UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

Rasmus Kløvgaard Stavenuiter

Poincaré inequalities, non-linear spectral calculus and superexpanders

Advisor:

Magdalena Musat

Master thesis

Department of Mathematical Sciences

September 29, 2020

Abstract

We study coarse geometric properties of expander graph sequences in terms of Poincaré inequalities, which form an obstruction for coarsely embedding such sequences into Hilbert space. This leads to a generalization of expanders, namely the notion of superexpander sequences, which do not embed coarsely into uniformly convex Banach space. We prove that Schreier coset graphs of residually finite groups with Kazhdan's property (T) are expander sequences. We then prove that being an expander with respect to a Banach space is invariant under sphere equivalence of Banach spaces, as well as a generalization of Matoušek's extrapolation theorem. Moreover, we prove that expander sequences do not coarsely embed into uniformly curved Banach spaces, which are contained in the class of uniformly convex Banach spaces. Finally, we follow [MN14] to prove that uniformly convex Banach spaces admit a non-linear spectral calculus which will lead to a combinatorial construction of superexpanders through the zig-zag product of graphs.

Contents

Introduction

Expander graphs are often informally described as *sparse* yet *highly connected* graphs. They have been widely studied within computer science for their pseudorandom properties. For example, the so-called *expander mixing lemma* implies that the number of edges between two vertex subsets are close to the expected number of edges between them in a random graph on the same vertex set. Moreover, one can prove Chernoff-like bounds for random walks in expander graphs, and due to the edge sparsity, sampling from a random walk on an expander is more efficient than sampling the vertices independently.

We study expander *sequences*, i.e., sequences of graphs on vertex sets of increasing size with uniform bounds on both their sparsity and expansion (read connectivity) properties. These seemingly contradictory properties make it an obvious question whether they even exist. This was first established probabilistically by Pinsker, who proved that random graphs are expanders *with overwhelming probability*, i.e., asymptotically, most graphs are expanders! Despite this fact, it has taken a great effort by mathematicians and computer scientists alike to give explicit constructions of expander sequences. We will describe Margulis' method to obtain expanders as *Schreier coset graphs* of groups with Kazhdan's property (T), which is a rigidity property for unitary representations of groups on Hilbert spaces.

More recently, expander graphs have been gaining interest from mathematicians for their analytic and *coarse* geometric properties. Expanders satisfy a certain set of inequalities known as *Poincaré inequalities* which prevent them from being *coarsely embedded* into Hilbert space. Loosely speaking, a coarse embedding is an embedding which looks like an isometry when viewed from increasing distance. Examples of coarse embeddings are bi-Lipschitz maps and, of course, actual isometries. The rather exotic geometric properties of expanders have found applications within operator algebras and geometric group theory. For example, Osajda [Osa18] used the existence of expanders to construct groups whose Cayley graphs contain such expanders isometrically. These provided new examples of non-exact groups. The method of constructing groups whose Cayley graphs contain expanders originates from [Osa14], where such groups are used to provide counterexamples to versions of the Baum-Connes conjecture.

On a related note, Gromov suggested in [FRR95, p. 67, problems (4) and (5)] to use coarse embeddings into Hilbert space or only *uniformly convex* Banach space in the study of the Novikov and Baum-Connes conjectures. This was carried out by Kasparov and Yu in [KY06] to prove that the coarse geometric Novikov conjecture holds for spaces which admit a coarse embedding into a uniformly convex Banach space. Gromov also asked whether any finitely generated or finitely presented group admits a coarse embedding into any Hilbert space or even into any uniformly convex Banach space. The work of Osajda mentioned above uses the fact that expanders exist to disprove this for Hilbert spaces. This leads naturally to the question whether there exist graph sequences which cannot be coarsely embedded into any uniformly convex Banach space. The focus of this thesis is this question rather than the applications mentioned above.

We will describe how the notion of expander sequences generalizes; one can introduce the notion of being an expander sequence *with respect to* a given metric space. We can then define *superexpanders* to be graph sequences which are expanders with respect to every uniformly convex Banach space. This makes any superexpander coarsely non-embeddable into uniformly convex Banach space. Again, it is a non-trivial task to construct explicit sequences of superexpanders, and giving one such construction will be the main objective of this thesis. The first construction of superexpanders was given in 2008 by Lafforgue as part of his work on the Novikov and Baum-Connes conjectures. He introduced several strengthenings of Kazhdan's property (T) which led to a refinement of Margulis' construction. We shall follow a more recent construction given by Mendel and Naor in 2014 using *zig-zag products* of graphs. The zig-zag product takes two compatible graphs and gives a combinatorial way to construct from them a new graph on a larger vertex set. It was first introduced in [RVW02] in order to produce (classical) expander sequences. The construction of Mendel and Naor involves iteratively producing a graph sequence through succesive zig-zag products as well as other combinatorial graph operations - in other words, it will involve some counting arguments. We will also introduce the notion of *non-linear spectral calculus*, a quantitative property for metric spaces which generalizes (in a very broad sense, admittedly) the spectral calculus known from linear algebra. This is an essential prerequisite for the zig-zag construction to work, and establishing it for uniformly convex Banach spaces will involve some Banach space theory, which we introduce appropriately. Moreover, the construction relies on the existence of a *base graph* with certain analytic properties. Demonstrating that such a base graph exists will involve further Banach space theory and analytic methods.

This exposition is intended to be as self-contained as possible. We will, however, assume familiarity with measure and integration theory as well as functional analysis. Also a basic knowledge of groups and graphs will be required.

Acknowledgements

First of all, I wish to thank my advisor, Magdalena Musat, for introducing me to this topic, for her valuable (and ongoing) criticism of my work, as well as for pushing me outside of my comfort zone. I hope that working under her supervision in the coming three years will prove to be as interesting and enlightening as the past one has. I am also very grateful to Emilie Mai Elkiær, Anna Munk Ebbesen and Mikkel Strunge, all of whom are former students of Magdalena Musat, for the pleasure of reading their Master theses. Since it is customary to refer to original sources or canonical text books, they are not cited anywhere in this thesis, but I have learned several of its topics from theirs. Finally, let me thank my partner, Rikke, for her support and patience with me during the final months of this process, and our kids, Ilja and Luma, for keeping me focused and helping me unlearn the bad habit of procrastinating.

1 Preliminaries

In this section we shall introduce the key notions of the thesis, such as the *coarse geometry* of metric spaces. We will introduce the notion of *expander graphs* and prove results concerning their coarse geometry when viewed as metric spaces. Since vector valued integration will be an indispensable tool for us, we will also introduce this theory and define the *Bochner spaces*, which are the vector valued analogue of the Lebesgue spaces. The treatment will be quite brief and cursory and is only intended to give an overview of the necessary prerequisites of this thesis. We will be giving references to the reader interested in obtaining a deeper understanding of these matters.

1.1 Graphs

Recall that a graph G is given by a pair (V, E), where V is any set and $E \subset V \times V$. We think of V as vertices, and of E as edges connecting certain pairs of vertices. Our graphs will always be undirected, which means that $(u, v) \in E$ if and only if $(v, u) \in E$. We will allow self-loops (i.e., $(u, u) \in E$) and multiple edges. The latter can be formalized by allowing E to be a multiset or by letting $E \subset V \times V \times \mathbb{N}_0$, where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, satisfying that for each pair $(u, v) \in V \times V$ there is exactly one $n \in \mathbb{N}_0$ such that $(u, v, n) \in E$, specifying the number of edges from u to v. We shall denote this number by $E(u, v)$ and observe that for undirected graphs $E(u, v) = E(v, u)$. This notation will be convenient at times, but oftentimes we shall omit the third coordinate of an edge. For the reader remotely familiar with graphs this should lead to no confusion; for example, when we wish to sum over the edges of a graph we shall usually write something along the lines of $(u, v) \in E$, by which of course we mean to sum over all pairs $(u, v) \in V \times V$ weighting each term with $E(u, v)$, and by 'there exists an edge (u, v)' we mean $E(u, v) > 0$.
A graph is called *finite* if its vertex set is finite. For $u, v \in V$, a finite sequence of edges $[(u_1, v_1), \cdots, (u_m, v_m)]$, $m \in \mathbb{N}$, such that $u_1 = u$, $v_m = v$ and $v_k = u_{k+1}$ for every $k = 1, \cdots, m-1$ is referred to as a *path* from u to v. We denote by $l(\gamma)$ its lenght, m. A path from $v \in V$ to itself is called a *cycle*. A graph G is called *connected* if every $u, v \in V$ admit a path between them. Any graph can be uniquely partitioned into connected subgraphs. The members of this partition are called the *path components* of the graph. If $G = (V, E)$ is a connected graph, we can equip V vith a metric, referred to as the *shortest path metric*, by letting $d_G(u, v) = \min \{l(\gamma) \mid \gamma \text{ is a path from } u \text{ to } v\}$.

Unless otherwise stated the reader can assume that our graphs are finite, undirected and connected, in which case we can identify V with $[n]$ where $n = |V|$, and $[n]$ denotes the set $\{1, \cdots, n\}$. We then have the following very useful piece of notation:

Definition 1.1. Let $G = (V, E)$ be a finite graph. By the *adjaceny matrix* of G we mean the matrix $(a_{ij})_{i,j\in[n]}$ where $a_{ij} = E(i, j)$, i.e., a_{ij} is the number of edges from i to j.
Clearly G is undirected if and only if its adjacency matrix is symmetric.

A finite graph is completely determined by its adjacency matrix and we shall often identify a graph with its adjacency matrix and write G and its adjacency matrix interchangeably. A graph is called *d-regular* for $d \in \mathbb{N}$ if every vertex has exactly d edges emmanating from it, i.e., $\sum_{v\in V} E(u, v) = d$ for every $u \in V$. The quantity $\sum_{v\in V} E(u, v)$ is referred to as the *degree* or *valence* of u, so d-regularity just means that every vertex has degree exactly d. In terms of the adjacency matrix this reads that every row, and hence every column, sums to d. There are precisely two connected 1-regular graphs; one with one vertex and one with two vertices, so we will always assume that the degree of regularity is at least 2 when considering regular graphs. When a graph is d-regular (as is almost surely the case in this thesis), we shall often consider the normalized adjacency matrix, $(a_{ij})/d$. This matrix is symmetric and stochastic, and can be viewed as the transition matrix of a random walk on G.

Definition 1.2. Let $G = (V, E)$ be a finite, connected graph. For a subset U of V, let $\partial U \subset E$ denote the edges in G with one endpoint in U and the other in U^c; ∂U is referred to as the *boundary* of U. We define the *Cheeger constant* of G to be

$$h(G) = \min \left\{ \frac{|\partial U|}{|U|} \mid U \subset V,\ 0 < |U| \leq |V|/2 \right\}.$$

In the literature, the Cheeger constant is also sometimes found under the name of *expanding constant* or *conductance*. It gives a lower bound for the cost, in terms of the number of edges one has to cut, of disconnecting a (non-empty) subset of V from the rest of the graph, relative to the size of that subset. Thus a connected graph with a large Cheeger constant is considered well-connected, since there are no virtually isolated parts of the graph.

Definition 1.3. A sequence of finite, connected, d-regular graphs $(G_n)_{n\geq 1}$ is called a sequence or family of *edge expanders* if $|V_n| \to \infty$ as $n \to \infty$, where V_n denotes the vertex set of G_n, and there exists $\varepsilon > 0$ such that $h(G_n) \geq \varepsilon$, for all $n \in \mathbb{N}$.

The assumption about d-regularity ensures that the number of edges grows only proportionally to the number of vertices, and the uniform bound away from zero of the Cheeger constants ensures that the graphs are reasonably well-connected, thus giving rise to the popular description of expander graphs as being 'sparse graphs with strong connectivity properties'. This description can be quantified by the spectral properties of the adjacency matrix as well, as we shall see now. Being symmetric and stochastic, it follows from elementary linear algebra that the normalized adjacency matrix $A = (a_{ij})$ of a d-regular graph has only real eigenvalues. These eigenvalues contain (not surprisingly) a lot of information about G. We have the following result:

Proposition 1.4. *Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of A, counted with muliplicity and put in descending order.*

1. *$|\lambda_i| \leq 1$, for all $i = 1, \cdots, n$, and $\lambda_1 = 1$.*
2. *$\lambda_2 < 1$ if and only if G is connected.*
3. *$\lambda_n = -1$ if and only if G has a bipartite connected component.*

Proof. $1:$ Let $1_V \in \mathbb{R}^n$ be the all ones vector. Then, since A is stochastic we have $A1_V = 1_V$, so 1 is an eigenvalue. Moreover, any vector $x \in \mathbb{R}^n$ satisfies $|(Ax)_i| \leq \max_{1\leq j\leq n} |x_j|$ for any $1 \leq i \leq n$, so that no eigenvalue can have absolute value larger than 1.

2 : Assume that G is disconnected and let $U \subsetneq V$ be a connected component of G. Then $A1_U = 1_U$ and $A1_{U^c} = 1_{U^c}$ and since 1_U and 1_{U^c} are orthogonal, and hence linearly independent, we have $\lambda_2 = 1$.
Now assume that G is connected and let $x = (x_i)_i \in \mathbb{R}^n$ be an eigenvector of A with eigenvalue 1. Then, for all $i \in [n]$, we have $\sum_{j=1}^n a_{ij}x_j = x_i$. In particular this holds for any $i_0 \in [n]$ realizing $\max_{i\in[n]} x_i$, which implies that $x_j = x_{i_0}$, for all j such that $a_{i_0 j} \neq 0$. Finally connectedness of G means precisely that for any $j \in [n]$ there is a sequence $i_0, i_1, \cdots, i_k = j$ such that $a_{i_l, i_{l+1}} \neq 0$, for all $l = 0, \cdots, k-1$, so the above argument iterated l times yields $x_j = x_{i_0}$. This demonstrates that x is a constant vector and hence a multiple of 1_V.
3 : Suppose G has a bipartite component $U = X \cup Y \subset V$, i.e., X and Y are nonempty and disjoint, each vertex of X is adjacent only to vertices from Y and vice versa. Consider the vector $x = 1_X - 1_Y \in \mathbb{R}^n$. For $i \in X$ we have

$$(Ax)_i = \sum_{j=1}^n a_{ij}(1_X(j) - 1_Y(j)) = \sum_{j=1}^n -a_{ij} = -1 = -x_i,$$

since $a_{ij} = 0$ whenever $j \notin Y$. It can be seen analogously that $(Ax)_i = 1 = -x_i$ whenever $i \in Y$. Since also clearly $(Ax)_i = 0 = -x_i$ when $i \in (X \cup Y)^c$ we have $Ax = -x$ and -1 is an eigenvalue of A with eigenvector x.
Suppose conversely that -1 is an eigenvalue of A with eigenvector x. This means that for every $i \in [n]$ we have $\sum_{j=1}^n a_{ij}x_j = -x_i$. In particular, this holds for any $i_0 \in [n]$ realizing $\max_{1\leq i\leq n} |x_i| > 0$. The above equality then implies that $x_j = -x_{i_0}$ for each j such that $a_{ij} \neq 0$, so any such j also realizes $\max_{1\leq i\leq n} |x_i|$. This also means that any j in the connected component of i_0 has $x_j = \pm x_{i_0}$ and adjacent vertices have opposite signs. Let $X = \{j \sim_p i_0 \mid x_j = x_{i_0}\}$ and $Y = \{j \sim_p i_0 \mid x_j = -x_{i_0}\}$, where $j \sim_p i_0$ denotes that j and i_0 are path connected in G. Then $X \cup Y$ form a bipartition of the path component of i_0. □

Definition 1.5. In the above notation, the quantity $\lambda(G) := 1 - \lambda_2 > 0$ is referred to as the *spectral gap* of G and $\lambda_+(G) := 1 - \max\{|x_2|, |x_n|\}$ as the *absolute* spectral gap of G.

By the above proposition we then have $\lambda(G) > 0$ if and only G is connected and $\lambda_+(G) > 0$ if and only if G is connected and not bipartite.

Definition 1.6. A sequence of finite, connected, d-regular graphs $(G_n)_{n\geq 1}$ is called a sequence or family of *spectral expanders* if $|V_n| \to \infty$ as $n \to \infty$ and there exists $\lambda > 0$ such that $\lambda(G_n) \geq \lambda$, for all $n \in \mathbb{N}$.

The following theorem, the content of which is known as the Cheeger inequalities, implies that being a family of edge expanders is equivalent to being a family of spectral expanders.

Theorem 1.7. *Let G be a finite, d-regular graph. Then*

$$\frac{d\lambda(G)}{2} \leq h(G) \leq 8d\sqrt{2\lambda(G)}.$$

In particular a sequence of connected, d-regular graphs is a sequence of edge expanders if and only if it is a sequence of spectral expanders.

We will prove Theorem 1.7 in Section 2 after introducing the necessary prerequisites. The existence of families of expanders is by no means obvious. The first existence proof is probabilistic (and hence unconstructive) and due to Pinsker (see [Pin73]). We shall describe explicit constructions of expander families in subsequent sections. We end this section by introducing a very important class of graphs, namely *Cayley graphs*:

Definition 1.8. Let Γ be a group and $\Sigma \subset \Gamma$. The Cayley graph of Γ with respect to Σ, denoted by $\mathrm{Cay}(\Gamma, \Sigma)$, is the graph whose vertex set is Γ and where (g, h) is an edge if there exists $s \in \Sigma$ such that $h = sg$.

By the cancellation rule there are no multiple edges, and there are self loops if and only if $e \in \Sigma$, in which case there is a self loop at every vertex. Moreover $\mathrm{Cay}(\Gamma, \Sigma)$ is undirected if and only if Σ is symmetric (i.e., $s \in \Sigma$ if and only if $s^{-1} \in \Sigma$) and connected if and only if Σ generates Γ. If Σ is finite then $\mathrm{Cay}(\Gamma, \Sigma)$ is $|\Sigma|$-regular.

Much more can be said about (edge and spectral) expanders, and we shall do so in Sections 2 and 3. This cursory introduction is by and large based on [Ost13]. We also refer to [AM85] or [RVW02] for more properties and applications. For more on the interconnections between groups and graphs we refer to [Mei08].

1.2 Coarse geometry and Poincaré inequalities

The key idea behind the notion of coarse geometry is viewing objects 'from far away' and consider them equal if they 'look the same' when seen from a large distance.

Definition 1.9. Let (X, d_X) and (Y, d_Y) be metric spaces. A map $f : X \to Y$ is called a *coarse map* if

1. There exists a non-decreasing function $\rho_+ : \mathbb{R}_+ \to \mathbb{R}_+$ such that
$$d_Y(f(x), f(x')) \leq \rho_+(d_X(x, x')),$$
for all $x, x' \in X$.
2. f is metrically proper, i.e., if $B \subset Y$ is bounded then so is $f^{-1}(B)$.

Condition 1 can be rephrased as: For every $R > 0$ there exists an $S > 0$ such that $d_Y(f(x), f(x')) \leq S$ whenever $d(x, x') \leq R$. Note the difference between *coarse* and *continuous*: we are not requiring that $f(x)$ and $f(x')$ be close when x and x' are, but only that they do not get 'too far apart'. Also condition 1 is a uniform condition, whereas continuity is a local condition. Condition 2 ensures that points that end up close could not have started too far apart.

Definition 1.10. Let (X, d_X) be a metric space. A subset $S \subset X$ is called a net, if there exists some $N \in \mathbb{N}$ such that for any $x \in X$ we can find $y \in S$ such that $d_X(x, y) < N$.

Again the idea is to encapture the notion of things looking similar from far away; when viewed from increasing distance, X will look like any net in itself.

Example 1.11. $n\mathbb{Z} \subset \mathbb{R}$ is a net for any $n \in \mathbb{N}$ with $N = \lceil n/2 \rceil + 1$.

Definition 1.12. A map $f : X \to Y$ is called a *coarse embedding* if there exist non-decreasing maps $\rho_\pm : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{t\to\infty} \rho_-(t) = \infty$ and, for all $x, x' \in X$,

$$\rho_-(d_X(x, x')) \leq d_Y(f(x), f(x')) \leq \rho_+(d_X(x, x')).$$

If furthermore $f(X)$ is a net in Y we call f a *coarse equivalence*. The maps ρ_- and ρ_+ will be referred to as the control functions of f.

Example 1.13. If Γ is a finitely generated group and Σ and Σ' are both symmetric generating subsets then the Cayley graphs $\mathrm{Cay}(\Gamma, \Sigma)$ and $\mathrm{Cay}(\Gamma, \Sigma')$ are coarsely equivalent via the identity map $\mathrm{id} : \Gamma \to \Gamma$, when both graphs are equiped with the path metric.
Also any two finite metric spaces are coarsely equivalent, and in particular any finite metric space embeds coarsely into any other metric space.

We now introduce the *coarse disjoint union* of metric spaces:

Definition 1.14. Let $(X_i, d_i)_{i\geq 1}$ be a sequence of finite metric spaces. The *coarse disjoint union* of these metric spaces is the space

$$\bigsqcup_{i\geq 1} X_i \quad \text{(the disjoint union of } \{X_i\}_{i\geq 1})$$

with a metric d such that $d(x, y) = d_i(x, y)$ whenever $x, y \in X_i$, for some $i \geq 1$, and such that $\mathrm{dist}(X_i, X_j) \to \infty$, as $i + j \to \infty$ and $i \neq j$.

Putting $d(x, y) = \max \{i + j, \mathrm{diam}(X_i), \mathrm{diam}(X_j)\}$ when $x \in X_i$ and $y \in X_j$ with $i \neq j$ defines a metric on X with the desired properties.
The coarse disjoint union makes a sequence of finite metric spaces into one big (infinite) metric space. One can think of this as putting the metric spaces in the sequence on a string passing through a designated point in each space, such that the further out we go in the sequence, the further apart they lie on the string. This means that not only are two spaces which are far apart in the sequence also far apart in the coarse disjoint union, but also two 'adjacent' spaces are far apart, provided they lie far out in the sequence. This idea, which is clearly formalized in the above definition, leads to the following observation:

Remark 1.15. Let $(X_i, d_i)_{i\geq 1}$ be a sequence of finite metric spaces and (Y, d_Y) be yet another metric space. Then there exists a coarse embedding from the coarse disjoint union of the X_i into Y if and only if there exist functions $f_i : X_i \to Y$ for each $i \geq 1$ and maps $\rho_\pm : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying the conditions of Definition 1.12 for each $i \geq 1$. That is, $\rho_\pm$ are control functions for all the f_i simultaneously. For this reason we will simply say that *the sequence* coarsely embeds into Y when we mean its coarse disjoint union.

We now introduce (discrete) Poincaré inequalities and describe their role as obstructions to coarse embeddability:

Definition 1.16. Let (X, d_X) and (Y, d_Y) be metric spaces, $(a_{x,y})_{x,y\in X}$ and $(b_{x,y})_{x,y\in X}$ be arrays of non-negative real numbers indexed over X with only finitely many non-zero entries, and $\Psi : \mathbb{R}_+ \to \mathbb{R}_+$ a non-decreasing function. If for any function $f : X \to Y$ the inequality

$$\sum_{x,y\in X} a_{x,y}\Psi(d_Y(f(x), f(y))) \geq \sum_{x,y\in X} b_{x,y}\Psi(d_Y(f(x), f(y))) \tag{1}$$

is satisfied, we say the Y-valued functions on X satisfy the above *Poincaré* inequality.

Many a reader is probably quite puzzled by this definition. First of all it is not difficult to find arrays and a function Ψ such that (1) is satisfied. The arrays can be chosen to be identical, in which case (1) is trivially satisfied and completely uninteresting. Moreover, the metric structure of X appears to be redundant - X is merely a labelling set. The metric structure of X does however play a more subtle role: it usually determines the choice of the arrays $(a_{x,y})$ and $(b_{x,y})$ and the function Ψ in order to get inequalities that yield interesting results relating the coarse geometries of X and Y.
We shall be interested in the following type of Poincaré inequalities for connected graphs:

Definition 1.17. Let $G = (V, E)$ be a finite, connected graph equipped with the shortest path metric d_G. Let (Y, d_Y) be any metric space. We say that Y-valued functions on the graph satisfy a p-Poincaré inequality with respect to the adjacency matrix if there exists a $\gamma > 0$ such that

$$\frac{\gamma}{|V|}\sum_{u,v\in V} d_Y(f(x), f(y))^p \leq \sum_{u,v\in V} a_{u,v} d_X(f(x), f(y))^p,$$

for all $f : V \to Y$.

In terms of Definition 1.16, p-Poincaré inequalities are Poincaré inequalities for functions on connected graphs with $\Psi(t) = t^p$, $b_{u,v} \equiv \gamma/|V|$ and $(a_{u,v})$ is the adjacency matrix of G. The reason why we are interested in these types of Poincaré inequalities is because of the following proposition which appears to be attributed to Gromov:

Proposition 1.18. *Let (Y, d_Y) be a metric space and $p > 0$. Suppose $G_n = (V_n, E_n)$ is a sequence of finite, connected, d-regular graphs such that $|V_n| \to \infty$ and there is $\gamma > 0$ such that for all $n \in \mathbb{N}$ and $f : V_n \to Y$ we have*

$$\frac{\gamma}{|V_n|}\sum_{u,v\in V_n} d_Y(f(u), f(v))^p \leq \sum_{u,v\in V_n} a^n_{u,v} d_Y(f(u), f(v))^p, \tag{2}$$

where $(a^n_{u,v})$ is G_n's adjacency matrix. Then $(G_n)_{n\geq 1}$ does not coarsely embed into Y.

Proof. Suppose there were functions $f_n : V_n \to Y$ and control functions $\rho_\pm : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying the conditions of Definition 1.12. Together with the assumed inequality (2) this yields for every $n \in \mathbb{N}$ that

$$\begin{aligned}\sum_{u,v\in V_n} \frac{\gamma}{|V_n|}\rho_-(d_{G_n}(u, v))^p &\leq \sum_{u,v\in V_n} a^n_{u,v} d_Y(f(u), f(v))^p \\ &\leq \sum_{u,v\in V_n} a^n_{u,v}\rho_+(d_{G_n}(u, v))^p \\ &\leq \sum_{u,v\in V_n} a^n_{u,v}\rho_+(1)^p = d|V_n|\rho_+(1)^p,\end{aligned}$$

using that $d_{G_n}(u, v) \leq 1$ whenever (u, v) is an edge in G_n, and the fact that $|E_n| = d|V_n|$ since G_n is d-regular. Now by d-regularity the number of vertices of distance at most $D \geq 0$ to a given vertex is no more than $d^D + 1$. Taking $D = \log_d(|V_n|/2 - 1)$ this reads that at least

$|V_n|^2/2$ of the $|V_n|^2$ terms on the left of the above inequality have $d_{G_n}(u,v) \geq \log_d(|V_n|/2-1)$. Because ρ_- is non-decreasing this yields

$$\frac{\gamma}{|V_n|}\frac{|V_n|^2}{2}\rho_-(\log_d(|V_n|/2-1))^p \leq d|V_n|\rho_-(1)^p,$$

and since $|V_n| \to \infty$ this clearly contradicts $\lim_{t\to\infty}\rho_-(t) = \infty$. □

In this thesis we will be concerned with introducing terminology and establishing Poincaré inequalities with respect to the adjacency matrix with uniform constants, as those assumed in Proposition 1.18. By the statement of this proposition we are then also establishing non-coarse embeddability results between metric spaces.

For a deeper treatment of coarse geometry and its role in Banach space theory, we refer to [NY12]. Poincaré inequalities as obstructions to coarse embeddings, including Proposition 1.18, are treated in [Ost13].

1.3 Superreflexive Banach spaces

Throughout this exposition, unless otherwise stated, vector spaces are over $\mathbb{K}$, where $\mathbb{K}$ is either $\mathbb{C}$ or $\mathbb{R}$. If a result holds only for one of these cases, this will be elaborated explicitly.

Definition 1.19. A normed vector space, $(X, \|\cdot\|)$, is called *uniformly convex* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that whenever $x, y \in X$ satisfy $\|x\|, \|y\| \leq 1$ and $\|x-y\| \geq \varepsilon$ we have $\frac{\|x+y\|}{2} \leq 1-\delta$.

Remark 1.20. It can be seen that if we replace the condition $||x||, ||y|| \leq 1$ with $||x|| = ||y|| = 1$ we get an equivalent definition of uniform convexity. So if we define the *modulus of (uniform) convexity* of B by

$$\delta_B(\varepsilon) = \inf\left\{1 - \frac{\|x+y\|}{2} \mid \|x\| = \|y\| = 1,\ \|x-y\| \geq \varepsilon\right\},$$

then B is uniformly convex precisely when $\delta_B(\varepsilon) > 0$, for all $0 < \varepsilon \leq 2$. We say that B has modulus of convexity *of power type* $p \geq 2$ if there exists a $C > 0$ such that $\delta_B(\varepsilon) \geq C\varepsilon^p$, for all $0 < \varepsilon \leq 2$. The modulus of uniform convexity, including its connection to other geometric properties of Banach spaces, will be treated more thoroughly in Section 6.

Uniform convexity is a geometric notion for Banach spaces, but the following proposition connects it to a functional analytical property:

Proposition 1.21. *If B is a uniformly convex Banach space then B is reflexive.*

Proof. We wish to prove that the isometric evaluation map $\Lambda : B \to B^{**}$ given by $\Lambda(x)(f) = f(x)$, $f \in B^*$ and $x \in B$, is surjective. The proof is an application of Goldstine's theorem which states that the image under the evaluation map of the (norm) closed unit ball of B is w^* dense in the (norm) closed unit ball of the bidual, i.e., $\overline{\Lambda(B_1)}^{w*} = (B^{**})_1$.
Take $\varphi \in B^{**}$ and observe that by linearity we can assume that φ has unit norm so that there exists, by Goldstine's theorem, a net $(x_i)_{i\in I} \subset B_1$ such that $\Lambda(x_i) \xrightarrow{w^*} \varphi$, i.e., $\Lambda(x_i)(f) \to \varphi(f)$ for every $f \in B^*$. It follows from uniform convexity that (x_i) is Cauchy in norm. Indeed, let

$0 < \varepsilon < 1$ and pick $\delta > 0$ according to uniform convexity, i.e., (by contraposing) whenever $x, y \in B$ satisfy $\|x\|, \|y\| \leq 1$ and $\left\|\frac{x+y}{2}\right\| > 1 - \delta$ then $\|x - y\| < \varepsilon$. Since $\|\varphi\| = 1$ there is $f \in B^*$ of unit norm such that $|\varphi(f)| > 1 - \delta$ and by scaling with a phase we may assume that $\varphi(f)$ is real and positive. Now since $f(x_i) = \Lambda(x_i)(f) \to \varphi(f)$ we can pick $i_0 \in I$ such that $\Re f(x_i) > 1 - \delta$ whenever $i \succeq i_0$. Hence for $i, j \succeq i_0$ we have

$$\left\|\frac{x_i + x_j}{2}\right\| \geq |\frac{1}{2} f(x_i + x_j)| \geq \frac{1}{2} \Re f(x_i + x_j) > 1 - \delta$$

and hence $||x_i - x_j|| < \varepsilon$, demonstrating that (x_i) is Cauchy in norms. By completeness of B we can therefore find $x \in B^*$ such that $x_i \overset{||\cdot||}{\to} x$ and since the evaluation map is an isometry we obtain $\Lambda(x_i) \overset{||\cdot||}{\to} \Lambda(x)$ implying $\Lambda(x_i) \overset{w^*}{\to} \Lambda(x)$. Since the w^* topology is Hausdorff we obtain $\varphi = \Lambda(x)$ demonstrating the desired. □

Example 1.22. Clearly uniform convexity passes to subspaces (and, with a little more work, quotients), but it is not stable under isomorphisms. This can be seen by considering ℓ_2 equipped with the norm $\|(x_n)\|' = \max\left\{2|x_1|, \left(\sum_{n\geq 2} |x_n|^2\right)^{1/2}\right\}$, which is equivalent to the usual 2-norm on ℓ_2, i.e., $(\ell_2, \|\cdot\|_2) \cong (\ell_2, \|\cdot\|')$. We will see later in this subsection that $(\ell_2, \|\cdot\|_2)$ is uniformly convex, but $(\ell_2, \|\cdot\|')$ is not. To see the latter claim consider $x = (1/2, 0, \cdots)$ and $y = (1/2, 1/2, 0, \cdots)$. Then $\|x\|' = \|y\|' = 1$ and $\|x - y\|' = 1/2$, whereas $\|x + y\|' /2 = 1$, so that $\delta_{(\ell_2, \|\cdot\|')}(1/2) = 0$.

