of Clarkson, Lemma 1.1.13, to get that

$$\|x+y\|^2 + \|x-y\|^2 \le 2(\|x\|^p + \|y\|^p)^{2/p}.$$

An application of Hölder's inequality shows that

$$\|x\|^p + \|y\|^p \le 2^{1-p/2}(\|x\|^2 + \|y\|^2)^{p/2},$$

and substitution of this into the previous expression shows that

$$\frac{\|x+y\|^2 + \|x-y\|^2}{2(\|x\|^2 + \|y\|^2)} \le 2^{2/p-1}.$$

We now show that the James constant of ℓ_p is $2^{1/p}$. Indeed, suppose that there are points $x, y \in S_{\ell_p}$ such that $\|x+y\|$ and $\|x-y\|$ are both larger than $2^{1/p}$, i.e., $\|x+y\|^p + \|x-y\|^p > 4$. Applying the fourth generalised Clarkson inequality from Theorem 1.1.16, we see that

$$\|x+y\|^p + \|x-y\|^p \le 2(\|x\|^p + \|y\|^p) = 4,$$

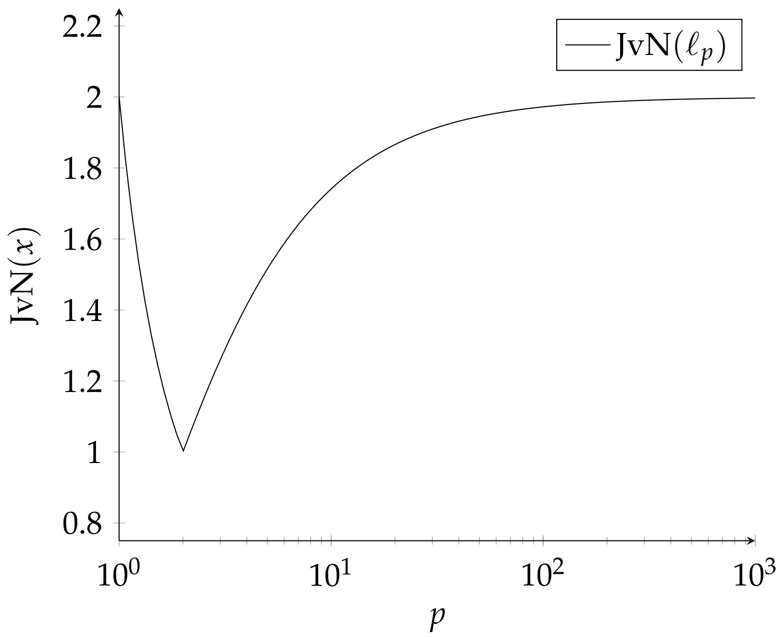
a contradiction.

