Geometric and Functional Analysis
Vol. 2, No. 2 (1992)
1016-443X/92/0200137-36$1.50+0.20/0

MARKOV CHAINS, RIESZ TRANSFORMS AND LIPSCHITZ MAPS

K. Ball

Abstract

It is shown that a version of Maurey's extension theorem holds for Lipschitz maps between metric spaces satisfying certain geometric conditions, analogous to type and cotype. As a consequence, a classical Theorem of Kirszbraun can be generalised to include maps into L_p, $1 < p < 2$. These conditions describe the wandering of symmetric Markov processes in the spaces in question. Estimates are obtained for the root-mean-square wandering of such processes in the L_p spaces. The duality theory for these geometric conditions (in normed spaces) is shown to be closely related to the behavior of the Riesz transforms associated to Markov chains. Several natural open problems are collected in the final chapter.

Introduction

A classical theorem of Kirszbraun states that if H and K are Hilbert spaces, Z is a subset of H and $f : Z \to K$ is Lipschitz, then there is an extension $\tilde{f} : H \to K$ of f whose Lipschitz norm is no more than that of f. Theorem 4.4, an application of the principal result of this paper, generalises Kirszbraun's theorem as follows: for each p in the interval (1,2) there is a constant C_p so that if H is Hilbert space, Z is a subset of H and $f : Z \to L_p$ is Lipschitz, then there is an extension $\tilde{f} : H \to L_p$ of f with $\|\tilde{f}\|_{\text{lip}} \leq C_p \|f\|_{\text{lip}}$.

For Kirszbraun's theorem (and other classical results described in [WW]) it is critical that the Lipschitz norm of the extension can be taken to be the same as that of the original function; i.e. the theorems are "isometric". This means that the extension can be performed one point at a

Supported in part by NSF DMS-8807243.

time. As soon as the codomain K is not Hilbert space (or an L_∞-space) the extension certainly cannot be performed isometrically. For this reason, an approach is required which deals with arbitrarily many points at once. (Note: the situation here is very different from that of linear maps. Any linear map defined on a subspace of Hilbert space, extends trivially, by composition with the orthogonal projection, and there is no increase in norm.)

The methods used in this paper are inspired by the theory of type and cotype of Banach spaces (see below for definitions, or [LT] for more information). The type and cotype properties describe the behaviour of sums of independent random variables in Banach spaces. In the non-linear setting, such sums do not make sense. The analogues of type and cotype introduced in this paper describe the behaviour of Markov chains in the spaces in question. Independence in the linear theory is replaced by the Markov property in the non-linear.

It has been known for some time that the problem of extending *linear* maps is closely related to type and cotype properties of normed spaces. A well-known theorem of Maurey [M] states that if X and Y are Banach spaces, X having type 2 and Y having cotype 2, Z is a subspace of X and $T : Z \to Y$ is a bounded linear map, then there is an extension $\widetilde{T} : X \to Y$ of T and $\|\widetilde{T}\|$ is bounded by $T_2(X)C_2(Y)\|T\|$ ($T_2(X)$ and $C_2(Y)$ being the type 2 and cotype 2 constants of X and Y respectively). In Chapter 1 of this paper it is shown that an analogue of this theorem holds for extensions of Lipschitz maps, provided that the domain and codomain satisfy geometric conditions that describe the speed at which symmetric Markov chains can wander. These conditions will be called Markov type 2 and Markov cotype 2. Neither of these conditions has previously been studied even in the context of the linear theory. It is hoped that this paper will provide a stimulus to the further investigation of Markov chains in normed spaces. It also seems reasonable to hope that the Markov type and cotype properties will have applications to the theory of more general manifolds, where the study of Markov chains already plays an important role.

As mentioned above, the basic (abstract) extension theorem is proved in Chapter 1. In Chapter 2 it is shown that the Markov type and cotype 2 properties are stronger than their linear (or Rademacher) counterparts. Chapter 3 contains a discussion of uniform convexity and uniform smoothness and in Chapter 4 it is shown that 2-uniformly convex spaces have Markov cotype 2. This implies, in particular, that for $1 < p \leq 2$, L_p has

Markov cotype 2 and completes the proof of the theorem described in the first paragraph of this introduction.

The principal problem raised by this paper is whether L_q has Markov type 2 for $2 < q < \infty$. Chapter 5 contains a discussion of the behaviour of Markov chains in uniformly smooth spaces with some suggestions as to how the Markov type 2 problem might be tackled. An estimate on the wandering of Markov chains is proved, that is slightly weaker than the Markov type 2 condition.

Chapter 6 is intended to suggest future lines of research on the problems studied here. There is a well-developed duality theory for type and cotype (due principally to Maurey and Pisier): see e.g. [MS] for details. The problem of duality for the Markov properties is considered in Chapter 6 and shown to be closely related to the theory of Riesz transforms. Some open problems raised by this exposition are also collected in Chapter 6.

It is perhaps worth recalling some of the background on Lipschitz maps and type and cotype, before embarking upon the main discussion. If (X, d_X) and (Y, d_Y) are metric spaces, a Lipschitz function $f : X \to Y$ has Lipschitz norm defined to be

$$\|f\|_{\mathrm{lip}} = \sup\left\{\frac{d_Y\big(f(x), f(y)\big)}{d_X(x,y)} : x, y \in X \ , \ x \neq y\right\} .$$

The general extension problem is as follows. Suppose X and Y are metric spaces, Z is a subset of X and $f : Z \to Y$ is Lipschitz. Under what conditions on X and Y (and perhaps Z and f) does there exist a Lipschitz extension $\tilde{f} : X \to Y$ of f, and what can be said about $\|\tilde{f}\|_{\mathrm{lip}}$ in terms of $\|f\|_{\mathrm{lip}}$? There are two classical theorems which provide answers in special cases. Kirszbraun's theorem, stated above, provides a positive answer if X and Y are Hilbert spaces and the non-linear Hahn-Banach theorem gives extensions for arbitrary X if Y is an L_∞-space. Lindenstrauss [L1], proved a less isometric result, stating that the extension problem can always be solved if Y for example, is c_0, with $\|\tilde{f}\|_{\mathrm{lip}} \le 2\|f\|_{\mathrm{lip}}$.

In the papers [MarP], [JL] and [JLS], estimates are obtained for extensions of Lipschitz maps that are initially defined on finite sets. For example, Johnson and Lindenstrauss showed that if Y is a Hilbert space (X arbitrary), there is a constant K so that any f defined initially on an n-point subset of X has an extension $\tilde{f}$ with

$$\|\tilde{f}\|_{\mathrm{lip}} \le K\sqrt{\log n}\,\|f\|_{\mathrm{lip}} .$$

Type and cotype play an important role in the theory of normed spaces, providing a link between geometry and probability. Let $(\varepsilon_i)_1^\infty$ be an iid sequence of Bernoulli random variables on some probability space. For $1 < p \leq 2$, a normed space X is said to have type p (or Rademacher type p) if there is a constant K so that for all $n \in \mathbb{N}$ and all sequences $(x_i)_1^n$ in X,

$$E\left\|\sum_1^n \varepsilon_i x_i\right\|^p \leq K^p \sum_1^n \|x_i\|^p .$$

Cotype q, $2 \leq q < \infty$, is defined similarly but with the inequality reversed. (All normed spaces have type 1 and cotype ∞ if the definitions are extended in the obvious way.) There have been several suggested definitions of type for general metric spaces: examples of results using these may be found in [E], [BMW] and [G]. The paper by Gromov uses a type 2 property in connection with the eigenvalues of Laplacians on graphs. Given the close connection between Markov chains and differential operators, this would suggest that the Markov type 2 property should be explicitly studied for manifolds.

On the other hand, no very convenient definition of cotype was found for general metric spaces. In the context of the Lipschitz extension problem one can perhaps see why. For maps into a metric space Y to have extensions, it is necessary that Y should consist of more than a few isolated points. So one expects that a cotype property, appropriate to the extension problem, will involve some existential assertion concerning points in Y. To simplify the present exposition, it will be assumed that the codomain Y is a normed linear space (or convex subset thereof). Some remarks concerning the generalisation of cotype to arbitrary metric spaces are included in Chapter 6.

I am indebted to several people for their suggestions concerning the problems discussed here, in particular, W.B. Johnson, J. Lindenstrauss, G. Pisier and G. Schechtman. I am especially grateful to Prof. Johnson, without whose ideas and encouragement, this work would not have taken place.

1. The General Extension Problem

This chapter begins with a lemma which provides a necessary and sufficient condition for Lipschitz maps to have extensions. This lemma is a variant of one used by Maurey. A related lemma was found earlier by Johnson, Lindenstrauss and Schechtman: their result actually characterises extensions which factor through subsets of Hilbert space, a problem much closer to Maurey's argument. Their lemma provided much of the stimulus for the present work.

The second lemma in this chapter, provides a further, slight, reformulation of the problem. This is not really necessary but simplifies the succeeding arguments. After this, the Markov type and cotype properties are motivated and introduced and some equivalent forms of the Markov type property are described. Finally, the basic extension theorem for Lipschitz maps is stated and proved.

LEMMA 1.1. *Let (X,d) be a metric space and Y a normed space, Z a subset of X and $f:Z\to Y$ Lipschitz. Then, there is an extension $\tilde{f}:X\to Y^{**}$ of f with $\|\tilde{f}\|_{\text{lip}}\le K$ if (and only if), for every $n\in\mathbb{N}$, $n\times n$ symmetric matrix $H=(h_{ij})$ with non-negative entries and sequence $(x_i)_1^n$ in X, there is a map*

$$\hat{f}=\hat{f}_H:\{x_1,\dots,x_n\}\longrightarrow Y^{**}$$

which agrees with f on $Z\cap\{x_1,\dots,x_n\}$ and satisfies

$$\sum_{ij}h_{ij}\left\|\hat{f}(x_i)-\hat{f}(x_j)\right\|^2\le K^2\sum_{ij}h_{ij}d(x_i,x_j)^2\ . \tag{1.1}$$

Proof: Plainly, if $\tilde{f}$ exists, its restriction to a finite set $\{x_1,\dots,x_n\}$ will satisfy (1.1) for any H.