Definition 1.23. A Banach space, $(B, || \cdot ||)$, is called *superreflexive* if it admits a uniformly convex norm which is equivalent to $|| \cdot ||$. In other words, B is superreflexive if it is isomorphic to a uniformly convex Banach space.

Obviously, uniform convexity implies superreflexivity which is stable under isomorphism. Superreflexivity, however, does not imply uniform convexity, as is seen by example 1.22; $(\ell_2, \|\cdot\|')$ is superreflexive, being isomorphic to $(\ell_2, \|\cdot\|_2)$, but not uniformly convex. Since reflexivity is preserved under isomorphisms, Proposition 1.21 yields the following corollary:

Corollary 1.24. *If B is superreflexive then B is also reflexive.*

We will provide an example below of a reflexive Banach space which is not superreflexive, so the latter is really a stronger requirement. Let us proceed to give some examples of superreflexive Banach spaces.

Proposition 1.25. *Let $1 < p < \infty$ and let $(\Omega, \mathscr{A}, \mu)$ be a σ-finite measure space. Then $L_p(\Omega)$ is superreflexive.*

Indeed, $L_p(\Omega)$ equipped with the usual p-norm is even uniformly convex - a consequence of Clarkson's inequalities (see [Cla36]):

Lemma 1.26. *Let $f, g \in L_p(\Omega)$.*

1. *If $1 < p < 2$ then* $\|f + g\|_p^{p/(p-1)} + \|f - g\|_p^{p/(p-1)} \leq 2\left(\|f\|_p^p + \|g\|_p^p\right)^{1/(p-1)}$
2. *If $2 \leq p < \infty$ then* $\|f + g\|_p^p + \|f - g\|_p^p \leq 2^{p-1}\left(\|f\|_p^p + \|g\|_p^p\right)$

Taking $p = 2$ we obtain that Hilbert space is uniformly convex, a fact which also follows from the parallelogram identity. It is easy to see directly that $L_1(\Omega)$ and $L_\infty(\Omega)$ with the usual norms are not uniformly convex. But since they are not reflexive it follows from Corollary 1.24 that they are not even superreflexive, i.e., there is no equivalent norm on these spaces, which is uniformly convex. The definition of superreflexivity given in Definition 1.23 is not the original one, it is in fact a theorem that this definition is equivalent to the original one. We will proceed to discuss super-properties for Banach spaces in order to give the original definition of superreflexivity.

Definition 1.27. Let A and B be Banach spaces. We say that A is *finitely representable* in B if the following holds:
For every $\varepsilon > 0$ and every finite-dimensional subspace $V \subset A$, there exists a finite-dimensional subspace $W \subset B$ and an isomorphism $T : V \to W$ such that $\|T\| \left\|T^{-1}\right\| < 1 + \varepsilon$.

If $\mathscr{P}$ is some property defined for Banach spaces, then B is said to have *super-*$\mathscr{P}$ if every Banach space which is finitely representable in B has $\mathscr{P}$. Since obviously B is finitely representable in itself super-$\mathscr{P}$ implies $\mathscr{P}$. Also note that, since finite representability is obviously transitive, super-super-$\mathscr{P}$ is equivalent to super-$\mathscr{P}$. Now in this terminology 'superreflexivity' of B means that any Banach space which is finitely representable in B is reflexive. This is the original definition due to R. C. James, and it is a deep theorem of Enflo [Enf72] that this is equivalent to the one given in Definition 1.23. Using martingale techniques, Pisier [Pis75] improved Enflo's results and proved:

Theorem 1.28. *Let B be a Banach space. Then the following are equivalent*

1. *B is isomorphic to a uniformly convex Banach space (i.e., B is superreflexive in the sense of Definition 1.23).*
2. *Any Banach space which is finitely representable in B is reflexive (i.e., B is superreflexive in the sense of James' original definition).*
3. *B is isomorphic to a uniformly convex Banach space with modulus of uniform convexity of power type p for some $p \geq 2$.*

Superreflexive Banach spaces satisfy the following permanence properties, some of which have already been discussed

Proposition 1.29. *Suppose B is a superreflexive Banach space.*

1. *If $M \subset B$ is a closed subspace then M is superreflexive.*
2. *If E is another Banach space which is finitely representable in B then E is superreflexive.*
3. *If E is another Banach space which is isomorphic to B then E is superreflexive.*
4. *If $M \subset B$ is a closed subspace then the quotien B/M is superreflexive.*

Now we provide an example of a Banach space which is reflexive but not superreflexive:

Example 1.30. Consider the n-dimensional spaces, ℓ_1^n, equipped with the 1-norm. Since these are finite-dimensional they are all reflexive. For $1 \leq p < \infty$, we define

$$\left(\bigoplus_{n\geq 1} \ell_1^n\right)_p = \left\{ (f_n)_{n\geq 1} \mid f_n \in \ell_1^n, \ \sum_{n=1}^{\infty} \|f_n\|_{\ell_1^n}^p < \infty \right\}$$

which, equipped with the norm $\|(f_n)_{n\geq 1}\|_p = (\sum_{n=1}^{\infty} \|f_n\|_{\ell_1^n}^p)^{1/p}$, is a Banach space. Moreover, for any $1 < p < \infty$ we have $\left(\bigoplus_{n\geq 1} \ell_1^n\right)_p^* = \left(\bigoplus_{n\geq 1} (\ell_1^n)^*\right)_q$ where q is the Hölder conjugate of p. In particular reflexivity of ℓ_1^n for each $n \geq 1$ implies that $B := \left(\bigoplus_{n\geq 1} \ell_1^n\right)_2$ is reflexive. However, as we shall see below, ℓ_1 is finitely representable in B, yielding, since ℓ_1 is not reflexive, that B is not superreflexive. To see the above claim let V be a finite-dimensional subspace of ℓ_1 and let $\varepsilon > 0$. Take an algebraic basis $\xi_1, \cdots, \xi_m$ of V such that $\|\xi_i\|_{\ell_1} = 1$ for each $i = 1, \cdots, m$. Since any two norms on a finite-dimensional vector space are equivalent we can find $C > 0$ such that $\sum_{i=1}^m |a_i| \leq C^{-1} \|\sum_{i=1}^m a_i \xi_i\|_{\ell_1}$ for every sequence of scalars $a_1, \cdots, a_m$. Now let $0 < \alpha < C$ be such that $\frac{1+\alpha/C}{1-\alpha/C} < 1 + \varepsilon$. Since $\|\xi_i\|_{\ell_1} < \infty$ we can find $M \in \mathbb{N}$ such that $\sum_{j=M+1}^{\infty} |\xi_i(j)| < \alpha$ for each $i = 1, \cdots, m$. So if we define $\xi_1', \cdots, \xi_m' \in \ell_1$ by $\xi_i'(j) = \xi_i(j)$ if $j \leq M$ and $\xi_i'(j) = 0$ if $j > M$ we have $\|\xi_i - \xi_i'\|_{\ell_1} < \alpha$ for every $i = 1, \cdots, m$. Now we define $T : V \to \operatorname{span}\{\xi_1', \cdots, \xi_m'\}$ by $T(\xi_i) = \xi_i'$ for $i = 1, \cdots, m$. Then for every sequence of scalars $a_1, \cdots, a_m$ we have

$$\begin{aligned}\left\| T\left(\sum_{i=1}^m a_i \xi_i\right)\right\|_{\ell_1} &\leq \left\|\sum_{i=1}^m a_i(\xi_i' - \xi_i)\right\|_{\ell_1} + \left\|\sum_{i=1}^m a_i \xi_i\right\|_{\ell_1} \\ &\leq \alpha \sum_{i=1}^m |a_i| + \left\|\sum_{i=1}^m a_i \xi_i\right\|_{\ell_1} \leq (1 + \frac{\alpha}{C}) \left\|\sum_{i=1}^m a_i \xi_i\right\|_{\ell_1},\end{aligned}$$

so $\|T\| \leq 1 + \alpha/C$. It is seen in a similar fashion that

$$\left\|\sum_{i=1}^m a_i \xi_i\right\|_{\ell_1} \leq \frac{\alpha}{C} \left\|\sum_{i=1}^m a_i \xi_i\right\|_{\ell_1} + \left\|\sum_{i=1}^m a_i \xi_i'\right\|_{\ell_1},$$

which yields simultaneously that T is invertible and that $\left\|T^{-1}\right\| \leq (1 - \alpha/C)^{-1}$.
Since $\xi_1', \cdots, \xi_m'$ sit isometrically in ℓ_1^M which in turn sits isometrically in B, we obtain, by composing T with these isometries, an isomorphism $T' : V \to W$, where W is a finite-dimensional subspace of B such that $\|T'\| \left\|(T')^{-1}\right\| \leq \frac{1+\alpha/C}{1-\alpha/C} < 1 + \varepsilon$.

This subsection is based on [Bea82], which provides an excellent introduction to superreflexivity as a geometric property for Banach spaces. This includes even more characterizations of superreflexive Banach spaces, leading to proofs of the permanence properties stated in Proposition 1.29.

1.4 Vector valued integration

Recall that we may equip a topological space with the Borel σ-algebra, generated by the open sets. For finite dimensional (real or complex) vector spaces, this gives rise to the familiar theory of Lebesgue integration of functions with values in $\mathbb{K}^n$. This theory hinges on the fact that any Borel measurable, scalar valued function may be approximated pointwise by simple functions. This fact, however, does not hold for general Banach spaces, so to develop an integration theory for Banach space-valued functions we will need to make it an assumption. This turns out to be intimately connected to the σ-algebra generated by the dual of a Banach space, allowing for the use of functional analytical techniques.

Definition 1.31. Let B be a Banach space.

- Let $\mathscr{B}(B)$ denote the Borel σ-algebra on B.
- For $Y \subset B^*$ let $\sigma(Y)$ denote the σ-algebra generated by Y, i.e., the smallest σ-algebra making every $\varphi \in Y$ measurable (here the codomain of φ, $\mathbb{R}$ or $\mathbb{C}$, is of course equipped with the good old Borel σ-algebra).

Since any element of B^* is continuous we clearly have $\sigma(Y) \subset \sigma(B^*) \subset \mathscr{B}(B)$.

It is not difficult to see that $\sigma(Y)$ is generated by set of the form $((\varphi_1(x), \cdots, \varphi_n(x)) \in A)$, where $n \geq 1$, $\varphi_1, \cdots, \varphi_n \in Y$ and $A \in \mathscr{B}(\mathbb{K}^n)$. One might wonder whether the above inclusions are in fact equalities. They are not in general but the following proposition shows that they are if we assume separability of B:

Proposition 1.32. *Assume that B is separable and Y is a weak* dense subspace of B^*. Then $\sigma(Y) = \sigma(B^*) = \mathscr{B}(B)$.*

Proof. We prove $\sigma(B^*) = \mathscr{B}(B)$, the other equality can be seen by applying (a corollary of) the Krein-Šmulian theorem. We trivially have $\sigma(B^*) \subset \mathscr{B}(B)$ and for the other inclusion, choose, by separability, a sequence of functionals $(\varphi_n) \subset B^*$ of unit norm which is norming for B. Then for any open ball, $\mathcal{B}(x_0, r) \subset B$, we have

$$\mathcal{B}(x_0, r) = \{x \in B \mid \|x - x_0\| < r\} = \left\{x \in B \mid \sup_{n \geq 1} |\varphi_n(x - x_0)| < r\right\} \in \sigma(B^*).$$

Since the open balls generate the Borel σ-algebra we obtain $\mathscr{B}(B) \subset \sigma(B^*)$ as desired. □

Definition 1.33. Let $(S, \mathscr{A})$ be a measurable space, B a Banach space and $f : S \to B$ a function.

- If f is $\mathscr{A} - \mathscr{B}(B)$ measurable we say that it is *Borel measurable*, in the scalar valued case (i.e., $B = \mathbb{K}$) we shall oftentimes simply say that f is measurable.
- We say that f is *weakly* measurable if it is $\mathscr{A} - \sigma(B^*)$ measurable. This is clearly equivalent to $\varphi \circ f$ being measurable for each $\varphi \in B^*$.

We introduce one more notion of measurability:

Definition 1.34. Let $(S, \mathscr{A})$ be a measurable space. For a function $f : S \to \mathbb{K}$ and $x \in B$ we denote by $f \otimes x : S \to B$ the function given by $f \otimes x(s) := f(s)x$. Functions $f : S \to B$ of the form $f = \sum_{i=1}^n 1_{A_i} \otimes x_i$, where $A_i \in \mathscr{A}$ and $x_i \in B$, are called *simple*.

Definition 1.35. A function $f : S \to B$ is called *strongly measurable* if it is the pointwise limit of a sequence of simple functions.

Remark 1.36. We have to decorate the above definition with the prefix 'strongly', because this notion does not in general coincide with Borel measurability. This is easily seen by considering the case $(S, \mathscr{A}) = (B, \Sigma)$, where Σ is any σ-algebra which contains the Borel σ-algebra, and the identity map $I : B \to B$. I is clearly Borel measurable, but if we assume B to be non-separable then I cannot be approximated by a sequence of simple functions. Indeed, if this were the case, then the values taken by the functions in such a sequence would constitute a countable, dense subset of B.

We shall see that strong measurability implies Borel measurability. We now have three different notions of measurability and their interconnections are described in the following theorem and its corollaries:

Theorem 1.37. *Let $(S, \mathscr{A})$ be a measurable space and B a Banach space. Then for a function $f : S \to B$ the following are equivalent:*

1. *f is strongly measurable.*
2. *f is separably valued and weakly measurable.*

where by f being separably valued we mean that there exists a closed, separable subspace $B_0 \subset B$ such that $f(S) \subset B_0$.

Proof. (1) $\implies$ (2): Let (f_n) be a sequence of simple functions converging pointwise to f. Then since each f_n takes only finitely many values the set $\{f_n(s) \mid n \geq 1,\ s \in S\} \subset B$ is countable, and since, by assumption, any element of $f(S)$ can be approximated by elements of this set, we see that $f(S)$ is contained in its closed linear span, which is separable. Moreover, since simple functions are clearly Borel measurable, we have that $\varphi \circ f_n$ is Borel measurable for each $\varphi \in B^*$, yielding, since φ is continuous, that $\varphi \circ f$ is Borel measurable, being the pointwise limit of a sequence of Borel measurable functions (we are here back in scalar valued case). This demonstrates that f is weakly measurable.
(2) $\implies$ (1): Let B_0 be as above, and choose a sequence $(\varphi_n) \subset B^*$ of unit norm which is norming for B_0 (such a sequence exists since B_0 is separable). Also, let $(x_n) \subset B_0$ be a dense sequence, and for each $n \geq 1$ define the function $g_n : B_0 \to \{x_1, \cdots, x_n\}$ in the following way: For $y \in B_0$ let $k(n, y)$ be the smallest $k \in [n]$ such that $\|y - x_k\| = \min_{i \in [n]} \|y - x_i\|$. Now put $g_n(y) = x_{k(n,y)}$, and observe that, by construction and since $(x_n) \subset B_0$ is dense, we have $\|g_n(y) - y\| \to 0$, for all $y \in B_0$. Hence if we define $f_n : S \to \{x_1, \cdots x_n\} \subset B$ by $f_n = g_n \circ f$ we have $\|f_n(s) - f(s)\| \to 0$, for all $s \in S$, so if we can demonstrate that f_n is simple for each $n \in \mathbb{N}$ we are done. To see that f_n is simple we need to show that $(f_n = x_k) \in \mathscr{A}$ for every $k \in [n]$. To this end observe that

$$(f_n = x_k) = \left(\|f - x_k\| = \min_{i \in [n]} \|f - x_i\| < \min_{i \in [k-1]} \|f - x_i\| \right)$$

and, since for any $i \geq 1$ we have $\|f - x_i\| = \sup_{j \geq 1} |\varphi_j(f - x_i)|$ and f is assumed weakly measurable, this yields that $s \mapsto \|f(s) - x_i\| : S \to \mathbb{R}$ is measurable. This implies that $(f_n = x_k) \in \mathscr{A}$ as desired. $\square$

Note that the prove also yields that if f takes its values in a closed subspace $X \subset B$, then f is strongly measurable as a function with values in B if and only if it is strongly measurable considered as a function with values in X. Also by a slight modification of the definition of the φ_n's we can choose f_n such that $\|f_n\| \le \|f\|$, for all $n \ge 1$, and it is easy to see that a pointwise limit of strongly measurable functions is again strongly measurable. With Proposition 1.32 and Theorem 1.37 at hand we are now able to connect strong measurability to Borel measurability:

Corollary 1.38. *Let $f : S \to B$ be a function. Then f is strongly measurable if and only if it is separably valued and Borel measurable.*

Proof. If f is strongly measurable then by Theorem 1.37 f is weakly measurable and there is a closed, separable subspace $B_0 \subset B$ such that $f(S) \subset B_0$. By referring to Hahn-Banach again we obtain that f is weakly measurable considered as a B_0 valued function. By Proposition 1.32 we then infer that f is Borel measurable as a B_0 valued function. Now for $A \in \mathscr{B}(B)$ we have $A_0 := A \cap B_0 \in \mathscr{B}(B_0)$ and hence

$$(f \in A) = (f \in A_0) \in \mathscr{A}$$

demonstrating that f is Borel measurable considered as a B valued function. Now if we assume f to be separably valued and Borel measurable, we have in particular that $\varphi \circ f$ is measurable for any $\varphi \in B^*$. Hence f is weakly measurable and separably valued and therefore stongly measurable by Theorem 1.37. □

If B is separable (e.g. if it is equal to $\mathbb{C}$ or $\mathbb{R}$), the above corollary gives the characterization of Borel measurability that we know and love from scalar valued integration theory, but as the above examples show we have to make it an assumption in the general case. This is what we were alluding to, when we wrote in the beginning of this subsection that the Borel σ-algebra is too large to be useful in the general setting. Having introduced the relevant notions of measurability we now move to generalizing integration theory to vector valued functions.

Definition 1.39. Let $(S, \mathscr{A}, \mu)$ be a measure space. A μ-simple function $f : S \to B$ is a function of the form $f = \sum_{i=1}^n 1_{A_i} \otimes x_i$, where $n \ge 1$, $x_i \in B$ and $A_i \in \mathscr{A}$ with $\mu(A_i) < \infty$. A function is called strongly μ-measurable if it is the pointwise limit of a sequence of μ-simple functions μ-almost everywhere.

Proposition 1.40. *Let $f : S \to B$. Then*

1. *If f is strongly μ-measurable then it is μ-almost everywhere equal to a strongly measurable function.*

2. *If f is almost everywhere equal to a strongly measurable function and $(S, \mathscr{A}, \mu)$ is σ-finite then f is strongly μ-measurable.*

Proof. (1): Suppose f is strongly μ-measurable. Let f_n be μ-simple and $N \in \mathscr{A}$ be a null set such that $f_n \to f$ pointwise on N^c. Then $1_{N^c} f_n$ is simple for each $n \in \mathbb{N}$ and $1_{A^c} f_n \to 1_{N^c} f$ pointwise on S, so $1_{N^c} f$ is strongly measurable and clearly equals f μ-almost everywhere.
(2): Suppose $(S, \mathscr{A}, \mu)$ is σ-finite and f is μ-almost everywhere equal to a strongly measurable function $\tilde{f}$. Let $N \in \mathscr{A}$ be a null set such that $f = \tilde{f}$ an A^c. Now take a sequence $A_1 \subset A_2 \subset \cdots$ of measurable sets of finite measure such that $S = \cup_{n \ge 1} A_n$. Moreover let $\tilde{f}_n$ be a sequence of simple functions converging pointwise to $\tilde{(f)}$. Then $f_n := 1_{A_n} \tilde{f}_n$ is μ-simple for every $n \in \mathbb{N}$ and $f_n(s) \to f(s)$ whenever $s \in N^c$, so f is strongly μ-measurable. □

Remark 1.41. Considering any constant function $S \to B$ it is clear that σ-finiteness is really a necessary assumption in item (2) above. In particular, also for scalar-valued functions, being strongly μ-measurable is not the same as being Borel/strongly measurable unless we assume σ-finiteness for μ.

Definition 1.42. A function $f : S \to B$ is called weakly μ-measurable if $\varphi \circ f$ is strongly μ-measurable for every $\varphi \in B^*$.

Theorem 1.43. *A function $f : S \to B$ is called μ-essentially separably valued if there exists a closed, separable subset $B_0 \subset B$ such that $f(s) \in B_0$ for μ-almost all $s \in S$. Then f is strongly μ-measurable if and only if f is μ-essentially separably valued and weakly μ-measurable.*

Proof. The 'only if' part is proven similarly to the analogous statement in Theorem 1.37. For the other implication let B_0 be a closed, separable subspace of B in which f takes its values almost surely, and let $(\varphi_k)_{k\geq 1} \subset B^*$ be a norming sequence for B_0. Then $\varphi_k \circ f : S \to \mathbb{K}$ is strongly μ-measurable for every $k \in \mathbb{N}$. Let $(f_n^{(k)})_{n\geq 1}$ be a μ-almost surely approximating sequence of μ-simple functions. Then $\mu(f_n^{(k)} \neq 0) < \infty$, so if we put $S_{1,k} = \cup_{n\geq 1}(f_n^{(k)} \neq 0)$ and $S_{0,k} = S_{1,k}^c$ then $\varphi_k \circ f \equiv 0$ almost surely on $S_{0,k}$, μ is σ-finite on $S_{1,k}$ and S is the disjoint union of $S_{0,k}$ and $S_{1,k}$. Hence if we put $S_0 = \cap_{k\geq 1} S_{0,k}$ and $S_1 = S_0^c$ we have, since (φ_k) is norming for B_0, that $f \equiv 0$ on S_0 and μ is σ-finite on S_1. So if we can approximate f almost surely by μ-simple functions on S_1 we are done. This shows that we can assume that μ is σ-finite on S. Hence if we let $(x_j)_{j\geq 1} \subset B_0$ be a dense subset then the constant function $1_S \otimes x_j$ is strongly μ-measurable for every $j \in \mathbb{N}$, so by the 'only if' part of the theorem it is weakly μ-measurable. This implies that $\varphi_k \circ (f - x_j) : S \to \mathbb{K}$ is strongly μ-measurable for every $j, k \in \mathbb{N}$. By item (1) of Proposition 1.40 we can now find an array of stongly measurable functions $g_{jk} : S \to \mathbb{K}$ and a μ-null set such that $g_{jk} = \varphi_k(f - x_j)$ on N^c. Hence, by passing to the g_{jk}-sequence, we may assume that $\varphi_k(f - x_j)$ is strongly measurable. Since $(\varphi_k)_{k\geq 1}$ is norming for B_0 and $(x_j)_{j\geq 1}$ is dense in B_0 we now find ourselves in the same situation as in the proof of (2) $\implies$ (1) in Theorem 1.37, so we construct a sequence of simple functions $(f_n)_{n\geq 1}$ in the same way such that $f_n \to f$ pointwise, yielding that f is stongly measurable. Finally, since we reduced to the case where μ is σ-finite, we use item (2) of Proposition 1.40 to conclude that f is strongly μ-measurable as desired. $\square$

Let f be a μ-simple functions, and write $f = \sum_{i=1}^n 1_{A_i} \otimes x_i$, where $\mu(A_i) < \infty$. We define its *Bochner integral* in the obvious way:

$$\int_S f d\mu = \sum_{i=1}^n \mu(A_i) x_i \in B.$$

Note that $\int_S f d\mu$ is independent of the specific choice of A_i and x_i to represent f, and the well known formulae $\left\| \int_S f d\mu \right\| \leq \int_S \|f\| \, d\mu$ and $\int_S f d\mu + \int_S g d\mu = \int_S f + g d\mu$ hold.
If $f : S \to B$ is strongly μ-measurable and g is a μ-simple function then g is Borel measurable and f is almost everywhere equal to a strongly, hence Borel, measurable function so that $\|f - g\| \in \mathcal{M}^+(S)$. Hence the quantity $\int_S \|f - g\| \, d\mu$ makes sense in the following definition:

Definition 1.44. A strongly μ-measurable function $f : S \to B$ is called *Bochner integrable* with respect to μ if there exists a sequence of μ-simple functions $g_n : S \to B$ such that

$$\lim_{n\to\infty} \int_S \|f - g_n\| \, d\mu = 0.$$

Using the definition of Bochner integrals of μ-simple functions, it is easily seen that if f and (g_n) satisfy the above definition then $(\int_S g_n d\mu)$ is a Cauchy sequence in B. Hence, it converges to an element in B, which we denote by $\int_S f d\mu$. This notation suggests that the choice of sequence, (g_n), in the above definition does not matter, which is readily seen to be the case using the triangle inequality. An important characterization of the Bochner integrable functions is the following:

Proposition 1.45. *A strongly μ-measurable function $f : S \to B$ is Bochner integrable if and only if $\int_S \|f\| \, d\mu < \infty$.*

Proof. First assume $\int_S \|f\| \, d\mu < \infty$ and take a sequence of μ-simple functions (f_n) such that $f_n \to f$ and $\|f_n\| \leq \|f\|$ almost everywhere. It then follows by dominated convergence that $\int_S \|f_n - f\| \, d\mu \to 0$. On the other hand, if f is assumed Bochner integrable we get by observing that $\|f\| \leq \|f - f_n\| + \|f_n\|$ that $\int_S \|f\| \, d\mu < \infty$. □

A number of classical results from scalar valued integration theory carry over to Bochner integrals *mutatis mutandis*, These count dominated convergence, substitution, Jensen's inequality (whenever μ is a probability measure) and Fubini's theorem (see [Hyt+16, sec. 1.2]). Also, by first considering μ-simple functions and then an approximation argument, we get a triangle inequality for Bochner integrals: $\left\| \int_S f d\mu \right\| \leq \int_S \|f\| \, d\mu$. Finally the Bochner integral is linear and more generally:

Theorem 1.46. *Let $f : S \to B$ and $T : V \to Y$ a linear map, where $V \subset B$ is a closed subspace and Y is a Banach space.*

- *If f is Bochner integrable, $V = B$ and T is bounded then $Tf : S \to Y$ is Bochner integrable and $\int_S Tf d\mu = T\left(\int_S f d\mu\right)$. In particular, for all $\varphi \in B^*$, we have $\int_S \varphi f d\mu = \varphi\left(\int_S f d\mu\right)$.*
- *If f is Bochner integrable, $f \in V$ μ-almost surely, T is closed and $Tf : S \to Y$ (which is almost everywhere defined) is Bochner integrable, then f is Bochner integrable as a V valued function and $\int_S Tf d\mu = T\left(\int_S f d\mu\right)$.*

The proof of the first item (which is the one we will need) follows directly from the definition of the Bochner integral. We may now define the *Bochner spaces* $L_p(S; B)$ by

- For $1 \leq p < \infty$, a function $f : S \to B$ is in $L_p(S; B)$ if and only if f is strongly μ-measurable and $\int_S \|f\|^p \, d\mu < \infty$.
- A function $f : S \to B$ is in $L_\infty(S; B)$ if and only if f is strongly μ-measurable and there is $r \geq 0$ such that $\mu(\|f\| > r) = 0$.

where we of course identify functions that are almost surely equal. We can equip these spaces with the norms $\|f\|_p = \left(\int_S \|f\|^p\right)^{1/p}$ for $1 \leq p < \infty$, and $\|f\|_\infty = \inf\{r \geq 0 \mid \mu(\|f\| > r) = 0\}$ (we shall occasionally write $\|f\|_{L_p(S;B)}$ if there is any risk of ambiguity). The proofs of the facts that these are norms and that $(L_p(S; B), \|\cdot\|_p)$ is a Banach space for every $1 \leq p \leq \infty$, are analogous to the ones for the scalar valued case.

Proposition 1.47. *Let $(S_1, \mathscr{A}_1, \mu_1)$ and $(S_2, \mathscr{A}_2, \mu_2)$ be measure spaces and B a Banach space. For all $1 \leq p \leq q < \infty$, the map*

$$1_A \otimes (1_B \otimes x) \mapsto 1_B \otimes (1_A \otimes x)$$

extends uniquely to a contractive embedding $L_p(S_1; L_q(S_2; B)) \to L_q(S_2; L_p(S_1; B))$.

The proof of Proposition 1.47 essentially follows from the triangle inequality for Bochner integrals and the fact that the μ-simple functions are dense in the Bochner spaces. We will omit it. Simplifying to the case $p = q$ we obtain the following version of Fubini's theorem:

Corollary 1.48. *Let $(S_1, \mathscr{A}_1, \mu_1)$ and $(S_2, \mathscr{A}_2, \mu_2)$ be measure spaces and B a Banach space. For all $1 \leq p < \infty$, the map*

$$1_A \otimes (1_B \otimes x) \mapsto 1_B \otimes (1_A \otimes x)$$

extends uniquely to an isometric isomorphism $L_p(S_1; L_p(S_2; B)) \cong L_p(S_2; L_p(S_1; B))$.
If, moreover, we assume that μ and ν are σ-finite then the map

$$1_A \otimes (1_B \otimes x) \mapsto (1_A \otimes 1_B) \otimes x$$

extends uniquely to an isometric isomorphism $L_p(S_1; L_p(S_2; B)) \cong L_p(S_1 \times S_2; B)$.

Example 1.49. Consider the function $f : S \to B$ given by $f = f' \otimes x$ for some non-zero $x \in B$ and measurable $f' : S \to \mathbb{K}$. Then f is Bochner integrable if and only if f' is integrable and in this case $\int_S f d\mu = \int_S f' d\mu \cdot x$. In particular $f \in L_p(S; B)$ if and only if $f' \in L_p(S)$ and $\|f\|_{L_p(S;B)} = \|f'\|_{L_p(S)} \|x\|$.