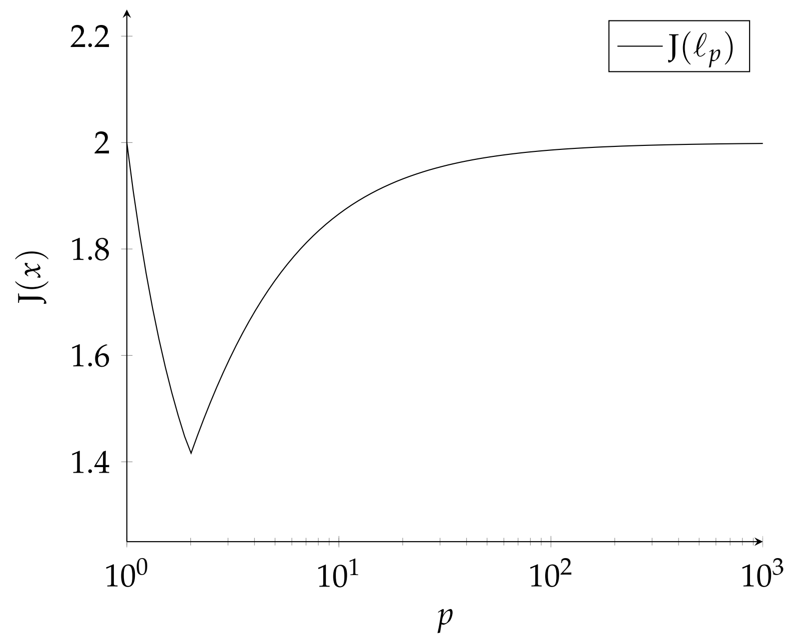
$2 < p < \infty$

In the case $2 < p < \infty$ the Jordan-von Neumann constant and the James constant are both attained with the unit vectors $x = (e_1 + e_2)/2^{1/p}$ and $y = (e_1 - e_2)/2^{1/p}$. In this case we have that $\|x + y\| = \|x - y\| = 2^{1-1/p}$ and thus $J(\ell_p) = 2^{1-1/p}$ and $JvN(\ell_p) = 2^{1-2/p}$.

The proofs here are entirely the same as the previous case, except we apply the relevant Clarkson inequality where needed (noting that Hölder's inequality still works identically.) □

We will include plots here of the Jordan-von Neumann constant of ℓ_p as p varies, and of the James constant of ℓ_p as p varies. The key property of these two graphs is that we have two isometric invariants of spaces that are minimized for $p = 2$.





Our strategy of proof in what follows is, for each p, to find strictly convex spaces X isomorphic to ℓ_p such that one of the constants $JvN(X)$ or $J(X)$ is equal to the corresponding constant for ℓ_p, but that the supremum in the definition is not attained. This is then enough to show the result. Indeed, this is the contents of the following theorem:

Theorem 4.5.4. *Suppose that X is a space isomorphic to ℓ_p such that either $JvN(X) = JvN(\ell_p)$ or $J(X) = J(\ell_p)$, but the supremum in the definition of the constant is not attained. Then there is a five-point subset of ℓ_p that does not embed isometrically into X.*

Proof. Suppose $p < 2$. In this case the set $U = \{0, \pm e_1, \pm e_2\}$ does not embed isometrically into X. Indeed, suppose there was an isometric embedding $f : U \to X$. Then, by the strict convexity of X (and Lemma 1.2.5) we have that $f(-e_1) = -f(e_1)$ (and the same for e_2). Denote $f(e_1) = x$ and $f(e_2) = y$. Then we have that $\|x + y\|_X = \|x - y\|_X = \|e_1 \pm e_2\|_p = 2^{1/p}$. This is a contradiction of the fact that the supremum in the definition is not attained.

If $2 < p < \infty$ the set $V = \{0, \pm(e_1 + e_2)/2^{1/p}, \pm(e_1 - e_2)/2^{1/p}\}$ takes the role of U above, and the proof is the same. □

Remark 4.5.5. A consequence of Krivine's theorem is that if X is isomorphic to ℓ_p then $J(X) \geq J(\ell_p)$ and $JvN(X) \geq JvN(\ell_p)$. Indeed, Krivine's Theorem (see Theorem 1.1.32) implies that for any $\epsilon > 0$ we have that $\ell_p^2 \overset{1+\epsilon}{\hookrightarrow} X$. We can thus see that, for any $\epsilon > 0$ we have that $J(X) \geq (1+\epsilon)^{-1} J(\ell_p^2)$. Since $J(\ell_p^2) = J(\ell_p)$, and we can take ϵ as small as we please, we have that $J(X) \geq J(\ell_p)$. A similar computation shows us that this holds for $JvN(X)$ as well.

4.5.2 The easy case

We can now give an easy proof of the following:

Theorem 4.5.6. *For each $p \neq 2$ there is a Banach space X such that ℓ_p is finitely represented in X, and a five point subset M of ℓ_p such that M does not isometrically embed into X.*

In this section we are denoting the ℓ_p norm by $\|.\|_p$

Proof. Fix some $p \neq 2$. We define a sequence p_n to be a sequence converging to p with the following property:

- If $p > 2$ we take $2 < p_n < p$ and $p_n \nearrow p$.
- If $p < 2$ we take $p < p_n < 2$ and $p_n \searrow p$.