Conversely, suppose the condition holds. The argument is in two steps. The first is to show that for a finite set $S=\{x_1,\dots,x_n\}$ in X, there is a map $f_s:S\to Y$ which agrees with f on $Z\cap\{x_1,\dots,x_n\}$ and has Lipschitz norm at most K. Consider the set C of $n\times n$ matrices $M=(m_{ij})$ of the form $m_{ij}=\left\|\hat{f}(x_i)-\hat{f}(x_j)\right\|^2$ for some map $\hat{f}$ agreeing with f on $Z\cap\{x_1,\dots,x_n\}$. Let D be the set of matrices of the form $M+M'$ where $M\in C$ and M' has non-negative entries. Finally, let T be the matrix $t_{ij}=K^2d(x_i,x_j)^2$. The aim is to show that $T\in D$.

The set D is convex, for suppose $\hat{f}_1$ and $\hat{f}_2$ are "extensions" of f to S, M_1 and M_2 their corresponding matrices and M_1' and M_2' have non-negative entries. For a fixed $\lambda\in(0,1)$, define an "extension" $\hat{f}:\{x_1,\dots,x_n\}\to Y$ by

$$\hat{f}(x_i)=\lambda\hat{f}_1(x_i)+(1-\lambda)\hat{f}_2(x_i)\ ,\qquad 1\le i\le n$$

and let M be the corresponding element of C. Then

$$\begin{aligned} m_{ij} &\le \left(\lambda\left\|\hat{f}_1(x_i)-\hat{f}_1(x_j)\right\|+(1-\lambda)\left\|\hat{f}_2(x_i)-\hat{f}_2(x_j)\right\|\right)^2 \\ &\le \lambda\left\|\hat{f}_1(x_i)-\hat{f}_1(x_j)\right\|^2+(1-\lambda)\left\|\hat{f}_2(x_i)-\hat{f}_2(x_j)\right\|^2 \\ &= \lambda(M_1)_{ij}+(1-\lambda)(M_2)_{ij} \ . \end{aligned}$$

Hence, there is a matrix M' with non-negative entries so that

$$M+M'=\lambda(M_1+M_1')+(1-\lambda)(M_2+M_2') \ .$$

Now, suppose that $T \notin D$. Then there is a symmetric matrix $H=(h_{ij})$ and some $\alpha \in \mathbf{R}$ for which

$$\sum_{ij} h_{ij}m_{ij} \ge \alpha \quad \text{if} \quad (m_{ij}) \in D$$

but

$$\sum_{ij} h_{ij}t_{ij} < \alpha.$$

The definition of D ensures that each entry of H must be non-negative, so that these conditions violate the hypothesis.

The second step is to find $\tilde{f}$. Assume that Z is non-empty, $x_0 \in Z$ (say) and that $f(x_0)=0$. For each $x \in X$, let B_x be the topological space consisting of the set

$$\left\{y \in Y^{**} : \|y\| \le Kd(x,x_0)\right\} ,$$

equipped with the weak* topology inherited from Y^{**}. B_x is compact and hence, so is the Cartesian product $B=\prod_{x\in X} B_x$. For each finite subset S of X, containing x_0, there is a Lipschitz map $f_s : S \to Y^{**}$ "extending" f, with norm at most K. Let $b^{(s)}$ be the point of B given by

$$b_x^{(s)}=\begin{cases} f_s(x) & \text{if } x \in S \\ 0 & \text{otherwise} \end{cases}$$

for each $x \in X$. Note that the Lipschitz assumption on f_s guarantees that $f_s(x) \in B_x$, for each $x \in S$. If b is an accumulation point of the $b^{(s)}$'s along the net of finite subsets of X which contain x_0, then an extension $\tilde{f}: X \to Y^{**}$ is given by

$$\tilde{f}(x) = b_x, \quad x \in X. \qquad \square$$

By the homogeneity in Lemma 1.1, it would suffice to consider only those matrices H whose largest eigenvalue is 1. Let (h_{ij}) be such a matrix and $\theta = (\theta_i)_1^n$ be an eigenvector with eigenvalue 1, satisfying

$$\theta_i > 0, \quad 1 \leq i \leq n \quad \text{and} \quad \sum_1^n \theta_i^2 = 1.$$

There is a Markov semigroup associated with H as follows. Let $A = (a_{ij})$ be the matrix given by

$$a_{ij} = \theta_i^{-1} h_{ij} \theta_j \qquad 1 \leq i, j \leq n \ .$$

For each i,

$$\sum_j a_{ij} = \theta_i^{-1} \sum_j h_{ij} \theta_j = 1$$

while for each j,

$$\sum_i \theta_i^2 a_{ij} = \sum_i \theta_i h_{ij} \theta_j = \theta_j^2 \ .$$

So A has rows that sum to 1 and the vector $(\theta_1^2, \ldots, \theta_n^2)$ is a steady state vector for A. The powers $(A^k)_{k=0}^{\infty}$ form a Markov semigroup of operators on $\mathbb{R}^n$ which is said to be symmetric (or time-reversible) with respect to the probability which assigns mass θ_i^2 to the state i, $1 \leq i \leq n$. If $M_0, M_1, M_2, \ldots$ is the Markov chain satisfying

$$\begin{aligned} &P(M_0 = i) = \theta_i^2 \qquad 1 \leq i \leq n \\ &P(M_{k+1} = j \mid M_k = i) = a_{ij} \qquad 1 \leq i, j \leq n, \quad k = 0, 1, 2 \ldots \end{aligned}$$

then for any $m \in \mathbb{N}$ and sequence $i_0, \ldots, i_m$ of indices

$$\begin{aligned} P(M_0 &= i_0, M_1 = i_1, \ldots, M_m = i_m) \\ &= P(M_0 = i_m, \ldots, M_m = i_0) \ . \end{aligned}$$

The simplest examples of such Markov semigroups are those for which θ_i is independent of i: i.e. $H = A$ is a symmetric, stochastic matrix. Only these matrices will be discussed below: versions of all the results proved, hold for general symmetric semigroups, but the added complication of this generality is not needed. The next lemma explains why. From now on, a matrix (not necessarily square) will be called stochastic if it has non-negative entries and its rows sum to 1. The proof of the following lemma is easy but is very tedious to write carefully. Only a sketch is given. Similar refinement arguments appear in many places.

LEMMA 1.2. *Let (X,d) be a metric space, Y a reflexive normed space, Z a subset of X and $f : X \to Y$ Lipschitz. There is an extension $\tilde{f} : X \to Y$ of f with $\|\tilde{f}\|_{\text{lip}} \le K$ if and only if, for all $m,n \in \mathbb{N}$, $n \times n$ symmetric, stochastic A, $n \times m$ stochastic B, $\alpha \in (0,1)$ and sequences $(z_r)_1^m$ in Z, $(x_i)_1^n$ in X, there are points $(y_i)_1^n$ in Y satisfying*

$$\alpha \sum a_{ij}\|y_i - y_j\|^2 + 2(1-\alpha)\sum b_{ir}\big\|y_i - f(z_r)\big\|^2$$
$$\le K^2 \left(\alpha \sum a_{ij} d(x_i,x_j)^2 + 2(1-\alpha)\sum b_{ir} d(x_i,z_r)^2\right) . \quad (1.2)$$

Sketch of proof: Given a symmetric matrix H, as in the hypothesis of Lemma 1.1, reorder the indices so that the first m (say) correspond to points in Z and the remainder to points in $X \backslash Z$. Let H be

$$m\{ \begin{pmatrix} T & B^T \\ B & A \end{pmatrix}$$

for appropriate matrices A, B and T. By approximating H and scaling, assume that there is some $\alpha \in (0,1)$ so that each row of B has a sum which is an integral multiple of $(1-\alpha)$, say p_i times $(1-\alpha)$ for the i^{th} row, and then, by adding appropriate numbers to the diagonal of A, that the i^{th} row of A adds up to $p_i\alpha$ (for each i). Note that the diagonal of A is irrelevant for the inequality (1.1). The aim is to construct new matrices in place of A and B. First divide B by $(1-\alpha)$, and then, for each i, replace the i^{th} row of $\frac{1}{1-\alpha}B$ with p_i copies of $\frac{1}{p_i}$ times this row, to obtain B' (say). Divide A by α: replace the i^{th} row of $\frac{1}{\alpha}A$ with p_i copies of $\frac{1}{p_i}$ times this row to obtain A'. Now, replace the i^{th} column of A' with p_i copies of $\frac{1}{p_i}$ times this column. For example, if $\alpha = \frac{1}{2}$,

$$A = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad \begin{matrix} p_1 = 1 \\ p_2 = 2 \end{matrix}$$

then
$$\frac{1}{\alpha A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$
$$A' = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad A'' = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} .$$

The new matrices B' and A'' are stochastic and A'' is symmetric. Apply the hypothesis (1.2) with these matrices, the above value of α and, for each i, p_i copies of x_i. This yields a collection of y's: the y's corresponding to different indices in the i^{th} "block" may not be the same; but they can be chosen to be, just by selecting the average, in place of all of them. The resulting y_i's will then satisfy the hypothesis of Lemma 1.1 as the images $\hat{f}(x_i)$. □

Lemma 1.2 motivates the definition of Markov type. Suppose, to begin with, that the codomain, Y, is Hilbert space. Given A, B, α and $\left(f(x_r)\right)_1^m$, the aim is to choose $(y_i)_1^n$ in Y so that the quantity

$$\alpha \sum a_{ij} \|y_i - y_j\|^2 + 2(1-\alpha) \sum b_{ir} \left\|y_i - f(x_r)\right\|^2 \tag{1.3}$$

is small. But in Hilbert space, one can easily determine the minimum of the quadratic form (1.3) over all choices of $(y_i)_1^n$: namely

$$(1-\alpha) \sum_{rs} (B^T C B)_{rs} \left\|f(z_r) - f(z_s)\right\|^2 \tag{1.4}$$

where

$$\begin{aligned} C &= (1-\alpha)(I - \alpha A)^{-1} \\ &= (1-\alpha) \sum_{k=0}^{\infty} \alpha^k A^k . \end{aligned} \tag{1.5}$$

This minimum is attained when

$$y_i = \sum_r (CB)_{ir} f(x_r) \quad \text{for each} \quad i.$$

Note that, from (1.5), C is a symmetric, stochastic matrix: so each y_i is a convex combination of the $f(z_r)$'s. It is natural to estimate (1.4) by

$$(1-\alpha) \|f\|_{\text{lip}}^2 \sum_{rs} (B^T C B)_{rs} d(z_r, z_s)^2.$$

Hence, Lemma 1.2 guarantees extensions if the domain X has the property that for some K and for all $A, B, \alpha, (z_r)_1^n$ and $(x_i)_1^n$,

$$(1-\alpha)\sum(B^TCB)_{rs}d(z_r,z_s)^2 \\ \le K^2\left(\alpha\sum a_{ij}d(x_i,x_j)^2+2(1-\alpha)\sum b_{ir}d(x_i,z_r)^2\right).$$

This property could be taken to be the definition of Markov type 2. However, it is unnecessarily complicated. The proof of Theorem 1.7 below, will show that it is enough to consider only the cases $m = n$, $z_i = x_i$ $(1 \le i \le n)$ and B the identity matrix.