The above example motivates the notation $f \otimes x$ for scalar valued functions f, and shows that $\|\cdot\|_{L_p(S;B)}$ is a cross-norm on the algebraic tensor product $L_p(S) \otimes B \subset L_p(S; B)$. It is not difficult to see, using what we know from scalar valued integration theory and (the proof of) Proposition 1.45, that the μ-simple functions are dense in $L_p(S; B)$ for $1 \leq p < \infty$, so in particular $L_p(S) \otimes B$ is dense in $L_p(S; B)$. This is an important fact for formulating *the extension problem*: Let $(S_1, \mathscr{A}_1, \mu_1)$ and $(S_2, \mathscr{A}_2, \mu_2)$ be measure spaces, $1 \leq p_1, p_2 \leq \infty$ and $T : L_{p_1}(S_1) \to L_{p_2}(S_2)$ a bounded linear operator. We may then define the linear operator $T \otimes I_B : L_{p_1}(S_1) \otimes B \to L_{p_2}(S_2) \otimes B$ by $(T \otimes I_B)(f \otimes x) = Tf \otimes x$. If this operator is bounded, it extends uniquely to an operator from $L_{p_1}(S_1; B)$ to $L_{p_2}(S_2; B)$ with norm $\|T \otimes I_B\|$. We will denote this operator by T_B. The question is, which assumptions on T (and/or on p and q) guarantee boundedness, and hence extendability, of $T \otimes I_B$. For positive operators, i.e., $Tf \geq 0$ whenever $f \geq 0$, we have the following nice result:

Proposition 1.50. *Suppose $1 \leq p_1, p_2 < \infty$ and $T : L_{p_1}(S_1) \to L_{p_2}(S_2)$ is a positive, bounded linear operator. Then $T \otimes I_B : L_{p_1}(S_1) \otimes B \to L_{p_2}(S_2) \otimes B$ is bounded for every Banach space B and hence extends to a bounded linear operator $T_B : L_{p_1}(S_1; B) \to L_{p_2}(S_2; B)$. Moreover, this extention satisfies $\|T_B\| = \|T \otimes I_B\| = \|T\|$.*

Proof. It suffices to show boundedness for μ_1-simple functions, so let $f = \sum_{i=1}^n 1_{A_i} \otimes x_i$ where $A_i \in \mathscr{A}_1$ are pairwise disjoint with $\mu_1(A_i) < \infty$ and $x_i \in B$. Then

$$\begin{aligned}
\|(T \otimes I_B)f\|_{L_{p_2}(S_2;B)} &= \left(\int_{S_2} \left\|\sum_{i=1}^n T1_{A_i} \otimes x_i\right\|_B^{p_2} d\mu_2\right)^{1/p_2} \\
&\leq \left(\int_{S_2} \left(\sum_{i=1}^n T1_{A_i} \|x_i\|_B\right)^{p_2} d\mu_2\right)^{1/p_2} \\
&= \left(\int_{S_2} \left(T\left(\sum_{i=1}^n \|x_i\|_B 1_{A_i}\right)\right)^{p_2} d\mu_2\right)^{1/p_2} \\
&= \left\|T\left(\sum_{i=1}^n \|x_i\|_B 1_{A_i}\right)\right\|_{L_{p_2}(S_2)} \\
&\leq \|T\| \left\|\sum_{i=1}^n \|x_i\|_B 1_{A_i}\right\|_{L_{p_1}(S_1)} = \|T\| \left\|\sum_{i=1}^n 1_{A_i} \otimes x_i\right\|_{L_{p_1}(S_1;B)},
\end{aligned}$$

where we used positivity of T to conclude that $\|T1_{A_i} \otimes x_i\|_B = |T1_{A_i}| \, \|x_i\|_B = T1_{A_i} \|x_i\|$. This demonstrates that $T \otimes I_B$ is bounded with $\|T \otimes I_B\| \leq \|T\|$. Since the reverse inequality is trivial this concludes the proof. □

Remark 1.51. If, in stead of positivity of T, we make the weaker assumption that there exists a positive $R : L_{p_1}(S_1) \to L_{p_2}(S_2)$ such that $|Tf| \leq R|f|$, for all $f \in L_{p_1}(S_1)$. Then the above proof carries over, with the estimate $\|T1_{A_i} \otimes x_i\|_B = |T1_{A_i}| \, \|x_i\|_B \leq R1_{A_i} \|x_i\|_B$, to demonstrate extendability of $T \otimes I_B$ with $\|T \otimes I_B\| \leq \|R\|$.

For Hilbert spaces we need no assumptions on the operator (we simplify to $p_1 = p_2$, since this is the version we will need, but the theorem holds also without this assumption):

Theorem 1.52. *Suppose $1 \leq p < \infty$ and let $T : L_p(S_1) \to L_p(S_2)$ be a bounded linear operator. Then $T \otimes I_H$ extends uniquely to a bounded linear operator $T_H : L_p(S_1; H) \to L_p(S_2; H)$, for every Hilbert space H. Its norm satisfies $\|T_H\| = \|T\|$.*

This theorem was originally proven by Paley, Marcinkiewicz and Zygmund, but we follow [Hyt+16]. Its proof requires the following proposition:

Proposition 1.53. *Let (Ω, μ) be a probability space and let $(\gamma_n)_{n\geq 1}$ be any sequence of mutually independent standard Gaussian random variables on Ω. Then ℓ_2 is isometrically isomorphic to $\overline{\operatorname{span}}\,\{\gamma_n \mid n \in \mathbb{N}\} \subset L_p(\Omega)$, for every $1 \leq p < \infty$.*

Proof. Take $\sum_{k=1}^n a_n\gamma_n \in \operatorname{span}\{\gamma_n \mid n \in \mathbb{N}\}$. Then by elementary properties of Gaussian random variables this is again a centered Gaussian with variance $\|(a_k)_{k=1}^n\|_{\ell_2}$. Hence

$$\left\|\sum_{k=1}^n a_n\gamma_n\right\|_{L_p(\Omega)} = \|\gamma\|_p \, \|(a_k)_{k=1}^n\|_{\ell_2}, \tag{3}$$

where $\|\gamma\|_p$ denotes the absolute pth moment of a standard Gaussian. Hence, by a density argument, the map

$$(a_n)_{n\geq 1} \mapsto \frac{1}{\|\gamma\|_p} \sum_{n=1}^{\infty} a_n \gamma_n$$

establishes an isometric isomorphism between ℓ_2 and $\overline{\operatorname{span}}\,\{\gamma_n \mid n \in \mathbb{N}\}$. □

Proof of Theorem 1.52. First suppose that H is separable. Then it is isometrically isomorphic to ℓ_2, so we may identify it with the subspace in $L_2(\Omega)$ from Proposition 1.53, which we denote by G. For example, we can take $\Omega = (0,1)$ with the Lebesgue measure; this probability space certainly supports a sequence of independent Gaussians. Take $f \in L_p(S_1; G)$. Then, as $(\gamma_n)_{n\geq 1}$ is an orthonormal basis for G, we can write $f(s) = \sum_{n=1}^{\infty} f_n(s)\gamma_n$, for every $s \in S_1$, where $f_n(s) = \langle f(s), \gamma_n \rangle = \int_\Omega f(s)\overline{\gamma_n} d\mu$. Then, since $\|f_n\|_{L_p(S_1)}^p \leq \|\gamma_n\|_2^p \|f\|_{L_p(S_1;G)} < \infty$, we have that $f_n \in L_p(S_1)$. This demonstrates that $L_p(S_1) \otimes \operatorname{span}\{\gamma_n \mid n \in \mathbb{N}\}$ is dense in $L_p(S_1; G)$, so it suffices to prove boundedness on elements of $L_p(S_1) \otimes \operatorname{span}\{\gamma_n \mid n \in \mathbb{N}\}$. To this end, take $f = \sum_{n=1}^{N} f_n \gamma_n \in L_p(S_1) \otimes \operatorname{span}\{\gamma_n \mid n \in \mathbb{N}\}$. By formula (3) and Corollary 1.48, we obtain for such f that

$$\begin{aligned}
\|(T \otimes I_G) f\|_{L_p(S_2;G)} &= \left\| \sum_{n=1}^{N} T f_n \gamma_n \right\|_{L_p(S_2;G)} = \frac{1}{\|\gamma\|_p} \left\| \sum_{n=1}^{N} T f_n \gamma_n \right\|_{L_p(S_2;L_p(\Omega))} \\
= \frac{1}{\|\gamma\|_p} \left\| \sum_{n=1}^{N} T f_n \gamma_n \right\|_{L_p(\Omega;L_p(S_2))} &\leq \frac{\|T\|}{\|\gamma\|_p} \left\| \sum_{n=1}^{N} f_n \gamma_n \right\|_{L_p(\Omega;L_p(S_1))} = \frac{\|T\|}{\|\gamma\|_p} \left\| \sum_{n=1}^{N} f_n \gamma_n \right\|_{L_p(S_1;L_p(\Omega))} \\
&= \frac{\|T\| \, \|\gamma\|_p}{\|\gamma\|_p} \left\| \sum_{n=1}^{N} f_n \gamma_n \right\|_{L_p(S_1;G)} = \|T\| \, \|f\|_{L_p(S_1;G)} .
\end{aligned}$$

This demonstrates that $\|T \otimes I_G\| = \|T\|$, since the other inequality is trivial. Hence $T \otimes I_G$ extends uniquely to $T_G : L_p(S_1; G) \to L_p(S_2; G)$ with $\|T_G\| = \|T\|$.
Now in the general case, suppose $T \otimes I_H$ is not bounded. Then we can find a sequence of norm one functions $f_n \in L_p(S_1) \otimes H$, $n \in \mathbb{N}$, such that $\|(T \otimes I_H) f_n\|_{L_p(S_2;H)} \to \infty$. But since each f_n takes values in a finite-dimensional subspace of H, all the f_n take values in a separable, closed subspace $K \subset H$. Then $T \otimes I_K$ is bounded with norm $\|T\|$ and hence

$$\|(T \otimes I_H) f_n\|_{L_p(S_2;H)} = \|(T \otimes I_K) f_n\|_{L_p(S_2;H)} \leq \|T\| ,$$

a contradiction, so $T \otimes I_H$ must be bounded. A reiteration of this argument shows that we also must have $\|T \otimes I_H\| \leq \|T\|$. This concludes the proof. □

1.5 Martingales

We will now move on to introduce the notion of *martingales*, which is an important tool in Banach space theory and will play a crucial role in parts of this thesis. We will write $L_p(S, \mathscr{A}; B)$ if we need to emphasize the σ-algebra, $\mathscr{A}$, with respect to which we are considering the Bochner space. Note that if $\mathscr{F}$ is a sub σ-algebra of $\mathscr{A}$ then $L_p(S, \mathscr{F}; B)$ is a closed subspace of $L_p(S, \mathscr{A}; B)$.

Definition 1.54. Let $(S, \mathscr{A}, \mu)$ be a probability space and B a Banach space. Suppose $\mathscr{F} \subset \mathscr{A}$ is a sub σ-algebra and $f \in L_1(S, \mathscr{A}; B)$. A function $g \in L_1(S, \mathscr{F}; B)$ is called a *conditional expectation* of f given $\mathscr{F}$ if

$$\int_F f d\mu = \int_F g d\mu, \tag{4}$$

for all $F \in \mathscr{F}$.

For the remainder of this subsection it will be a standing assumption that $(S, \mathscr{A}, \mu)$ is a probability space and $\mathscr{F} \subset \mathscr{A}$ a sub σ-algebra. We will prove that conditional expectations exist and are almost surely unique, but first an example:

Example 1.55. Suppose $\mathscr{F} = \sigma(F_1, \cdots, F_n)$ for mutually disjoint $F_1, \cdots, F_n \in \mathscr{A}$ forming a partition of S and such that $\mu(F_i) > 0$. Then for $f \in L_1(S; B)$ the conditional expectation of f given $\mathscr{F}$ exists and is given by

$$\sum_{i=1}^{n} 1_{F_i} \frac{\int_{F_i} f d\mu}{\mu(F_i)},$$

as can be seen by direct computation, using that any $F \in \mathscr{F}$ is a (possibly empty) union of a subsequence of $(F_i)_{i=1}^n$.

Lemma 1.56. *Let $f \in L_1(S)$ and assume that $g, g' \in L_1(S, \mathscr{F})$ are conditional expectations of f given $\mathscr{F} \subset \mathscr{A}$. Then $g = g'$ μ-almost surely.*

Proof. First assume $\mathbb{K} = \mathbb{R}$. Then since $(g > g'), (g' > g) \in \mathscr{F}$ by assumption, we have by (4) that

$$\int_{(g \neq g')} |g - g'| d\mu = \int_{(g > g')} g - g' d\mu + \int_{(g' > g)} g' - g d\mu$$
$$= \int_{(g > g')} f d\mu - \int_{(g > g')} f d\mu + \int_{(g' > g)} f d\mu - \int_{(g' > g)} f d\mu = 0,$$

which, as above, implies $\mu(g \neq g') = 0$.
The case $\mathbb{K} = \mathbb{C}$ now follows by observing that if g is a conditional expectation of f then $\Re g$ and $\Im g$, are conditional expectations of $\Re f$ and $\Im f$ respectively. Indeed, $\Re g = \frac{g + \overline{g}}{2}$ is $\mathscr{F}$-measurable and for $F \in \mathscr{F}$

$$\int_F \Re g d\mu = \frac{\int_F g d\mu + \int_F \overline{g} d\mu}{2} = \frac{\int_F f d\mu + \overline{\int_F f d\mu}}{2} = \int_F \Re f d\mu$$

and similarly with the imaginary parts. □

Theorem 1.57. *Suppose $f \in L_1(S; B)$ and that $g, g' \in L_1(S, \mathscr{F}; B)$ are conditional expectations of f given $\mathscr{F}$. Then $g = g'$ μ-almost surely.*

Proof. It follows directly from Theorem 1.46 that $\varphi \circ g$ and $\varphi \circ g'$ are conditional expectations of $\varphi \circ f$ for every $\varphi \in B^*$ It follows from Lemma 1.56 that $\varphi \circ g = \varphi \circ g'$ almost surely for every $\varphi \in B^*$. The Theorem now follows from the following claim:

Claim 1. Suppose $g, g' : S \to B$ are strongly μ-measurable functions such that $\varphi \circ g = \varphi \circ g'$ μ-almost surely for every $\varphi \in B^*$. Then $g = g'$ μ-almost surely.

Proof of claim. Since g and g' are μ-essentially separably valued, we can find a separable, closed subspace $B_0 \subset B$ and a null set $N \in \mathscr{A}$ such that $g(s), g'(s) \in B_0$, for all $s \in N^c$. Take a sequence $(\varphi_n)_{n\geq 1} \subset B^*$ which is norming for B_0. Then for each $n \geq 1$ there is a null set $N_n \in \mathscr{A}$ such that $\varphi_n(g(s) - g'(s)) = 0$, for all $s \in N_n^c$. Hence for each s outside the null set $N \cup \bigcup_{n\geq 1} N_n$ we have

$$\|g(s) - g'(s)\| = \sup_{n\geq 1} |\varphi_n(g(s) - g'(s))| = 0.$$

This proves the claim. $\square$

Theorem 1.57 allows us to introduce the notation $E(f \mid \mathscr{F})$ for the conditional expectation of f given $\mathscr{F}$, whenever it exists. We now turn to a few basic properties of conditional expectations after which we will adress existence.

Proposition 1.58. *Suppose $\mathbb{K} = \mathbb{R}$ and that $g_1, g_2 \in L_1(S, \mathscr{F})$ are conditional expectations of $f_1, f_2 \in L_1(S)$ respectively. If $f_1 \leq f_2$ almost surely then $g_1 \leq g_2$ almost surely.*

Proof. Since $(g_1 > g_2) \in \mathscr{F}$ we have by (4) that

$$\int_{(g_1>g_2)} g_2 - g_1 d\mu = \int_{(g_1>g_2)} f_2 d\mu - \int_{(g_1>g_2)} f_1 d\mu \geq 0,$$

by assumption of f_1 and f_2. Since obvisouly $\int_{(g_1>g_2)} g_2 - g_1 d\mu \leq 0$ we have that this integral is in fact equal to zero. Since obviously $g_2 - g_1 < 0$ on $(g_1 > g_2)$ we must have $\mu(g_1 > g_2) = 0$ $\square$

The next result is often referred to as a 'conditional triangle inequality'.

Proposition 1.59. *Let $f \in L_1(S; B)$ and suppose that there exist conditional expectations of f and $\|f\|$ given $\mathscr{F}$. Then*

$$\|E(f \mid \mathscr{F})\| \leq E(\|f\| \mid \mathscr{F}) \;\; a.s.$$

Proof. Since any complex Banach space is also a real Banach space we may assume that B is real. Furthermore, since f is μ-essentially separably valued we may also assume that B is separable (a conditional expectation of f as a function with values in a closed subspace $B_0 \subset B$ is also a conditional expectation of f considered as a functions with values in B). Take a norming sequence $(\varphi_n)_{n\geq 1} \subset B^*$. Then it is easy to see, using Theorem 1.46, that $\varphi_n(E(f \mid \mathscr{F}))$ is a conditional expectation of $\varphi_n(f)$ given $\mathscr{F}$. Moreover, since $\varphi_n(f) \leq \|f\|$, Proposition 1.58 yields that $\varphi_n(E(f \mid \mathscr{F})) \leq E(\|f\| \mid \mathscr{F})$. Hence

$$\|E(f \mid \mathscr{F})\| = \sup_{n\geq 1} \varphi_n(E(f \mid \mathscr{F})) \leq E(\|f\| \mid \mathscr{F}) \;\; a.s.,$$

which demonstrates the desired estimate. $\square$

Lemma 1.60. *Suppose $f \in L_2(S)$. Then $E(f \mid \mathscr{F})$ exists and $E(f \mid \mathscr{F}) \in L_2(S, \mathscr{F})$.*

Proof. Let $P : L_2(S) \to L_2(S, \mathscr{F})$ denote the orthogonal projection. Then since orthogonal projections are self-adjoint we get for any $F \in \mathscr{F}$ that

$$\int_F Pf d\mu = \int 1_F Pf d\mu = \int (P1_F) f d\mu = \int 1_F f d\mu = \int_F f d\mu,$$

where we used that $1_F \in L_2(S, \mathscr{F})$. □

Lemma 1.61. *Let $f \in L_1(S)$. Then $E(f \mid \mathscr{F})$ exists and satisfies $\|E(f \mid \mathscr{F})\|_1 \leq \|f\|_1$.*

Proof. First assume that $f \in L_1(S) \cap L_\infty(S)$. Then by Hölder $f, |f| \in L_2(S)$, so their conditional expectations exist. Since obviously $E(\alpha f + \beta g \mid \mathscr{F}) = \alpha E(f \mid \mathscr{F}) + \beta E(g \mid \mathscr{F})$ for any $f, g \in L_1(S) \cap L_\infty(S)$ and $\alpha, \beta \in \mathbb{K}$, we see that $E(\cdot \mid \mathscr{F}) : L_1(S) \cap L_\infty(S) \to L_1(S, \mathscr{F})$ is a linear operator. By Proposition 1.59

$$\|E(f \mid \mathscr{F})\|_1 = \int_S |E(f \mid \mathscr{F})| d\mu \leq \int_S E(|f| \mid \mathscr{F}) d\mu = \int_S |f| d\mu = \|f\|_1 ,$$

so this operator is contractive. Since $L_1(S) \cap L_\infty(S)$ is dense in $L_1(S)$ (e.g., because it contains the simple functions) it extends uniquely to a contractive operator $L_1(S) \to L_1(S, \mathscr{F})$. Denote this operator by G. It remains to see that Gf is a conditional expectation of f in the case where $f \in L_1(S) \setminus L_\infty(S)$. To this end set $f_n = 1_{(|f| \leq n)} f$. Then $f_n \in L_1(S) \cap L_\infty(S)$. Moreover, $f_n \to f$ pointwise and hence, by dominated convergence, $f_n \to f$ in $L_1(S)$. Hence for any $F \in \mathscr{F}$ we have

$$\int_F Gf d\mu = \lim_{n\to\infty} \int_F Gf_n d\mu = \lim_{n\to\infty} \int_F f_n d\mu = \int_F f d\mu,$$

demonstrating that Gf is a conditional expectation of f. □

By Lemma 1.61 and Theorem 1.57 we now have a unique contractive linear operator $E(\cdot \mid \mathscr{F}) : L_1(S) \to L_1(S, \mathscr{F})$ sending f to its conditional expectation given $\mathscr{F}$.

Theorem 1.62. *Suppose $f \in L_1(S; B)$. Then f has a unique conditional expectation given $\mathscr{F}$ and $\|E(f \mid \mathscr{F})\|_1 \leq \|f\|_1$.*

Proof. By Propositions 1.59 and 1.58 we get that $E(\cdot \mid \mathscr{F}) : L_1(S) \to L_1(S, \mathscr{F})$ is contractive and positive. Hence by Proposition 1.50 the contraction

$$E(\cdot \mid \mathscr{F}) \otimes I_B : L_1(S) \otimes B \to L_1(S, \mathscr{F}) \otimes B$$

extends to a unique contraction $G : L_1(S; B) \to L_1(S, \mathscr{F}; B)$. It follows from example 1.49 that $E(\cdot \mid \mathscr{F}) \otimes I_B(f \otimes x) = E(f \mid \mathscr{F}) \otimes x$ is the unique conditional expectation of $f \otimes x$ whenever $f \in L_1(S)$ and $x \in B$, i.e., $E(f \otimes x \mid \mathscr{F}) = E(f \mid \mathscr{F}) \otimes x$. Hence we can use an approximation argument similar to that from the proof of Lemma 1.61 to obtain that Gf is a conditional expectation of any $f \in L_1(S; B)$. □

Remark 1.63. The notion of conditional expectations can be defined, more generally, for μ-measurable functions on arbitrary measure spaces, and existence and uniqueness can be proven under weaker assumptions than integrability. We will only need the present level of generality in this exposition but a more detailed description, including limit theorems and more properties, can be found in [Hyt+16, Sec. 2.6]

We are now finally ready to define martingales:

Definition 1.64. Let $(S, \mathscr{A}, \mu)$ be a probability space, and suppose $(\mathscr{F}_n)_{n\geq 1}$ is a sequence of sub σ-algebras of $\mathscr{A}$ such that $\mathscr{F}_n \subset \mathscr{F}_{n+1}$ for every $n \in \mathbb{N}$ (such a sequence is referred to as a *filtration*). A sequence of functions $(f_n)_{n\geq 1} \subset L_1(S; B)$ is called a *martingale* if f_n is strongly μ-measurable with respect to $\mathscr{F}_n$ (we say that $(f_n)_{n\geq 1}$ is *adapted* to $(\mathscr{F}_n)_{n\geq 1}$), and $E(f_{n+1} \mid \mathscr{F}_n) = f_n$ for every $n \in \mathbb{N}$.

Martingales are widely studied in probability theory and analysis, as they satisfy many intersting inequalities and convergence theorems. In probability theory the following piece of intuition is often given: The conditional expectation of a random variable given a σ-algebra can be seen as a 'best guess' on the outcome of that variable, given the knowledge contained in the σ-algebra. Example 1.55 substantiates this line of thought. A martingale is then a random process, where the best guess on the outcome of the next variable based on the current knowledge is simply the outcome of the current variable.

This subsection about martingales as well as the preceding one about vector valued integration are based on [Hyt+16].

2 Poincaré inequalities for expanders

In this section we shall obtain p-Poincaré inequalities for functions on connected graphs taking values in different normed spaces with constants depending on either the Cheeger constant or the spectral gap of the graph. These will yield, by Proposition 1.18, that sequences of (edge or spectral) expanders do not coarsely embed into these normed spaces.

2.1 Poincaré inequalities connected to the Cheeger constant

For a connected, undirected graph G with the path metric and $Y = L_1(0,1)$ we have the following p-Poincaré inequality with respect to the adjacency matrix:

Theorem 2.1. *If $G = (V,E)$ has cheeger constant h then for any $f : V \to L_1(0,1)$ we have*

$$\sum_{u,v\in V} a_{u,v} \left\|f(u) - f(v)\right\|_1 \geq \sum_{u,v\in V} \frac{h}{2|V|} \left\|f(u) - f(v)\right\|_1 ,$$

where $(a_{u,v})_{u,v\in V}$ is the adjacency matrix of G.

The proof of Theorem 2.1 is found in [Ost13] and requires the following lemma:

Lemma 2.2. *For any function $f : V \to \mathbb{R}$ the following inequality holds*

$$\sum_{u\in V} |f(u) - M| \leq \sum_{(u,v)\in E} \frac{1}{h}|f(u) - f(v)|, \tag{5}$$

where M is a median of the set $\{f(u) \mid u \in V\}$.

Proof. First observe that we can assume $M = 0$ and, for notational ease, we also assume that the number of vertices is odd. Let $f_1 \leq f_2 \leq \cdots \leq f_k \leq 0 = f_{k+1} \leq f_{k+2} \leq \ldots f_{2k+1}$ be the function values of f. For $1 \leq i \leq k$, let $L_i^- = \{v \in V \mid f(v) < f_{i+1}\}$ and for $k+2 \leq i \leq 2k+1$ let $L_i^+ = \{v \in V \mid f(v) > f_{i-1}\}$. Then $|L_1^-| = 1 - \delta_{f_1,f_2}$ and, for $1 \leq i \leq k-1$ we have

$$|L_{i+1}^-| = |L_i^-| + (1 - \delta_{f_i,f_{i+1}})(i + 1 - \min\{j \leq i \mid f_i = f_j\}).$$

Also, $|L_{2k+1}^+| = 1 - \delta_{f_{2k},f_{2k+1}}$ and, for $k + 2 \leq i \leq 2k$, we have

$$|L_i^+| = |L_{i+1}^+| + (1 - \delta_{f_{i+1},f_i})(\max\{j \geq i \mid f_j = f_i\} - (i - 1)).$$

Using these equalities, we arrive at

$$\sum_{i=1}^{2k+1} |f_i| = \sum_{i=1}^{k} |L_i^-|(f_{i+1} - f_i) + \sum_{i=k+2}^{2k+1} |L_i^+|(f_i - f_{i-1}),$$

so L_i^- and L_i^+ just give a way of writing $\sum_{v\in V} |f(v) - M| = \sum_{i=1}^{2k+1} |f_i|$ as a sort of weighted telescopic sum. As they are all subsets of V of size at most $k \leq |V|/2$, we get by definition of the cheeger constant that $|L_i^-| \leq |\partial L_i^-|/h$ and $|L_i^+| \leq |\partial L_i^+|/h$, and hence

$$\sum_{v\in V} |f(v) - M| \leq \frac{1}{h}\left(\sum_{i=1}^{k} |\partial L_i^-|(f_{i+1} - f_i) + \sum_{i=k+2}^{2k+1} |\partial L_i^+|(f_i - f_{i-1})\right).$$

This sum is now indexed by the edges of G in stead of the vertices (since ∂L_i^- and ∂L_i^+ are subsets of the edge set) and to compute the contribution of some edge $(u,v) \in E$ assume without loss of generality that $f(u) \le f(v)$. First consider the case where there are $i_u < i_v \le k$ such that $u \in L_{i_u}^-$ and $v \in L_{i_v}^-$ and take both i_u and i_v to be minimal with this property. Then $f(u) = f_{i_u}$ and $f(v) = f_{i_v}$ and since this edge will make a contribution of $f_{i+1} - f_i$ if and only if $u \in L_i^-$ and $v \notin L_i^-$ for $1 \le i \le k$ (and it will not contribute for any $i \ge k+2$), we get a total contribution of $\sum_{j=i_u}^{i_v-1} f_{j+1} - f_j = f_{i_v} - f_{i_u} = |f(u) - f(v)|$. The other cases are checked similarly to obtain the same contribution (if, e.g., there is no i_v with the above property take $i_v \ge k+2$ maximal such that $v \in L_{i_v}^+$ and similarly with i_u). This now yields

$$\frac{1}{h}\left(\sum_{i=1}^{k} |\partial L_i^-|(f_{i+1} - f_i) + \sum_{i=k+2}^{2k+1} |\partial L_i^+|(f_i - f_{i-1})\right) = \frac{1}{h}\sum_{(u,v)\in E} |f(u) - f(v)|,$$

which concludes the proof. □

Remark 2.3. Before we proceed to the proof of Theorem 2.1 let us first observe that we could use the above lemma as an alternative definition of the Cheeger constant. Indeed, if a constant $h > 0$ satisfies the inequality from Lemma 2.2 we have for any non-empty $F \subset V$ of size at most $|V|/2$ that $h \le |\partial F|/|F|$. This can be seen by considering the function $1_F : V \to \mathbb{R}$. Since $|F| \le |V|/2$ this function has 0 as a median so inequality (5) reads $|F| \le |\partial F|/h$ as desired. Hence $h \le h(G)$ so we could define the Cheeger constant as the supremum over all h satisfying (5), for all functions $f : V \to \mathbb{R}$.

Proof of Theorem 2.1. Let $f : V \to L_1(0,1)$ and let f_u denote the L_1 function corresponding to $u \in V$. Furthermore, for each $t \in [0,1]$, let $M(t)$ be a median of $\{f_u(t) \mid u \in V\}$. Since the continuous functions are dense in $L_1(0,1)$ we can restrict ourselves to the case where f_u is continuous for each $u \in V$. In this case we can pick the medians such that $M(t)$ is a continuous function of t and hence $M \in L_1(0,1)$. By Lemma 2.2 we have

$$\sum_{u\in V} |f_u(t) - M(t)| \le \frac{1}{h}\sum_{(u,v)\in E} |f_u(t) - f_v(t)|,$$

for every $t \in [0,1]$. By integrating either side of this inequality, we obtain

$$\sum_{u\in V} h \left\| f_u - M \right\|_1 \le \sum_{(u,v)\in E} \left\| f_u - f_v \right\|_1 .$$

This yields

$$\begin{aligned}\sum_{u,v\in V} \frac{h}{2|V|} \left\| f_u - f_v \right\|_1 &\le \sum_{u,v\in V} \frac{h}{2|V|} \left(\left\| f_u - M \right\|_1 + \left\| f_v - M \right\|_1 \right) \\ &= \frac{h}{2}\left(\sum_{u\in V} \left\| f_u - M \right\|_1 + \sum_{u\in V} \left\| f_u - M \right\|_1 \right) \\ &\le \sum_{u,v\in V} a_{u,v} \left\| f_u - f_v \right\|_1 ,\end{aligned}$$

which is the desired estimate. □

Theorem 2.1 immediately yields, by Proposition 1.18, that a family of edge (and hence spectral) expanders does not coarsely embed into $L_1(0,1)$. Lemma 2.2 can also be used (with an analogous proof) to obtain similar Poincaré inequalities for $L_1(S)$ where $(S, \mathscr{A}, \mu)$ is an arbitrary measure space, in particular for ℓ_1. In [Mat97] Matoušek develops a way to extrapolate Poincaré inequalities to L_p-valued from the one obtained for L_1-valued functions (with constants depending on p). This is contained in the following theorem:

Theorem 2.4 (**Matoušek's extrapolation theorem**). *Let $G = (V, E)$ be a d-regular graph with Cheeger constant h. Then for any function $f : V \to L_p(S)$, where $(S, \mathscr{A}, \mu)$ is any measure space and $p > 1$, we have*

$$\sum_{u,v \in V} \left(\frac{c}{p}\right)^p \frac{1}{n} ||f(u) - f(v)||_p^p \leq \sum_{(u,v) \in E} ||f(u) - f(v)||_p^p, \tag{6}$$

where $c = h/4(2d)^{1-1/p}$.

The proof of (6) will follow roughly the same steps as that of Theorem 2.1. So we will first establish an inequality for real valued functions:

Lemma 2.5. *Let $f : V \to \mathbb{R}$ be any function and $p > 1$. Then*

$$\sum_{(u,v) \in E} |f(u) - f(v)|^p \geq \frac{(h/2p)^p}{(2d)^{p-1}} \sum_{v \in V} |f(v) - M|^p \geq \frac{(h/4p)^p}{n(2d)^{p-1}} \sum_{u,v \in V} |f(u) - f(v)|^p, \tag{7}$$

where M is a median of f.