We define the space $X = \left(\oplus_{n=1}^{\infty} \ell_{p_n}^n\right)_2$. Evidently ℓ_p is finitely representable in X: any finite dimensional subspace of ℓ_p is $(1+\epsilon)$-equivalent to a finite dimensional subspace of ℓ_p^N for sufficiently large N (by truncation), and ℓ_p^N can be $(1+\epsilon/2)$-embedded into $\ell_{p_M}^M$ for M large enough.

We claim that the Jordan-von Neumann constant of X is the same as the Jordan-von Neumann constant of ℓ_p, but that there are no points $x, y \in X$ such that $\frac{\|x+y\|^2+\|x-y\|^2}{2(\|x\|^2+\|y\|^2)}$ is equal to the Jordan-von Neumann constant of X.

For $x, y \in X$ we have that

$$\begin{aligned}\|x+y\|^2 + \|x-y\|^2 &= \sum_{n=1}^{\infty} \|x_n + y_n\|_{p_n}^2 + \|x_n - y_n\|_{p_n}^2 \\ &\leq \sum_{n=1}^{\infty} JvN(\ell_{p_n}) 2(\|x_n\|^2 + \|y_n\|^2).\end{aligned}$$

This is *strictly* less than

$$JvN(\ell_p) 2 \sum_{n=1}^{\infty} (\|x_n\|^2 + \|y_n\|^2) = JvN(\ell_p) 2(\|x\|^2 + \|y\|^2)$$

i.e., that

$$\frac{\|x+y\|^2 + \|x-y\|^2}{2(\|x\|^2 + \|y\|^2)} < JvN(\ell_p).$$

It is, however, obvious that $JvN(X) = JvN(\ell_p)$. Indeed, $JvN(X) \geq JvN(\ell_{p_n}^2) = JvN(\ell_{p_n})$ for each n, and thus (as $JvN(\ell_{p_n}) \to JvN(\ell_p)$) we have that $JvN(X) = JvN(\ell_p)$.

We are thus done by Theorem 4.5.4. □

4.5.3 Interlude

The results of the previous section give us the "easy" counter-example. This, in fact, answers our question about the Finite Isometric Krivine Property, one of the questions in Question 1.4.2 - indeed we have the following:

Theorem 4.5.7. *There is a superreflexive space with unconditional basis that fails to have the Finite Isometric Krivine Property (and thus fails Property AI-I).*

Proof. We have to demonstrate that $X = \left(\oplus_{n=1}^{\infty} \ell_{p_n}^n\right)_2$ is superreflexive (since the other constraints are obvious). Indeed, every uniformly convex space is superreflexive, and we only have to show that this space is uniformly convex. This follows from a result of Day - the interested reader can see the paper [9].[8] □

Remark 4.5.8. The space X constructed above has the property that, upon passing to an infinite dimensional subspace of X, we lose the property that ℓ_p is finitely representable in X. Thus the 'Krivine p' of X is not actually p. It is, in fact, 2. The space we construct later will not have this property.

We will have to work a little harder to create spaces X which are *isomorphic* to ℓ_p but fail to have the Finite Isometric Krivine Property, not just spaces X that almost-isometrically contain ℓ_p^n. The proof splits up naturally into two cases, the case $p < 2$ and the case $p > 2$.

It is worth giving a little intuition about what we are doing. We are attempting to perturb the ℓ_p norm "in the direction of Hilbert space". I think of this as "trying to make the ℓ_p norm slightly rounder". This explains why the proof ends up being different for $p > 2$ and $p < 2$, we have to push in different directions.

We will show the following theorem:

Theorem 4.5.9. *For $p \neq 2$ there is a strictly convex Banach space X such that X is isomorphic to ℓ_p, $J(X) = J(\ell_p)$ and the supremum in the definition of the James constant is not attained.*

We will then be done by Theorem 4.5.4.

The $p < 2$ case ends up being easier, and we will start with this. The main technical tool in this section will be Clarkson's inequalities.