DEFINITION 1.3. *A metric space (X, d) will be said to have Markov type 2 if there is a constant K so that if $n \in \mathbb{N}$, A is an $n \times n$ symmetric, stochastic matrix, $\alpha \in (0,1)$ and $(x_i)_1^n$ is a sequence in X,*

$$(1-\alpha)\sum_{ij}c_{ij}d(x_i,x_j)^2 \le K^2\alpha\sum_{ij}a_{ij}d(x_i,x_j)^2 \tag{1.6}$$

where $C = (1-\alpha)(I-\alpha A)^{-1}$. The least K for which this holds will be denoted $M_2(X)$. □

The matrix C plays the role of a Green's function for the difference equations defined by A: but as will be seen in Proposition 1.6, it is more useful to think of C as a transition matrix for a Markov chain stopped at an independent, geometrically distributed time. For $1 \le p < 2$, it makes sense to define Markov type p by replacing the exponent 2 in (1.6) by p. Some simple observations concerning Markov type are collected here.

PROPOSITION 1.4. i) *Every metric space has Markov type 1.*
ii) *Hilbert spaces have Markov type 2.*
iii) *If $1 < p < 2$, L_p has Markov type p.*

Proof: i) Let $(x_i)_1^n$ be points in a metric space (X, d). With A, α and C as above, $C = (1-\alpha)I + \alpha CA$ so

$$\begin{aligned}\sum c_{ij}d(x_i,x_j) &= \alpha\sum_{ikj}c_{ik}a_{kj}d(x_i,x_j)\\ &\le \alpha\sum_{ikj}c_{ik}a_{kj}\big(d(x_i,x_k)+d(x_k,x_j)\big)\\ &= \alpha\sum_{ik}c_{ik}d(x_i,x_k)+\alpha\sum_{kj}a_{kj}d(x_k,x_j)\ .\end{aligned}$$

Subtracting $\alpha \sum c_{ij} d(x_i, x_j)$ from both sides gives

$$(1-\alpha) \sum c_{ij} d(x_i, x_j) \leq \alpha \sum a_{ij} d(x_i, x_j) .$$

ii) This can be checked by expanding the inner product and using the fact that if λ is an eigenvalue of A, $\lambda \leq 1$ and $\frac{1-\alpha}{1-\alpha\lambda}$ is an eigenvalue of C with the same eigenspace.

The result also follows from the assertions made earlier. For, if $(x_i)_1^n$ is a sequence in Hilbert space, consider the quadratic form

$$\alpha \sum a_{ij} \|v_i - v_j\|^2 + 2(1-\alpha) \sum \|v_i - x_i\|^2 . \tag{1.7}$$

If $v_i = x_i$ $(1 \leq i \leq n)$ then (1.7) takes the value

$$\alpha \sum a_{ij} \|x_i - x_j\|^2 .$$

But, as was stated earlier, the minimum of (1.7) is

$$(1-\alpha) \sum c_{ij} \|x_i - x_j\|^2 .$$

iii) For $1 \leq p \leq 2$, L_p can be equipped with a new metric d by, $d(x,y) = \|x-y\|^{\frac{p}{2}}$, and the resulting metric space embeds isometrically in Hilbert space, by results of Schoenberg [S] (see e.g. [WW]). So iii) is immediate from ii). □

Now consider codomains Y other than Hilbert space. The hope is that Lipschitz maps defined on subsets of spaces with Markov type 2, into Y, should be extensible. So it is natural to ask that there be a K so that for any sequence $(x_r)_1^m$ in Y and all A, B and as above, there exist points $(y_i)_1^n$ in Y with,

$$\begin{aligned} \alpha \sum a_{ij} \|y_i - y_j\|^2 &+ 2(1-\alpha) \sum b_i \|y_i - x_r\|^2 \\ &\leq K^2 (1-\alpha) \sum (B^T C B)_{rs} \|x_r - x_s\|^2 i . \end{aligned}$$

Just as for type, it is enough to consider only the case $B =$ identity. Moreover, at least for normed spaces, it seems to be appropriate to insist that the y_i's can be chosen as in Hilbert space; i.e.

$$y_i = \sum c_{ij} x_j \qquad (1 \leq i \leq n) .$$

Definition 1.5. *A normed space X will be said to have Markov cotype 2 if there is a K so that if $n \in \mathbb{N}$, A is an $n \times n$ symmetric, stochastic matrix, $\alpha \in (0,1)$ and $(x_i)_1^n$ is a sequence in X,*

$$\alpha \sum a_{ij} \left\| \sum c_{ir} x_r - \sum c_{js} x_s \right\|^2 \le K^2 (1-\alpha) \sum c_{ij} \|x_i - x_j\|^2$$

where $C = (1-\alpha)(I - \alpha A)^{-1}$. The least K for which this holds will be denoted $N_2(X)$. □

Again, it is easy to check that Hilbert spaces have Markov cotype 2 with constant 1.

The next theorem is intended to explain the nomenclature; at least the phrase Markov type. If X is a metric space, a Markovian sequence $M_0, M_1, \ldots$ of X-valued random variables will be called a simple, symmetric Markov chain if there is an $n \in \mathbb{N}$, a sequence $(x_i)_1^n$ in X and an $n \times n$ symmetric stochastic matrix $A = (a_{ij})$ so that

$$P(M_0 = x_i) = \frac{1}{n} \qquad (1 \le i \le n)$$
$$P(M_{k+1} = x_j \mid M_k = x_i) = a_{ij} \quad (1 \le i, j \le n) \quad k = 0, 1, 2, \ldots,$$

Note: strictly speaking, it may not be possible to find such a Markov chain if the x_i's are not all different. What one really wants is an X-valued function of a Markov chain with state space $1, \ldots n$.

Theorem 1.6. *Let (X, d) be a metric space. The following are equivalent:*

i) *X has Markov type 2*

ii) *There is a constant K so that if $M_0, M_1, \ldots,$ is a simple, symmetric Markov chain in X and $m \in \mathbb{N}$,*

$$\begin{aligned} Ed(M_m, M_0)^2 &\le K^2 \sum_{k=0}^{m-1} Ed(M_{k+1}, M_k)^2 \\ &= K^2 m Ed(M_1, M_0)^2 \ . \end{aligned}$$

iii) *There is a constant K so that if $M_0, M_1, \ldots,$ is a simple, symmetric Markov chain in X and T is a random time independent of $(M_k : k \ge 0)$ then*

$$Ed(M_T, M_0)^2 \le K^2 ET.Ed(M_1, M_0)^2 \ .$$

Proof: ii) $\Rightarrow$ iii). For $m = 0, 1, 2, \ldots$ let $p_m = P(T = m)$. Then by independence

$$\begin{aligned} Ed(M_T, M_0)^2 &= \sum_{m=0}^{\infty} p_m Ed(M_m, M_0)^2 \\ &\leq K^2 \sum_{m=0}^{\infty} p_m m d(M_1, M_0)^2 \\ &= K^2 ET.Ed(M_1, M_0)^2 \ . \end{aligned}$$

iii) $\Rightarrow$ i) Take T to be distributed geometrically with $P(T = m) = (1-\alpha)\alpha^m$, $m = 0, 1, 2, \ldots$ and note that

$$ET = \frac{\alpha}{1-\alpha}.$$

i) $\Rightarrow$ ii) Suppose A, $(x_i)_1^n$ and $m \in \mathbb{N}$ are given. Let $\alpha = 1 - \frac{1}{m+1}$ and $C = (1-\alpha)(I - \alpha A)^{-1}$ and observe that

$$\begin{aligned} &\frac{1}{(m+1)^2} \sum_{k=0}^{m} \left(1 - \frac{1}{m+1}\right)^k \sum_{ij} (A^k)_{ij} d(x_i, x_j)^2 \\ &\leq (1-\alpha) \sum c_{ij} d(x_i, x_j)^2 \\ &\leq K^2 \alpha \sum a_{ij} d(x_i, x_j)^2 \\ &= \frac{K^2 m}{m+1} \sum a_{ij} d(x_i, x_j)^2 \ . \end{aligned}$$

Hence

$$\sum_0^m \sum_{ij} (A^k)_{ij} d(x_i, x_j)^2 \leq K^2 em(m+1) \sum_{ij} a_{ij} d(x_i, x_j)^2 \ ;$$

i.e.

$$\sum_{k=0}^{m} Ed(M_k, M_0)^2 \leq K^2 em(m+1) Ed(M_1, M_0)^2 \ .$$

Now for each k,

$$\begin{aligned} Ed(M_m, M_0)^2 &\leq E\big(d(M_m, M_k) + d(M_k, M_0)\big)^2 \\ &\leq 2\big(Ed(M_m, M_k)^2 + Ed(M_k, M_0)^2\big) \\ &= 2\big(Ed(M_{m-k}, M_0)^2 + Ed(M_k, M_0)^2\big) \ . \end{aligned}$$

Summing over k,

$$(m+1)Ed(M_m,M_0)^2 \le 4\sum_{k=0}^{m} Ed(M_k,M_0)^2$$

and so

$$Ed(M_m,M_0)^2 \le 4K^2emEd(M_1,M_0)^2 \ . \qquad \square$$

In view of the preceding theorem, Proposition 1.4 ii) is strengthened by the observation that if $M_0, M_1, \dots$ is a symmetric Markov chain in a Hilbert space

$$E\|M_m - M_0\|^2 \le mE\|M_1 - M_0\|^2, \quad m = 1,2,\dots \ .$$

Observe that in any metric space,

$$\begin{aligned} Ed(M_m,M_0)^2 &\le E\left(\sum_{k=0}^{m-1} d(M_{k+1},M_k)\right)^2 \\ &\le m\sum_{k=0}^{m-1} Ed(M_{k+1},M_k)^2 \\ &= m^2 Ed(M_1,M_0)^2 \ . \end{aligned} \tag{1.8}$$

The principal result of this chapter is the following.