Proof. Again we may assume that $M = 0$. Let $S = \sum_{(u,v) \in E} |f(u) - f(v)|^p$ and $T = \sum_{v \in V} |f(v)|^p$ and observe that 0 is also a median of the set $\left\{ f(u)|f(u)|^{p-1} \mid u \in V \right\}$ (we think of $(\cdot)|\cdot|^{p-1}$ as a sign preserving way of raising to the pth power). Hence by Lemma 2.2 applied to these values we obtain

$$hT \leq \sum_{(u,v) \in E} |f(u)|f(u)|^{p-1} - f(v)|f(v)|^{p-1}|. \tag{8}$$

We now claim the following:

Claim 2. For real numbers a and b and $p \geq 1$ we have the following inequality

$$|a^p - b^p| \leq p|a - b|(|a|^{p-1} + |b|^{p-1}). \tag{9}$$

Proof of claim: First observe that by passing to $-a$ and $-b$ if need be we may assume that $a \geq |b| > 0$ (if either a or b is 0 the claim trivially holds). Furthermore by rescaling we may assume $a = 1$, so it suffices to prove $1 - b^p \leq p(1 - b)$ which is equivalent to $1 + p(b - 1) \leq (1 + (b - 1))^p$, and since $b - 1 \geq -2$ this follows by Bernoulli's inequality. $\square$

Note that (9) is equally valid if we replace $(\cdot)^p$ by $(\cdot)|\cdot|^{p-1}$ (there are a few cases to check) so we can use the claim on the individual terms on the right hand side of (8) followed by

Hölder's inequality to obtain the bound

$$\begin{aligned} hT &\leq \sum_{(u,v)\in E} p|f(u)-f(v)|\left(|f(u)|^{p-1}+|f(v)|^{p-1}\right) \\ &\leq p\left(\sum_{(u,v)\in E}|f(u)-f(v)|^p\right)^{1/p}\left(\sum_{(u,v)\in E}\left(|f(u)|^{p-1}+|f(v)|^{p-1}\right)^q\right)^{1/q}, \end{aligned}$$

where $q = p/(p-1)$ is the Hölder conjugate of p. Since

$$(|f(u)|^{p-1}+|f(v)|^{p-1})^q \leq 2^q(|f(u)|^{q(p-1)}+|f(v)|^{q(p-1)}) = 2^q(|f(u)|^p+|f(v)|^p),$$

and by using the d-regularity of G we obtain

$$hT \leq pS^{1/p}2(2d)^{1/q}\left(\sum_{v\in V}|f(v)|^p\right)^{1/q} = pS^{1/p}2(2d)^{1/q}T^{1/q}.$$

By rearranging we obtain the first inequality in (7). The second is routine to establish using $|f(u)-f(v)|^p \leq 2^p(|f(u)|^p+|f(v)|^p)$. □

Theorem 2.4 now follows directly by using Lemma 2.5 pointwise on each $f(v)$, $v \in V$. Theorem 2.4 yields that sequences of expanders do not coarsely embed into any Hilbert space. This can also be established using the spectral characterization of expanders, i.e., we can obtain 2-Poincaré inequalities with respect to the adjacency matrix for Hilbert space valued functions with the constant depending on the spectral gap. Since this result will be necessary to see how the notion of expander graphs generalizes we will state and prove it in the following subsection.

2.2 Poincaré inequalities through the spectral gap

We will now establish p-Poincaré inequalities similar to the ones in the previous subsection but with constants depending on the spectral gap of G in stead of the Cheeger constant. This will lead to a proof of the Cheeger inequalities and play an important role in generalizing the notion of expander sequences to that of superexpanders. The setup is slightly more general: Let $G = (V, E)$ be a finite, connected, unoriented graph on n vertices, where each vertex is of finite (but not necessarily constant) degree. Let $(a_{xy})_{x,y\in V}$ denote its (non-normalized!) adjacency matrix. For each vertex $x \in V$ let $\nu(x)$ denote its degree. We can then consider the n-dimensional (weighted) L_2 space $L_2(V, \nu)$ consisting of the functions $f : V \to \mathbb{C}$ with inner product

$$\langle f, g\rangle = \frac{1}{|\nu|}\sum_{x\in V}\nu(x)f(x)g(x),$$

where $|\nu| = \sum_{x\in V}\nu(x)$. This space has an orthogonal basis $(\delta_x)_{x\in V}$ where $\delta_x(y) = \delta_{x,y}$ and we note that $\|\delta_x\|^2 = \langle\delta_x, \delta_x\rangle = \nu(x)/|\nu|$. We can now define the Laplacian as the positive operator $\Delta = D^*D/2$ where $D : L_2(V, \nu) \to L_2(E)$ is the differential operator given by $Df(x, y) = f(x)-f(y)$ for $(x, y) \in E$. By computing $\langle\Delta f, f\rangle = \|Df\|^2_{L_2(V,\nu)}/2$ it can easily be

checked that $\Delta f(x) = f(x) - \sum_{y\in V} a_{x,y} f(y)/\nu(x)$, which yields that $\Delta = I - [a_{xy}/\nu(x)]_{x,y\in V}$ and in particular $\mathbb{C}\cdot 1_V \subset \ker(\Delta)$. Note that the rows of the matrix $[a_{xy}/\nu(x)]_{x,y\in V}$ add up to one (in the case where G is d-regular this is just the normalized adjacency matrix) so, since G is connected, the proof of Proposition 1.4 carries over to show that $\ker(\Delta) = \mathbb{C}\cdot 1_V$. Now if we let $0 = \alpha_0 \leq \alpha_1 \leq \cdots \leq \alpha_{n-1}$ denote the eigenvalues of Δ, counted with multiplicity, the statement that $\ker(\Delta) = \mathbb{C}\cdot 1_V$ means precisely that $\alpha_1 > 0$.
Finally we observe that

$$\langle \Delta\delta_x, \delta_y\rangle = \begin{cases} \frac{\nu(x)-1}{|\nu|} & \text{if } y = x \text{ and } (x,y)\in E \\ \nu(x)/|\nu| & \text{if } y = x \text{ and } (x,y)\notin E \\ -1/|\nu| & \text{if } y \neq x \text{ and } (x,y)\in E \\ 0 & \text{else} \end{cases}.$$

Now let H be a Hilbert space and $f : V \to H$ be a function. Denote by $(f_x)_{x\in V} \in H^n$ its function values. We can view f as an element of the Hilbert space $L_2(V,\nu)\otimes H$ through the correspondence $f \leftrightarrow f' = \sum_{x\in V} \delta_x \otimes f_x$. From [BO07], we have the following Poincaré inequality for f:

Lemma 2.6. *Let $f : V \to H$ be a function with function values $(f_x)_{x\in V} \in H^n$ and let f' be its representation in $L_2(V,\nu)\otimes H$. Then*

$$\frac{\alpha_1}{2|\nu|^2} \sum_{x,y\in V} \nu(x)\nu(y)||f_x - f_y||_H^2 = \alpha_1 \left(||f'||^2_{L_2(V,\nu)\otimes H} - ||m(f)||_H^2 \right)$$
$$\leq \frac{1}{2|\nu|} \sum_{x,y\in V} a_{x,y} ||f_x - f_y||_H^2,$$

where $m(f) = \sum_{x\in V} \nu(x) f_x/|\nu| \in H$ is the average of f.

Proof. The equality to the left is straightforward to check by expanding the inner products and rearranging terms. For the inequality to the right, let $E : L_2(V,\nu) \to L_2(V,\nu)$ be the orthogonal projection onto the constant functions, i.e., $E(g) = \langle g, 1_V\rangle \cdot 1_V$ for $g \in L_2(V,\nu)$. Then if I_H denotes the identity on H we have

$$E\otimes I_H(f') = \sum_{x\in V} \langle \delta_x, 1_V\rangle 1_V \otimes f_x = \sum_{x\in V} \frac{\nu(x)}{|\nu|} 1_V \otimes f_x = 1_V \otimes \sum_{x\in V} \frac{\nu(x)}{|\nu|} f_x = 1_V \otimes m(f),$$

and since the tensor product norm is a cross norm and $1_V \in L_2(V,\nu)$ has unit norm we obtain $||m(f)||_H^2 = ||E\otimes I_H(f')||^2_{L_2(V,\nu)\otimes H}$. Moreoever, since $f' = f' - E\otimes I_H(f') + E\otimes I_H(f')$ and the two terms on the right are mutually orthogonal we get by Pythagoras that

$$||f'||^2_{L_2(V,\nu)\otimes H} - ||m(f)||_H^2 = ||f'||^2_{L_2(V,\nu)\otimes H} - ||E\otimes I_H(f')||^2_{L_2(V,\nu)\otimes H}$$
$$= ||f' - E\otimes I_H(f')||^2_{L_2(V,\nu)\otimes H}.$$

Now, since Δ is positive, we can find an orthonormal basis $(g_j)_{j=0}^{n-1}$ for $L_2(V,\nu)$ such that $\Delta g_j = \alpha_j g_j$. Note that $\alpha_0 = 0$ and by the above observations we have $g_0 = 1$. Then by standard results on tensor products of Hilbert spaces we can find $\xi_0, \dots, \xi_{n-1} \in H$ such that $f' = \sum_{j=0}^{n-1} g_j \otimes \xi_j$ and so

$$\langle \Delta \otimes I_H f', f' \rangle = \sum_{j=1}^{n-1} \alpha_j ||\xi_j||^2 \geq \alpha_1 \sum_{j=1}^{n-1} ||\xi_j||^2 = \alpha_1 ||f' - E \otimes I_H(f')||^2,$$

where the last equality follows since $E \otimes I_H(f') = g_0 \otimes \xi_0$ and by orthogonality. To complete the proof we should check that $\langle \Delta \otimes I_H(f'), f' \rangle = (2|\nu|)^{-1} \sum_{x \in V} a_{x,y} ||f_x - f_y||^2$ which is readily done using the computations of $\langle \Delta \delta_x, \delta_y \rangle$ above. □

When the graph is d-regular, we have $\alpha_1 = \lambda(G)$, the spectral gap of G, and the Poincaré inequality above simplifies to

$$\frac{\lambda(G)}{n} \sum_{u,v \in V} ||f_u - f_v||_H^2 \leq \sum_{u,v \in V} \frac{a_{x,y}}{d} ||f_u - f_v||_H^2,$$

an inequality very similar to that for L_1-valued functions obtained above. The significance of the above Poincaré inequality is that it gives an alternative characterization of the spectral gap which better lends it self to generalizations. The statement of the following proposition is found in [MN14]:

Proposition 2.7. *Let $G = (V, E)$ be a finite, connected d-regular graph on n vertices and let $A = (a_{u,v})_{u,v \in V}$ be its normalized adjacency matrix with eigenvalues (in descending order and counted with multiplicity) $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n \geq -1$. Then $1 - \lambda_2$ (the spectral gap of G) is equal to the supremum over all γ satisfying*

$$\frac{\gamma}{n} \sum_{u,v \in V} (x_u - x_v)^2 \leq \sum_{u,v \in V} a_{u,v} (x_u - x_v)^2, \tag{10}$$

for every $x = (x_u)_{u \in V} \in \mathbb{R}^n$.

Proof. First observe that since $\mathbb{R}$ is a Hilbert space the spectral gap satisfies (10), for all $x \in \mathbb{R}^n$, by Lemma 2.6. Now let $\gamma > 0$ be such that inequality (10) holds for any V-indexed array of real numbers. Then the all-ones vector $1_V \in \mathbb{R}^n$ is the eigenvector of λ_1, and since A is symmetric, we get by the min-max theorem that $\lambda_2 = \sup_{||x||=1,\ x \perp 1_V} \langle Ax, x \rangle$ and hence

$$1 - \lambda_2 = \inf_{||x||=1,\ x \perp 1_V} (1 - \langle Ax, x \rangle).$$

So take $x \in \mathbb{R}^n$ such that $||x|| = 1$ and $x \perp 1_V$ and note that $x \perp 1_V$ means exactly that $\sum_i x_i = 0$, where $x_1, \dots, x_n$ are the entries of x. Then the entries of this vector satisfy (10). The left hand side of (10) simplifies to

$$\frac{\gamma}{n} \sum_{i,j} (x_i - x_j)^2 = \frac{\gamma}{n} \sum_{i,j} (x_i^2 + x_j^2 - 2x_i x_j) = 2\gamma,$$

using that $\sum_{i,j} x_i^2 = \sum_j \sum_i x_i^2 = n||x||^2 = n$ and that $\sum_{i,j} x_i x_j = \sum_j x_j \sum_i x_i = \sum_j 0 = 0$ by the assumptions on x. The right hand side equals

$$\sum_{i,j} a_{i,j} (x_i - x_j)^2 = \sum_{i,j} a_{i,j} x_i^2 + \sum_{i,j} a_{i,j} x_j^2 - 2 \sum_{i,j} a_{i,j} x_i x_j = 2 (1 - \langle Ax, x \rangle),$$

using that A is normalized, whence $\sum_i \sum_j a_{i,j} x_i^2 = \sum_i x_i^2 = 1$, and similarly with the second term. This yields $\gamma \leq 1 - \langle Ax, x\rangle$ from which we conclude

$$\gamma \leq \inf_{||x||=1,\ x \perp 1_V} (1 - \langle Ax, x\rangle) = 1 - \lambda_2.$$

The conclusion now follows. □

We are now ready to give the proof of the Cheeger inequalities:

Proof of Theorem 1.7. By Lemma 2.5 with $p = 2$ and Proposition 2.7 we obtain

$$\lambda(G) \geq \frac{1}{2}\left(\frac{h(G)}{8d}\right)^2,$$

which is precisely the right hand side of the inequality in Theorem 1.7. For the left hand side let $A = (a_{uv})_{u,v\in V}$ be the normalized adjacency matrix of G with eigenvalues $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n \geq -1$. Observe that for $U, W \subset V$ the standard $\mathbb{R}^n$-inner product of $(dA)1_U$ with 1_W is precisely the number of edges between U and W, i.e., $\langle A1_U, 1_W\rangle = |E(U, W)|/d$, where $E(U, W)$ denotes the edges of G with one end point in U and one in W. This follows directly from the definition of the adjacency matrix. Now take $F \subset V$ such that $|F| \leq |V|/2$ and $|\partial F|/|F| = h(G)$. We may decompose 1_F into an orthogonal sum of a constant vector (i.e., an element of the eigenspace of 1) and a mean zero vector (an element of the direct sum of the remaining eigenspaces). Specifically we have

$$1_F = \frac{|F|}{|V|}1_V + x,$$

where $\sum_{i=1}^n x_i = 0$. Since $\|1_F\|_2^2 = |F|$ and $\left\|\frac{|F|}{|V|}1_V\right\|_2^2 = |F|^2/|V|$ we have, by orthogonality, that $\|x\|_2^2 = |F| - |F|^2/|V|$. Moreover, since x is an orthogonal sum of eigenvectors with eigenvalues at most λ_2, we get

$$\begin{aligned}\langle A1_F, 1_F\rangle &= \left\langle \frac{|F|}{|V|}1_V, \frac{|F|}{|V|}1_V \right\rangle + \langle Ax, x\rangle \leq \frac{|F|^2}{|V|} + \lambda_2 \|x\|_2^2 \\ &= \lambda_2|F| + (1-\lambda_2)\frac{|F|^2}{|V|} \leq \lambda_2|F| + \frac{1-\lambda_2}{2}|F| = \frac{1+\lambda_2}{2}|F|.\end{aligned}$$

By the initial observation we also have

$$\langle A1_F, 1_F\rangle = \langle A1_F, 1_V\rangle - \langle A1_F, 1_{F^c}\rangle = |F| - |\partial F|/d = |F| - |F|\frac{h(G)}{d}.$$

So we obtain

$$1 - \frac{h(G)}{d} \leq \frac{1+\lambda_2}{2} = \frac{-\lambda(G)}{2} + 1,$$

which yields the desired inequality. □

Remark 2.8. The inequality we have proven can be improved. In fact the multiplicative constant 8 on the right hand side can be omitted without it losing its validity. This stronger version is what is usually referred to as the Cheeger inequalities, and its proof is based on a clever application of Lemma 2.2 (see e.g. [Ost13, Th. 5.7]). Since our inequality with suboptimal constants suffices for demonstrating equivalence of edge expansion and spectral expansion we will settle for this version.

We now have four equivalent ways to characterize sequences of expanders, which we sum up in the following corollary:

Corollary 2.9. *Let $(G_n = (V_n, E_n))$ be a sequence of connected, d-regular graphs such that $|V_n| \to \infty$. Then the following are equivalent:*

1. *G_n is a sequence of spectral expanders, i.e., if $\lambda_2(G_n)$ denotes the second largest eigenvalue of the normalized adjacency matrix of G_n then $\sup_{n\geq 1} \lambda_2(G_n) < 1$.*
2. *G_n is a sequence of edge expanders, i.e., the Cheeger constants satisfy $\inf_{n\geq 1} h(G_n) > 0$.*
3. *There is $\gamma > 0$ such that, for all $n \geq 1$ and $f : V_n \to \mathbb{R}$, we have*
$$\frac{\gamma}{n} \sum_{u,v\in V_n} (f(u) - f(v))^2 \leq \sum_{u,v\in V_n} a^n_{u,v}(f(u) - f(v))^2,$$
where $(a^n_{u,v})$ is the normalized adjacency matrix of G_n.
4. *For any Hilbert space H there is $\gamma > 0$ such that, for all $n \geq 1$ and $f : V_n \to H$, we have*
$$\frac{\gamma}{n} \sum_{u,v\in V_n} \|f(u) - f(v)\|^2 \leq \sum_{u,v\in V_n} a^n_{u,v} \|f(u) - f(v)\|^2,$$
where $(a^n_{u,v})$ is the normalized adjacency matrix of G_n.

If any (and hence all) of these conditions are satisfied we shall simply call (G_n) a sequence of expanders.

The last condition in the above corollary plays an important role in generalizations of expander families, and we deliberately formulated it this way, even though Lemma 2.6 demonstrates that the same γ, namely the spectral gap, does the trick for any Hilbert space. This topic will be further explored in subsequent sections. For now we turn our attention to the thus far neglected question of existence of expander sequences.

3 Constructing expander sequences

We now introduce Kazhdan's property (T), an approximation property for representations of groups. Residually finite groups with property (T) will result in sequences of expanders, thus reducing the construction of such sequences to finding groups with these properties.

3.1 Kazhdan's property (T)

Our discussion of Kazhdan's property (T) for groups below is based on [BO07] and will be somewhat cursory and only intended to introduce the concepts we need for constructing sequences of expander graphs. First some terminology: Let H be a Hilbert space. We denote by $B(H)$ the bounded linear operators on H, i.e., linear maps T such that there exists a constant $C \geq 0$ satisfying $||T\xi|| \leq C||\xi||$, for all $\xi \in H$. By $U(H)$ we denote the unitary operators, i.e., the operators $U \in B(H)$ such that $UU^* = U^*U = I$, where I is the identity operator on H and U^* is the unique operator such that $\langle U\xi, \eta\rangle = \langle \xi, U^*\eta\rangle$, for all $\xi, \eta \in H$. Note that $U(H)$ is a group with composition of maps as multiplication.

Let Γ be a group. A *unitary representation* of Γ on a Hilbert space, H, is a group homomorphism $\pi : \Gamma \to U(H)$ for some Hilbert space H. A non-zero vector $\xi \in H$ is said to be Γ-*invariant* if $\pi(t)\xi = \xi$, for all $t \in \Gamma$. A net $(\xi_i)_{i\in I} \subset H$ of unit vectors is called *almost* Γ-invariant if $||\pi(t)\xi_i - \xi_i|| \to 0$, for all $t \in \Gamma$. We are now ready to introduce Kazhdan's property for groups:

Definition 3.1. A group, Γ, has Kazhdan's property (T) if any unitary representation with an almost Γ-invariant net also has a Γ-invariant vector.

Recall that a group Γ is *amenable* if and only if the left regular representation has an almost Γ-invariant net. Moreover, if $\xi \in \ell^2(\Gamma)$ is a non-zero Γ-invariant vector for λ, then

$$\langle \xi, \delta_t\rangle = \langle \xi, \lambda(t)\delta_e\rangle = \langle \lambda(t^{-1})\xi, \delta_e\rangle = \langle \xi, \delta_e\rangle,$$

for all $t \in \Gamma$, where $(\delta_t)_{t\in\Gamma}$ is the canoncial orthonormal basis for $\ell^2(\Gamma)$. Hence, we obtain

$$||\xi||^2 = \sum_{t\in\Gamma} |\langle \xi, \delta_t\rangle|^2 = \sum_{t\in\Gamma} |\langle \xi, \delta_e\rangle|^2 = |\Gamma||\langle \xi, \delta_e\rangle|^2,$$

so that, since $\xi \neq 0$, Γ must be finite. This demonstrates that if Γ is amenable *and* has property (T), then Γ must be finite. Since all abelian groups are amenable, no *infinite* abelian group can have property (T), so in particular $\mathbb{Z}$ does not have (T). Since property (T) passes to quotients (see below), and $\mathbb{Z}$ is a quotient group of $\mathbb{F}_2$, the free group on two generators, the latter cannot have property (T). It is also a fact that $\mathbb{F}_2$ is not amenable, so there exist infinite groups which are neither amenable nor have (T). To see that property (T) passes to quotients consider the quotient map $q : \Gamma \to \Gamma/N$. For any unitary representation π of Γ/N with an almost invariant net, we get a representation of Γ with an almost invariant net by taking $\pi \circ q$. Property (T) now follows by surjectivity of q.

Definition 3.2. Let $S \subset \Gamma$, $k > 0$ and let π be a unitary representation of Γ on H. If $\sup_{s\in S} ||\pi(s)\xi - \xi|| < k||\xi||$ for some $\xi \in H$ we say that ξ is an (S, k)-invariant vector for π. A pair (S,k) with $S \subset \Gamma$ and $k > 0$ is called a Kazhdan pair if any unitary representation with an (S, k)-invariant vector also has a Γ-invariant vector.

Proposition 3.3. *Suppose (S,k) is a Kazhdan pair for Γ and let $\pi : \Gamma \to U(H)$ be a unitary representation such that there exists $\xi \in H$ with $\pi(s)\xi = \xi$, for all $s \in S$. Then ξ is Γ-invariant for π.*

Proof. Let $H_0 \subset H$ be the closed subspace of all Γ-invariant vectors and let $K = H_0^{\perp}$. Then both H_0 and K are invariant under $\pi(s)$, for all $s \in \Gamma$. Let $\pi_1 = \pi|_K : \Gamma \to U(K)$, i.e., $\pi_1(s) = \pi(s)|_K$. Then by definition of K, π_1 has no Γ-invariant vectors. Since (S,k) is a Kazhdan pair, for any $\eta \in K$ there exists $s \in S$ such that $||\pi(s)\eta - \eta|| \geq k||\eta||$. Now write $\xi = \xi_0 + \eta$ uniquely with $\xi_0 \in H_0$ and $\eta \in K$. Then, for all $s \in S$, we have

$$\xi = \pi(s)\xi = \pi(s)\xi_0 + \pi(s)\eta = \xi_0 + \pi(s)\eta,$$

and hence by uniqueness of the decomposition we get $\pi(s)\eta = \eta$, for all $s \in S$. In particular, this holds for the s such that $\|\pi(s)\eta - \eta\| \geq k\,\|\eta\|$, which implies that $k||\eta|| \leq 0$ and hence $\eta = 0$, so that $\xi \in H_0$ as desired. □

Lemma 3.4. *If (S,k) is a Kazhdan pair for Γ, then S generates Γ.*

Proof. Let $\Lambda < \Gamma$ be the subgroup generated by S, and let Γ/Λ be the set of left cosets of Λ. Let $\pi : \Gamma \to B(\ell^2(\Gamma/\Lambda))$ be given by $\pi(t)\delta_{s\Lambda} = \delta_{ts\Lambda}$, where $(\delta_{s\Lambda})_{s\Lambda\in\Gamma/\Lambda}$ is the canonical basis for $\ell^2(\Gamma/\Lambda)$. This is well-defined and a unitary representation of Γ on $\ell^2(\Gamma/\Lambda)$. Now observe that $\pi(s)\delta_\Lambda = \delta_{s\Lambda} = \delta_\Lambda$, for any $s \in S$, and hence by Proposition 3.3 we get that $\pi(t)\delta_\Lambda = \delta_\Lambda$, for all $t \in \Gamma$. Hence $t\Lambda = \Lambda$, for all $t \in \Gamma$, so we must have $\Lambda = \Gamma$ as desired. □

Proposition 3.5. *A group Γ has property (T) if and only if there exists a finite subset $S \subset \Gamma$ and $k > 0$ such that (S,k) is a Kazhdan pair for Γ.*

Proof. Assume that there exist a finite subset $S \subset \Gamma$ and $k > 0$ such that (S,k) is a Kazhdan pair for Γ. Let $\pi : \Gamma \to U(H)$ be a unitary representation of Γ with a net $(\xi_i)_{i\in I} \subset H$ of almost Γ-invariant vectors. Then for any $s \in S$ there is an $i_s \in I$ such that $||\pi(s)\xi_i - \xi_i|| < k$ whenever $i \geq i_s$. Since S is finite and I is directed there exists i_0 such that $i_0 \geq i_s$, for all $s \in S$. Putting $\xi = \xi_{i_0} \neq 0$ we obtain $||\pi(s)\xi - \xi|| < k$, for all $s \in S$, and hence, since (S,k) is a Kazhdan pair, we infer that there is a Γ-invariant vector, so Γ has property (T).

Now assume that there is no Kazhdan pair (S,k) with S finite. Define an index set by $I = \{(S,k) \mid S \subset \Gamma,\ \#S < \infty,\ k > 0\}$, with the ordering $(S_1,k_1) \preceq (S_2,k_2)$ if $S_1 \subset S_2$ and $k_2 \leq k_1$, making I a directed set. For each $i = (S_i,k_i) \in I$ we can find a unitary representation $\pi_i : \Gamma \to U(H_i)$ and a unit vector $\xi_i \in H_i$ such that $||\pi_i(s)\xi_i - \xi_i|| \leq k_i$, for all $s \in S_i$, but such that there are no non-zero Γ-invariant vectors for π_i. This is because, by assumption, (S_i,k_i) is not a Kazhdan pair. Let $\pi : \Gamma \to U(H)$, where $H = \bigoplus_{i\in I} H_i$, be given by $\pi = \bigoplus_{i\in I}\pi_i$. Then, since $H_i \subset H$, we can consider $(\xi_i)_{i\in I}$ as a net of unit vectors in H. Moreover, $||\pi(t)\xi_i - \xi_i|| \to 0$, for all $t \in \Gamma$, which can be seen by considering $i_0 = (\{t\},\varepsilon) \in I$ for each $\varepsilon > 0$. Now assume $\xi \in H$ is Γ-invariant for π, and for each $i \in I$ let $P_i : H \to H_i$ be the orthogonal projection. Then $\pi(t)P_i = P_i\pi(t)$, since H_i is invariant under $\pi(t)$, and hence

$$\pi(t)P_i\xi = P_i\pi(t)\xi = P_i\xi, \quad \text{for all } t \in \Gamma.$$

Demonstrating that $P_i\xi$ is Γ-invariant for π_i and hence, by assumption, $P_i\xi = 0$. Since $i \in I$ was arbitrary we obtain $\xi = 0$. This demonstrates that there are no non-zero Γ-invariant vectors for π and hence that Γ does not have property (T). □

It is an important fact that $(\Gamma, \sqrt{2})$ is a Kazhdan pair for any group Γ. In light of the above proposition, this implies that all finite groups have property (T), so the groups which are amenable and have (T) are exactly the finite ones.

Remark 3.6. Let Γ be a group with property (T). By the above results we can find a Kazhdan pair (Σ, k) with Σ finite such that Σ generates Γ. Furthermore, by adding the inverses, we may assume that Σ is symmetric.

3.2 Expander sequences from groups with property (T)

Recall that for a group Γ and $\Sigma \subset \Gamma$ the *Cayley graph* $\mathrm{Cay}(\Gamma, \Sigma)$ is the graph whose vertices consist of the group elements, i.e., $V = \Gamma$, and has edge set $E = \{(t, st) \mid t \in \Gamma,\ s \in \Sigma\}$. This notion generalizes to quotients as follows: For $\Lambda < \Gamma$, consider the graph $\mathrm{Sch}(\Gamma, N, \Sigma)$ whose vertex set is $V := \Gamma/N$ and whose edge set is $E = \{(tN, s, stN) \mid tN \in \Gamma/N,\ s \in \Sigma\}$. The second coordinate is to distinguish distinct edges with the same source and range (which will occur whenever $s's^{-1} \in N$ for distinct $s', s \in \Sigma$). Such graphs are referred to in the literature as *Schreier coset graphs*, and obviously any Cayley graph is a Schreier coset graph with $N = \{e\}$. Consistent with this fact, we may sometimes abuse notation slightly and write $\mathrm{Sch}(\Gamma, N, \Sigma) = \mathrm{Cay}(\Gamma/N, \Sigma)$. We will now use the Schreier coset graphs of groups with property (T) to construct expander graphs. This idea is originally due to Margulis, but we follow [AM85].

Theorem 3.7. *Let Γ be a group with property (T) (in particular it is finitely generated). Let (Σ, k) be a Kazhdan pair with Σ a finite, symmetric, generating set for Γ.*
If $N \lhd \Gamma$ is of finite index then $Sch(\Gamma, N, \Sigma)$, which is d-regular with $d = |\Sigma|$ on n vertices with $n = |\Gamma/N|$, has Cheeger constant at least $k^2/2$.

Proof. Consider the Hilbert space $H = \ell^2(\Gamma/N)$. Then for $t \in \Gamma$ the map $f \mapsto tf$, where $tf(sN) = f(st^{-1}N)$ as in the proof of Lemma 3.4 is a unitary representation. The subspace $H_0 = \ell^2_0(\Gamma/N)$ of functions, f, such that $\sum_{s \in \Gamma/N} f(s) = 0$ is invariant under this representation, i.e., if $f \in H_0$ then so is tf for any $t \in \Gamma$. Hence the left action of Γ is also a unitary representation on H_0. Furthermore if $f \in H_0$ and $tf = f$, for all $t \in \Gamma$, then in particular $f(t^{-1}) = f(e)$, for all $t \in \Gamma$, so f is constant. Since we also have $\sum_v f(v) = 0$ we conclude that $f = 0$, demonstrating that there are no non-zero Γ-invariant vectors for this representation. Hence, (Σ, k) being a Kazhdan pair for Γ, for any $f \in H_0$, there must exist some $s \in \Sigma$ such that $||sf - f|| \geq k||f||$. Now let $S \subset \Gamma/N$ such that $|S| \leq \frac{n}{2}$. Let $f \in H_0$ be given by $f(x) = 1_S(x)|S^c| - 1_{S^c}|S|$. Then

$$||f||^2 = |S^c|^2|S| + |S|^2|S^c| = |S||S^c|(|S| + |S^c|) = |S||S^c|n.$$

Moreover, for any $s \in \Sigma$, we have

$$|sf(x) - f(x)| = \begin{cases} |S| + |S^c| & \text{if } x \in S,\ xs^{-1} \in S^c \text{ or } x \in S^c,\ xs^{-1} \in S \\ 0 & \text{if } x, xs^{-1} \in S \text{ or } x, xs^{-1} \in S^c \end{cases},$$

so if we define $\partial_s S = \left\{(x, y) \in \partial S \mid y = xs \text{ or } y = xs^{-1}\right\}$ a straightforward computation yields $||sf - f||^2 = (|S| + |S^c|)^2 |\partial_s S| = n^2 |\partial_s S|$. Now pick $s \in \Sigma$ such that $||sf - f||^2 \geq k^2||f||^2$.

Then

$$\frac{|\partial S|}{|S|} \geq \frac{|\partial_s S|}{|S|} = \frac{||sf - f||^2}{n^2|S|} \geq \frac{k^2||f||^2}{n^2|S|} = \frac{2k^2|S^c|}{2n} \geq \frac{k^2}{2},$$

because $2|S^c| \geq n$. Since S was arbitrary of cardinality at most $n/2$, this yields that the Cheeger constant is at least $\frac{k^2}{2}$ as desired. □

Recall that a group Γ is residually finite if there exists a *filtration*, i.e., a sequence $N_1 \supset N_2 \supset \dots$ of normal, finite index subgroups such that $\cap_{i\geq 1} N_i = \{e\}$, where $e \in \Gamma$ is the neutral element. In view of Theorem 3.7 we now have the following corollary:

Corollary 3.8. *Suppose Γ is a residually finite group with Kazhdan's property (T). Let (Σ, k) be a Kazhdan pair for Γ with Σ a finite, symmetric, generating set for Γ and $(N_i)_{i\geq 1}$ a filtration. Then the graph sequence $G_i = Sch(\Gamma, N_i, \Sigma)$ is a sequence of expanders.*

Consider the special linear group $SL_n(\mathbb{Z})$. The kernels of the surjective homomorphisms $SL_n(\mathbb{Z}) \twoheadrightarrow SL_n(\mathbb{Z}/i\mathbb{Z})$, where $i \in \mathbb{N}$, are of finite index and intersect trivially, so the special linear group is residually finite. Moreover, it is an important theorem of Kazhdan himself that $SL_n(\mathbb{Z})$ has property (T) whenever $n \geq 3$ (we refer the reader to Shalom's proof in [Sha99]). Hence, by the above corollary, $SL_n(\mathbb{Z})$ can be used to construct sequences of expanders whenever $n \geq 3$:

Corollary 3.9. *Expander sequences exist and can be realized as Schreier graphs of the special linear group of order at least* 3*.*

4 Superexpanders

In this section we will generalize the spectral gap of a graph and use this to define the notion of being an expander with respect to a given metric space. We will then see that being an expander sequence in the sense of Section 2 is precisely being an expander with respect to the usual Euclidean metric on $\mathbb{R}$. Following [MN14], we will prove that certain of the properties of the euclidean spectral gap from Section 1 pass over to the general version. The generalization of the notion of expander sequences will also lead to the definition of *superexpander* sequences, being sequences which are expanders with respect to any superreflexive Banach space. Superexpanders are in particular expanders, but whether the reverse implication holds is an open question. Following [Pis10], we will give partial answers to this question. We will then proceed, following [Mim18], to prove certain invariance properties for sequences being expanders with respect to Banach spaces.