[8] Establishing this is a little non-trivial. The reader should have a gut feeling that the result is true (as the author does) but may find proving it harder than expected.

4.5.4 The $1 < p < 2$ case

For this case, we will recall the definition and some basic properties of Orlicz spaces.

Orlicz spaces

For more details on Orlicz spaces, the interested reader can see [26, Section 4.a] where a lot of information is given. To make this section self contained, we will recall the definition and give different properties. Orlicz spaces are examples of spaces that are rearrangement invariant sequence spaces, i.e., they have a basis e_n (canonically identified as the sequence $(0, \ldots, 0, 1, 0, \ldots)$ with 1 in the n'th place) such that $\| \sum a_n e_n \| = \| \sum a_{\pi(n)} e_n \|$ for all permutations π of the natural numbers.

We define the norm on an Orlicz sequence space in terms of a function $r : [0, \infty) \to [0, \infty)$[9] that is convex, continuous, non-decreasing, $r(0) = 0$ and $\lim_{t\to\infty} r(t) = \infty$. We call such a function an *Orlicz function*.

We now define the *Orlicz space* associated to the Orlicz function r as follows: ℓ_r is the collection of sequences $x = (x_1, x_2, \ldots)$ such that for some $\rho > 0$ we have that $\sum r(|x_i|/\rho) < \infty$. We then define $\|x\|_r$ by:

$$\|x\|_r = \inf\{\rho > 0 : \quad \sum_{i=1}^{\infty} r\left(\frac{|x_i|}{\rho}\right) \leq 1\}.$$

We shall not prove this is a norm - essentially it comes down to the conditions imposed on the function r, e.g., the convexity of r gives the triangle inequality for $\|.\|_r$. We observe that if $r(t) = t^p$ then we recover the ℓ_p spaces - i.e., the Orlicz spaces are generalizations of the ℓ_p spaces.

The main theorem on Orlicz spaces we will need is the following simple result:

Theorem 4.5.10 (Proposition 4.a.5 from [26]). *If r and s are two Orlicz functions, then ℓ_r is isomorphic to ℓ_s with the identity map being an isomorphism if there exists constant $k > 0$, $K > 0$ and $t_0 > 0$ such that*

$$K^{-1} r(k^{-1} t) \leq s(t) \leq K r(kt)$$

for all $0 \leq t \leq t_0$.

[9] The traditional letter for this function is M, however we have been consistently using M for a metric space.

Fix some $q \in (p, 2)$. We define $r(t) = t^p + t^q$ and observe that ℓ_r is evidently isomorphic to ℓ_p (by the previous theorem).

We will need to use the fact that ℓ_r is strictly convex. This follows from some classical theory - indeed this follows from the classical theory of Orlicz spaces. The interested reader can see [38, Chapter VII].

ℓ_r has the same James constant as ℓ_p but does not attain the supremum

Since ℓ_r is isomorphic to ℓ_p, by Remark 4.5.5, we have that $J(\ell_r) \geq J(\ell_p)$. We will now show that, in fact, $J(\ell_r) = J(\ell_p)$ by showing that there do not exist points $x, y \in S_{\ell_r}$ such that $\|x+y\|$ and $\|x-y\|$ are both $\geq 2^{1/p}$.

Let us spell out what it means for a vector x to have norm 1 in the space ℓ_r. In this case (by the continuity of the function r) we have that $\sum r(|x_i|) = 1$, which is the same as $\sum |x_i|^p + \sum |x_i|^q = 1$, i.e.,

$$\|x\|_p^p + \|x\|_q^q = 1.$$

Similarly, we have that

$$\|y\|_p^p + \|y\|_q^q = 1.$$

Adding these together we get that

$$\|x\|_p^p + \|x\|_q^q + \|y\|_p^p + \|y\|_q^q = 2.$$

Using Clarkson's inequality (namely, the third inequality of 1.1.16) we immediately see that for any points $x, y \in S_{\ell_r}$ that