THEOREM 1.7. *Let (X,d) be a metric space with Markov type 2, Y a reflexive normed space with Markov cotype 2, Z a subset of X and $f : Z \to Y$ Lipschitz. Then, there is an extension $\tilde{f} : X \to Y$ of f with*

$$\|\tilde{f}\|_{\mathrm{lip}} \le 3M_2(X)N_2(Y)\|f\|_{\mathrm{lip}} \ .$$

Remark: It seems likely that normed spaces with Markov cotype 2 are automatically reflexive (and hence, superreflexive). In Chapter 2 it is shown that they at least have non-trivial Rademacher type.

Proof: Let $(z_i)_1^m$ be a sequence in Z, $(x_i)_1^n$ a sequence in X, A an $n \times n$ symmetric, stochastic matrix, B an $n \times m$ stochastic matrix and $\alpha \in (0,1)$. The aim is to show that there are points $(y_i)_1^n$ in Y with

$$\begin{aligned} &\alpha \sum a_{ij}\|y_i - y_j\|^2 + 2(1-\alpha)\sum b_{ir}\|y_i - f(z_r)\|^2 \\ &\quad \le K^2\left(\alpha\sum a_{ij}d(x_i,x_j)^2 + 2(1-\alpha)\sum b_{ir}d(x_i,z_r)^2\right) \ , \end{aligned}$$

where $K = 3M_2(X).N_2(Y)\|f\|_{\text{lip}}$.

Let

$$y_i = \sum_r (CB)_{ir} f(z_r) \ , \qquad 1 \le i \le n \ .$$

Then

$$\begin{aligned} &\alpha \sum a_{ij} \|y_i - y_j\|^2 + 2(1-\alpha) \sum b_{ir} \|y_i - f(z_r)\|^2 \\ &\quad \le N_2(Y)^2 (1-\alpha) \sum c_{ij} \Big\| \sum_r b_{ir} f(z_r) - \sum_s b_{js} f(z_s) \Big\|^2 \\ &\quad\quad + 2(1-\alpha) \sum b_{ir} \|y_i - f(z_r)\|^2 \end{aligned}$$

by the definition of the Markov cotype 2 property, and the latter is at most

$$(N_2(Y)^2 + 2)(1-\alpha) \sum (B^T CB)_{rs} \|f(z_r) - f(z_s)\|^2$$

by convexity. The hypothesis on f ensures that this is at most

$$3N_2(Y)^2 \|f\|_{\text{lip}}^2 (1-\alpha) \sum (B^T CB)_{rs} d(z_r, z_s)^2 \ .$$

Now,

$$\begin{aligned} &(1-\alpha) \sum (B^T CB)_{rs} d(z_r, z_s)^2 \\ &\quad \le (1-\alpha) \sum_{ijrs} c_{ij} b_{ir} b_{js} \big(d(z_r, x_i) + d(x_i, x_j) + d(x_j, z_s) \big)^2 \\ &\quad \le 3(1-\alpha) \Big[\sum c_{ij} d(x_i, x_j)^2 + 2 \sum b_{ir} d(x_i, z_r)^2 \Big] \end{aligned}$$

since B and C are stochastic. The first term is at most

$$3M_2(X)^2 \alpha \sum a_{ij} d(x_i, x_j)^2$$

and so the whole is at most

$$3M_2(X)^2 \Big(\alpha \sum a_{ij} d(x_i, x_j)^2 + 2(1-\alpha) \sum b_{ir} d(x_i, z_r)^2 \Big) \ . \qquad \square$$

2. Connections With (Rademacher) Type and Cotype

The principal result in this chapter shows that Markov type implies (Rademacher) type and that Markov cotype is stronger than (Rademacher) cotype. In fact, if X is a normed space and either X has Markov type 2 or X^* has Markov cotype 2, then X has type 2. This shows that ℓ_1 does not have Markov cotype 2 (since c_0 does not have type 2) even though L_1-spaces do have cotype 2. It is almost certain that there are spaces of type 2 which do not have Markov type 2: very probably, any space with Markov type 2 can be equipped with an equivalent 2-smooth norm (see Chapter 3).

The main tool used in this chapter is the following duality lemma.

LEMMA 2.1. *Let X be a normed space and let K be either the M_2 constant of X or the N_2 constant of X^*. Then if A is a $n \times n$ symmetric, stochastic matrix, $\alpha \in (0,1)$, $(x_i)_1^n$ is a sequence in X and $C = (1-\alpha)(I-\alpha A)^{-1}$,*

$$(1-\alpha)\sum_i \left\| x_i - \sum_j c_{ij}x_j \right\|^2 \leq K^2\alpha \sum_{ij} a_{ij}\|x_i - x_j\|^2. \tag{2.1}$$

Proof: If $K = M_2(X)$, (2.1) is obvious because C is stochastic so that by convexity

$$\left\| x_i - \sum_j c_{ij}x_j \right\|^2 \leq \sum_j c_{ij}\|x_i - x_j\|^2 \quad \text{for each} \quad i\ .$$

Now suppose that $K = N_2(X^*)$. For $1 \leq i \leq n$, choose a functional $\phi_i \in X^*$ satisfying

$$\|\phi_i\|^2 = \left\| x_i - \sum_j c_{ij}x_j \right\|^2 = \phi_i\Big(x_i - \sum_j c_{ij}x_j\Big)\ . \tag{2.2}$$

Then,

$$\begin{aligned}(1-\alpha)\sum_i \left\| x_i - \sum_j c_{ij}x_j \right\|^2 &= (1-\alpha)\sum_i \phi_i\Big(x_i - \sum_j c_{ij}x_j\Big) \\ &= (1-\alpha)\sum_{ij}(I-C)_{ij}\phi_i(x_j)\ .\end{aligned}$$

It is easy to check that $(1-\alpha)(I-C) = \alpha C(I-A)$ so that the last expression is

$$\alpha \sum_{ij} c_{ik}(I-A)_{kj}\phi_i(x_j) = \alpha \sum_{kj}(I-A)_{kj}\psi_k(x_j)$$

where for each k, $\psi_k = \sum_i c_{ki}\phi_i$. This last expression is

$$\alpha \sum_{jk} a_{kj}\psi_k(x_k - x_j) = \frac{1}{2}\alpha \sum_{kj} a_{kj}(\psi_k - \psi_j)(x_k - x_j)$$
$$\leq \frac{1}{2}\left(\alpha \sum_{kj} a_{kj}\|\psi_k - \psi_j\|^2\right)^{\frac{1}{2}} \left(\alpha \sum a_{kj}\|x_k - x_j\|^2\right)^{\frac{1}{2}}.$$

The hypothesis that X^* has Markov cotype 2 implies exactly that

$$\alpha \sum a_{kj}\|\psi_k - \psi_j\|^2 \leq K^2(1-\alpha)\sum c_{kj}\|\phi_k - \phi_j\|^2$$

and the latter is at most

$$4K^2(1-\alpha)\sum_k \|\phi_k\|^2$$

by the triangle inequality. Hence

$$(1-\alpha)\sum_i \left\|x_i - \sum_j c_{ij}x_j\right\|^2$$
$$\leq K\left((1-\alpha)\sum_k \|\phi_k\|^2\right)^{\frac{1}{2}} \left(\alpha \sum a_{kj}\|x_k - x_j\|^2\right)^{\frac{1}{2}}$$
$$= K\left((1-\alpha)\sum_k \left\|x_k - \sum_j c_{kj}x_j\right\|^2\right)^{\frac{1}{2}} \left(\alpha \sum a_{kj}\|x_k - x_j\|^2\right)^{\frac{1}{2}}.$$

by 2.2. This reduces to 2.1. □

The implications from the Markov properties to the Rademacher properties are now straightforward.

THEOREM 2.2. *Let X be a normed space and K be either $M_2(X)$ or $N_2(X^*)$. Then the (Rademacher) type 2 constant of X is at most $2\sqrt{2}\,K$.*

Proof: Let $(x_i)_1^k$ be a sequence in X and for each $\varepsilon \in \{-1,1\}^k$ let $x_\varepsilon = \sum_i \varepsilon_i x_i$. The problem is to show that

$$2^{-k} \sum_{\varepsilon} \|x_\varepsilon\|^2 \leq 8K^2 \sum_i \|x_i\|^2 . \tag{2.3}$$

Let A be the transition matrix of the symmetric random walk on the k-cube. Thus A is a $2^k \times 2^k$ matrix whose entries are indexed by pairs (ε, δ) of elements of $\{-1,1\}^k$ and

$$a_{\varepsilon\delta} = \begin{cases} k^{-1} & \text{if } \varepsilon \text{ and } \delta \text{ differ in exactly one place} \\ 0 & \text{otherwise .} \end{cases}$$

If $\alpha_{\varepsilon\delta} \neq 0$ and ε and δ differ in the i^{th} place then $\|x_\varepsilon - x_\delta\| = 2\|x_i\|$. Hence

$$\sum a_{\varepsilon\delta}\|x_\varepsilon - x_\delta\|^2 = 2^k \cdot \frac{4}{k} \sum_i \|x_i\|^2 .$$

If $\alpha \in (0,1)$ and $C = (1-\alpha)(I - \alpha A)^{-1}$ then, by symmetry, $\sum c_{\varepsilon\delta} x_\delta$ is a multiple of x_ε for each ε. It is not hard to check that if $\alpha = \frac{k}{k+2}$

$$\sum c_{\varepsilon\delta} x_\delta = \frac{1}{2} x_\delta \quad \text{for each} \quad \varepsilon .$$

So for this value of α,

$$\begin{aligned}(1-\alpha) \sum_{\varepsilon} \left\| x_\varepsilon - \sum_{\delta} c_{\varepsilon\delta} x_\delta \right\|^2 &= \frac{2}{k+2} \cdot \frac{1}{4} \sum_{\varepsilon} \|x_\varepsilon\|^2 \\ &= \frac{1}{2(k+2)} \sum_{\varepsilon} \|x_\varepsilon\|^2 .\end{aligned}$$

By Lemma 2.1,

$$\begin{aligned}\frac{1}{2(k+2)} \sum_{\varepsilon} \|x_\varepsilon\|^2 &\leq K^2 \alpha \frac{2^k}{k} \cdot 4 \sum_i \|x_i\|^2 \\ &= \frac{K^2 2^k \cdot 4}{k+2} \sum_i \|x_i\|^2\end{aligned}$$

giving (2.3). □

It is well known (and easily checked) that the type 2 constant of ℓ_∞^n is at least

$$\sqrt{1 + \lfloor \log_2 n \rfloor} > \sqrt{\log n} .$$

Hence, $N_2(\ell_1^n)$ is at least $\sqrt{\frac{\log n}{8}}$.