4.1 Non-linear spectral gaps

Let $G = (V, E)$ be a d-regular graph on n vertices and $A = (a_{ij})$ its normalized adjacency matrix. The following definition is from [MN14]:

Definition 4.1. Let X be a set and $K : X \times X \to \mathbb{R}_+$ a kernel, i.e., $K(x, y) = K(y, x)$, for all $x, y \in X$, we define the *reciprocal spectral gap* of G with respect to K as the infimum over all $\gamma > 0$ such that

$$\frac{1}{n^2}\sum_{i,j=1}^{n} K(x_i, x_j) \leq \frac{\gamma}{n}\sum_{i,j=1}^{n} a_{ij}K(x_i, x_j), \tag{11}$$

for all $x_1, \ldots, x_n \in X$, and we denote it by $\gamma(G, K)$.
Analogously we define the reciprocal *absolute* spectral gap $\gamma_+(G, K)$ as the infimum over all $\gamma > 0$ such that

$$\frac{1}{n^2}\sum_{i,j=1}^{n} K(x_i, y_j) \leq \frac{\gamma}{n}\sum_{i,j=1}^{n} a_{ij}K(x_i, y_j), \tag{12}$$

for all $x_1, \ldots, x_n, y_1, \ldots, y_n \in X$.

Note, that clearly $\gamma_+(G, K) \geq \gamma(G, K) \geq 1$. Examples of kernels are powers of norms or, more generally, metrics. These will be our main interest, but we will state some results in the general context of kernels.

Proposition 4.2. *Let (X, d_X) be a metric space with at least two points and $p \geq 1$. Then $\gamma(G, d_X^p)$ is finite if and only G is connected and $\gamma_+(G, d_X^p)$ is finite if and only if G is connected and not bipartite.*

Proof. If G is not connected then any $f : V \to X$ which is constant on each path component of G and attains at least two distinct values makes the left hand side of (11) strictly positive and the right hand side zero. This demonstrates that no γ satisfies (11) for all $f : V \to X$ and hence $\gamma(G, d_X^p) = \infty$. By the observation preceding the proposition we also have $\gamma_+(G, d_X^p) = \infty$ in this case.

If, conversely, G is connected then, for any $x, y \in V$, let $\gamma_{x,y}$ be any shortest path between x and y. In particular, $l(\gamma_{x,y}) \leq n$ and hence, for any $f : V \to X$, we have, by Hölder's inequality, the estimate

$$d_X(f(x), f(y))^p \leq l(\gamma_{x,y})^{p-1} \sum_{(u,v)\in\gamma_{x,y}} d_X(f(u), f(v))^p \leq n^{p-1} \sum_{(u,v)\in\gamma_{x,y}} d_X(u, v)^p.$$

And from this we obtain

$$\begin{aligned}\frac{1}{n^2} \sum_{x,y\in V} d_X(f(x), f(y))^p &\leq \frac{n^{p-1}}{n^2} \sum_{x,y\in V} \sum_{(u,v)\in\gamma_{x,y}} d_X(f(u), f(v))^p \\ &\leq n^{p-1} \sum_{(u,v)\in E} d_X(f(u), f(v))^p = \frac{dn^p}{nd} \sum_{(u,v)\in E} d_X(f(u), f(v))^p,\end{aligned}$$

since no edge can occur twice in a shortest path, so each edge occurs at most n^2 times in the sum after the first inequality. This demonstrates that $\gamma(G, d_X^p) \leq dn^p < \infty$.
If G is connected and bipartite, let $U \cup W$ be a bipartition of V. Let $x \neq y$ be two distinct elements of X. Define $f : V \to X$ to be constantly equal to x on U and y on W, and define $g : V \to X$ to be constantly equal to y on U and x on W. Then for these f and g the left hand side of (12) is strictly positive, whereas the right hand side is zero. As before this yields that $\gamma_+(G, d_X^p) = \infty$.
If G is connected and not bipartite, then G contains a cycle of odd length. Indeed, if all cycles were of even length then, for any $u, v \in V$, either all paths between u and v are of even length or all paths between u and v are of odd length. Hence, if we fix some $v_0 \in V$ and let $U \subset V$ consist of the vertices which are connected to u_0 by paths of even length, then $U \cup U^c$ form a bipartition of V; a contradiction. Let C^o denote such a cycle of odd length. For any $u, v \in V$, let γ be any shortest path from u to C^o, and let γ' be any shortest path from v to C^o. The terminal vertices of γ and γ' respectively partition C^o into a subpath of even length and a subpath of odd length. Hence by concatenating γ and γ' through either the odd length or the even length subpath we can obtain a path from u to v whose length is odd and at most $2n$, and such that each edge occurs at most twice in this path. For each $u, v \in V$ let $[(u, u_1^{u,v}), (u_1^{u,v}, u_2^{u,v}), \cdots, (u_{2m_{u,v}}^{u,v}, v)]$, with $m_{u,v} \leq n - 1$, be such a path. Then for any $f, g : V \to X$ we have

$$\begin{aligned}&\frac{1}{n^2} \sum_{u,v\in V} d_X(f(u), g(v))^p \\ \leq& \sum_{u,v\in V} \frac{(2n)^{p-1}}{n^2} \left(d_X(f(u), g(u_1^{u,v}))^p + \sum_{i=1}^{m_{u,v}} \left(d_X(g(u_{2i-1}^{u,v}), f(u_{2i}^{u,v}))^p + d_X(f(u_{2i}^{u,v}), g(u_{2i+1}^{u,v}))^p \right) \right) \\ &\leq (2n)^{p-1} 2 \sum_{(u,v)\in E} d_X(f(u), g(v))^p = \frac{(2n)^p 2d}{dn} \sum_{(u,v)\in E} d_X(f(u), g(v))^p,\end{aligned}$$

which yields $\gamma_+(G, d_X^p) \leq (2n)^p 2d < \infty$, as desired. □

We saw in Proposition 2.7 that in the case $X = \mathbb{R}$ and $K(x, y) = (x - y)^2$ we have $\gamma(G, K) = \frac{1}{\lambda(G)} = \frac{1}{1-\lambda_2}$, where $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq -1$ are the eigenvalues of A (here we interpret $\frac{1}{1-\lambda_2}$ as ∞ if $\lambda_2 = 0$). This, together with the above proposition, explains the terminology. In the euclidean case the absolute spectral gap is also given in terms of the eigenvalues of A. This proposition is also stated in [MN14], we follow the proof from [Ost13].

Proposition 4.3. *Let $X = \mathbb{R}$ and $K : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ be given by $K(x, y) = (x - y)^2$. Then $\gamma_+(G, K) = \frac{1}{\lambda_+(G)}$ where $\lambda_+(G) = 1 - \max\{|\lambda_2|, |\lambda_n|\}$ as in Section 1. Again, we interpret $\frac{1}{\lambda_+(G)}$ as ∞ if $\lambda_+(G) = 0$.*

Proof. If $\lambda_+(G) = 0$ then by Proposition 1.4 G is either disconnected or bipartite. In both these cases $\gamma_+(G, K) = \infty$ by Proposition 4.2, so we may assume that $\lambda_+(G) > 0$. Let $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n) \in \mathbb{R}^n$. By subtracting $\sum_i y_i/n$ from the y_i's and the x_i's we may assume that $\sum_i y_i = 0$. The left hand side of (12) computes

$$\begin{aligned} \frac{1}{n^2} \sum_{i,j=1}^{n} (x_i - y_j)^2 &= \frac{1}{n^2} \sum_{i,j=1}^{n} x_i^2 + y_j^2 - 2x_i y_j \\ &= \frac{1}{n} \left(\|x\|_2^2 + \|y\|_2^2 \right). \end{aligned} \tag{13}$$

The right hand side of (12) yields

$$\begin{aligned} \frac{\gamma}{n} \sum_{i,j=1}^{n} a_{ij}(x_i - y_j)^2 &= \frac{\gamma}{n} \sum_{i,j=1}^{n} a_{ij} \left(x_i^2 + y_j^2 - 2x_i y_j \right) \\ &= \frac{\gamma}{n} \left(\|x\|_2^2 + \|y\|_2^2 - 2\langle Ay, x\rangle \right), \end{aligned} \tag{14}$$

using that A is symmetric and stochastic. By rearranging, this yields that $\gamma_+(G, K)$ is the minimal $\gamma > 0$ such that

$$\langle Ay, x\rangle \le \left(1 - \frac{1}{\gamma}\right) \frac{\|x\|_2^2 + \|y\|_2^2}{2}, \tag{15}$$

for all $x, y \in \mathbb{R}^n$ with $\sum_i y_i = 0$. So if we demonstrate that

$$1 - \lambda_+(G) = \sup \frac{2\langle Ay, x\rangle}{\|x\|_2^2 + \|y\|_2^2}, \tag{16}$$

where the supremums runs over all $x, y \in \mathbb{R}^n$, which are not both zero and such that $\sum_i y_i = 0$, we are done. To see this, observe that $\sum_i y_i = 0$ means precisely that y is orthogonal to $\mathbf{1}$, the all ones vector, which is the eigenvector of A corresponding to the eigenvalue $\lambda_1 = 1$. Let A_0 denote the restriction of A to the subspace of mean zero vectors (which is invariant under A). Since A is symmetric we have that the operator norm $\|A_0\| = 1 - \lambda_+(G)$, Hence for $x, y \in \mathbb{R}^n$ as above we have

$$\langle Ay, x\rangle \le |\langle Ay, x\rangle| \le \|Ay\|_2 \|x\|_2 \le (1 - \lambda_+(G)) \|y\|_2 \|x\|_2 \le (1 - \lambda_+(G)) \frac{\|x\|_2^2 + \|y\|_2^2}{2},$$

demonstrating that $1 - \lambda_+(G) \ge \sup 2\langle Ay, x\rangle/(\|x\|_2^2 + \|y\|_2^2)$. To obtain equality take $x = y$ to be an eigenvector of A orthogonal to $\mathbf{1}$ such that the corresponding eigenvalue has absolute value $1 - \lambda_+(G)$. Then (passing to $-y$ if needed) we get

$$\langle Ay, x\rangle = (1 - \lambda_+(G))\langle y, x\rangle = (1 - \lambda_+(G)) \frac{\|x\|_2^2 + \|y\|_2^2}{2},$$

demonstrating the desired. $\square$

Propositions 4.3 and 2.7 show that Definition 4.1 is really a generalization of the (euclidean) spectral gaps as described in Section 1. Proposition 4.2 shows that, if we restrict ourselves to metric spaces, the properties of Proposition 1.4 carry over. If we further restrict ourselves to normed spaces we get the following analogy: For a Banach space B, let $L_p^n(B)$ denote the Bochner space with values in B on the probability space $([n], \mathcal{P}([n]), \tau/n)$, where τ denotes the counting measure. For a symmetric, stochastic $n \times n$ matrix $A = (a_{ij})$ the linear operator $A \otimes I_B^n : L_p^n(B) \to L_p^n(B)$ is given by $(A \otimes I_B^n) f(i) := \sum_{j=1}^n a_{ij} f(j)$, for $f \in L_p^n(B)$ and $i \in [n]$. If $B = \mathbb{K}$ this is just the usual action of A on $\mathbb{K}^n$. As usual, let $(L_p^n(B))_0$ denote the mean-zero functions, which is invariant under $A \otimes I_B^n$ since A is symmetric and stochastic. Let $\lambda_B^{(p)}(A) := 1 - \|A \otimes I_B^n\|_{(L_p^n(B))_0 \to (L_p^n(B))_0}$ which is non-negative, again since A is symmetric and stochastic. Observe that by elementary linear algebra $\lambda_{\mathbb{K}}^{(2)}(A) = \lambda_+(A)$, so that $\gamma_+(A, |\cdot|^2) = 1/\lambda_{\mathbb{K}}^{(2)}(A)$. We cannot give a precise characterization of $\gamma_+(A, \|\cdot\|_B^p)$ in terms of $\lambda_B^{(p)}(A)$ as in the euclidean case, but a lower bound on $\lambda_B^{(p)}(A)$ gives an upper bound on $\gamma_+(A, \|\cdot\|_B^p)$:

Proposition 4.4. *In the above set-up we have*

$$\gamma_+(A, \|\cdot\|_B^p) \leq \left(1 + \frac{4}{\lambda_B^{(p)}(A)}\right)^p,$$

where the right hand side is interpreted as ∞ if $\lambda_B^{(p)}(A) = 0$.

Proof. For ease of notation let $\lambda = \lambda_B^{(p)}(A)$ and assume $\lambda > 0$. For $f, g : [n] \to B$ let $\overline{f} = \frac{1}{n}\sum_{i=1}^n f(i)$ and $\overline{g} = \frac{1}{n}\sum_{i=1}^n g(i)$, making $f_0 = f - \overline{f}$ and $g_0 = g - \overline{g}$ elements of $(L_p^n(B))_0$. Hence $\|(A \otimes I_B^n) f_0\|_{L_p^n(B)} \leq (1-\lambda)\|f_0\|_{L_p^n(B)}$ and similarly for g_0. Let $h \in L_p^{2n}(B)$ be given by

$$h(i) := \begin{cases} f_0(i), & \text{if } i \in [n] \\ g_0(i), & \text{if } i \in [2n] \setminus [n] \end{cases}.$$

We then have the estimate

$$\begin{aligned}
\lambda \|h\|_{L_p^{2n}(B)} &= \|h\|_{L_p^{2n}(B)} - (1-\lambda)\|h\|_{L_p^{2n}(B)} \\
&= \|h\|_{L_p^{2n}(B)} - \left(\frac{(1-\lambda)^p \|f_0\|_{L_p^n(B)}^p + (1-\lambda)^p \|g_0\|_{L_p^n(B)}^p}{2}\right)^{1/p} \\
&\leq \|h\|_{L_p^{2n}(B)} - \left(\frac{\|(A \otimes I_B^n) f_0\|_{L_p^n(B)}^p + \|(A \otimes I_B^n) g_0\|_{L_p^n(B)}^p}{2}\right)^{1/p} \\
&= \|h\|_{L_p^{2n}(B)} - \left\|(B \otimes I_B^{2n}) h\right\|_{L_p^{2n}(B)},
\end{aligned}$$

where $B = \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}$. Using the reverse triangle inequality, stochasticity of A and convexity of $t \mapsto t^p$ we then obtain

$$\begin{aligned}
\lambda \|h\|_{L_p^{2n}(B)} &\leq \left\| \left(I_{L_p^{2n}(B)} - B \otimes I_B^{2n} \right) h \right\|_{L_p^{2n}(B)} \\
&= \left(\frac{1}{2n} \sum_{i=1}^n \left(\left\| \sum_{j=1}^n a_{ij}(f_0(i) - g_0(j)) \right\|_B^p + \left\| \sum_{j=1}^n a_{ij}(g_0(i) - f_0(j)) \right\|_B^p \right) \right)^{1/p} \\
&\leq \left(\frac{1}{n} \sum_{i,j=1}^n a_{ij} \|f_0(i) - g_0(j)\|_B^p \right)^{1/p} \\
&\leq \left\| \overline{f} - \overline{g} \right\|_B + \left(\frac{1}{n} \sum_{i,j=1}^n a_{ij} \|f(i) - g(j)\|_B^p \right)^{1/p} \\
&= \left\| \frac{1}{n} \sum_{i,j=1}^n a_{ij}(f(i) - g(j)) \right\|_B + \left(\frac{1}{n} \sum_{i,j=1}^n a_{ij} \|f(i) - g(j)\|_B^p \right)^{1/p} \\
&\leq 2 \left(\frac{1}{n} \sum_{i,j=1}^n a_{ij} \|f(i) - g(j)\|_B^p \right)^{1/p} .
\end{aligned}$$

Finally, similar computations yield

$$\begin{aligned}
\left(\frac{1}{n^2} \sum_{i,j=1}^n \|f(i) - g(j)\|_B^p \right)^{1/p} &\leq \left\| \overline{f} - \overline{g} \right\|_B + \left(\frac{1}{n^2} \sum_{i,j=1}^n \|f_0(i) - g_0(j)\|_B^p \right)^{1/p} \\
&\leq \left\| \overline{f} - \overline{g} \right\|_B + \left(\frac{1}{n^2} \sum_{i,j=1}^n 2^{p-1} \left(\|f_0(i)\|_B^p + \|g_0(j)\|_B^p \right) \right)^{1/p} \\
&\leq \left(\frac{1}{n} \sum_{i,j=1}^n a_{ij} \|f(i) - g(j)\|^p \right)^{1/p} + 2 \|h\|_{L_p^{2n}(B)} \\
&\leq \left(1 + \frac{4}{\lambda} \right) \left(\frac{1}{n} \sum_{i,j=1}^n a_{ij} \|f(i) - g(j)\|_B^p \right)^{1/p} ,
\end{aligned}$$

which yields the desired estimate for $\gamma_+(A, \|\cdot\|_B^p)$. □

Definition 4.5. For a metric space (X, d_X) we call a sequence of d-regular graphs $(G_n)_{n \geq 1}$ an expander sequence *with respect to* X if $|V(G_n)| \to \infty$ and if there exists a $p > 0$ such that $\sup_{n \geq 1} \gamma(G_n, d_X^p) < \infty$.

In the euclidean case this is equivalent to $\inf_n \lambda(G_n) > 0$ so our definition of sequences of expanders from Section 2 is just that of being an expander with respect to $\mathbb{R}$ equipped with the kernel given by the square of the euclidean metric.

Remark 4.6. If (G_n) is an expander sequence with resepct to (X, d_X) then, by Proposition 1.18, (G_n) does not coarsely embed into X. Some authors take this latter, weaker, property as the definition of being a sequence of expanders with respect to (X, d_X).

Definition 4.7. A sequence of connected, d-regular graphs, $G_n = (V_n, E_n)$, is called a sequence of *superexpanders* if $|V_n| \to \infty$ and if for every superreflexive Banach space B, there is some $p > 0$ such that $\sup_n \gamma(G_n, \|\cdot\|^p) < \infty$.

Two questions present themselves: First of all, do superexpanders exist? To answer this question in the affirmative, we shall give an explicit construction in Section 5.

Secondly, one might ask whether the class of expanders and that of superexpanders coincide. In light of the results obtained in Section 2 we see (recall, Hilbert spaces are superreflexive) that any sequence of superexpanders is in particular a sequence of expanders, but whether the converse holds is still unresolved. We have already seen Matoušek's extrapolation theorem yield in Section 2 that any sequence of expanders is als a sequence of expanders with respect to the L_p spaces, which are superreflexive for $1 < p < \infty$. We shall explore this topic further in the rest of this section.

4.2 Regular operators and uniformly curved Banach spaces

Uniformly curved Banach spaces were first introduced by Lafforgue and the term was coined by Pisier in [Pis10].

Definition 4.8. Let $1 \leq p_1, p_2 < \infty$. A linear operator $T : L_{p_1}(\Omega_1) \to L_{p_2}(\Omega_2)$, where $(\Omega_1, \mathscr{A}_1, \mu_1)$ and $(\Omega_2, \mathscr{A}_2, \mu_2)$ are measure spaces, is called *regular* if there exists a positve, linear operator $S : L_{p_1}(\mu_1) \to L_{p_2}(\mu_2)$ such that $|Tf| \leq S|f|$, for all $f \in L_{p_1}(\mu_1)$. We define the *regular norm* of T by $\|T\|_{\text{reg}} = \inf\{\|S\| \mid S \geq 0,\ |T(\cdot)| \leq S(|\cdot|)\}$ (it is easily checked that this is really a norm).

Example 4.9. Every linear operator on ℓ_p^n is regular. Indeed, for some linear operator $A = (a_{ij})$ on ℓ_p^n, there is a positive operator majorizing A, which realizes the regular norm of A. This is given by $\text{abs}\,(A) := (|a_{ij}|)$, as is readily seen from the definition.

We have the following alternative characterizations of regularity of an operator:

Proposition 4.10. *For $1 \leq p_1, p_2 < \infty$ and an operator $T : L_{p_1}(\Omega_1) \to L_{p_2}(\Omega_2)$ the following are equivalent:*

1. *T is regular*
2. *There exists $C \geq 0$ such that for all $x_1, \cdots, x_n \in L_{p_1}(\mu_1)$ we have*

$$\left\| \sum_{i=1}^{n} |Tx_i| \right\|_{L_{p_2}(\Omega_2)} \leq C \left\| \sum_{i=1}^{n} |x_i| \right\|_{L_{p_1}(\Omega_1)} \tag{17}$$

3. *T is the difference between two positive operators.*

In this case the optimal constant in (17) is $\|T\|_{reg}$.

The proof of Proposition 4.10 can be found in [Hyt+16, Prop. 2.7.3] and we omit it here. If T is itself positive then clearly $|Tf| \leq T|f|$ for all $f \in L_{p_1}(\Omega_1)$ and hence T is regular with $\|T\|_{\mathrm{reg}} = \|T\|$. Recall also from Proposition 1.50 that positive operators satisfy the extention problem for every Banach space. This is no coincidence; in fact regular operators are precisely those that satisfy the extention problem for every Banach space:

Theorem 4.11. *Let $1 \leq p_1, p_2 < \infty$ and $T : L_{p_1}(\Omega_1) \to L_{p_2}(\Omega_2)$. Then the following are equivalent:*

1. *$T \otimes I_B$ is bounded for every Banach space B.*
2. *$T \otimes I_{\ell_1}$ is bounded.*
3. *T is regular.*

and for every Banach space B we have $\|T \otimes I_B\| \leq \|T\|_{reg} \leq \|T \otimes I_{\ell_1}\|$. In particular $\|T\|_{reg} = \|T \otimes I_{\ell_1}\| = \sup_B \|T \otimes I_B\|$ (the first equality also follows from Proposition 4.10).

Proof. That (1) $\implies$ (2) is trivial.

To show that (2) $\implies$ (3) let $x_1, \cdots, x_n \in L_{p_1}(\mu_1)$. Putting $x_i = 0$ for $i > n$ we have $(x_i)_i \in L_{p_1}(\mu_1; \ell_1)$ and clearly $(T \otimes I_{\ell_1})(x_i)_i = (Tx_i)_i$ (since $(x_i)_i = \sum_{i=1}^n x_i \otimes e_i$, where $(e_i)_i$ denotes the standard basis of ℓ_1). Hence

$$\begin{aligned}\left\| \sum_{i=1}^n |Tx_i| \right\|_{L_{p_2}(\Omega_2)} &= \left\| \|(Tx_i)_i\|_{\ell_1} \right\|_{L_{p_2}(\Omega_2)} = \|(Tx_i)_i\|_{L_{p_2}(\Omega_2;\ell_1)} = \|(T \otimes I_{\ell_1})(x_i)_i\|_{L_{p_2}(\Omega_2;\ell_1)} \\ &\leq \|T \otimes I_{\ell_1}\| \, \|(x_i)_i\|_{L_{p_1}(\Omega_1;\ell_1)} = \|T \otimes I_{\ell_1}\| \left\| \sum_{i=1}^n |x_i| \right\|_{L_{p_1}(\Omega_1)}\end{aligned}$$

which demonstrates, by Proposition 4.10, that T is regular and that $\|T\|_{\mathrm{reg}} \leq \|T \otimes \ell_1\|$.
(3) $\implies$ (1) follows from (the remark succeeding) Proposition 1.50, from which it also follows that $\|T \otimes I_B\| \leq \|T\|_{\mathrm{reg}}$. ☐

Definition 4.12. Let B be a Banach space. For $\varepsilon > 0$, we define

$$\Delta_B(\varepsilon) := \sup \{\|T_B\|\},$$

where the supremum is taken over all pairs of measure spaces $(S_1, \mathscr{A}_1, \mu_1)$, $(S_2, \mathscr{A}_2, \mu_2)$ and $T : L_2(S_1) \to L_2(S_2)$ with $\|T\|_{\mathrm{reg}} \leq 1$ and $\|T\| \leq \varepsilon$. We say that B is *uniformly curved* if $\Delta_B(\varepsilon) \to 0$ as $\varepsilon \to 0$.

The following proposition, found in [Pis10], simplifies the definition of Δ_B significantly:

Proposition 4.13. *Let B be a Banach space, $0 < \varepsilon < 1$ and $\delta > 0$. Then the following are equivalent:*

- *For any pair of measure spaces (Ω_1, S_1, μ_1) and (Ω_2, S_2, μ_2) and any linear operator $T : L_2(\Omega_1) \to L_2(\Omega_2)$ with $\|T\|_{reg} \leq 1$ and $\|T\| \leq \varepsilon$ we have $\|T_B\| \leq \delta$.*
- *For any $n \in \mathbb{N}$ and $n \times n$ matrix A with $\|\mathrm{abs}(A)\| \leq 1$ and $\|A\| \leq \varepsilon$ we have $\|A_B\| \leq \delta$.*

Remark 4.14. If we define

$$\Delta_B^{\text{fin}}(\varepsilon) = \sup \{\|A_B\| \mid n \in \mathbb{N},\ A : \ell_2^n \to \ell_2^n,\ \|\text{abs}(A)\| \le 1,\ \|A\| \le \varepsilon\},$$

then, by Proposition 4.13, we have $\Delta_B^{\text{fin}} = \Delta_B$.

By Theorem 1.52, we have that $\Delta_H(\varepsilon) = \varepsilon$, for every Hilbert space H, and hence Hilbert spaces are uniformly curved. More generally we have the following: A *strictly θ-Hilbertian* Banach space, $0 < \theta \le 1$, is a Banach space which arises as a complex interpolation space $(B, H)_\theta$, where B is any Banach space and H is a Hilbert space.

Proposition 4.15. *If B is a strictly θ-Hilbertian Banach space then $\Delta_B(\varepsilon) \le \varepsilon^\theta$. In particular any strictly θ-Hilbertian Banach space is uniformly curved.*

The proof of this result (found in [Pis10]) requires the Riesz-Thorin interpolation theorem for Bochner spaces (see [BL76, Th.'s 4.1.2 & 5.1.2]):

Theorem 4.16. *Let B_0 and B_1 be a compatible couple of complex Banach spaces, let $(S_0, \mathscr{A}_0, \mu_0)$ and $(S_1, \mathscr{A}_1, \mu_1)$ be measure spaces, and let $1 \le p_0, p_1, q_0, q_1 \le \infty$. Suppose that*

$$T : L_{p_0}(S_0; B_0) + L_{p_1}(S_0; B_1) \to L_{q_0}(S_1; B_0) + L_{q_1}(S_1; B_1)$$

is a linear operator, such that $T : L_{p_j}(S_0; B_j) \to L_{q_j}(S_1; B_j)$, for $j = 0, 1$, with norm at most M_j. Then, for every $0 \le \theta \le 1$, the operator T maps $L_{p_\theta}(S_0; (B_0, B_1)_\theta)$ into $L_{q_\theta}(S_1; (B_0, B_1)_\theta)$ with norm at most $A_0^{1-\theta} A_1^\theta$, where p_θ and q_θ are defined by the relation

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q_\theta} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

Proof of Proposition 4.15. Suppose B is a strictly θ-Hilbertian Banach space, and let V be a Banach space and H a Hilbert space such that $B = (V, H)_\theta$. Let $(\Omega_1, \mathscr{A}_1, \mu_1)$ and $(\Omega_2, \mathscr{A}_2, \mu_2)$ be measure spaces and $T : L_2(\Omega_1) \to L_2(\Omega_2)$ a linear operator with $\|T\|_{\text{reg}} \le 1$ and $\|T\| \le \varepsilon$. By Theorem 1.52 we have that $\|T_H\|_{L_2(\Omega_1;H) \to L_2(\Omega_2;H)} = \|T\|_{L_2(\Omega_1) \to L_2(\Omega_2)} \le \varepsilon$, and by definition of the regular norm we have $\|T_V\|_{L_2(\Omega_1;V) \to L_2(\Omega_2;V)} \le \|T\|_{\text{reg}} \le 1$. Since, in the above notation, we have $p_0 = p_1 = q_0 = q_1 = 2$, and hence $p_\theta = q_\theta = 2$, Theorem 4.16 yields $\|T_B\|_{L_2(\Omega_1;B) \to L_2(\Omega_2;B)} \le 1^{1-\theta} \varepsilon^\theta = \varepsilon^\theta$. This demonstrates that $\Delta_B(\varepsilon) \le \varepsilon^\theta$, as desired. $\square$

Proposition 4.15 implies that the L_p spaces are uniformly curved whenever $1 < p < \infty$. For $p = 1$ or $p = \infty$ they are not, however, since these spaces are not reflexive, and, as we shall see, uniformly curved Banach spaces are superreflexive. To see that uniform curvedness implies superreflexivity we will need a criterion for superreflexivity from [Pis75] which we state without proof:

Proposition 4.17. *For $n \in \mathbb{N}$, let $\Gamma(n)$ be the $n \times n$ matrix given by*

$$\Gamma(n)_{ij} := \begin{cases} \frac{1}{n-i-j} & \text{if } i + j \ne n \\ 0 & \text{if } i + j = n \end{cases}.$$

Let B be a Banach space and suppose that $\|\Gamma(n)_B\| / \log(n) \to 0$. Then B is superreflexive.