$$\|x+y\|_p^p + \|x-y\|_p^p + \|x+y\|_q^q + \|x-y\|_q^q \leq 4. \tag{4.3}$$

Now suppose that there are points x, y of norm 1 such that $\|x+y\|_r$ and $\|x-y\|_r$ are both larger than or equal to $2^{1/p}$. Let us now write out what this means. This means that

$$\frac{\|x+y\|_p^p}{2} + \frac{\|x+y\|_q^q}{2^{q/p}} \geq 1$$

and

$$\frac{\|x-y\|_p^p}{2} + \frac{\|x-y\|_q^q}{2^{r/p}} \geq 1.$$

Adding these together, and multiplying by 2, we get that

$$\|x+y\|_p^p + \|x-y\|_p^p + 2^{1-q/p}\|x+y\|_q^q + 2^{1-q/p}\|x-y\|_q^q \geq 4.$$

Since $2^{1-q/p} < 1$ this contradicts Equation (4.3). Thus, if $x, y \in S_{\ell_r}$ we have that at least one $\|x+y\|$ and $\|x-y\|$ is *less* than $2^{1/p}$.

Conclusion

We are now done, i.e., we have constructed the space ℓ_r that is isomorphic to ℓ_p, but the James constant of ℓ_r is not attained. By Theorem 4.5.4 ℓ_r is an example of a space isomorphic to ℓ_p that does not contain every subset of ℓ_p isometrically.

4.5.5 The $2 < p < \infty$ case

For this case, we will recall the definition and some basic properties of modular sequence spaces, which are more complicated versions of Orlicz sequence spaces. In the previous section, we perturbed ℓ_p in the direction ℓ_2 by considering the Orlicz function $t^p + t^q$ where q lay between p and 2. However, if we do this for $p > 2$ we would create a space isomorphic to ℓ_q and not ℓ_p. It is for this reason that we have to be slightly cleverer.

Modular sequence spaces

For more details on modular sequence spaces, again, the reader can consult [26, Section 4.d]. The snappy motto of modular sequence spaces is "they are Orlicz spaces, except you can have different Orlicz functions for each n".

Suppose we have a sequence of Orlicz functions $(r_i)_{i=1}^{\infty}$. We define the *modular sequence space* associated to the sequence r_i as follows: $\ell_{(r_i)}$ is the collection of sequences $x = (x_1, x_2, \dots)$ such that for some $\rho > 0$ we have that $\sum r_i(|x_i|/\rho) < \infty$. We then define $\|x\|_{(r_i)}$ by:

$$\|x\|_{(r_i)} = \inf\{\rho > 0 : \quad \sum_{i=1}^{\infty} r_i\left(\frac{|x_i|}{\rho}\right) \leq 1\}.$$

The main theorem we will need is the analogue of Theorem 4.5.10 for modular sequence spaces, which is as follows:

Theorem 4.5.11. *If $(r_i)_{i=1}^{\infty}$ and $(s_i)_{i=1}^{\infty}$ are two collections of Orlicz functions, then $\ell_{(r_i)}$ is isomorphic to $\ell_{(s_i)}$ with the identity map being an isomorphism if there exist constants $K > 0$, $t_n \geq 0$ and an integer n_0 such that:*

- *$K^{-1} r_n(t) \leq s_n(t) \leq K r_n(t)$ for all $n \geq n_0$ and $t \geq t_n$.*
- *$\sum_{n=1}^{\infty} r_n(t_n) < \infty$.*

Fix some sequence $2 < p_i$ such that $p_i \to p$ very quickly and set $r_i(t) = t^p + t^{p_i}$. If p_i converges to p fast enough, then the previous criterion applied with $s(t) = t^p$, $K = 3$ and $n_0 = 1$ shows that the generated space $\ell_{(r_i)}$ is isomorphic to ℓ_p.

The space $\ell_{(r_i)}$ is strictly convex

Our first goal is to show that $\ell_{(r_i)}$ is strictly convex. This is not as easily deducible from classical theory (results on strict convexity of modular sequence spaces tend to not be well known), so we prefer to give an explicit proof.