3. Uniform Convexity and Smoothness

This chapter contains a brief account of the theory of uniform convexity and uniform smoothness as it will be used in the sequel. In the standard texts, e.g. [LT], it is usual to define the modulus of convexity $\delta : [0,1] \to [0,1]$ of a space X by

$$\delta(\varepsilon) = \inf \left\{ 1 - \frac{1}{2}\|x+y\| : x, y \in X \ , \ \|x\| = \|y\| = 1 \text{ and } \|x-y\| = \varepsilon \right\} .$$

The space X is said to have modulus of convexity of power type q (for some $q \in [2,\infty)$) if there is a constant K for which

$$\delta(\varepsilon) \geq \frac{\varepsilon^q}{K} , \qquad 0 \leq \varepsilon \leq 2 .$$

For $q = 2$ there is an equivalent definition which is more convenient for many purposes.

DEFINITION. *A normed space X is said to have modulus of convexity of power type 2, or to be 2-uniformly convex, if there is a constant K so that for all $x, y \in X$*

$$2\|x\|^2 + \frac{2}{K^2}\|y\|^2 \leq \|x+y\|^2 + \|x-y\|^2 . \tag{3.1}$$

The least K for which this holds will be called the 2-uniform convexity constant of X.

Clearly, Hilbert space is 2-uniformly convex with constant 1 and no space can have a constant smaller than 1. If $1 < p \leq 2$, L_p is 2-uniformly convex. The constant can be taken to be $1/\sqrt{p-1}$. G. Pisier showed me an argument by which this may be deduced from hypercontractivity results of Beckner [Be].

Inequality (3.1) can be extended to cover X-valued random vectors. A generalisation of this fact was proved by Pisier [P1], using results from [BG] and [F].

LEMMA 3.1. *Let X be a 2-uniformly convex space with constant K, and U a random vector in $L_2(X)$. Then*

$$\|EU\|^2 + \frac{1}{K^2}E\|U - EU\|^2 \leq E\|U\|^2 . \tag{3.2}$$

Proof: By the triangle inequality $E\|U\|^2 \geq \|EU\|^2$. So there is some non-negative θ (possibly 0) which is maximal with the property that for any random vector V in $L_2(X)$,

$$\theta E\|V - EV\|^2 \leq E\|V\|^2 - \|EV\|^2 .$$

The problem is to show that $\theta \geq K^{-2}$. For a given $\phi > \theta$, choose V non-constant with

$$\phi E\|V - EV\|^2 \geq E\|V\|^2 - \|EV\|^2 .$$

Now, pointwise on the probability space

$$2\left\|\frac{1}{2}V + \frac{1}{2}EV\right\|^2 + \frac{2}{K^2}\left\|\frac{1}{2}V - \frac{1}{2}EV\right\|^2 \leq \|V\|^2 + \|EV\|^2 .$$

Hence,

$$\begin{aligned}\phi E\|V - EV\|^2 &\geq E\|V\|^2 - \|EV\|^2\\ &\geq 2\left(E\left\|\frac{1}{2}V + \frac{1}{2}EV\right\|^2 - \|EV\|^2\right) + \frac{2}{K^2}E\left\|\frac{1}{2}V - \frac{1}{2}EV\right\|^2\\ &\geq 2\theta E\left\|\frac{1}{2}V - \frac{1}{2}EV\right\|^2 + \frac{2}{K^2}E\left\|\frac{1}{2}V - \frac{1}{2}EV\right\|^2\end{aligned}$$

by definition of θ and the fact that

$$\frac{1}{2}V + \frac{1}{2}EV - E\left(\frac{1}{2}V + \frac{1}{2}EV\right) = \frac{1}{2}V - \frac{1}{2}EV .$$

So $\phi \geq \frac{\theta}{2} + \frac{1}{2K^2}$ and, taking limits as $\phi \to \theta$, one obtains $\theta \geq \frac{1}{K^2}$. □

Pisier used Lemma 3.1 to analyse the behaviour of martingales in uniformly convex spaces. Although the result will not be used in the sequel, it is stated for completeness.

PROPOSITION 3.2. *Let X be a 2-uniformly convex space with constant K and $M_0, M_1, \ldots, M_m$ a martingale in $L_2(X)$. Then*

$$\sum_{k=0}^{m-1} E\|M_{k+1} - M_k\|^2 \leq K^2 E\|M_m - M_0\|^2 .$$ □

It follows immediately from Proposition 3.2 that 2-uniformly convex spaces have cotype 2 (in the ordinary sense). In the next chapter, Lemma 3.1 will be used to show that such spaces also have Markov cotype 2.

The property of normed spaces, dual to that of 2-uniform convexity is 2-uniform smoothness.

DEFINITION. *A normed space X is said to be 2-uniformly smooth (or 2-smooth) if there is a constant K so that for all $x, y \in X$,*

$$\|x+y\|^2 + \|x-y\|^2 \leq 2\|x\|^2 + 2K^2\|y\|^2 \ .$$

The least K for which this holds will be called the 2-smoothness constant of K.

The 2-smoothness constant of a space X is equal to the 2-uniform convexity constant of its dual X^* (and vice-versa), [L2]. Hence, for $2 \leq q < \infty$, L_q is 2-smooth with constant $\sqrt{q-1}$.

There are analogues of Lemma 3.1 and Proposition 3.2 valid in 2-smooth spaces [P1].

PROPOSITION 3.3. *Let X be a 2-smooth space with constant K.*

i) *If U is a random vector in $L_2(X)$*

$$E\|U\|^2 \leq \|EU\|^2 + K^2 E\|U - EU\|^2 \ .$$

ii) *If $(M_k)_0^m$ is a martingale in $L_2(X)$*

$$E\|M_m - M_0\|^2 \leq K^2 \sum_{k=0}^{m-1} E\|M_{k+1} - M_k\|^2 \ . \tag{3.3}$$

□

The proofs are the same as for the 2-uniformly convex results. Note that (3.3) is exactly the condition for Markov type 2 stated in Theorem 1.6 (except that it refers to martingales).

4. Markov Cotype 2

The main result in this chapter states that L_p has Markov cotype 2 for $1 < p \leq 2$. This is deduced from Lemma 3.1 for 2-uniformly convex spaces. The lemma is applicable because the Markov cotype property involves the Green's matrix, C, inside the norm. The situation for Markov type 2 is rather different: martingale methods do not seem to work directly.

L_1 is not 2-uniformly convex and, as was shown in Chapter 2, does not have Markov cotype 2. The estimate given for L_p, $1 < p \leq 2$ implies that the lower bound $N_2(\ell_1^n) \geq \sqrt{\log n}$ gives the correct order of growth of $N_2(\ell_1^n)$. This leaves open the question of whether Lipschitz maps from subsets of L_2 into L_1, extend to the whole of L_2. There is a dearth of genuinely non-linear examples which might shed some light on this matter.

THEOREM 4.1. *Let X be a normed space with 2-uniform convexity constant K. Then*

$$N_2(X) \le 2K .$$

Proof: Let A, α and C be as in Definition 1.5 and $(x_i)_1^n$ in X. The problem is to show that

$$\alpha \sum a_{ij} \|y_i - y_j\|^2 \le 4K^2(1-\alpha) \sum c_{ij} \|x_i - x_j\|^2 ,$$

where, for each i,

$$y_i = \sum c_{ir} x_r .$$

Observe that

$$C = (1-\alpha)I + \alpha AC . \tag{4.1}$$

Fix i and j and define a random vector $U \in L_2(X)$ with

$$P(U = x_i - y_j) = 1 - \alpha$$

and

$$P(U = y_r - y_j) = \alpha a_{ir} , \qquad 1 \le r \le n .$$

Then

$$\begin{aligned} EU &= (1-\alpha)(x_i - y_j) + \alpha \sum_r a_{ir}(y_r - y_j) \\ &= (1-\alpha)x_i + \alpha \sum_r a_{ir} y_r - y_j \\ &= \sum_r \left[(1-\alpha)I + \alpha AC\right]_{ir} x_r - y_j \\ &= \sum c_{ir} x_r - y_j = y_i - y_j . \end{aligned}$$

So, by Lemma 3.1,

$$\begin{aligned} \|y_i - y_j\|^2 + \frac{1}{K^2}(1-\alpha)\|x_i - y_i\|^2 + \frac{1}{K^2}\alpha \sum_r a_{ir} \|y_r - y_i\|^2 \\ \le (1-\alpha)\|x_i - y_j\|^2 + \alpha \sum_r a_{ir} \|y_r - y_j\|^2 . \end{aligned}$$

Hence (ignoring the second term on the left),

$$\begin{aligned} \|y_i - y_j\|^2 + \frac{1}{K^2}\alpha \sum_r a_{ir} \|y_i - y_r\|^2 \\ \le (1-\alpha)\|x_i - y_j\|^2 + \alpha \sum_r a_{ir} \|y_r - y_j\|^2 . \end{aligned}$$

Multiply by c_{ij} and sum over i and j using the fact that C is stochastic:

$$\sum_{ij} c_{ij}\|y_i - y_j\|^2 + \frac{1}{K^2}\alpha \sum_{ir} a_{ir}\|y_i - y_r\|^2$$
$$\leq (1-\alpha)\sum_{ij} c_{ij}\|x_i - y_j\|^2 + \sum_{rj}(\alpha AC)_{rj}\|y_r - y_j\|^2 .$$

From (4.1), the second term on the right is $\sum_{rj} c_{rj}\|y_r - y_j\|^2$ and so cancels with the first term on the left. Hence

$$\alpha \sum_{ir} a_{ir}\|y_i - y_r\|^2 \leq K^2(1-\alpha)\sum_{ij} c_{ij}\|x_i - y_j\|^2 .$$

To complete the proof, observe that, by convexity

$$\begin{aligned}\sum c_{ij}\|x_i - y_j\|^2 &= \sum c_{ij}\|x_i - \sum_k c_{jk}x_k\|^2 \\ &\leq \sum c_{ij}c_{jk}\|x_i - x_k\|^2 \\ &\leq 2\sum c_{ij}c_{jk}\left(\|x_i - x_j\|^2 + \|x_j - x_k\|^2\right) \\ &= 4\sum c_{ij}\|x_i - x_j\|^2 .\end{aligned}$$ □

COROLLARY 4.2. *For $1 < p \leq 2$, the Markov cotype 2 constant of L_p is at most*

$$\frac{2}{\sqrt{p-1}} .$$ □

COROLLARY 4.3. *For $n \in \mathbf{N}$, the Markov cotype 2 constant of ℓ_1^n is at most a constant multiple of $\sqrt{1+\log n}$.*

Proof: The ℓ_1 and ℓ_p norms are equivalent on $\mathbf{R}^n$ (up to a constant independent of n) provided

$$p < 1 + \frac{1}{1+\log n} .$$ □

The most important consequence of Theorems 4.1 and 1.3 is the following.