It was shown in [Tit26] that $\sup_{n\geq 1} \|\Gamma(n)\|_{B(\ell_2^n)} < \infty$. To give an idea of how these matrices look, let us print $\Gamma(5)$:

$$\Gamma(5) = \begin{pmatrix} 1/3 & 1/2 & 1 & 0 & -1 \\ 1/2 & 1 & 0 & -1 & -1/2 \\ 1 & 0 & -1 & -1/2 & -1/3 \\ 0 & -1 & -1/2 & -1/3 & -1/4 \\ -1 & -1/2 & -1/3 & -1/4 & -1/5 \end{pmatrix}.$$

The following proposition is stated in [Pis10]:

Proposition 4.18. *Suppose B is uniformly curved. Then $\|\Gamma(n)_B\| / \log(n) \to 0$, as $n \to \infty$, and hence B is superreflexive.*

Proof. As observed above, there exists $C \geq 1$ such that $\|\Gamma(n)\| \leq C$, for all $n \in \mathbb{N}$. Recall from Example 4.9 that the regular norm of $\Gamma(n)$ is given by the operator norm of $\mathrm{abs}(\Gamma(n))$. Observe that $\mathrm{abs}(\Gamma(n)) = \sum_{k=2}^{2n} \Gamma_k(n)$, where $\Gamma_n(n) = 0$ and $\Gamma_k(n)$ is given by

$$\Gamma_k(n)_{ij} = \begin{cases} \frac{1}{|n-k|} & \text{if } i+j=k \\ 0 & \text{if } i+j \neq k \end{cases},$$

when $k \neq n$. When $k \neq n$, we see that $\Gamma_k(n)$ has all zero entries, except on the $(k-1)$st North-East to South-West diagonal, all of whose entries are equal to $1/|n-k|$. It follows, e.g. by applying $\Gamma_k(n)$ to the standard orthonormal basis of ℓ_2^n, that $\|\Gamma_k(n)\| = 1/|n-k|$. Hence, by the triangle inequality, we obtain

$$\|\mathrm{abs}(\Gamma(n))\| \leq \left(\frac{1}{n-2} + \frac{1}{n-3} + \cdots + 1 + 0 + 1 + \frac{1}{2} + \cdots + \frac{1}{n}\right) \leq 2\sum_{k=1}^{n} \frac{1}{k} \in O(\log(n)).$$

Together with the initial observation this allows us to conlclude that there exists $C' \geq 1$ and $N \in \mathbb{N}$ such that $\|\Gamma(n)\| \leq C'$ and such that $\|\Gamma(n)\|_{\mathrm{reg}} \leq C' \log(n)$, whenever $n \geq N$. Hence, for such n, we have $\left\|\frac{1}{C'\log(n)}\Gamma(n)\right\| \leq 1/\log(n)$ and $\left\|\frac{1}{C'\log(n)}\Gamma(n)\right\|_{\mathrm{reg}} \leq 1$, which implies, by definition of Δ_B, that

$$\|\Gamma(n)_B\| \leq C' \log(n) \Delta_B \left(\frac{1}{\log(n)}\right).$$

Since $\Delta_B(\varepsilon) \to 0$ as $\varepsilon \to 0$, by assumption, the conclusion follows. □

So far we have seen that uniformly curved Banach spaces constitute a rich subclass of the class of superreflexive Banach spaces, and it is in fact not known whether it exhausts this class or not. In terms of embeddability of expander sequences, we have the following result:

Proposition 4.19. *Suppose $(G_n)_{n\geq 1}$ is a sequence of finite, connected, d-regular graphs such that $|V_n| \to \infty$ and such that $\inf_n \lambda_+(G_n) > 0$, where $\lambda_+(G_n)$ was defined in Section 1 (in particular (G_n) is a sequence of expanders). Let B be a uniformly curved Banach space. Then $\sup_n \gamma(G_n, \|\cdot\|^2) < \infty$, i.e., (G_n) is an expander sequence with respect to B.*

Proof. Let $A_n = (a^n_{u,v})$ denote the (normalized) adjacency matrix of G_n. Then we can decompose $\ell_2(V_n) = \ell_2^0(V_n) \oplus \ell_2^c(V_n)$ into a direct sum of the mean-zero functions and the constant functions. Clearly $\ell_2^c(V_n) = \mathbb{C} \cdot 1_{V_n}$ is the eigenspace of the eigenvalue 1, so we have $1 - \lambda_+(G_n) = \|A_n\|_{B(\ell_2^0(V_n))}$. For $f \in \ell_2(V_n)$ let $E_n f = \sum_{u \in V_n} f(u)/|V_n|$ denote the average of f. We can view this as an element of $\mathbb{C} \cdot 1_{V_n} \subset \ell_2(V_n)$, and $E_n : \ell_2(V_n) \to \mathbb{C} \cdot 1_{V_n}$ is the orthogonal projection. Let $\gamma = \sup_n(1 - \lambda_+(G_n)) = 1 - \inf_n \lambda_+(G_n) < 1$. Then $\|A_n\|_{B(\ell_2^0(V_n))} \leq \gamma < 1$, for every $n \geq 1$, and by a straightforward computation, we have $A_n E_n = E_n A_n = E_n$. Hence

$$\|A_n f - E_n f\|_2 \leq \gamma \|f - E_n f\|_2 \leq \gamma \|f\|_2 ,$$

for all $n \geq 1$ and $f \in \ell_2(V_n)$, demonstrating that $\|A_n - E_n\|_{B(\ell_2(V_n))} \leq \gamma$. We also see, since A_n clearly also commutes with $I_n - E_n$, that $A_n^k - E_n = A_n^k(I_n - E_n) = (A_n - E_n)^k$, and hence $\left\|A_n^k - E_n\right\|_{B(\ell_2(V_n))} \leq \gamma^k$, for every $k \in \mathbb{N}$. Note that $\|A_n\|_{\mathrm{reg}} = \|A_n\|_{B(\ell_2(V_n))} \leq 1$ and $\|E_n\|_{\mathrm{reg}} = \|E_n\|_{B(\ell_2(V_n))} \leq 1$, since A_n and E_n are both positive contractions, and hence $\left\|A_n^k - E_n\right\|_{\mathrm{reg}} \leq 2$, for every $k \in \mathbb{N}$. Since $\gamma < 1$ and we have by assumption that $\Delta_B(\varepsilon) \to 0$ as $\varepsilon \to 0$, we can choose $k \in \mathbb{N}$ such that $\delta := 2\Delta_B(\gamma^k/2) < 1$. Then $\left\|\frac{1}{2}(A_n^k - E_n)\right\|_{B(\ell_2(V_n))} \leq \gamma^k/2$ and $\left\|\frac{1}{2}(A_n^k - E_n)\right\|_{\mathrm{reg}} \leq 1$, and hence, by definition of Δ_B, we infer that $\left\|A_n^k - E_n\right\|_{B(\ell_2(V_n;B))} \leq \delta < 1$. This yields

$$\|f - E_n f\|_2 \leq \left\|f - A_n^k f\right\|_2 + \left\|A_n^k f - E_n f\right\|_2 \leq \left\|f - A_n^k f\right\|_2 + \delta \|f - E_n f\|_2 ,$$

for any $f : V_n \to B$, and hence $\|f - E_n f\|_2 \leq \frac{1}{1-\delta} \left\|f - A_n^k f\right\|_2$. Since

$$\frac{1}{|V_n|} \sum_{u,v \in V_n} \|f(u) - f(v)\|_B^2 \leq 4 \sum_{u \in V} \|f(u) - E_n f\|_B^2 = 4 \|f - E_n f\|_2^2 ,$$

we obtain from the above that

$$\begin{aligned}
\frac{1}{|V_n|} \sum_{u,v \in V_n} \|f(u) - f(v)\|_B^2 &\leq \left(\frac{2}{1-\delta}\right)^2 \left\|f - A_n^k f\right\|_2^2 \\
&= \left(\frac{2}{1-\delta} \left\|\sum_{i=0}^{k-1} A_n^i (f - A_n f)\right\|_2\right)^2 \\
&\leq \left(\frac{2}{1-\delta} \sum_{i=0}^{k-1} \|f - A_n f\|_2\right)^2 \\
&= \left(\frac{2k}{1-\delta}\right)^2 \|f - A_n f\|_2^2 \\
&\leq \left(\frac{2k}{1-\delta}\right)^2 \sum_{u,v \in V_n} a^n_{u,v} \|f(u) - f(v)\|^2 .
\end{aligned}$$

Since k and δ are independent of n this demonstrates that $\gamma(G_n, \|\cdot\|^2) \leq 4k^2/(1-\delta)^2$ for all $n \geq 1$ which is what we wanted. $\square$

Remark 4.20. By Proposition 1.18 we get as a consequence of the above proposition that if a graph sequence satisfies $\inf_{n\geq 1} \lambda_+(G_n) > 0$, then $(G_n)_{n\geq 1}$ does not coarsely embed into uniformly curved Banach space. It still remains to consider graph sequences which satisfy $\inf_n \lambda(G_n) > 0$ but such that $\inf_n \lambda_+(G_n) = 0$. This can only happen if $\inf_n \lambda_n = -1$, where λ_n denotes the smallest eigenvalue of the adjacency matrix of G_n. We can think of such sequences as 'approximately bipartite' expander sequences (recall Proposition 1.4).

The above piece of intuition (obtained in [Ost13, chap. 6]) yields the following proposition:

Proposition 4.21. *Suppose (G_n) is a sequence of connected, d-regular graphs satisfying $\inf_n \lambda(G_n) > 0$. Then there exists a sequence of $(d + d^2)$-regular graphs (H_n) on the same vertex sets satisfying $\inf_n \lambda_+(H_n) > 0$ such that H_n embeds coarsely into G_n for each $n \geq 1$ with control functions independent of n.*

Proof. The edge set of H_n is constructed in the following way: Between each pair of vertices $u, v \in V(H_n) = V(G_n)$ we draw an edge in H_n for each edge *and* for each path of length 2 in G_n connecting u and v. Clearly each H_n is connected and $(d + d^2)$-regular and if A_n is the non-normalized adjacency matrix of G_n then $B_n = \frac{1}{d+d^2}(A_n + A_n^2)$ is the normalized adjacency matrix of H_n. Denote the eigenvalues of A_n by $d = \lambda_1^{(n)} > \lambda_2^{(n)} \geq \cdots \geq \lambda_{k(n)}^{(n)} \geq -d$ and let $k(n) = |V(H_n)| = |V(G_n)|$. Our assumption that $\inf_n \lambda(G_n) > 0$ then means precisely that $\sup_n \lambda_2^{(n)} < d$. By the spectral theorem the eigenvalues of B_n are

$$1 = \frac{\lambda_1^{(n)} + (\lambda_1^{(n)})^2}{d + d^2}, \frac{\lambda_2^{(n)} + (\lambda_2^{(n)})^2}{d + d^2}, \cdots, \frac{\lambda_{k(n)}^{(n)} + (\lambda_{k(n)}^{(n)})^2}{d + d^2},$$

and hence

$$\lambda_+(H_n) = 1 - \max_{2\leq i\leq k(n)} \frac{|\lambda_i^{(n)} + (\lambda_i^{(n)})^2|}{d + d^2}.$$

By the monotonicity properties of the function $t \mapsto t + t^2$ there are three cases to consider:

- If $d \geq \lambda_i^{(n)} > 0$ then also $\lambda_2^{(n)} > 0$ and
$$|\lambda_i^{(n)} + (\lambda_i^{(n)})^2| = \lambda_i^{(n)} + (\lambda_i^{(n)})^2 \leq \lambda_2^{(n)} + (\lambda_2^{(n)})^2 \leq \sup_n \lambda_2^{(n)} + (\sup_n \lambda_2^{(n)})^2 < d + d^2.$$
- If $-1 \leq \lambda_i^{(n)} \leq 0$ then $|\lambda_i^{(n)} + (\lambda_i^{(n)})^2| \leq 1/4 < d + d^2$.
- If $-d \leq \lambda_i^{(n)} < -1$ then
$$|\lambda_i^{(n)} + (\lambda_i^{(n)})^2| = \lambda_i^{(n)} + (\lambda_i^{(n)})^2 \leq -d + d^2 < d + d^2.$$

Hence

$$\lambda_+(H_n) \geq 1 - \frac{\max\left\{\sup_n \lambda_2^{(n)} + (\sup_n \lambda_2^{(n)})^2, 1/4, -d + d^2\right\}}{d + d^2} > 0,$$

for all $n \geq 1$, demonstrating that $\inf_n \lambda_+(H_n) > 0$. For the coarse embeddability of H_n into G_n observe that, since G_n is a subgraph of H_n we have $d_{G_n}(u,v) \geq d_{H_n}(u,v)$ for every $u, v \in V(H_n)$. Moreover every edge in a path in H_n between vertices u and v is either an edge in G_n or can be replaced by two edges in G_n forming a path between its end points. So every such path admits a path in G_n between u and v of at most double the lenght of the original path. Hence $d_{G_n}(u,v) \leq 2d_{H_n}(u,v)$. This demonstrates that the identity map $V(H_n) \to V(G_n)$ forms a coarse embedding with control functions $\rho_-(t) = t/2$ and $\rho_+(t) = t$, which are independent of n. □

Since the control functions of the above proposition were independent of n and compositions of coarse embeddings are again coarse embeddings we obtain:

Corollary 4.22. *Suppose (X, d_X) is a metric space such that any sequence (G_n) of finite, connected, regular graphs with $|V(G_n)| \to \infty$ and $\inf_n \lambda_+(G_n) > 0$ does not coarsely embed into X. Then also any sequence with $|V(G_n)| \to \infty$ and $\inf_n \lambda(G_n) > 0$ does not coarsely embed into X. In particular sequences of expanders do not coarsely embed into uniformly curved Banach space.*

The above corollary shows that an expander sequence is an expander sequence with respect to any uniformly curved Banach space in the weaker sense of Remark 4.6. Furthermore, if we want to prove coarse non-embeddability results for graph sequences with $\sup_{n\geq 1} \lambda(G_n) < 1$ then it always suffices to prove non-embeddability only for sequences satisfying $\sup_{n\geq 1} \lambda_+(G_n) < 1$.

4.3 Sphere equivalence and invariance

Definition 4.23. Two Banach spaces B and V are called sphere equivalent if there exists a uniformly continuous homeomorphism $\varphi : S(B) \to S(V)$, i.e., a uniformly continuous bijection with uniformly continuous inverse. Here $S(B)$ and $S(V)$ denote the unit spheres of B and V respectively.

We write $B \sim_S V$ to signify that B and V are sphere equivalent, and it is clear that this relation is an equivalence relation. We denote the class of spaces which are sphere equivalent to B by $[B]_S$. The following classical theorem due to [Maz29] (see also [BL00, Th. 9.1]) gives an important class of examples of sphere equivalent spaces:

Theorem 4.24. *Let (Ω, μ) be a measure space and $1 < p, q < \infty$. Then the Mazur map $M_{p,q} : L_p(\Omega) \to L_q(\Omega)$, given by $M_{p,q}(f) = |f|^{p/q-1} f$ if $f \neq 0$ and $M_{p,q}(0) = 0$, restricts to a Lipschitz function between $S(L_p(\Omega))$ and $S(L_q(\Omega))$ - in particular it is uniformly continuous. Since $M_{q,p}$ is a two-sided inverse of $M_{p,q}$ we have in particular that $L_p(\Omega) \sim_S L_q(\Omega)$.*

This theorem generalizes to Bochner spaces:

Lemma 4.25. *[LS17, Lemma 3.10] Let (Ω, μ) be a measure space, B a Banach space and $1 < p, q < \infty$. Define the Mazur map $M_{p,q} : L_p(\Omega; B) \to L_q(\Omega; B)$ by $M_{p,q}(f) = \|f\|_B^{p/q-1} \cdot f$ if $f \neq 0$ and $M_{p,q}(0) = 0$. Then there exists a constant $C = C_{p,q} > 0$ such that*

$$\|M_{p,q}(f_1) - M_{p,q}(f_2)\|_{L_q(\Omega;B)} \leq C \|f_1 - f_2\|_{L_p(\Omega;B)}^{\min\{p/q,1\}},$$

whenever $\|f_1\|_{L_p(\Omega;B)}, \|f_2\|_{L_p(\Omega;B)} \leq 1$. In particular $L_p(\Omega; B) \sim_S L_q(\Omega; B)$.

Proof. Write $f_i = r_i\varphi_i$ with $r_i : \Omega \to \mathbb{R}_+$ and $\varphi_i : \Omega \to S(B)$. Then $r_i \in S(L_p(\Omega))$ (since $r_i = \|f_i\|_B$) and hence, by Theorem 4.24, we have $\left\|r_1^{p/q} - r_2^{p/q}\right\|_{L_q(\Omega)} \le C\|r_1 - r_2\|_{L_p(\Omega)}$, for some $C = C_{p,q} > 0$. Moreover, the reverse triangle inequality yields that

$$\|r_1 - r_2\|_{L_p(\Omega)} = \left(\int |\,\|f_1\|_B - \|f_2\|_B\,|^p d\mu\right)^{1/p} \le \left(\int \|f_1 - f_2\|_B^p\, d\mu\right)^{1/p} = \|f_1 - f_2\|_{L_p(\Omega;B)}.$$

Since $M_{p,q}(f_1) - M_{p,q}(f_2) = r_1^{p/q}\varphi_1 - r_2^{p/q}\varphi_2 = (r_1^{p/q} - r_2^{p/q})\varphi_1 + r_2^{p/q}(\varphi_1 - \varphi_2)$ we obtain that

$$\begin{aligned}\|M_{p,q}(f_1) - M_{p,q}(f_2)\|_{L_q(\Omega;B)} &\le \left\|r_1^{p/q} - r_2^{p/q}\right\|_{L_q(\Omega)} + \left\|r_2^{p/q}(\varphi_1 - \varphi_2)\right\|_{L_q(\Omega;B)} \\ &\le C\|f_1 - f_2\|_{L_p(\Omega;B)} + \left\|r_2^{p/q}(\varphi_1 - \varphi_2)\right\|_{L_q(\Omega;B)}. \qquad (18)\end{aligned}$$

If $q \le p$ then Hölder's inequality yields the following estimate for the second term above:

$$\begin{aligned}\left\|r_2^{p/q}(\varphi_1 - \varphi_2)\right\|_{L_q(\Omega;B)}^q &= \int r_2^p \|\varphi_1 - \varphi_2\|_B^q\, d\mu = \int \|\varphi_1 - \varphi_2\|_B^q\, 1_\Omega d(r_2^p \cdot \mu) \\ &\le \left(\int \left(\|\varphi_1 - \varphi_2\|_B^q\right)^{p/q} d(r_2^p \cdot \mu)\right)^{q/p} \left(\int 1_\Omega d(r_2^p \cdot \mu)\right)^{(p-q)/p} \\ &= \|r_2(\varphi_1 - \varphi_2)\|_{L_p(\Omega;B)}^q \cdot \|r_1\|_{L_p(\Omega)}^{p-q} \\ &\le \|r_2(\varphi_1 - \varphi_2)\|_{L_p(\Omega;B)}^q \\ &\le \left(\|r_2\varphi_1 - r_1\varphi_1\|_{L_p(\Omega;B)} + \|r_1\varphi_1 - r_2\varphi_2\|_{L_p(\Omega;B)}\right)^q \\ &\le \left(2\|f_1 - f_2\|_{L_p(\Omega;B)}\right)^q,\end{aligned}$$

so (18) yields

$$\|M_{p,q}(f_1) - M_{p,q}f(2)\|_{L_q(\Omega;B)} \le (C+2)\|f_1 - f_2\|_{L_p(\Omega;B)} = (C+2)\|f_1 - f_2\|_{L_p(\Omega;B)}^{\min\{p/q,1\}}.$$

If $q > p$ then, since

$$\|\varphi_1 - \varphi_2\|_B^q \le \|\varphi_1 - \varphi_2\|_{L_\infty(\Omega, r_2^p\cdot\mu;B)}^{q-p} \|\varphi_1 - \varphi_2\|_B^p, \quad (r_2^p \cdot \mu) - \text{a.s.},$$

we obtain

$$\begin{aligned}\|r_2^p(\varphi_1 - \varphi_2)\|_{L_q(\Omega;B)} &\le \|\varphi_1 - \varphi_2\|_{L_q(\Omega, r_2^p\cdot\mu;B)} \\ &\le \|\varphi_1 - \varphi_2\|_{L_\infty(\Omega, r_2^p\cdot\mu;B)}^{1-p/q} \cdot \|\varphi_1 - \varphi_2\|_{L_p(\Omega, r_2^p\cdot\mu;B)}^{p/q} \\ &\le 2^{1-p/q}\|r_2(\varphi_1 - \varphi_2)\|_{L_p(\Omega;B)}^{p/q} \\ &\le 2^{1-p/q}\left(2\|f_1 - f_2\|_{L_p(\Omega;B)}\right)^{p/q} = 2\|f_1 - f_2\|_{L_p(\Omega;B)}^{p/q}.\end{aligned}$$

This estimate together with (18) yields

$$\begin{aligned}\|M_{p,q}(f_1) - M_{p,q}(f_2)\|_{L_q(\Omega;B)} &\le (C2^{1-p/q} + 2)\|f_1 - f_2\|_{L_p(\Omega;B)}^{p/q} \\ &= (C2^{1-p/q} + 2)\|f_1 - f_2\|_{L_p(\Omega;B)}^{\min\{p/q,1\}},\end{aligned}$$

concluding the proof. □

We now follow Mimura's proof in [Mim18] of the following theorem and, using Lemma 4.25, its corollary:

Theorem 4.26. *Let B and V be Banach spaces, $p \geq 1$ and (G_n) a sequence of finite, connected, d-regular graphs. If B is sphere equivalent to V, then $\sup_n \gamma(G_n, \|\cdot\|_B^p) < \infty$ if and only if $\sup_n \gamma(G_n, \|\cdot\|_V^p) < \infty$.*

Corollary 4.27. *Let B and (G_n) be as above and $p, q > 1$. Then $\sup_n \gamma(G_n, \|\cdot\|_B^p) < \infty$ if and only if $\sup_n \gamma(G_n, \|\cdot\|_B^q) < \infty$.*

These two results serve as a generalization of Matoušek's extrapolation theorem (Theorem 2.4) since, by Theorem 4.24, $L_p(\Omega) \sim_S L_2(\Omega)$ for any $p > 1$.

Definition 4.28. Let B be a Banach space, $p \geq 1$ and $G = (V, E)$ a finite, connected, d-regular graph. We define the Banach space spectral gap of G with respect to (B, p) by

$$\lambda_1(G, B, p) := \frac{1}{2} \inf \frac{\sum_{(u,v)\in E} \|f(u) - f(v)\|^p}{\sum_{u\in V} \|f(u) - m(f)\|^p}, \tag{19}$$

where $m(f) := \sum_{u\in V} f(u)/|V|$ denotes the mean of f and the infimum is taken over all non-constant functions $f : V \to B$.

Using convexity of $t \mapsto t^p$ and Hölder's inequality, it is not difficult to see that the Banach space spectral gap of [Mim18] relates to the reciprocal spectral gap of [MN14] by

$$\frac{d}{2\lambda_1(G, B, p)} \leq \gamma(G, \|\cdot\|^p) \leq \frac{2^p d}{2\lambda_1(G, B, p)}.$$

Hence $\sup_n \gamma(G_n, \|\cdot\|^p) < \infty$ is equivalent to $\inf_n \lambda_1(G_n, B, p) > 0$. This demonstrates that the above definition and terminology is consistent with the theory we have developed so far and we pass to the Banach space spectral gap in this subsection for notational convenience. It follows from the definition that if $E \subset B$ is a closed subspace then $\lambda_1(G, E, p) \geq \lambda_1(G, B, p)$ and so in particular $\lambda_1(G, B, p) \geq \lambda_1(G, L_p(\Omega; B), p)$ for any measure space $(\Omega, \mathscr{A}, \mu)$. Using the estimate from (19) pointwise and integrating yields the reverse inequality. We summarize this observation in a lemma:

Lemma 4.29. *For any measure space $(\Omega, \mathscr{A}, \mu)$ we have $\lambda_1(G, B, p) = \lambda_1(G, L_p(\Omega; B), p)$. In particular, if we let $B_p := \ell_p(\mathbb{N}; B)$, we have $\lambda_1(G, B, p) = \lambda_1(G, B_p, p)$.*

If F is any countable set then the symmetric group on F, $\mathrm{Sym}(F)$, acts by isometries on $\ell_r(F; B)$ for any $r \geq 1$ by $\sigma.\xi(a) = \xi(\sigma^{-1}(a))$ for $\sigma \in \mathrm{Sym}(F)$, $\xi \in \ell_r(F; B)$ and $a \in F$. A map $\varphi : S \to \ell_q(F; B)$, where $S \subset \ell_p(F; B)$ and $p, q \geq 1$, is then called $\mathrm{Sym}(F)$-invariant if $\sigma.\varphi(\xi) = \varphi(\sigma.\xi)$ for all $\xi \in S$ and $\sigma \in \mathrm{Sym}(F)$.
When F is finite then $\xi \in \ell_r(F; B)$ is constant precisely when $\sigma.\xi = \xi$ for all $\sigma \in \mathrm{Sym}(F)$, and hence if $\xi \in S \subset \ell_r(F; B)$ is constant and $\varphi : S \to \ell_q(F; B)$ is $\mathrm{Sym}(F)$-invariant, we have for any $\sigma \in \mathrm{Sym}(F)$ that $\sigma.\varphi(\xi) = \varphi(\sigma.\xi) = \varphi(\xi)$, demonstrating that $\varphi(\xi)$ is constant. Moreover, it is clear that if φ is bijective and $\mathrm{Sym}(F)$-invariant then φ^{-1} is $\mathrm{Sym}(F)$-invariant as well. These elementary observations will come in handy in the subsequent arguments.

Definition 4.30. For Banach spaces B and V and a map $\varphi : S(B) \to S(V)$ we define the class of upper moduli of continuity $\mathcal{M}_\varphi$ consisting of all functions $\delta : [0,2] \to [0,2]$ that are non-decreasing such that $\lim_{\varepsilon \searrow 0} \delta(\varepsilon) = 0$, and which satisfy $\|\varphi(x) - \varphi(y)\|_V \leq \delta(\|x-y\|_B)$ for all $x, y \in B$. Then φ is uniformly continuous if and only if $\mathcal{M}_\varphi \neq \emptyset$.

For $\varphi : S(B) \to S(V)$ we may also define its canonical extension $\bar{\varphi} : B \to V$ by $\bar{\varphi}(0) = 0$ and $\bar{\varphi}(x) = \|x\|_B \, \varphi(x/\|x\|_B)$ for $x \neq 0$.

Proposition 4.31. *Suppose $\varphi : S(B) \to S(V)$ is a uniformly continuous map. Then the map $\Phi_p : S(B_p) \to S(V_p)$ given by $\Phi_p((x_i)_{i \geq 1}) = (\bar{\varphi}(x_i))_{i \geq 1}$ is also uniformly continuous. Moreover it is (by construction) Sym$(\mathbb{N})$-invariant.*

Proof. First observe that, by definition of $\overline{\varphi}$, we do have $\Phi_p(x) \in S(V_p)$ whenever $x \in S(B_p)$. To prove uniform continuity, take $\delta \in \mathcal{M}_\varphi$ and observe that we can assume that δ is concave. Furthermore, it can be checked that we can even assume that there exists $D > 0$ such that $\delta(t)^p \leq D\delta(t^p)$ whenever $t \in [0, 2^{1/p}]$. Now take $x = (x_i)_{i\geq 1}, y = (y_i)_{i \geq 1} \in S(B_p)$. First assume that $\|x_i\|_B = \|y_i\|_B$, for every $i \in \mathbb{N}$. Let $r_i = \|x_i\|_B$ and choose ε_i such that $\varepsilon_i r_i = \|x_i - y_i\|_B$. Note that it is possible to pick ε_i this way, since if $r_i = 0$ then so is $\|x_i - y_i\|_B$. Then $\sum_i r_i^p = 1$ and hence by concavity of δ and Jensen's inequality we obtain

$$\|\Phi_p(x) - \Phi_p(y)\|_{V_p}^p = \sum_{i : r_i \neq 0} r_i^p \, \|\varphi(x_i/r_i) - \varphi(y_i/r_i)\|_V^p$$
$$\leq \sum_{i=1}^\infty r_i^p \delta(\varepsilon_i)^p \leq D \sum_{i=1}^\infty r_i^p \delta\left(\varepsilon_i^p\right) \leq \delta\left(\sum_{i=1}^\infty (r_i \varepsilon_i)^p\right) = \delta\left(\|x-y\|_{B_p}^p\right),$$

whenever $\|x - y\|_{B_p} \leq 2^{1/p}$. For general $x, y \in S(B_p)$ define $z = (z_i)_{i\geq 1}$ by $z_i = \frac{\|x_i\|_B}{\|y_i\|_B} y_i$ if $y_i \neq 0$ and $z_i = x_i$ otherwise. By the reverse triangle inequality we have

$$\|z_i - y_i\|_B = \|y_i\|_B \, | \, \|x_i\|_B / \|y_i\|_B - 1| = | \, \|x_i\|_B - \|y_i\|_B \, | \leq \|x_i - y_i\|_B \, ,$$

when $y_i \neq 0$, and by the way the z_i's were defined, we have equality when $y_i = 0$. In any case, we obtain, since $\|z_i\|_B = \|x_i\|_B$, for every $i \in \mathbb{N}$, that

$$\|\Phi_p(x) - \Phi_p(z)\|_{V_p}^p \leq \delta\left(\|x - z\|_{B_p}^p\right) \leq \delta\left(2^{p-1}\left(\|x-y\|_{B_p} + \|y - z\|_{B_p}\right)\right) \leq \delta\left(2^p \, \|x-y\|_{B_p}^p\right),$$

whenever $\|x - y\|_{B_p} \leq 2^{1/p - 1}$. Moreover, by definition of z and Φ_p, we can check that $\|\Phi_p(y) - \Phi_p(z)\|_{V_p} \leq \|x - y\|_{B_p}$, and hence

$$\|\Phi_p(x) - \Phi_p(y)\|_{V_p} \leq \delta\left(\left(2 \, \|x-y\|_{B_p}\right)^p\right)^{1/p} + \|x - y\|_{B_p} \, ,$$

whenever $\|x - y\|_{B_p} \leq 2^{1/p-1}$. This demonstrates that Φ_p is uniformly continuous. $\square$

We now turn our attention to Schreier coset graphs (see Section 3) which turn out, as we shall see, to constitute quite a large subclass of all d-regular graphs.

Lemma 4.32. *Let Γ be a finitely generated group, $\Lambda < \Gamma$ of finite index, and S a finite, symmetric generating subset of Γ not containing the identity. For a Banach space B and $p \geq 1$ define the p-displacement constant by*

$$\kappa = \kappa_{B,p}(\Gamma, \Lambda, S) := \inf_{\xi \notin \ell_p^c(\Gamma/\Lambda; B_p)} \max_{s \in S} \frac{\|s.\xi - \xi\|_p}{\|\xi - m(\xi)\|_p}, \tag{20}$$

where $\ell_p^c(\Gamma/\Lambda; B_p) \subset \ell_p(\Gamma/\Lambda; B_p)$ denotes the subspace of constant functions. Then if G is the Schreier coset graph $Sch(\Gamma, \Lambda, S)$, we have $\kappa^p \leq \lambda_1(G, B, p) \leq \frac{|S|}{2}\kappa^p$.

Proof. Recall that $\lambda_1(G, B, p) = \lambda_1(G, B_p, p)$. Then by definition of $\mathrm{Sch}(\Gamma, \Lambda, S)$ we have

$$\begin{aligned}
\lambda_1(G, B_p, p) &= \inf_{f \text{ non-constant}} \frac{1}{2} \frac{\sum_{(u,v) \in E} \|f(u) - f(v)\|_{B_p}^p}{\sum_{u \in \Gamma/\Lambda} \|f(u) - m(f)\|_{B_p}^p} \\
&= \inf_{\xi \notin \ell_p^c(\Gamma/\Lambda, B_p)} \frac{1}{2} \frac{\sum_{x\Lambda \in \Gamma/\Lambda} \sum_{s \in S} \left\|\xi(s^{-1}x\Lambda) - \xi(x\Lambda)\right\|_{B_p}^p}{\sum_{x\Lambda \in \Gamma/\Lambda} \|\xi(x\Lambda) - m(\xi)\|_{B_p}^p} \\
&= \inf_{\xi \notin \ell_p^c(\Gamma/\Lambda, B_p)} \frac{1}{2} \sum_{s \in S} \left(\frac{\|s.\xi - \xi\|_p}{\|\xi - m(\xi)\|_p} \right)^p.
\end{aligned}$$

Since $\|s.\xi - \xi\|_p = \left\|s^{-1}.\xi - \xi\right\|_p$ and S is symmetric we have

$$\frac{|S|}{2} \max_{s \in S} \left(\frac{\|s.\xi - \xi\|_p}{\|\xi - m(\xi)\|_p} \right)^p \geq \frac{1}{2} \sum_{s \in S} \left(\frac{\|s.\xi - \xi\|_p}{\|\xi - m(\xi)\|_p} \right)^p \geq \max_{s \in S} \left(\frac{\|s.\xi - \xi\|_p}{\|\xi - m(\xi)\|_p} \right)^p,$$

which gives the desired estimate. □

Lemma 4.33. *Let B and V are Banach spaces and $p, q \geq 1$. Suppose $\Phi : S(B_p) \to S(V_q)$ is a $Sym(\mathbb{N})$-invariant uniform homeomorphism. For Γ, Λ and S as above we then have*

$$\kappa_{B,p}(\Gamma, \Lambda, S) \geq \delta_1^{-1} \left(\frac{1}{2} \delta_2^{-1} \left(\frac{1}{2} \right) \kappa_{V,q}(\Gamma, \Lambda, S) \right), \tag{21}$$

for every $\delta_1 \in \mathcal{M}_\Phi$ and $\delta_2 \in \mathcal{M}_{\Phi^{-1}}$. Here, if δ_1 or δ_2 are not bijective we interpret $\delta_1^{-1}(s)$ as $\inf\{t \in [0,2] \mid \delta_1(t) \geq s\}$ and similarly with δ_2.