We take vectors $x, y \in S_{\ell_{(r_i)}}$, with $x \neq y$. We need to show that $\|x + y\| < 2$.

Let us first look at what $x \in S_{\ell_{(r_i)}}$ means, it means that $\sum r_i(|x_i|) = 1$. This is the same as $\sum |x_i|^p + |x_i|^{p_i} = 1$, i.e.,

$$\|x\|_p^p + \sum |x_i|^{p_i} = 1.$$

Similarly we have that

$$\|y\|_p^p + \sum |y_i|^{p_i} = 1.$$

Adding these together we get that

$$\|x\|_p^p + \|y\|_p^p + \sum (|x_i|^{p_i} + |y_i|^{p_i}) = 2.$$

We now use Clarkson's inequality for $p > 2$ on this to get that,

$$2^{1-p}(\|x + y\|_p^p + \|x - y\|_p^p) + \sum_{i=1}^{\infty} 2^{1-p_i}(|x_i + y_i|^{p_i} + |x_i - y_i|^{p_i}) \leq 2, \tag{4.4}$$

with this inequality being strict unless $x_i = y_i$ for all i.

We know that $\|x + y\| \leq 2$, so for contradiction assume that $\|x + y\| = 2$. This is the statement that

$$2^{-p}\|x + y\|_p^p + \sum_{i=1}^{\infty} 2^{-p_i}|x_i + y_i|^{p_i} = 1.$$

Multiplying this by 2 and comparing with Equation (4.4) is a contradiction, i.e., $\|x + y\| < 2$.

$\ell_{(r_i)}$ has the same James constant as ℓ_p but does not attain the supremum

Since $\ell_{(r_i)}$ is isomorphic to ℓ_p, by Remark 4.5.5, we have that $J(\ell_{(r_i)}) \geq J(\ell_p)$. We will now show that, in fact, $J(\ell_{(r_i)}) = J(\ell_p)$ by showing that there do not exit points $x, y \in S_{\ell_{(r_i)}}$ such that $\|x + y\|$ and $\|x - y\|$ are both $\geq 2^{1-1/p}$.

Indeed, suppose that x, y of norm 1 such that $\|x+y\|$ and $\|x-y\|$ are both larger than or equal to $2^{1-1/p}$. Writing this out, we have that

$$\frac{\|x+y\|_p^p}{2^{p-1}} + \sum_{i=1}^{\infty} \frac{|x_i+y_i|^{p_i}}{2^{p_i(1-\frac{1}{p})}} \geq 1$$

and

$$\frac{\|x-y\|_p^p}{2^{p-1}} + \sum_{i=1}^{\infty} \frac{|x_i-y_i|^{p_i}}{2^{p_i(1-\frac{1}{p})}} \geq 1.$$

Adding and rearranging we get that

$$2^{1-p}(\|x+y\|_p^p + \|x-y\|_p^p) + \sum_{i=1}^{\infty} 2^{p_i(\frac{1}{p}-1)}(|x_i+y_i|^{p_i} + |x_i-y_i|^{p_i}) \geq 2.$$

Since $2 < p_i < p$, this equation contradicts Equation (4.4). Thus, if $x, y \in S_{\ell_{(r_i)}}$ one of $\|x+y\|, \|x-y\|$ is strictly less than $2^{1-1/p}$.

Conclusion

We are now done, i.e., we have constructed the space $\ell_{(r_i)}$ that is isomorphic to ℓ_p, for which the James constant of $\ell_{(r_i)}$ is not attained. By Theorem 4.5.4 $\ell_{(r_i)}$ is an example of a space isomorphic to ℓ_p that does not contain every subset of ℓ_p isometrically.

4.5.6 Some concluding remarks

It is, perhaps, of interest to note that the examples of subsets that do not embed isometrically in the previous section are *small*.