THEOREM 4.4. *Let $1 < p \leq 2$, Z be a subset of L_2 and $f : Z \to L_p$ Lipschitz. Then there is an extension $\tilde{f} : L_2 \to L_p$ of f with*

$$\|\tilde{f}\|_{\text{lip}} \leq \frac{6}{\sqrt{p-1}}\|f\|_{\text{lip}} .$$ □

Remark: As $p \to 2$ from below, the constant in Theorem 4.4 does not approach 1, the constant in Kirszbraun's theorem. This situation can be remedied if one chooses not to make the simplification $B = I$ in the definitions of Markov type and cotype. The cost is that the formulae become unintelligible.

5. Markov Type 2

This chapter contains a discussion of the Markov type 2 property for normed spaces. The situation here is not as clear as for the cotype property. It would be natural to conjecture that 2-smooth spaces have Markov type 2; in particular, that this is true for L_p, $2 < p < \infty$. At present I am unable to prove this.

An analogue of the proof of Theorem 4.1 easily yields the following.

PROPOSITION 5.1. *Let A be an $n \times n$ symmetric, stochastic matrix and $(x_i)_1^n$ a sequence in a normed space X with 2-smoothness constant K. For $k = 0, 1, 2, \ldots$ set*

$$x_i^{(k)} = \sum_j (A^k)_{ij} x_j \ .$$

Then for $m \in \mathbb{N}$,

$$\sum_{ij} (A^{2m})_{ij} \|x_i - x_j\|^2 \le 16K^2 \sum_{k=0}^{m-1} \sum_{ij} a_{ij} \|x_i^{(k)} - x_j^{(k)}\|^2 \ . \qquad \square$$

Thus, in a 2-smooth space, Markov type 2 would be implied by an estimate on the behaviour of the images of $(x_i)_1^n$ under the (discrete) Markov semigroup generated by A. It would be enough to find a constant M so that for all A, $(x_i)_1^n$ and $k \in \mathbb{N}$

$$\sum_{ij} a_{ij} \|x_i^{(k)} - x_j^{(k)}\|^2 \le M^2 \sum_{ij} a_{ij} \|x_i - x_j\|^2 \ . \tag{5.1}$$

Equally, it would suffice to obtain a similar upper bound, uniformly in α, on

$$\sum_{ij} a_{ij} \left\| \sum_r c_{ir} x_r - \sum_s c_{js} x_s \right\|^2 , \tag{5.2}$$

with $C = (1-\alpha)(I-\alpha A)^{-1}$. It is easy to see that such estimates hold (with $M = 1$) if X is Hilbert space. In general, such estimates may be related to the UMD property of normed spaces: this matter will be taken up again at the end of Chapter 6.

For many natural matrices, estimates like (5.1) hold for all normed spaces, with $M = 1$. An example is the discrete Laplacian matrix described in Chapter 6. However, not all symmetric, stochastic matrices are so well-behaved. The simplest example seems to be the symmetric random walk on the dodecahedron. For this 20×20 matrix A, there is a sequence $(x_i)_1^{20}$ in ℓ_∞^{30} for which

$$\sum a_{ij} \|x_i^{(1)} - x_j^{(1)}\|^2 = \frac{16}{9} \sum a_{ij} \|x_i - x_j\|^2 .$$

The main result in this chapter is an estimate on the wandering of Markov chains in 2-smooth spaces which is slightly weaker than that needed for Markov type 2 but is much stronger than the trivial (1.8), valid in all normed spaces. It will be convenient for the proof of this theorem, to work with the continuous semigroup of matrices with infinitesimal generator $I - A$.

For a symmetric, stochastic A, let $\{S^t : t \geq 0\}$ be the semigroup of matrices

$$S^t = \exp\big(-t(I-A)\big) , \qquad t \geq 0 .$$

It is easily checked that X has Markov type 2 if and only if, there is a constant K so that for all A, S^t as above and $(x_i)_1^n$ in X,

$$\sum_{ij} (S^t)_{ij} \|x_i - x_j\|^2 \leq K^2 t \sum_{ij} a_{ij} \|x_i - x_j\|^2 .$$

THEOREM 5.2. *Let $K > 1$ and $0 < \theta < 1$. There is a constant $M = M(K,\theta)$ so that if X is a normed space with 2-smoothness constant K, $(x_i)_1^n$ is a sequence in X and A and S^t are as above,*

$$\sum_{ij} (S^t)_{ij} \|x_i - x_j\|^2 \leq (t + M t^{1+\theta}) \sum_{ij} a_{ij} \|x_i - x_j\|^2 . \tag{5.3}$$

Remark: The proof below gives an estimate on M of the form $M \le K^{\frac{2}{\theta}}$ so that (by optimising over θ) one could replace $Mt^{1+\theta}$ by $t\exp\left(\sqrt{8\log K}\sqrt{\log t}\right)$.

The crucial step in the proof of Theorem 5.2 is a lemma which is analogous to Lemma 2.1, but stated for continuous semigroups. While the result could be deduced from Lemma 2.1 and Theorem 4.1, a direct proof is given for simplicity. It will be convenient to abuse notation slightly: for a sequence $x = (x_i)_1^n \in \ell_2^n(X)$ denote by $S^t x$ the sequence

$$\left(\sum_j (S^t)_{ij} x_j\right)_{i=1}^n$$

and by $S^t x_i$ the i^{th} term of this sequence.

LEMMA 5.3. *Let X be a 2-smooth normed space with constant K, $(x_i)_1^n$ in X and A and S^t as above. Then for each $t \ge 0$,*

$$\sum_i \|x_i - S^t x_i\|^2 \le K^2 t \sum_j a_{ij} \|x_i - x_j\|^2 .$$

Proof: Since S^0 is the identity, it suffices to prove that for each $t \ge 0$,

$$\frac{d}{du} \sum_i \|x_i - S^u x_i\|^2 \Big|_{u=t} \le K^2 \sum a_{ij} \|x_i - x_j\|^2 . \tag{5.4}$$

Now, by convexity

$$\sum_i \|x_i - S^{t+2u} x_i\|^2 \le \sum_{ij} S_{ij}^u \|x_i - S^{t+u} x_j\|^2 . \tag{5.5}$$

Fix j and consider a random vector U in $L_2(X)$ with

$$P(U = x_i - S^{t+u} x_j) = S_{ij}^u , \qquad 1 \le i \le n .$$

By Proposition 3.3 i),

$$\begin{aligned} &\sum_i S_{ij}^u \|x_i - S^{t+u} x_j\|^2 \\ &\quad \le \|S^u x_j - S^{t+u} x_j\|^2 + K^2 \sum_i S_{ij}^u \|x_i - S^u x_j\|^2 . \end{aligned}$$

The first term on the right involves only the index j. So, by convexity, the whole is at most

$$\sum_k S^u_{jk}\|x_k - S^t x_k\|^2 + K^2 \sum_{ik} S^u_{ij} S^u_{jk} \|x_i - x_k\|^2 \ .$$

Sum over j (and use 5.5) to get

$$\sum_i \|x_i - S^{t+2u}x_i\|^2 \le \sum_k \|x_k - S^t x_k\|^2 + K^2 \sum_{ik} S^{2u}_{ik}\|x_i - x_k\|^2 \ .$$

Hence

$$\begin{aligned}
&\frac{d}{du}\sum_i \|x_i - S^u x_i\|^2\Big|_{u=t} \\
&= \lim_{u\to 0}\frac{1}{2u}\left(\sum_i \|x_i - S^{t+2u}x_i\|^2 - \sum_k \|x_k - S^t x_k\|^2\right) \\
&\le \lim_{u\to 0}\frac{1}{2u}K^2 \sum_{ik} S^{2u}_{ik}\|x_i - x_k\|^2 \\
&= K^2 \sum a_{ik}\|x_i - x_k\|^2 \quad \text{giving (5.4)} \ .
\end{aligned}$$

□

Proof of Theorem 5.2: By homogeneity, it may be assumed that $\sum a_{ij}\|x_i - x_j\|^2 = 1$. For $t \ge 0$, set

$$F(t) = \sum_{ij} S^t_{ij}\|x_i - x_j\|^2 \ .$$

It is not hard to deduce from the trivial (1.8) that $F(t) \le t + t^2$ for all t. So, if $M > 1$, $F(t) < t + Mt^{1+\theta}$ for $t \le 1$. For such an M, suppose that it is possible to choose $t > 1$ so that $F(t) \ge t + Mt^{1+\theta}$ and choose the least t with this property. It will be shown below that for any $\lambda \in \left(0, \frac{1}{2}\right)$

$$F(t) \le \left(1 + \frac{\lambda}{2}\right) F\big((1-\lambda)t\big) + 3K^2 t + K^2 F(2\lambda t) \ . \tag{5.6}$$

Then, by the minimality of t,

$$\begin{aligned}
t + Mt^{1+\theta} &\le \left(1 + \frac{\lambda}{2}\right)\left((1-\lambda)t + M(1-\lambda)^{1+\theta}t^{1+\theta}\right) \\
&\quad + 3K^2 t + 2K^2\lambda t + 4K^2 M\lambda^{1+\theta}t^{1+\theta} \\
&< t + M\left(1 - \frac{\lambda}{2}\right)t^{1+\theta} + 4K^2 t + 4K^2 M\lambda^{1+\theta}t^{1+\theta} \ .
\end{aligned}$$

This implies that

$$\frac{M\lambda}{2}(1-8K^2\lambda^\theta)t^\theta < 4K^2$$

and (since $t > 1$) that

$$\frac{M\lambda}{2}(1-8K^2\lambda^\theta) < 4K^2 .$$

Choose λ sufficiently small, to get an estimate for M.