Proof. First observe that B_p is isometrically isomorphic to $\ell_p(\Gamma/\Lambda, B_p)$ (since Γ/Λ is finite) and similarly with V_q so we may regard $\Phi : S(\ell_p(\Gamma/\Lambda, B_p)) \to S(\ell_q(\Gamma/\Lambda, V_q))$ a $\mathrm{Sym}(\Gamma/\Lambda)$-invariant uniform homeomorphism. For notational convenience, put $\kappa_p = \kappa_{B,p}(\Gamma, \Lambda, S)$ and $\kappa_q = \kappa_{V,q}(\Gamma, \Lambda, S)$. Let $\xi \in S(\ell_p(\Gamma/\Lambda, B_p))$ be non-constant and observe that, by passing to $\xi - m(\xi)$, we may assume that ξ has mean zero. Put $\eta = \Phi(\xi) \in S(\ell_q(\Gamma/\Lambda, V_q))$. Observe that by $\mathrm{Sym}(\Gamma/\Lambda)$-invariance we have $\Phi(S(\ell_p^c(\Gamma/\Lambda, B_p))) \subset S(\ell_q^c(\Gamma/\Lambda, V_q))$, and similarly for Φ^{-1}. This in particular implies that $\eta \notin \ell_q^c(\Gamma/\Lambda, V_q)$. Now for any $\zeta \in \ell_p^c(\Gamma/\Lambda, B_p)$ we have

$$\begin{aligned}
\|\zeta\|_p^p = \sum_{u \in \Gamma/\Lambda} \|\zeta(u)\|_{B_p}^p &= \sum_{u \in \Gamma/\Lambda} \left\| \frac{1}{|\Gamma/\Lambda|} \sum_{v \in \Gamma/\Lambda} (\zeta(v) - \xi(v)) \right\|_{B_p}^p \\
&\leq \sum_{v \in \Gamma/\Lambda} \|\zeta(v) - \xi(v)\|_{B_p}^p = \|\zeta - \xi\|_p^p,
\end{aligned}$$

and hence $1/2 = \|\xi\|_p/2 \leq (\|\zeta-\xi\|_p + \|\zeta\|_p)/2 \leq \|\zeta-\xi\|_p$, which demonstrates that $\operatorname{dist}(\xi, \ell_p^c(\Gamma/\Lambda, B_p)) \geq 1/2$. Hence, by uniform continuity of Φ^{-1} we obtain the estimate $\operatorname{dist}(\eta, S(\ell_q^c(\Gamma/\Lambda, V_q))) \geq \delta_2^{-1}(1/2)$. If $m(\eta) \neq 0$ we can view it as an element of $\ell_q^c(\Gamma/\Lambda, V_q)$ and put $\eta' = m(\eta)/\|m(\eta)\|_p \in S(\ell_q^c(\Gamma/\Lambda, V_q))$. By the above estimate, we then have $\|\eta-\eta'\|_q \geq \delta_2^{-1}(1/2)$. Also by the reverse triangle inequality we obtain

$$\left\|m(\eta)-\eta'\right\|_q = |\,\|m(\eta)\|_q - 1| = |\,\|m(\eta)\|_q - \|\eta\|_q\,| \leq \|\eta - m(\eta)\|_q\,,$$

which yields

$$\|\eta-m(\eta)\|_q \geq \frac{\|\eta-m(\eta)\|_q + \|m(\eta)-\eta'\|_q}{2} \geq \frac{\|\eta-\eta'\|_q}{2} \geq \frac{\delta_2^{-1}(1/2)}{2}.$$

If $m(\eta)=0$ similar (but even simpler) computations may be carried through to obtain $\|\eta\|_q \geq \delta_2^{-1}(1/2)/2$, by just letting η' be any element of $S(\ell_q^c(\Gamma/\Lambda, V_q))$. It follows, by definition of the q-displacement constant of V, that

$$\max_{s\in S}\|s.\eta-\eta\|_q \geq \|\eta-m(\eta)\|_q\,\kappa_q \geq \frac{1}{2}\delta_2^{-1}\left(\frac{1}{2}\right)\kappa_q.$$

Moreover, by $\mathrm{Sym}(\Gamma/\Lambda)$-invariance of Φ, we have $\|s.\eta-\eta\|_q = \|\Phi(s.\xi)-\Phi(\xi)\|_q \leq \delta_1\left(\|s.\xi-\xi\|_p\right)$, which together with the above inequality yields

$$\max_{s\in S}\|s.\xi-\xi\| \geq \delta_1^{-1}\left(\frac{1}{2}\delta_2^{-1}\left(\frac{1}{2}\right)\kappa_V\right).$$

The desired estimate now follows. □

The reason why it suffices for us to study Schreier coset graphs is the following theorem due to Gross [Gro77]. Its very beautiful proof is an application of Petersen's 2-factor theorem, a classical result, which is often referred to as one of the first theorems in graph theory. Nevertheless, we shall omit it here.

Theorem 4.34. *Every finite, connected, regular graph of even degree can be realized as (i.e., is isomorphic to) a Schreier coset graph.*

A general d-regular graph can be $2d$ edge completed. This completion is then, by Theorem 4.34, isomorphic to a Schreier coset graph. As we saw in Section 5, spectral gaps behave nicely under edge completion, and this holds for Mimura's Banach space spectral gaps as well: it follows from the definition that $\lambda_1(\mathscr{C}_{2d}(G), B, p) = 2\lambda_1(G, B, p)$ whenever G is d-regular. Hence for expansion properties it is often enough to argue for Schreier coset graphs. This method is known as *the Gross trick*.

Lemma 4.35. *Suppose $G = Sch(\Gamma, \Lambda, S)$ is a Schreier coset graph and $B \sim_S V$. Then*

$$\lambda_1(G, B, p) \geq \delta_1^{-1}\left(\frac{1}{2}\left(\frac{2}{|S|}\right)^{1/p}\delta_2^{-1}\left(\frac{1}{2}\right)\lambda_1(G, V, p)^{1/p}\right)^p$$

where δ_1 and δ_2 are as above.

Proof. By Proposition 4.31 we have a uniform homeomorphism $\Phi_p : S(B_p) \to S(V_p)$. Hence, by Lemma 4.33 we have

$$\lambda_1(G, B, p) \geq \kappa_B^p \geq \delta_1^{-1}\left(\frac{1}{2}\delta_2^{-1}\left(\frac{1}{2}\right)\kappa_V\right)^p,$$

where κ_B and κ_V are the p-displacement constants of B and V, respectively. By Lemma 4.32 we have $\kappa_V \geq \left(\frac{2}{|S|}\lambda_1(G, V, p)\right)^{1/p}$. Combining these two facts yields the desired. □

Corollary 4.36. *Suppose G is a finite, connected, d-regular graph and $B \sim_S V$. Then*

$$\lambda_1(G, B, p) \geq \frac{1}{2}\delta_1^{-1}\left(\frac{1}{2}\left(\frac{2}{d}\right)^{1/p}\delta_2^{-1}\left(\frac{1}{2}\right)\lambda_1(G, V, p)^{1/p}\right)^p$$

Proof. If d is even then $G = \mathrm{Sch}(\Gamma, \Lambda, S)$ with $d = |S|$ and the conclusion follows from Lemma 4.35. If not then $\lambda_1(G, B, p) = \lambda_1(\mathscr{C}_{2d}(G), B, p)/2$ and $\lambda_1(\mathscr{C}_{2d}(G), V, p) = 2\lambda_1(G, V, p)$, from which the desired estimate follows. □

Proof of Theorem 4.26. Interchanging the roles of B and V in the corollary above gives the desired conclusion. □

Proof of Corollary 4.27. By Lemma 4.25 the Mazur map $M_{p,q} : B_p \to B_q$ restricts to a sphere equivalence. Since it is defined coordinatewise it is also $\mathrm{Sym}(\mathbb{N})$-invariant. Hence, Lemma 4.33 allows us to conclude that

$$\kappa_{B,p}(\Gamma, \Lambda, S) \geq \delta_1^{-1}\left(\frac{1}{2}\delta_2^{-1}\left(\frac{1}{2}\right)\kappa_{B,q}(\Gamma, \Lambda, S)\right)^p,$$

for every $\delta_1 \in \mathcal{M}_{M_{p,q}}$ and $\delta_2 \in \mathcal{M}_{M_{q,p}}$. Finally, Lemma 4.32 and the Gross trick can be applied again to obtain

$$\lambda_1(G, B, p) \geq \delta_1^{-1}\left(\frac{1}{2}\left(\frac{2}{d}\right)^{1/q}\delta_2^{-1}\left(\frac{1}{2}\right)\lambda_1(G, B, q)^{1/q}\right)^p,$$

for every finite, connected, d-regular graph G. Interchanging p and q concludes the proof. □

5 Constructing superexpander sequences

In this section we will retrace the steps of [MN14] for an iterative construction of a sequence of superexpanders through zig zag products of graphs. The construction is iterative in the sense that it takes a base graph on which we perform certain graph operations iteratively in order to produce a sequence of larger and larger, constant degree graphs. For each superreflexive Banach space we can then choose a base graph such that the corresponding graph sequence is an expanding sequence with respect to that space (to be defined later). From this array of graphs we shall distill one sequence to rule them all - a sequence of superexpanders.

5.1 Graph operations

Our construction of superexpanders involves performing several types of graph operations, which we now introduce.

Definition 5.1. Let $G = (V, E)$ be a finite, regular graph with normalized adjacency matrix A. For $t \in \mathbb{N}$ the graph power G^t is the graph whose normalized adjacency matrix is A^t, i.e., G^t has the same vertex set as G and an edge is drawn between vertices $u, v \in V$ for each path from u to v in G of length t.

In the Euclidean case we have control over spectral gaps when taking graph powers in the following way: recall that $\gamma_+(G, |\cdot|^2) = \frac{1}{\lambda_+(G)}$ and since the adjacency matrix is self-adjoint we get by elementary spectral calculus that

$$\gamma_+(G^t, |\cdot|^2) = \frac{1}{\lambda_+(G^t)} = \frac{1}{1-(1-\lambda_+(G))^t},$$

which decays rather rapidly to 1 with t. This computation, however, is very specific to the spectral characterization of $\gamma_+(G, |\cdot|^2)$, and there is no reason to believe that it applies to more general kernels such as (powers of) norms on superreflexive Banach spaces - [MN14] even present counter examples, demonstrating that it does not. We can prove similar 'decay of the spectral gap' with respect to certain kernels though, if we pass to a different graph operation namely that of the *Cesáro average*:

Definition 5.2. In the same setup as above we define for $t \in \mathbb{N}$ the *Cesáro average* $\mathscr{A}_t(G)$ to be the graph whose normalized adjacency matrix is $\frac{1}{t}\sum_{m=0}^{t-1} A^m$, i.e., for $u, v \in V$ and $m \in \{0, \ldots, t-1\}$ we draw d^{t-1-m} edges between u and v for each path from u to v in G of length m.

If G is d-regular then $\mathscr{A}_t(G)$ is td^{t-1}-regular, as can be easily verified. Obtaining decay of the spectral gap under some appropriate graph operation, as the one for $\gamma_+(G^t, |\cdot|^2)$ above, for non-euclidean kernels is referred to as a *non-linear spectral calculus* and we shall treat this in the subsequent section.

Definition 5.3. Let $D \geq d \geq 2$ and $G = (V, E)$ a d-regular graph. The D-edge completion of G, denoted $\mathscr{C}_D(G)$ is the D-regular graph with vertex set V obtained by writing $D = md + r$, with $m \in \mathbb{N}$ and $0 \leq r \leq d-1$, and duplicating each edge in E m times followed by adding r self-loops to each vertex, i.e., if $E(G)(u, v)$ denotes the number of edges in G between u and v then $E(\mathscr{C}_D(G))(u, v) = mE(G)(u, v) + r\delta_{uv}$.

Spectral gaps behave in the following way when taking edge completions:

Lemma 5.4. *In the above setup we have for any kernel $K : X \times X \to \mathbb{R}_+$ that*

$$\gamma_+(\mathscr{C}_D(G), K) \leq 2\gamma_+(G, K) \qquad \textit{and} \qquad \gamma(\mathscr{C}_D(G), K) \leq 2\gamma(G, K)$$

Proof. We prove the statement for γ, the argument for γ_+ is analogous. So take any function $f : V \to X$ and observe that

$$\begin{aligned}
\frac{1}{n} \sum_{(u,v) \in E(\mathscr{C}_D(G))} \frac{1}{D} K(f(u), f(v)) &= \frac{1}{n} \sum_{u,v \in V} \frac{mE(G)(u,v) + r\delta_{uv}}{md + r} K(f(u), f(v)) \\
&\geq \frac{1}{n} \sum_{u,v \in V} \frac{mE(G)(u,v)}{(m+1)d} K(f(u), f(v)) \\
&\geq \frac{1}{n} \sum_{u,v \in V} \frac{1}{2} \frac{E(G)(u,v)}{d} K(f(u), f(v)) \\
&\geq \frac{1}{n^2} \sum_{u,v \in V} \frac{1}{2\gamma(G,K)} K(f(u), f(v)).
\end{aligned}$$

Hence, by definition of the spectral gap we obtain $\gamma(\mathscr{C}_D(G), K) \leq 2\gamma(G, K)$ as desired. $\square$

The operations we have introduced so far arguably do not seem very applicable to the task of constructing expander families; they preserve spectral properties but they dramatically increase the degree of the graph thereby making it less sparse. This is where the notion of the zig-zag product enters the picture. It takes as arguments two graphs, G_1 and G_2, and we think of G_1 as being a highly dense graph with certain spectral properties and G_2 as being a sparse base graph. The resulting graph is larger than both, inherits the spectral properties of G_1 and the sparsity of G_2. It is a rather cumbersome definition but it is possible to obtain some intuition about it, making it actually quite reasonable to work with.

Definition 5.5. Let $G_1 = (V_1, E_1)$ be a d_1-regular graph on n_1 vertices and $G_2 = (V_2, E_2)$ be a d_2-regular graph on d_1 vertices. The zig-zag product of G_1 and G_2, denoted by $G_1 ⓩ G_2$, is a graph whose vertex set is $V_1 \times V_2$ and whose edge set is determined in the following way:

- Enumerate the vertices of G_2 so that we may identify V_2 with $[d_1] = \{1, \ldots, d_1\}$.
- For each $u \in V_1$ fix an enumeration, $e_u^1, \ldots, e_u^{d_1}$, of the edges emanating from u.
- Two vertices $(u_1, a_1), (u_2, a_2) \in V_1 \times [d_1]$ are adjoined by an edge if there are $i, j \in [d_1]$ such that
  1. $(a_1, i) \in E_2$
  2. $e_{u_1}^i = e_{u_2}^j$
  3. $(j, a_2) \in E_2$.

This makes $G_1 ⓩ G_2$ a d_2^2-regular graph.

The intuition behind this definition is that we have a copy, often referred to as a *cloud*, of V_2 at each vertex of G_1, and each G_2-vertex in the cloud based at some G_1-vertex corresponds to an edge emanating from that vertex. Two vertices are then adjoined in $G_1 ⓩ G_2$ if the bases of their clouds are connected by an edge in G_1, and if they are connected in G_2 to the vertex in their respective clouds corresponding to that edge. This is often referred to as taking a 'zig' in G_2, followed by a 'zag' in G_1, followed by another 'zig' in G_2, which yields an obvious way to find the neighbours of a given vertex in $G_1 ⓩ G_2$, given that one can find the neighbours of vertices in G_1 and G_2 easily. Since there are d_2 possible 'zigs' from a given vertex in a given cloud, and then d_2 'zigs' again in the cloud where we end up following the 'zag', we see that $G_1 ⓩ G_2$ is indeed d_2^2-regular. The specific enumerations chosen do matter in the sense that different enumerations can lead to non-isomorphic graphs. For our purpose, however, they are of no importance, since any choice of enumerations will lead to the same behaviour of spectral gaps:

Theorem 5.6 (**Submultiplicativity of the zig-zag product**).
Let G_1 and G_2 be as above and $K : X \times X \to \mathbb{R}_+$ an arbitrary kernel on a set X. Then

$$\gamma_+(G_1 ⓩ G_2, K) \leq \gamma_+(G_1, K)\gamma_+(G_2, K)^2$$

under any choice of enumerations.

Proof. Let $f, g : V_1 \times [d_1] \to X$ be functions. For a vertex v (in G_1 or G_2) let $N(v)$ denote the set of neighbours of v and for each $u \in V_1$ let $\pi_u : N(u) \to [d_1]$ be an arbitrary (but fixed) enumeration of $N(u)$, i.e., a bijection to the vertex set of G_2. By definition of the absolute spectral gap we have for any fixed pair $a, b \in [d_1]$ that

$$\frac{1}{n_1^2} \sum_{u,v \in V_1} K(f(u,a), g(v,b)) \leq \frac{\gamma_+(G_1, K)}{n_1 d_1} \sum_{(u,v) \in E_1} K(f(u,a), g(v,b)).$$

Hence

$$\begin{aligned} T &:= \frac{1}{|V_1 \times [d_1]|^2} \sum_{(u,a),(v,b) \in V_1 \times [d_1]} K(f(u,a), g(v,b)) \\ &= \frac{1}{d_1^2} \sum_{a,b \in [d_1]} \frac{1}{n_1^2} \sum_{u,v \in V_1} K(f(u,a), g(v,b)) \\ &\leq \frac{1}{d_1^2} \sum_{a,b \in [d_1]} \frac{\gamma_+(G_1, K)}{n_1 d_1} \sum_{(u,v) \in E_1} K(f(u,a), g(v,b)) \\ &= \frac{\gamma_+(G_1, K)}{n_1 d_1} \sum_{u \in V_1} \sum_{b \in [d_1]} \frac{1}{d_1^2} \sum_{a \in [d_1]} \sum_{v \in N(u)} K(f(u,a), g(v,b)). \end{aligned} \tag{22}$$

Since $|N(u)| = d_1$ we can view $(g(v,b))_{v \in N(u)}$ as a sequence indexed over $[d_1]$ so, again by definition of the absolute spectral gap, we obtain for each $u \in V_1$ and $b \in [d_1]$ that

$$\frac{1}{d_1^2} \sum_{a \in [d_1]} \sum_{v \in N(u)} K(f(u,a), g(v,b)) \leq \frac{\gamma_+(G_2, K)}{d_1 d_2} \sum_{v \in N(u)} \sum_{a \in N(\pi_u(v))} K(f(u,a), g(v,b)).$$

Inserting this in (22)) we obtain

$$T \leq \frac{\gamma_+(G_1,K)\gamma(G_2,K)}{n_1 d_1^2 d_2} \sum_{u\in V_1} \sum_{b\in[d_1]} \sum_{v\in N(u)} \sum_{a\in N(\pi_u(v))} K(f(u,a),g(v,b)).$$

Now for $(u,v) \in E_1$ let $k^1_{u,v},\dots,k^{d_2}_{u,v}$ denote the neighbours of $\pi_u(v)$ in G_2. Then by rearranging and using the same idea as above we obtain

$$\begin{aligned}
T &\leq \frac{\gamma_+(G_1,K)\gamma(G_2,K)}{n_1 d_1^2 d_2} \sum_{u\in V_1} \sum_{b\in[d_1]} \sum_{v\in N(u)} \sum_{j=1}^{d_2} K(f(u,k^j_{u,v}),g(v,b)) \\
&= \frac{\gamma_+(G_1,K)\gamma_+(G_2,K)}{n_1 d_2} \sum_{v\in V_1} \sum_{j=1}^{d_2} \frac{1}{d_1^2} \sum_{u\in N(v)} \sum_{b\in[d_1]} K(f(u,k^j_{u,v}),g(v,b)) \\
&\leq \frac{\gamma_+(G_1,K)\gamma_+(G_2,K)^2}{n_1 d_2 d_1 d_2} \sum_{v\in V_1} \sum_{j=1}^{d_2} \sum_{u\in N(v)} \sum_{i=1}^{d_2} K(f(u,k^j_{u,v}),g(v,k^i_{v,u})) \\
&= \frac{\gamma_+(G_1,K)\gamma_+(G_2,K)^2}{|V_1\times[d_1]|d_2^2} \sum_{(u,v)\in E_2} \sum_{j=1}^{d_2} \sum_{i=1}^{d_2} K(f(u,k^j_{u,v}),g(v,k^i_{v,u})).
\end{aligned}$$

By the 'zig-zag intuition' this last sum amounts exactly to summing over all the edges of $G_1 ⓩ G_2$, so by definition of the absolute spectral gap we now get

$$\gamma_+(G_1 ⓩ G_2,K) \leq \gamma_+(G_1,K)\gamma_+(G_2,K)^2.$$

This concludes the proof. □

Back in the euclidean case the zig-zag submultiplicativity theorem gives, after rearranging, a zig-zag supermultiplicativity for the absolute spectral gap:

$$\lambda_+(G_1 ⓩ G_2) \geq \lambda_+(G_1)\lambda_+(G_2)^2.$$

Supermultiplicative properties for the zig-zag product were first proven in [RVW02] and used to construct a sequence of expanders. We shall briefly sketch their construction: Suppose H is a graph on n_0 vertices of degree d_0 such that $\lambda_+(H) > 0$. Pick $\varepsilon,\delta \in (0,1)$ such that $s\lambda_+(H)^2 \geq \varepsilon$ whenever $s \geq \delta$ (such ε and δ trivially exist). Also find t_0 such that

$$\min\left\{1-(1-\lambda_+(H))^{2t_0}, 1-(1-\varepsilon)^{t_0}\right\} \geq \delta.$$

We will assume that $n_0 = d_0^{2t_0}$ (the existence of a base graph with all these properties is of course by no means obvious; for a treatment of this matter see [RVW02]). Now define $G_1 = H^2$, and inductively define $G_{i+1} = G_i^{t_0} ⓩ H$ (this is a legal zig-zag product by the assumption that $n_0 = d_0^{2t_0}$). Then G_i is a graph on $n_0^i = d_0^{2it_0}$ vertices, which is regular of degree d_0^2. Moreover, we have $\lambda_+(G_i) \geq \min\left\{1-(1-\lambda_+(H))^2, \varepsilon\right\} > 0$, for each $i \in \mathbb{N}$. Indeed, this holds for $i = 1$ since $\lambda_+(G_1) = \lambda_+(H^2) = 1-(1-\lambda_+(H))^2$. And if it holds true for $i \geq 1$ then

$$\lambda_+(G_i^{t_0}) = 1-(1-\lambda_+(G_i))^{t_0} \geq \min\left\{1-(1-\lambda_+(H))^{2t_0}, 1-(1-\varepsilon)^{t_0}\right\} > \delta,$$

and hence $\lambda_+(G_{i+1}) \geq \lambda_+(G_i^{t_0})\lambda_+(H)^2 \geq \varepsilon$, as claimed. Our goal is to generalize this construction to produce a sequence of superexpanders, but since the above construction relies heavily on the spectral characterization of $\gamma_+(G, |\cdot|^2)$, it does not immediately generalize. This is what motivates the need for a non-linear spectral calculus.

5.2 A non-linear spectral calculus

To set the framework we first introduce the notion of *metric Markov cotype*, a property which, once the necessary foot work has been done, makes the proof of non-linear spectral calculus for Cesáro averages come out quite neatly.

Definition 5.7. A metric space (X, d_X) has *metric Markov cotype* p with exponent q, where $p, q \in (0, \infty)$ if there exists $C \in (0, \infty)$ such that for every $n, t \in \mathbb{N}$, every $n \times n$ symmetric, stochastic matrix $A = (a_{ij})$ and every $x_1, \ldots, x_n \in X$ there exist $y_1, \ldots, y_n \in X$ such that

$$\sum_{i=1}^{n} d_X(x_i, y_i)^q + t^{q/p} \sum_{i,j=1}^{n} a_{ij} d_X(y_i, y_j)^q \leq C^q \sum_{i,j=1}^{n} \mathscr{A}_t(A)_{ij} d_X(x_i, x_j)^q. \tag{23}$$

We let $C_p^{(q)}(X, d_X)$ denote the infimum over C satisfying this.

Let us immediately proceed to discuss why this definition is relevant to us.

Theorem 5.8. *Let B be a superreflexive Banach space. Then there exists $p = p(B) \in [2, \infty)$ such that $C_p^{(2)}(B, ||\cdot||) < \infty$.*

One might object that it seems redundant to set our theory of spectral calculus in the framework of metric spaces (hence the 'non-linear' in the headline) to then apply it to Banach spaces - thus taking us back in the linear setting. We do believe though that it is of great interest to treat this matter in the non-linear setting since it uncovers its true identity (as with the zig-zag submultiplicativity, which even holds for general kernels) as a non-linear phenomenon. The study of such phenomena was coined by Bourgain in 1985 as *the Ribe program*, and is an active area of research with many interesting open problems (see [Nao12] for a good introduction). The proof of Theorem 5.8 will be postponed to section 6, for now we will focus on the following important result.

Theorem 5.9. *Suppose (X, d_X) has metric Markov cotype p with exponent q. Then for any $C \geq C_p^{(q)}(X, d_X)$ we have*

$$\gamma_+(\mathscr{A}_t(A), d_X^q) \leq (45C)^q \max\left\{1, \frac{\gamma_+(A, d_X^q)}{t^{q/p}}\right\},$$

whenever $n, t \in \mathbb{N}$ and A is an $n \times n$ symmetric, stochastic matrix.

Remark 5.10. The name 'metric Markov cotype' suggests that there is also a notion of 'metric Markov type', which is of course the case (see [Ost13, sec. 8]). Metric Markov type serves as a generalization of the well known fact that the expected distance to the origin at time t of a standard random walk on $\mathbb{Z}$ is at most $\sqrt{t}$. Metric Markov cotype is intended to be dual to that of Markov type in the sense that the inequality is reversed, but there are also some new elements introduced: the power of the transition matrix is replaced by the Cesáro average and we pass to the approximating points $y_1, \cdots, y_n$ (hence the initial error term in (23) ensuring that, on q-average, these are close to the initial points $x_1, \cdots, x_n$). One might object that this makes Definition 5.7 rather unsatisfactory as a dual notion to metric Markov type, and referring to [Nao12, sec. 4.1] we could also motivate the notion of metric Markov cotype simply because it yields non-linear spectral calculus in the sense of Theorem 5.9.

The proof of Theorem 5.9 requires several lemmas. The first one being a spectral calculus inequality for the (non-absolute) spectral gap.

Lemma 5.11. *Suppose (X, d_X) has metric Markov cotype p with exponent $q \geq 1$. Then for any $C \geq C_p^{(q)}(X, d_X)$ we have*

$$\gamma(\mathscr{A}_t(A), d_X^q) \leq (3C)^q \max\left\{1, \frac{\gamma(A, d_X^q)}{t^{q/p}}\right\},$$

whenever A is a symmetric, stochastic $n \times n$ matrix and $t \in \mathbb{N}$.

Proof. For notational ease, write $\mathscr{A}_t(A) = (b_{ij})$, $A = (a_{ij})$ and suppose $\gamma(\mathscr{A}_t(A), d_X^q) > (3C)^q$. Take $\gamma \in \big((3C)^q, \gamma(\mathscr{A}_t(A), d_X^q)\big)$, i.e., there exist $x_1, \ldots, x_n \in X$ such that

$$\frac{1}{n^2} \sum_{i,j=1}^{n} d_X(x_i, x_j)^q > \frac{\gamma}{n} \sum_{i,j=1}^{n} b_{ij} d_X(x_i, x_j). \tag{24}$$

Since $C \geq C_p^{(q)}(X, d_X)$, there exist $y_1, \ldots, y_n \in X$ such that

$$\sum_{i=1}^{n} d_X(x_i, y_i)^q + t^{q/p} \sum_{i,j=1}^{n} a_{ij} d_X(y_i, y_j)^q \leq C^q \sum_{i,j=1}^{n} b_{ij} d_X(x_i, x_j)^q. \tag{25}$$

Finally, by the triangle and Hölder inequalities we have

$$d_X(x_i, x_j)^q \leq 3^{q-1}(d_X(x_i, y_i)^q + d_X(y_i, y_j)^q + d_X(y_j, x_j)^q), \tag{26}$$

for all $i, j \in \{1, \ldots, n\}$. Combining (26) with (24), (25) and the definition of the spectral gap we obtain

$$\begin{aligned}
\frac{\gamma(A, d_X^q)}{n} \sum_{i,j=1}^{n} a_{ij} d_X(y_i, y_j)^q &\geq \frac{1}{n^2} \sum_{i,j=1}^{n} d_X(y_i, y_j)^q \\
&\geq \frac{1}{3^{q-1} n^2} \sum_{i,j=1}^{n} d_X(x_i, x_j)^q - \frac{2}{n} \sum_{i=1}^{n} d_X(x_i, y_i)^q \\
&> \frac{\gamma}{3^{q-1} n} \sum_{i,j=1}^{n} b_{ij} d_X(x_i, x_j)^q - \frac{2}{n} \sum_{i=1}^{n} d_X(x_i, y_i)^q \\
&\geq \frac{\gamma t^{q/p}}{3^{q-1} C^q n} \sum_{i,j=1}^{n} a_{ij} d_X(y_i, y_j)^q + \left(\frac{\gamma}{3^{q-1} C^q n} - \frac{2}{n}\right) \sum_{i=1}^{n} d_X(x_i, y_i)^q \\
&> \frac{\gamma t^{q/p}}{(3C)^q n} \sum_{i,j=1}^{n} a_{ij} d_X(y_i, y_j)^q,
\end{aligned}$$

where the last inequality is due to the assumption that $\gamma > (3C)^q$. Cancellation yields $\gamma(A, d_X^q) > \gamma t^{q/p}/(3C)^q$ and letting $\gamma \nearrow \gamma(\mathscr{A}_t(A), d_X^q)$ we obtain the desired. □

Lemma 5.12. *Let (X, d_X) be a metric space, $1 \leq q < \infty$ and A a stochastic, symmetric $n \times n$ matrix. Then*

$$\frac{2\gamma\left(\left(\begin{smallmatrix} 0 & A \\ A & 0 \end{smallmatrix}\right), d_X^q\right)}{2^q+1} \leq \gamma_+(A, d_X^q) \leq 2\gamma\left(\left(\begin{smallmatrix} 0 & A \\ A & 0 \end{smallmatrix}\right), d_X^q\right)$$

Proof. For notational ease, put $\gamma = \gamma\left(\left(\begin{smallmatrix} 0 & A \\ A & 0 \end{smallmatrix}\right), d_X^q\right)$. Let $f, g : [n] \to X$ be arbitrary and define $h : [2n] \to X$ by $h(i) = f(i)$, for $i \in [n]$, and $h(i) = g(i-n)$, for $i \in [2n] \setminus [n]$. Then

$$\begin{aligned}
\frac{1}{n^2}\sum_{i,j=1}^{n} d_X(f(i), g(j))^q &= \frac{1}{n^2}\sum_{i,j=1}^{n} d_X(h(i), h(j+n))^q \leq \frac{1}{2n^2}\sum_{i,j=1}^{2n} d_X(h(i), h(j))^q \\
&\leq \frac{2\gamma}{2n}\sum_{i,j=1}^{2n}\left(\begin{smallmatrix} 0 & A \\ A & 0 \end{smallmatrix}\right)_{ij} d_X(h(i), h(j))^q = \frac{2\gamma}{n}\sum_{i,j=1}^{n} a_{ij} d_X(f(i), g(j))^q,
\end{aligned}$$

demonstrating the inequality to the right. For the other one let $h : [2n] \to X$ be arbitrary and define $f, g : [n] \to X$ by $f(i) = h(i)$ and $g(i) = h(i+n)$. Then by the triangle inequality we obtain

$$\begin{aligned}
\sum_{i,j=1}^{n} d_X(h(i), h(j))^q &\leq \frac{1}{n}\sum_{i,j,l=1}^{n} 2^{q-1}\left(d_X(h(i), h(l+n))^q + d_X(h(j), h(l+n))^q\right) \\
&= 2^q \sum_{i,j=1}^{n} d_X(f(i), g(j))^q.
\end{aligned}$$

By a similar computation we also get $\sum_{i,j=1}^{n} d_X(h(i+n), h(j+n))^q \leq 2^q \sum_{i,j=1}^{n} d_X(f(i), g(j))^q$. Using these two estimates, we obtain

$$\begin{aligned}
\frac{1}{(2n)^2}\sum_{i,j=1}^{2n} d_X(h(i), h(j))^q &= \frac{1}{(2n)^2}\sum_{i,j=1}^{n} d_X(h(i), h(j))^q + \frac{1}{(2n)^2}\sum_{i,j=1}^{n} d_X(h(i+n), h(j+n))^q \\
&\quad + \frac{1}{(2n)^2}\sum_{i,j=1}^{n} d_X(h(i), h(j+n))^q + \frac{1}{(2n)^2}\sum_{i,j=1}^{n} d_X(h(i+n), h(j))^q \\
&\leq \frac{2^q+1}{2n^2}\sum_{i,j=1}^{n} d_X(f(i), g(j))^q \\
&\leq \frac{(2^q+1)\gamma_+(A, d_X^q)}{2n}\sum_{i,j=1}^{n} a_{ij} d_X(f(i), g(j))^q \\
&\leq \frac{(2^q+1)\gamma_+(A, d_X^q)}{2} \cdot \frac{1}{2n}\sum_{i,j=1}^{2n}\left(\begin{smallmatrix} 0 & A \\ A & 0 \end{smallmatrix}\right)_{ij} d_X(h(i), h(j))^q,
\end{aligned}$$

which yields the inequality to the left. □

The following lemma has a proof as well. It is similar in flavour to the above argument, only it involves even more tedious computations. For this reason, we omit it and refer the interested (and stubborn) reader to [MN14, Lemma 2.4].