In fact, the examples constructed in the previous section are as small as possible. By Lemma 1.2.3, all three point metric spaces embed into every infinite-dimensional Banach space. However, in this case we have constructed a four point subset of ℓ_p that does not isometrically embed into a Banach space isometric to ℓ_p![10]

The space we constructed in 4.5.4 also gives an example of a space that lacks Property AI-I that has a collection of very nice properties:

[10] The reader may observe that the subsets from Theorem 4.5.4 consist of five points.[11] However, we may also notice that if we omit any single non-zero point from the space the proofs work identically since if an Equation of the form (4.4) is true for x, y it is automatically true for $-x, -y$.

[11] And that four does not equal five

Theorem 4.5.12. *There is a superreflexive Banach space X with symmetric, unconditional basis that lacks Property AI-I.*

The reader should now be convinced that Property AI-I is a little weird as a Banach space theory property. It bears very little relation to any niceness properties of Banach spaces: the properties above are about as nice as one could reasonably ask for. However, the weakenings of Property AI-I remain interesting. We end the thesis with the following three conjectures that the author finds very interesting.

Conjecture 4.5.13. *The original Tsirelson space is finitely isometrically universal.*

Conjecture 4.5.14. *The dual of the original Tsirelson space (constructed by Figiel and Johnson) isometrically contains every finite subset of ℓ_1.*

Conjecture 4.5.15. *Every infinite-dimensional Banach space contains every finite subset of ℓ_2 isometrically.*

Bibliography

[1] F. Albiac and N. Kalton. *Topics in Banach space theory*, volume 233 of *Graduate Texts in Mathematics*. Springer, [Cham], second edition, 2016. With a foreword by Gilles Godefory.

[2] D. Amir and V. D. Milman. A quantitative finite-dimensional Krivine theorem. *Israel J. Math.*, 50(1-2):1–12, 1985.

[3] K. Ball. Isometric embedding in l_p-spaces. *European J. Combin.*, 11(4):305–311, 1990.

[4] Yoav Benyamini and Joram Lindenstrauss. *Geometric nonlinear functional analysis. Vol. 1*, volume 48 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2000.

[5] B. Bollobás. *Linear analysis*. Cambridge University Press, Cambridge, second edition, 1999. An introductory course.

[6] N. L. Carothers. *A short course on Banach space theory*, volume 64 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2005.

[7] J. A. Clarkson. The von Neumann-Jordan constant for the Lebesgue spaces. *Ann. of Math. (2)*, 38(1):114–115, 1937.

[8] James A. Clarkson. Uniformly convex spaces. *Trans. Amer. Math. Soc.*, 40(3):396–414, 1936.

[9] Mahlon M. Day. Some more uniformly convex spaces. *Bull. Amer. Math. Soc.*, 47:504–507, 1941.

[10] F. Delbaen, H. Jarchow, and A. Peł czyński. Subspaces of L_p isometric to subspaces of l_p. *Positivity*, 2(4):339–367, 1998.

[11] A. Dvoretzky. Some results on convex bodies and Banach spaces. In *Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960)*, pages 123–160. Jerusalem Academic Press, Jerusalem; Pergamon, Oxford, 1961.

[12] H. G. Eggleston. *Convexity*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 47. Cambridge University Press, New York, 1958.

[13] Per Enflo. Banach spaces which can be given an equivalent uniformly convex norm. *Israel J. Math.*, 13:281–288 (1973), 1972.

[14] D. J. H. Garling. *Inequalities: a journey into linear analysis*. Cambridge University Press, Cambridge, 2007.

[15] A. Grothendieck. Sur certaines classes de suites dans les espaces de Banach et le théorème de Dvoretzky-Rogers. *Bol. Soc. Mat. São Paulo*, 8:81–110 (1956), 1953.

[16] Nik Weaver (https://mathoverflow.net/users/23141/nik weaver). Term for a metric space for which the triangle inequality is strict? MathOverflow. URL:https://mathoverflow.net/q/232337 (version: 2016-02-28).

[17] Alexandre Eremenko (https://mathoverflow.net/users/25510/alexandre eremenko). Continuity of mapping sending a function to its (brouwer) fixed point. MathOverflow. URL:https://mathoverflow.net/q/277063 (version: 2017-07-23).