To establish (5.6), fix i and k and apply Proposition 3.3 with U a random vector satisfying

$$P(U = x_i - x_j) = S^{\lambda t}_{kj} , \qquad 1 \le j \le n .$$

This gives

$$\sum_j S^{\lambda t}_{kj}\|x_i - x_j\|^2 \le \|x_i - S^{\lambda t}x_k\|^2 + K^2 \sum_j S^{\lambda t}_{kj}\|S^{\lambda t}x_k - x_j\|^2 .$$

Multiply by $S^{(1-\lambda)t}_{ik}$ and sum over i and k:

$$\begin{aligned}\sum S^t_{ij}\|x_i - x_j\|^2 &\le \sum S^{(1-\lambda)t}_{ik}\|x_i - S^{\lambda t}x_k\|^2 + K^2\sum S^{\lambda t}_{kj}\|S^{\lambda t}x_k - x_j\|^2\\ &\le \sum S^{(1-\lambda)t}_{ik}\left(\|x_i - x_k\| + \|x_k - S^{\lambda t}x_k\|\right)^2\\ &\quad + K^2\sum S^{2\lambda t}_{ij}\|x_i - x_j\|^2\end{aligned}$$

by convexity. The second term is $K^2F(2\lambda t)$. To estimate the first term, use the fact that for any a and b,

$$(a+b)^2 \le \left(1+\frac{\lambda}{2}\right)a^2 + \left(1+\frac{2}{\lambda}\right)b^2$$

to get

$$\begin{aligned}&\left(1+\frac{\lambda}{2}\right)\sum S^{(1-\lambda)t}_{ik}\|x_i - x_k\|^2 + \left(1+\frac{2}{\lambda}\right)\sum_{ik} S^{(1-\lambda)t}_{ik}\|x_k - S^{\lambda t}x_k\|^2\\ &= \left(1+\frac{\lambda}{2}\right)F((1-\lambda)t) + \left(1+\frac{2}{\lambda}\right)\sum_k \|x_k - S^{\lambda t}x_k\|^2 .\end{aligned}$$

Finally observe that by Lemma 5.3, the second term is at most

$$\left(1+\frac{2}{\lambda}\right)K^2\lambda t\sum a_{ij}\|x_i - x_j\|^2 \le 3K^2 t. \qquad \square$$

6. Open Problems

The aim of this chapter is to describe some of the lines along which one would like to develop a full non-linear analogue of the linear theory of type and cotype.

The first part of the chapter raises the questions of duality between the Markov type and cotype properties for normed spaces: i.e., under what circumstances is it possible to estimate $M_2(X)$ by $N_2(X^*)$ and vice versa? For Rademacher type and cotype, the situation is well understood thanks to the Maurey-Pisier theorem [MP] and Pisier's K-convexity theorem [P2]. It is always true that $C_2(X) \leq T_2(X^*)$ but a reverse inequality holds (if and) only if X (or equivalently X^*) is K-convex. (A space X is said to be K-convex if the Rademacher projections on $L_2(X)$ are uniformly bounded.) Moreover, X is K-convex if and only if it does not contain subspaces uniformly close to ℓ_1^n for all n. For the Markov properties, the situation seems to be more symmetric. There is an analogue of the Rademacher projection, for each symmetric, stochastic matrix A, and the boundedness of these projections seems to be needed to estimate $N_2(X)$ by $M_2(X^*)$ as well as the other way around. The fact that L_1 does not have Markov cotype 2 may indicate that this symmetry is to be expected.

The second part of the chapter contains a brief discussion of the Markov type 2 problem. For applications of the theory to Lipschitz extensions, this is clearly the most pressing open problem.

The third part of this chapter asks whether Theorem 4.4 can be extended to include domains which are not uniformly convex.

The first lemma describes the combinatorial aspect of the duality theory for Markov type and cotype and opens the way to a discussion of the analogues of the Rademacher projection.

LEMMA 6.1. *Let X be a normed space and assume that there is a constant K so that for each A, and each $(x_i)_1^n$ in X, there is a sequence $(\phi_i)_1^n$ in X^* satisfying*

$$\frac{1}{K^2}\sum_{ij} a_{ij}\|\phi_i-\phi_j\|^2 \leq \sum_{ij} a_{ij}\|x_i-x_j\|^2 = \sum_{ij} a_{ij}(\phi_i-\phi_j)(x_i-x_j)\ . \qquad (6.1)$$

Then,

$$M_2(X) \leq K M_2(X^*) \qquad (6.2)$$

$$N_2(X) \le KM_2(X^*) \ . \tag{6.3}$$

Remark: The assumption on X is that expressions like $\sum a_{ij}\|x_i - x_j\|^2$ can be normed by expressions of the same type. In general, one could only assume the existence of some antisymmetric matrix (ϕ_{ij}) with entries in X^*, satisfying

$$\sum a_{ij}\|\phi_{ij}\|^2 = \sum a_{ij}\|x_i - x_j\|^2 = \sum a_{ij}\phi_{ij}(x_i - x_j) \ .$$

Proof of Lemma 6.1: The argument below proves (6.3); the proof of (6.2) is similar but simpler. Let A, α and $(x_i)_1^n$ be given and $C = (1-\alpha)(I-\alpha A)^{-1}$. For each i, let $y_i = \sum_j c_{ij}x_j$. Choose $(\phi_i)_1^n$ in X^* so that

$$\begin{aligned}\frac{1}{K^2}\alpha\sum a_{ij}\|y_i - y_j\|^2 &\le \alpha\sum a_{ij}\|y_i - y_j\|^2 \\ &= \alpha\sum a_{ij}(\phi_i - \phi_j)(x_i - x_j) \ .\end{aligned} \tag{6.4}$$

As in Lemma 2.1, observe that

$$\alpha(I - A)C = (1 - \alpha)(I - C) \ .$$

Hence

$$\begin{aligned}\alpha\sum a_{ij}\|y_i - y_j\|^2 &= \alpha\sum a_{ij}(\phi_i - \phi_j)(y_i - y_j) \\ &= 2\alpha\sum a_{ij}\phi_i(y_i - y_j) \\ &= 2\alpha\sum (I - A)_{ij}\phi_i(y_j) \\ &= 2\alpha\sum_{ik} \left[(I - A)C\right]_{ik}\phi_i(x_k) \\ &= 2(1 - \alpha)\sum_{ik}(I - C)_{ik}\phi_i(x_k) \\ &= 2(1 - \alpha)\sum_{ik} c_{ik}\phi_i(x_i - x_k) \\ &= (1 - \alpha)\sum c_{ik}(\phi_i - \phi_k)(x_i - x_k) \\ &\le \left[(1 - \alpha)\sum c_{ik}\|\phi_i - \phi_k\|^2\right]^{\frac{1}{2}} \\ &\quad \cdot \left[(1 - \alpha)\sum c_{ik}\|x_i - x_k\|^2\right]^{\frac{1}{2}} \ .\end{aligned}$$

From the definition of Markov type 2, the first factor is at most

$$M_2(X^*)\left[\alpha\sum a_{ik}\|\phi_i-\phi_k\|^2\right]^{\frac{1}{2}}$$

which is at most

$$K.M_2(X^*)\left[\alpha\sum a_{ik}\|y_i-y_k\|^2\right]^{\frac{1}{2}}$$

by (6.4). Hence

$$\left[\alpha\sum a_{ij}\|y_i-y_j\|^2\right]^{\frac{1}{2}}\le K.M_2(X^*)\left[(1-\alpha)\sum c_{ij}\|x_i-x_j\|^2\right]^{\frac{1}{2}}$$

as required. □

The property of X assumed in Lemma 6.1 is characterized by the boundedness of certain projections. For a symmetric, stochastic A, let L_2^A be the space of antisymmetric $(n\times n)$ real-valued matrices (u_{ij}) with norm

$$\|(u_{ij})\|=\left(\frac{1}{2}\sum_{ij}a_{ij}u_{ij}^2\right)^{\frac{1}{2}}.$$

Let $G: L_2^A\to L_2^A$ be the orthogonal projection onto the subspace consisting of matrices of the form (u_i-u_j) for some sequence $(u_i)_1^n$ of reals.

For X a normed space, $L_2^A(X)$ is defined to be the space of X-valued, antisymmetric matrices with norm

$$\|(x_{ij})\|=\left(\frac{1}{2}\sum a_{ij}\|x_{ij}\|^2\right)^{\frac{1}{2}}.$$

This may be regarded as a tensor product $L_2^A\otimes X$ in the usual way and the map $G\otimes I_X$ is a projection on $L_2^A(X)$: (I_X being the identity on X). The image of a matrix (x_{ij}) under $G\otimes I_X$ is the unique matrix of the form (x_i-x_j) with the property that for any sequence $(\phi_i)_1^n$ in X^*,

$$\sum_{ij}a_{ij}(\phi_i-\phi_j)(x_{ij})=\sum_{ij}a_{ij}(\phi_i-\phi_j)(x_i-x_j).$$

The proof of the following lemma is standard.

LEMMA 6.2. *For any normed space X and symmetric, stochastic A,*

$$\|G \otimes I_X\| = \|G \otimes I_{X^*}\|$$

and this number is the least K for which the hypothesis of Lemma 6.1 holds (for the matrix A). □

The projections $G \otimes I_X$ can be regarded as projections onto gradients of potentials. They are well-known to probabilists in connection with Riesz transforms associated to Markov semigroups. (This was pointed out to me by G. Pisier.) Let $\tilde{\ell}_2^n$ be the subspace of ℓ_2^n consisting of sequences $(u_i)_1^n$ satisfying $\sum u_i = 0$. Given an $n \times n$ symmetric, stochastic matrix A, define

$$\begin{aligned} &\text{grad} : \tilde{\ell}_2^n \to L_2^A \quad \text{by} \\ &\text{grad}\ (u_i) = (u_i - u_j)_{ij} \end{aligned}$$

and

$$\begin{aligned} &\text{div} : L_2^A \longrightarrow \tilde{\ell}_2^n \quad \text{by} \\ &\text{div}(u_{ij})_{ij} = \left(- \sum_j a_{ij} u_{ij} \right)_{i=1}^n . \end{aligned}$$

Let $L : \tilde{\ell}_2^n \to \tilde{\ell}_2^n$ be the operator –div.grad : i.e. the matrix $I - A$. The Riesz transform $R : \tilde{\ell}_2^n \to L_2^A$ is defined formally by

$$R = \text{“grad } \frac{1}{\sqrt{L}}\text{”}.$$

To be more precise, for each sequence $u = (u_i)_1^n$, the limit

$$\lim_{\alpha \to 1^-} \left(\sum_r (I - \alpha A)_{ir}^{-\frac{1}{2}} u_r - \sum_s (I - \alpha A)_{js}^{-\frac{1}{2}} u_s \right)$$

exists for each i and j and the matrix of these is an element Ru of L_2^A. The resulting map R is a linear isometry of $\tilde{\ell}_2^n$ into L_2^A. The adjoint $R^* : L_2^A \to \tilde{\ell}_2^n$ is defined formally by

$$R^* = \text{“}\frac{1}{\sqrt{L}}\text{div”} .$$

It is easy to see that $R^* R : \tilde{\ell}_2^n \to \tilde{\ell}_2^n$ is the identity on $\tilde{\ell}_2^n$ and not difficult to check that $RR^* : L_2^A \to L_2^A$ is the projection G defined above.

Although Riesz transforms have been extensively studied, relatively little seems to be known in the very general setting described here. Particular examples, such as the Hilbert transform, are well understood. A discrete version of the Hilbert transform arises in the above setting if the matrix A is taken to be the transition matrix of a symmetric random walk around a cycle: i.e. if there is an n-cycle σ belonging to the group of permutations of $\{1,\ldots,n\}$ with

$$a_{ij} = \begin{cases} \frac{1}{2} & \text{if } \sigma(i)=j \text{ or } \sigma(j)=i \\ 0 & \text{otherwise .} \end{cases}$$

In this case $L = I - A$ is a discrete version of the Laplacian on the circle. The discrete Hilbert transform associated with this A is bounded on $\tilde{\ell}_2^n(X)$ if and only if X has the so-called U.M.D. property. This is proved in [Bu] and [B]. However, for this matrix A, the projection $G \otimes I_X$ is bounded on $L_2^A(X)$ for all normed spaces X.

The projection G can be approximated using the inverses $(I - \alpha A)^{-1}$ as $\alpha \to 1$ from below: the formal statement

$$G = \text{“grad}\frac{1}{L}\text{div”}$$

translates as follows. For every matrix $u = (u_{ij}) \in L_2^A$,

$$Gu = \lim_{\alpha \to 1^-} G_\alpha u$$

where, for each $\alpha \in (0,1)$

$$(G_\alpha u)_{ij} = \frac{\alpha}{1-\alpha}\left(\sum_{rs} c_{ir}a_{rs}u_{rs} - \sum_{pq} c_{jp}a_{pq}u_{pq}\right)$$

with $C = (1-\alpha)(I-\alpha A)^{-1}$ as usual. If X is a Hilbert space then $G_\alpha \otimes I_X$ is a contraction on $L_2^A(X)$ for all α. More interestingly, for any normed space X, an estimate on $\|G_\alpha \otimes I_X\|$ implies a corresponding estimate on the expression (5.2) since if $x_{ij} = x_i - x_j$ it is easy to check that

$$\begin{aligned} &\alpha^2 \sum a_{ij} \left\|\sum c_{ir}a_{rs}x_{rs} - \sum c_{jp}a_{pq}x_{pq}\right\|^2 \\ &= (1-\alpha)^2 \sum a_{ij}\left\|x_i - x_j - \left(\sum c_{ir}x_r - \sum c_{js}x_s\right)\right\|^2 \end{aligned}$$

and the latter is at least

$$(1-\alpha)^2\left(\frac{1}{2}\sum a_{ij}\left\|\sum c_{ir}x_r-\sum c_{js}x_s\right\|^2-\sum a_{ij}\|x_i-x_j\|^2\right) .$$

The principal problem raised by this discussion of duality is whether the projections $G\otimes I$ are bounded (uniformly in A) on $L_2^A(L_p)$ for each fixed $p\in(1,\infty)$. If this is true it implies that L_q has Markov type 2 for $2<q<\infty$. In the light of the linear theory it also makes sense to ask whether the $G\otimes I$ are uniformly bounded on $L_2^A(X)$ whenever X has Markov type 2 (or Markov cotype 2). In an abstract setting one would hope that boundedness of $G\otimes I$ holds for UMD spaces of type 2 and their duals.

Probably the most important open problem raised by this paper is whether L_q has Markov type 2 for $2<q<\infty$. As mentioned at the start of Chapter 5, this would follow from estimates of the form (5.2). My feeling is that 2-smoothness, by itself, is not enough to imply Markov type 2. There is some evidence that spaces with continuously twice differentiable norms might have Markov type 2. (It was brought to my attention by G. Godefroy that it is not known whether every space with C^2-norm is UMD.)

One natural question is whether Theorem 4.4 remains true if the codomain Y is not assumed to be 2-uniformly convex; in particular if $Y=L_1$. It is not known whether all Lipschitz maps from subsets of Hilbert space into normed spaces extend to the whole of Hilbert space. My feeling is that this is not true and I do not believe it even for $Y=L_1$.

It would be nice to determine the correct rate of growth of $M_2(\ell_\infty^n)$ with n. From Chapter 2 it follows that the constant is at least a fixed multiple of $\sqrt{\log n}$. An estimate of the same order from above would imply the result of Johnson and Lindenstrauss mentioned in the introduction because any n-point set embeds isometrically in ℓ_∞^n.

Finally, it is perhaps worth mentioning how Markov cotype may be defined for general metric spaces. In view of Lemma 1.2, one could say that a metric space (Y,d) has Markov cotype 2 if there is a K so that for every A,B,α and $(z_r)_1^m$ in Y, there are points $(y_i)_1^n$ in Y with

$$\begin{aligned}\alpha\sum a_{ij}d(y_i,y_j)^2&+2(1-\alpha)\sum b_{ir}d(y_i,z_r)^2\\&\le K^2(1-\alpha)\sum(B^TCB)_{rs}d(z_r,z_s)^2 .\end{aligned}$$

However, in line with the proof of Theorem 1.7 it might be more natural to split this statement into two. Say that (Y,d) is approximately convex if there is a K so that for every stochastic B, symmetric, stochastic C and $(z_r)_1^m$ in Y, there are points $(x_i)_1^n$ in Y with

$$\sum c_{ij} d(x_i, x_j)^2 + 2 \sum b_{ir} d(x_i, z_r)^2 \le K^2 \sum (B^T C B)_{rs} d(z_r, z_s)^2 .$$

Then, if (Y, d) is approximately convex, say that it has Markov cotype 2 if there is a K so that for every A, α and $(x_i)_1^n$ in Y, there are points $(y_i)_1^n$ in Y satisfying

$$\alpha \sum_{ij} a_{ij} d(y_i, y_j)^2 + 2(1-\alpha) \sum_i d(y_i, x_i)^2 \le K^2 (1-\alpha) \sum c_{ij} d(x_i, x_j)^2 . \quad (*)$$

Note that every normed space is approximately convex since one may take $x_i = \sum_r b_{ir} z_r$. Thus a normed space has Markov cotype 2 provided only that it satisfies (*). It is not clear whether this property of normed spaces is equivalent to the one used in the text.

References

[Be] W. BECKNER, Inequalities in Fourier Analysis, Ann. of Math. 102 (1975), 159-182.

[B] J. BOURGAIN, Some remarks on Banach spaces in which martingale difference sequences are unconditional, Arkiv für Math. 21 (1983), 163-168.

[BMW] J. BOURGAIN, V.D. MILMAN, H. WOLFSON, On the type of metric spaces, Trans. Amer. Math. Soc. 294 (1986), 295-317.

[Bu] D. BURKHOLDER, A geometrical condition that implies the existence of certain singular integrals of Banach-space-valued functions, Proc. Conf. Harmonic Analysis (in honor of A. Zygmund), Univeristy of Chicago, 1981.

[BG] BUI-MINH-CHI, V.I. GURARII, Some characteristics of normed spaces and their applications to the generalisation of Parseval's inequality for Banach spaces, Sbor. Theor. Funct. 8 (1969), 74-91 (Russian).

[E] P. ENFLO, Uniform homeomorphisms between Banach spaces, Séminaire Maurey-Schwartz 75-76. Exposé no. 18 Ecole Polytechnique, Paris.

[F] T. FIGIEL, On the moduli of convexity and smoothness, Studia Math. 56 (1976) 121-155.

[G] M. GROMOV, Filling Riemannian manifolds, J. Diff. Geom. 18 (1983), 1-147.

[JL] W.B. JOHNSON, J. LINDENSTRAUSS, Extensions of Lipschitz mappings into a Hilbert space, Conference in modern analysis and probability, Contemp. Math. 26 Amer. Math. Soc. (1984).

[JLS] W.B. JOHNSON, J. LINDENSTRAUSS, G. SCHECHTMAN, On Lipschitz embeddings of finite metric spaces into low dimensional normed spaces, Israel Seminar on G.A.F.A., Springer-Verlag, Lecture notes 1267, (1987).

[L1] J. LINDENSTRAUSS, On non-linear projections in Banach spaces, Michigan Math. J. 11 (1964), 263-287.

[L2] J. LINDENSTRAUSS, On the modulus of smoothness and divergent series in Banach spaces, Michigan Math. J. 10 (1963), 241-252.

[LT] J. LINDENSTRAUSS, L. TZAFRIRI, Classical Banach Spaces II, Ergebnisse 97, Springer-Verlag (1979).

[MarP] M.B. MARCUS, G. PISIER, Characterizations of almost surely continuous p-stable random Fourier series and strongly stationary processes, Acta Math. 152 (1984), 245-301.

[M] B. MAUREY, Un théorème de prolongement, C.R. Acad. Sci. Paris 279 (1974), 329-332.

[MP] B. MAUREY, G. PISIER, Séries de variables aléatoires vectorielles indépendentes et propriétés géométriques des espaces de Banach, Studia Math. 58 (1976), 45-90.

[MS] V.D. MILMAN, G. SCHECHTMAN, Asymptotic theory of finite dimensional normed spaces, Lecture Notes 1200, Springer-Verlag (1986).

[P1] G. PISIER, Martingales with values in uniformly convex spaces, Israel J. Math. 20 (1975), 326-350.

[P2] G. PISIER, Holomorphic semi-groups and the geometry of Banach spaces, Ann. Math. 115 (1982), 375-392.

[S] I. SCHOENBERG, Metric spaces and completely monotonic functions, Ann. of Math. 39 (1938), 811-841.

[WW] J.H. WELLS, L.R. WILLIAMS, Embeddings and Extensions in Analysis, Ergebnisse 84, Springer-Verlag (1975).

Keith Ball
Department of Mathematics and Department of Mathematics
T.A.M.U. U.C.L.
College Station, TX 77843 London WC1E
USA England

Submitted: October, 1991