[18] Brian R. Hunt, Tim Sauer, and James A. Yorke. Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces. *Bull. Amer. Math. Soc. (N.S.)*, 27(2):217–238, 1992.

[19] Robert C. James. Uniformly non-square Banach spaces. *Ann. of Math. (2)*, 80:542–550, 1964.

[20] W. B. Johnson and J. Lindenstrauss, editors. *Handbook of the geometry of Banach spaces. Vol. I.* North-Holland Publishing Co., Amsterdam, 2001.

[21] N. J. Kalton and G. Lancien. Best constants for Lipschitz embeddings of metric spaces into c_0. *Fund. Math.*, 199(3):249–272, 2008.

[22] J. Kilbane. On embeddings of finite subsets of ℓ_2. *ArXiv e-prints*, September 2016.

[23] J. Kilbane. On Embeddings of Finite Subsets of ℓ_p. *ArXiv e-prints*, April 2017.

[24] J. L. Krivine. Sous-espaces de dimension finie des espaces de Banach réticulés. *Ann. of Math. (2)*, 104(1):1–29, 1976.

[25] Casimir Kuratowski. Quelques problémes concernant les espaces métriques non-séparables. *Fundamenta Mathematicae*, 25(1):534–545, 1935.

[26] Joram Lindenstrauss and Lior Tzafriri. *Classical Banach spaces. I*. Springer-Verlag, Berlin-New York, 1977. Sequence spaces, Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 92.

[27] Manor Mendel and Assaf Naor. Metric cotype. *Ann. of Math. (2)*, 168(1):247–298, 2008.

[28] S. K Mercourakis and G. Vassiliadis. Isometric embeddings of a class of separable metric spaces into Banach spaces. *ArXiv e-prints*, May 2017.

[29] V. D.. Milman and G. Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.

[30] Vitali D. Milman and Gideon Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.

[31] Harold Willis Milnes. Convexity of Orlicz spaces. *Pacific J. Math.*, 7:1451–1483, 1957.

[32] A. Moltó, J. Orihuela, S. Troyanski, and V. Zizler. Strictly convex renormings. *J. Lond. Math. Soc. (2)*, 75(3):647–658, 2007.

[33] G Nordlander. The modulus of convexity in normed linear spaces. *Ark. Mat.*, 4:15–17 (1960), 1960.

[34] Edward Odell and Thomas Schlumprecht. The distortion problem. *Acta Math.*, 173(2):259–281, 1994.

[35] M. Ostrovskii. *Metric embeddings*, volume 49 of *De Gruyter Studies in Mathematics*. De Gruyter, Berlin, 2013. Bilipschitz and coarse embeddings into Banach spaces.

[36] J. R. Partington. Subspaces of certain Banach sequence spaces. *Bull. London Math. Soc.*, 13(2):163–166, 1981.

[37] F. V. Petrov, D. M. Stolyarov, and P. B. Zatitskiy. On embeddings of finite metric spaces in l_∞^n. *Mathematika*, 56(1):135–139, 2010.

[38] M. M. Rao and Z. D. Ren. *Theory of Orlicz spaces*, volume 146 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 1991.

[39] W. Rudin. *Functional Analysis*. McGraw-Hill Book Co., New York-Düsseldorf-Johannesburg, 1973. McGraw-Hill Series in Higher Mathematics.

[40] S. A. Shkarin. Isometric embedding of finite ultrametric spaces in Banach spaces. *Topology Appl.*, 142(1-3):13–17, 2004.

[41] K. J. Swanepoel. Equilateral sets in finite-dimensional normed spaces. In *Seminar of Mathematical Analysis*, volume 71 of *Colecc. Abierta*, pages 195–237. Univ. Sevilla Secr. Publ., Seville, 2004.

[42] Jarno Talponen. A natural class of sequential Banach spaces. *Bull. Pol. Acad. Sci. Math.*, 59(2):185–196, 2011.

[43] Wolfgang Weil. Ein Approximationssatz für konvexe Körper. *Manuscripta Math.*, 8:335–362, 1973.