Lemma 5.13. *In the same setup as above we have*

$$\gamma\left(\left(\begin{smallmatrix} 0 & \mathscr{A}_t(A) \\ \mathscr{A}_t(A) & 0 \end{smallmatrix}\right), d_X^q\right) \leq (2^{q+1}+1)\gamma\left(\mathscr{A}_t\left(\left(\begin{smallmatrix} 0 & A \\ A & 0 \end{smallmatrix}\right)\right), d_X^q\right)$$

We are now ready to give the proof of Theorem 5.9:

Proof of Theorem 5.9: By Lemmas 5.12 and 5.13 we have

$$\gamma_+(\mathscr{A}_t(A), d_X^q) \leq 2\gamma\left(\left(\begin{smallmatrix} 0 & \mathscr{A}_t(A) \\ \mathscr{A}_t(A) & 0 \end{smallmatrix}\right), d_X^q\right) \leq 2(2^{q+1}+1)\gamma\left(\mathscr{A}_t\left(\left(\begin{smallmatrix} 0 & A \\ A & 0 \end{smallmatrix}\right)\right), d_X^q\right).$$

Moreover, by Lemma 5.11 and Lemma 5.12 again we obtain

$$\begin{aligned}\gamma\left(\mathscr{A}_t\left(\left(\begin{smallmatrix} 0 & A \\ A & 0 \end{smallmatrix}\right)\right), d_X^q\right) &\leq (3C)^q \max\left\{1, \frac{\gamma\left(\left(\begin{smallmatrix} 0 & A \\ A & 0 \end{smallmatrix}\right), d_X^q\right)}{t^{q/p}}\right\} \\ &\leq (3C)^q \max\left\{1, \frac{2^q+1}{2}\cdot\frac{\gamma_+(A, d_X^q)}{t^{q/p}}\right\}.\end{aligned}$$

Using that $(2^q+1)(2^{q+1}+1) \leq 15^q$ the above two estimates yield the desired. □

5.3 Constructing superexpanders

We are still one ingredient short for the iterative construction of superexpanders using zig-zag products namely that of a base graph, but before we adress this let's see how the non-linear spectral calculus obtained above comes into play. Note how the assumptions imposed on G_0 (the base graph) in the following lemma are similar in flavour to the ones imposed on the base graph in the construction of [RVW02] sketched above. In addition, however, we have to make spectral calculus (i.e., decay of the absolute spectral gap) an assumption.

Lemma 5.14. *Let $d, m, t \in \mathbb{N}$ satisfy $td^{2(t-1)} \leq m$ and suppose $G_0 = (V_0, E_0)$ is a d-regular graph on m vertices. Then there exists a sequence of d^2-regular graphs, $F_j^t = (V_j^t, E_j^t)$, $j \in \mathbb{N}$, with $|V_j^t| = m^j$ satisfying the following: If $K : X \times X \to \mathbb{R}_+$ is a kernel such that*

- *$\gamma_+(G_0, K) \leq \gamma$ for some $\gamma \geq 1$,*
- *there exist $C \geq 1$ and $\varepsilon \in (0,1)$, satisfying $t \geq (2C\gamma^2)^{1/\varepsilon}$, such that for every finite, regular graph G we have $\gamma_+(\mathscr{A}_t(G), K) \leq C\max\{1, \gamma_+(G,K)/t^\varepsilon\}$,*

then $\sup_j \gamma_+(F_j^t, K) \leq 2\gamma^2 C$.

Proof. Define $F_1^t := \mathscr{C}_{d^2}(G_0)$ and recursively define $F_{j+1}^t := \mathscr{C}_m(\mathscr{A}_t(F_j^t))\textcircled{z}G_0$. Then F_1^t has m vertices and is d^2 regular, and if F_j^t has m^j vertices and is d^2-regular then $\mathscr{A}_t(F_j^t)$ has m^j vertices and is $td^{2(t-1)}$-regular which is no more than m by assumption. Hence we can perform the m-edge completion on it, making $\mathscr{C}_m(\mathscr{A}_t(F_j^t))$ an m-regular graph on m^j vertices enabling us to perform the zig-zag product on it with G_0. This yields a d^2-regular graph on $m \cdot m^j = m^{j+1}$ vertices as claimed. By Lemma 5.4 we have

$$\gamma_+(F_1^t, K) \leq 2\gamma_+(G_0, K) \leq 2\gamma \leq 2\gamma^2 C.$$

Moreover, if $\gamma_+(F_j^t, K) \leq 2\gamma^2 C$ we get, by using zig-zag submultiplicativity and Lemma 5.4 once more, that

$$\begin{aligned}
\gamma_+(F_{j+1}^t, K) &\leq \gamma_+(\mathscr{C}_m(\mathscr{A}_t(F_j^t)), K)\gamma_+(G_0, K)^2 \\
&\leq 2\gamma_+(\mathscr{A}_t(F_j^t), K)\gamma^2 \\
&\leq 2\gamma^2 C \max\left\{1, \frac{\gamma_+(F_j^t, K)}{t^\varepsilon}\right\} \\
&\leq 2C\gamma^2 \max\left\{1, \frac{2C\gamma^2}{t^\varepsilon}\right\} \leq 2C\gamma^2,
\end{aligned}$$

so the uniform bound on the absolute spectral gaps follows by induction. □

It follows from Theorems 5.8 and 5.9 that superreflexive Banach spaces satisfy the spectral calculus inequality from Lemma 5.14 for every $t \in \mathbb{N}$, but with constants depending on the space. Hence, this lemma can be used to construct expander sequences with respect to some fixed superreflexive Banach space provided we can find a base graph satisfying the remaining properties. The following lemma provides an entire continuum of graphs from which we can pick base graphs for any superreflexive Banach space.

Lemma 5.15. *There exists a strictly increasing sequence of natural numbers $(m_n)_{n\geq 1}$ such that $2^{n/10} \leq m_n \leq 2^n$ satisfying the following:*
For every $1 \geq \delta > 0$ there is $n_0(\delta) \in \mathbb{N}$ and a sequence of regular graphs $(H_n(\delta))_{n \geq n_0(\delta)}$ such that $|V(H_n(\delta))| = m_n$ for every $n \geq n_0(\delta)$.
The degree of $H_n(\delta)$, denoted $d_n(\delta)$, satisfies $d_n(\delta) \leq \exp(\log(m_n)^{1-\delta})$.
Finally, for every superreflexive Banach space B we have $\gamma_+(H_n(\delta), ||\cdot||^2) < \infty$ for every $\delta \in (0,1)$ and $n \geq n_0(\delta)$ and there exists $\delta_0(B) \in (0,1)$ such that

$$\sup_{0<\delta\leq\delta_0(B)} \sup_{n\geq n_0(\delta)} \gamma_+(H_n(\delta), ||\cdot||^2) < \infty.$$

We postpone the proof of this lemma to Section 7, where we will prove it for the larger class of *K-convex* Banach spaces, and for now soldier on to see how it applies to our problem:

Corollary 5.16. *For every $k \in \mathbb{N}$ there exists $d_k \in \mathbb{N}$ and a sequence of d_k-regular graphs, $(F_j(k))_{j\geq 1}$ such that $|V(F_j(k))| =: n_j(k) \nearrow \infty$, as $j \to \infty$, and the following holds:*

For every superreflexive Banach space B we have $\gamma_+(F_j(k), ||\cdot||^2) < \infty$ for every $j, k \in \mathbb{N}$, and there exist $k(B) \in \mathbb{N}$ such that

$$\sup_{\substack{j,k\in\mathbb{N} \\ k\geq k(B)}} \gamma_+(F_j(k), ||\cdot||^2) \leq k(B).$$

Proof. Let $k \in \mathbb{N}$. Since (following the notation from Lemma 5.15) $2^{n/10} \leq m_n \leq 2^n$ we can find $n(k) \geq n_0(1/k)$ such that $k \exp\left(2(k-1)\log(m_{n(k)})^{1-1/k}\right) \leq m_{n(k)}$. Now consider $H_{n(k)}(1/k)$ from Lemma 5.15. The degree, $d_{n(k)}(1/k)$, of $H_{n(k)}(1/k)$ satisfies

$$kd_{n(k)}(1/k)^{2(k-1)} \leq k \exp\left(2(k-1)\log(m_{n(k)})^{1-1/k}\right) \leq m_{n(k)} = |V(H_{n(k)}(1/k))|,$$

so by Lemma 5.14 there is a sequence of $\left(d_{n(k)}(1/k)\right)^2$-regular graphs on $m^j_{n(k)}$ vertices. Denote this sequence $(F_j(k))_{j\geq 1}$ and let $d_k := \left(d_{n(k)}(1/k)\right)^2$ and $n_j(k) := m^j_{n(k)}$. Now let B be a superreflexive Banach space, and pick $\delta_0(B)$ according to Lemma 5.15, i.e., there is $\gamma \geq 1$ such that $\gamma_+(H_{n(k)}(1/k), ||\cdot||^2) < \gamma$ whenever $k \geq 1/\delta_0(B)$. By Theorem 5.9 we infer the existence of $C = C(B) \in [0, \infty)$ and $\varepsilon = \varepsilon(B) \in (0,1)$ such that

$$\gamma_+(\mathscr{A}_t(A), ||\cdot||^2) \leq C \max\left\{1, \gamma_+(A, ||\cdot||^2)/t^{\varepsilon}\right\},$$

for every $t \in \mathbb{N}$ and stochastic, symmetric $n \times n$ matrix A.
So if we put $k(B) = \lceil \max\left\{1/\delta_0(B), (2C\gamma^2)^{1/\varepsilon}, 2C\gamma^2\right\}\rceil$ we have by Lemma 5.14 that

$$\sup_j \gamma_+(F_j(k), ||\cdot||^2) \leq 2C\gamma^2 \leq k(B), \text{ whenever } k \geq k(B),$$

which yields the desired. $\square$

This corollary settles the base graph issue for fixed target spaces, thus producing expander sequences with respect to a fixed superreflexive Banach space; for a superreflexive Banach space, B, we just take any $k \geq k(B)$, and then $(F_j(k))_{j\geq 1}$ from Corollary 5.16 is an expander sequence with respect to B. The trouble is that the degree of the graphs in this sequence might grow with k (and by the way the graphs in Lemma 5.15 are constructed, this is in fact the case) so it takes one more intricate construction to obtain from the array $(F_j(k))_{j,k}$ a sequence of constant degree graphs which is a sequence of expanders with respect to any superreflexive Banach space. This construction is contained in the following lemma:

Lemma 5.17. *Let $(d_k)_{k\geq 1}$ be a sequence of natural numbers and for each $k \in \mathbb{N}$ let $(n_j(k))_{j\geq 1}$ be a strictly increasing sequence of natural numbers, such that for every $j, k \in \mathbb{N}$ there is a d_k-regular graph, $F_j(k)$, on $n_j(k)$ vertices. Suppose that $\mathscr{K}$ is a family of kernels such that $\gamma_+(F_j(k), K) < \infty$, for all $K \in \mathscr{K}$ and $j, k \in \mathbb{N}$. Suppose furthermore that for every $K \in \mathscr{K}$ there are $k_1(K), k_2(K) \in \mathbb{N}$ such that*

$$\sup_{\substack{j,k\in\mathbb{N}\\ k\geq k_1(K)}} \gamma_+(F_j(k), K) < k_1(K) \tag{27}$$

and, for any finite, regular graph G and $t \in \mathbb{N}$, we have

$$\gamma_+(\mathscr{A}_t(G), K) \leq k_2(K) \max\left\{1, \frac{\gamma_+(G,K)}{t^{1/k_2(K)}}\right\}. \tag{28}$$

Then there exists $d \in \mathbb{N}$ and a sequence of d-regular graphs $(H_i)_{i\geq 1}$ such that $|V(H_i)| \to \infty$ and $\sup_{i\geq 1} \gamma_+(H_i, K) < \infty$, for all $K \in \mathscr{K}$.

Proof. We shall again be zig zagging our way through a great many graphs from which we will finally be able to extract our desired sequence. First we shall need to define some integer sequences. For every $k \in \mathbb{N}$ define $M_k := (2k^3)^k$, and

$$j(k) := \min\left\{j \in \mathbb{N} \mid n_j(k) > 2d_1^2 + M_{k+1}d_{k+1}^{2(M_{k+1}-1)}\right\}.$$

Now let $i \in \mathbb{N}$ and define $l(i) \in \mathbb{N}$ and a sequence of regular graphs $W_i^0, \cdots, W_i^{l(i)}$ in the following way: Put $W_1^0 := \mathscr{C}_{2d_1^2}(F_{j(1)}(1))$ and $l(1) := 0$. For $i > 1$, put $W_i^0 := F_{j(i)}(i)$ and define $h_1(i) := \min\left\{h \in \mathbb{N} \mid n_{j(h)}(h) \geq d_i\right\}$. Note that $i-1 \in \left\{h \in \mathbb{N} \mid n_{j(h)}(h) \geq d_i\right\}$ since otherwise we would have

$$d_i > n_{j(i-1)}(i-1) > 2d_1^2 + M_i d_i^{2(M_i-1)} \geq d_i,$$

This demonstrates both that $h_1(i)$ is well defined and that $h_1(i) < i$. By definition of $h_1(i)$ we have $n_{j(h_1(i))}(h_1(i)) \geq d_i$ so that we may take the edge completion $\mathscr{C}_{n_{j(h_1(i))}(h_1(i))}(W_i^0)$ and since the number of vertices of $F_{j(h_1(i))}(h_1(i))$ is exactly $n_{j(h_1(i))}(h_1(i))$ we may perform the zig zag product (followed by taking Césaro average)

$$W_i^1 := \mathscr{A}_{M_{h_1(i)}}\left(\mathscr{C}_{n_{j(h_1(i))}(h_1(i))}(W_i^0) \textcircled{z} F_{j(h_1(i))}(h_1(i))\right),$$

which is $M_{h_1(i)}d_{h_1(i)}^{2(M_{h_1(i)}-1)}$-regular. For $k > 1$ assume we have already defined $h_{k-1}(i) \in \mathbb{N}$ and W_i^{k-1} which is regular of degree $M_{h_{k-1}(i)}d_{h_{k-1}(i)}^{2(M_{h_{k-1}(i)}-1)}$. If $h_{k-1}(i) = 1$ we terminate and put $l(i) = k-1$. Otherwise put

$$h_k(i) = \min\left\{h \in \mathbb{N} \mid n_{j(h)}(h) \geq M_{h_{k-1}(i)}d_{h_{k-1}(i)}^{2(M_{h_{k-1}(i)}-1)}\right\}.$$

In the same fashion as above it can be seen that the set we are minimizing contains $h_{k-1}(i)-1$ so that $h_k(i) < h_{k-1}(i)$. Then, as before, we may define

$$W_i^k := \mathscr{A}_{M_{h_k(i)}}\left(\mathscr{C}_{n_{j(h_k(i))}(h_k(i))}(W_i^{k-1}) \textcircled{z} F_{j(h_k(i))}(h_k(i))\right),$$

which is then $M_{h_k(i)}d_{h_k(i)}^{2(M_{h_k(i)}-1)}$-regular. Since the integer sequence $h_1(i), h_2(i), \cdots$ is strictly decreasing, this proces does terminate and we have our graph sequence $W_i^0, \cdots, W_i^{l(i)}$. Since $h_{l(i)}(i) = 1$ the degree of $W_i^{l(i)}$ is $2d_1^2$ so if we put $H_i = W_i^{l(i)}$ and $d = 2d_1^2$ we have, through a tour de force of integer sequences and subindexes, defined our d-regular graph sequence. Also, since by construction $|V(H_i)| \geq |V(W_i^{l(i)-1})| \geq \cdots \geq |V(W_i^0)| = n_{j(i)}(i) \geq M_{i+1}$ we have that $|V(H_i)| \to \infty$. It remains to prove the claim regarding the absolute spectral gaps. Note that we may assume, without destroying the argument, that all the absolute spectral gaps are at least 1 (all the bounds will be equally valid if we replace $\gamma_+(\,\cdot\,, K)$ with $\max\{1, \gamma_+(\,\cdot\,, K)\}$). Then by Lemma 5.4 and Theorem 5.6 we get for $i > 1$ and every $k \in \{1, \cdots, l(i)\}$ that

$$\begin{aligned}\gamma_+(W_i^k, K) &\leq k_2(K)\max\left\{1, \frac{2\gamma_+(W_i^{k-1}, K)\gamma_+(F_{j(h_k(i))}(h_k(i)), K)^2}{M_{h_k(i)}^{1/k_2(K)}}\right\}\\ &\leq 2k_2(K)\gamma_+(W_i^{k-1}, K)\gamma_+(F_{j(h_k(i))}(h_k(i)), K)^2. \qquad (29)\end{aligned}$$

Hence, by iterating over $l(i), l(i)-1, \cdots, 1$ we obtain

$$\gamma_+(H_i, K) = \gamma_+(W_i^{l(i)}, K) \leq (2k_2(K))^{l(i)}\gamma_+(F_{j(i)}(i), K) \prod_{k=1}^{l(i)} \gamma_+(F_{j(h_k(i))}(h_k(i)), K)^2 < \infty,$$

due to the assumptions of the lemma. Since $\gamma_+(H_1, K) < \infty$ we have demonstrated that $\gamma_+(H_i, K) < \infty$, for all $i \geq 1$. To prove a uniform bound put $k_3(K) = \max\{k_1(K), k_2(K)\}$ and fix $i > k_3(K)$. We now claim that it follows by induction on $k \in \{0, \cdots, l(i)\}$ that

$$h_k(i) > k_3(K) \implies \gamma_+(W_i^k, K) \leq k_3(K) \tag{30}$$

where we define $h_0(i) = i$. Indeed, for $k = 0$ we have $h_0(i) = i > k_3(K) \geq k_1(K)$ by assumption on i, and

$$\gamma_+(W_i^0, K) = \gamma(F_{j(i)}(i), K) \leq k_1(K) \leq k_3(K),$$

so the induction start holds. For $k > 0$, if $k-1$ satisfies (30) and $h_k(i) > k_3(K)$, then since $h_{k-1}(i) > h_k(i)$ we get by the induction hypothesis that $\gamma_+(W_i^{k-1}, K) \leq k_3(K)$. This yields, by a similar computation as above, that

$$\begin{aligned}\gamma_+(W_i^k, K) &\leq k_2(K) \max\left\{1, \frac{2k_3(K)\gamma_+(F_{j(h_k(i))}(h_k(i)), K)^2}{M_{h_k(i)}^{1/k_2(K)}}\right\} \\ &\leq k_3(K) \max\left\{1, \frac{2k_3(K)^3}{M_{k_3(K)}^{1/k_3(K)}}\right\} = k_3(K),\end{aligned}$$

where we used that $h_k(i) > k_3(K)$ and the way M_k was defined. This demonstrates the claim. Since $h_0(i) = i > k_3(K)$ we may define $k_0 = \max\{k \in \{0, \cdots, l(i)-1\} \mid h_k(i) > k_3(K)\}$. Then since $k_3(K) \geq h_{k_0+1}(i) > h_{k_0+2}(i) > \cdots > h_{l(i)}(i) = 1$ we conclude that the inclusion $\{h_{k_0+1}(i), \cdots, h_{l(i)}(i)\} \subset [k_3(K)]$ holds and so by an iterative application of (29) we obtain

$$\begin{aligned}\gamma_+(H_i, K) &\leq \gamma_+(W_i^{k_0}, K) \prod_{k=k_0+1}^{l(i)} 2k_2(K)\gamma_+(F_{j(h_k(i))}(h_k(i)))^2 \\ &\leq k_3(K)(2k_3(K))^{k_3(K)} \prod_{r=1}^{k_3(K)} \gamma_+(F_{j(r)}(r)) =: C < \infty.\end{aligned}$$

Note that C is independent of $i > k_3(K)$. Hence we have

$$\sup_{i\geq 1} \gamma_+(H_i, K) \leq \max\left\{\gamma_+(H_1, K), \cdots, \gamma_+(H_{k_3(K)}, K), C\right\} < \infty,$$

which demonstrates the lemma. □

Let $\mathscr{K}$ be the family of squares of norms on superreflexive Banach spaces. By corollary 5.16 we obtain the existence of an array of graphs such that condition (27) of the above lemma is satisfied. By Theorems 5.8 and 5.9 we infer that condition (28) is too. We obtain the following:

Theorem 5.18. *There exists $d \in \mathbb{N}$ and a sequence of d-regular graphs, which is a sequence of superexpanders.*

The proof of Lemma 5.15 from [MN14] gives an explicit construction of the $H_n(\delta)$'s through quotients of graphs on $\mathbb{F}_2^n$ so Theorem 5.18 is constructive, and each H_i even comes about through a finite number of basic graph operations. Performing these operations is arguably a tedious matter, especially since there are principally no bounds on the constants involved, i.e., the metric Markov cotype $p(B)$, the constant $k(B)$ from Lemma 5.16, the number of iterations $l(i)$ before we arrive at H_i, and the computation of the numbers in the auxiliary sequence $(h_k(i))_{k=0}^{l(i)}$.

6 Superreflexive Banach spaces have metric Markov cotype

The quantitative decay of the absolute spectral gap when taking Césaro averages, yielding the non-linear spectral calculus for superreflexive Banach spaces, hinges on the property of metric Markov cotype. In this section we will establish this property for superreflexive Banach spaces, following [MN14]. To this end we shall first introduce some preliminary notions.

6.1 Uniform convexity and uniform smoothness

Let B be a Banach space, and recall that its modulus of uniform convexity is given by

$$\delta_B(\varepsilon) = \inf\left\{1 - \frac{||x-y||}{2} \mid ||x|| = ||y|| = 1,\ ||x-y|| = \varepsilon\right\}, \quad \varepsilon \in [0,2],$$

and that B is uniformly convex if and only if $\delta_B(\varepsilon) > 0$ whenever $\varepsilon > 0$. Moreover, we say that B has modulus of convexity of power type $p \geq 2$ if there exists a constant $c > 0$ such that $\delta_B(\varepsilon) \geq c\varepsilon^p$, for all $\varepsilon \in [0,2]$. We have the following characterization of uniform convexity of power type p, due to [BCL94]:

Theorem 6.1. *A Banach space B has modulus of convexity of power type p if and only if there exists $K \geq 1$ such that for every $x, y \in B$ we have*

$$||x||^p + \frac{||y||^p}{K^p} \leq \frac{||x+y||^p + ||x-y||^p}{2}. \tag{31}$$

The infimum over such K is denoted by $K_p(B)$.

Definition 6.2. The *modulus of uniform smoothness* of B is given by

$$\rho_B(\tau) = \sup\left\{\frac{||x+\tau y|| + ||x-\tau y||}{2} - 1 \mid ||x|| = ||y|| = 1\right\}, \quad \tau > 0,$$

and we say that B is *uniformly smooth* if $\lim_{\tau\to 0}\rho_B(\tau)/\tau = 0$.

Similarly B has modulus of smoothness of power type $p \in (1,2]$ if there is $C > 0$ such that $\rho_B(\tau) \leq C\tau^p$, for all $\tau > 0$. We have, again due to [BCL94], the following characterization of uniform smoothness of power type p:

Theorem 6.3. *A Banach space B has modulus of smoothness of power type p if and only if there exists $S \geq 1$ such that for every $x, y \in B$ we have*

$$\frac{||x+y||^p + ||x-y||^p}{2} \leq ||x||^p + S^p||y||^p. \tag{32}$$

The infimum over such S is denoted $S_p(B)$.

From [Lin63] we have the following duality relation between the modulus of smoothness and convexity respectively:

Lemma 6.4. *For any Banach space B we have*

$$\rho_{B^*}(\tau) = \sup_{0\leq\varepsilon\leq 2}\left\{\frac{\tau\varepsilon}{2} - \delta_B(\varepsilon)\right\}.$$

Proof. Fix $\tau > 0$. For $0 \le \varepsilon \le 2$ let $x, y \in B$ with $||x|| = ||y|| = 1$ and $||x - y|| = \varepsilon$, and take, by the Hahn-Banach theorem, $f, g \in B^*$ with $||f|| = ||g|| = 1$ and $f(x + y) = ||x + y||$ and $g(x - y) = ||x - y|| = \varepsilon$. Then we have

$$\begin{aligned}
\rho_{B^*}(\tau) &\ge \frac{||f + \tau g|| + ||f - \tau g||}{2} - 1 \\
&\ge \frac{|f(x) + \tau g(x)| + |f(y) - \tau g(y)|}{2} - 1 \\
&\ge \frac{|f(x + y) + \tau g(x - y)|}{2} - 1 \\
&= \frac{||x + y||}{2} + \frac{\tau\varepsilon}{2} - 1 = \frac{\tau\varepsilon}{2} - \left(1 - \frac{||x + y||}{2}\right),
\end{aligned}$$

yielding $\rho_{B^*}(\tau) \ge \sup_{0\le\varepsilon\le 2} \{\tau\varepsilon/2 - \delta_B(\varepsilon)\}$. For the reverse inequality let $f, g \in B^*$ be of unit norm. For any $\eta > 0$ take $x, y \in B$ with $||x|| = ||y|| = 1$ such that $|f(x)+\tau g(x)| > ||f+\tau g||-\eta$ and $|f(y)-\tau g(y)| > ||f-\tau g||-\eta$. By multiplying by a phase we may assume that $f(x)+\tau g(x)$ and $f(y) - \tau g(y)$ are real and positive. Then we have

$$\begin{aligned}
\frac{||f + \tau g|| + ||f - \tau g||}{2} - 1 &< \frac{f(x) + \tau g(x) + f(y) - \tau g(y) + 2\eta}{2} - 1 \\
&= \frac{f(x + y) + \tau g(x - y) + 2\eta}{2} - 1 \\
&\le \frac{||x + y|| + \tau||x - y||}{2} - 1 + \eta \\
&= \frac{\tau||x - y||}{2} - \left(1 - \frac{||x + y||}{2}\right) + \eta \\
&\le \frac{\tau||x - y||}{2} - \delta_B\left(||x - y||\right) + \eta \le \sup_{0\le\varepsilon\le 2} \left\{\frac{\tau\varepsilon}{2} - \delta_B(\varepsilon)\right\} + \eta,
\end{aligned}$$

using that $||x - y|| \in [0, 2]$. Letting $\eta \searrow 0$, we obtain the desired. □

Using Lemma 6.4, it is not difficult to see that B is uniformly convex (resp. smooth) if and only if B^* is uniformly smooth (resp. convex). An application of Hölder's inequality yields the duality relation (see [BCL94, Lemma 5]):

$$K_p(B) = S_{p/(p-1)}(B^*). \tag{33}$$

For smoothness constants of Bochner spaces we have the following estimate

Lemma 6.5. *For every $1 < p \le 2$ and $q \ge p$ and any σ-finite measure space $(S, \mathscr{A}, \mu)$ we have*

$$S_p(L_q(S; B)) \le (5pq)^{1/p} S_p(B).$$

Proof. We will give the main idea of the proof. Suppose that there exists $K \ge 1$ such that for every $x, y \in B$ we have

$$\frac{||x + y||^q + ||x - y||^q}{2} \le \left(||x||^p + K||y||^p\right)^{q/p}. \tag{34}$$

Then for any $f, g \in L_q(\mu, B)$ we have

$$\begin{aligned}
\frac{\|f+g\|_{L_q}^p + \|f-g\|_{L_q}^p}{2} &\leq \left(\frac{\|f+g\|_{L_q}^q + \|f-g\|_{L_q}^q}{2} \right)^{p/q} \\
&= \left(\int_S \frac{\|f+g\|^q + \|f-g\|^q}{2} d\mu \right)^{p/q} \\
&\leq \left(\int_S (||f||^p + K||g||^p)^{q/p} \right)^{p/q} \\
&= \|(\|f\|^p + K\,\|g\|^p)\|_{L_{q/p}} \\
&\leq \|f\|_{L_q}^p + K\,\|g\|_{L_q}^p .
\end{aligned}$$

The first inequality follows from the fact that $t \mapsto t^{p/q}$ is concave (since $q \geq p$) the second is due to (34) and the third is the triangle inequality and the definition of norms on Bochner spaces. This yields $S_p(L_q(S;B))^p \leq K$. In [MN14] the estimate from (34) is proven with the constant $K = 5pqS_p(B)^p$ which yields the desired. $\square$

Remark 6.6. Recall the well-known duality relation $L_p(\Omega)^* = L_{p/(p-1)}(\Omega)$ for scalar valued L_p spaces. For Bochner spaces the relation $L_p(\Omega; B)^* = L_{p/(p-1)}(\Omega; B^*)$ unfortunately does not hold in general. It does hold, however, if B has the so-called *Radon-Nikodym property* (RNP), which allows one to extend the proof of the dual relation of scalar valued L_p spaces, using the Radon-Nikodym theorem, to Bochner spaces on B, see [Hyt+16, sec. 1.3]. Here it is also shown that reflexive spaces have the RNP, so in particular superreflexive spaces do.

The above remark and (33) now yield the following estimate for convexity constants:

Corollary 6.7. *Suppose B is reflexive. Then for $p \geq 2$ and $1 < q \leq p$ and any σ-finite measure space $(S, \mathscr{A}, \mu)$ we have*

$$K_p(L_q(S;B)) \leq \left(\frac{5pq}{(p-1)(q-1)} \right)^{1-1/p} K_p(B).$$

Proof. By (33) we get

$$K_p(L_q(S;B)) = S_{p/(p-1)}(L_q(S;B)^*) = S_{p/(p-1)}(L_{q/(q-1)}(S;B^*)).$$

Since $p \geq 2$ and $1 < q \leq p$ we have $1 \leq \frac{p}{p-1} \leq 2$ and $\frac{q}{q-1} \geq \frac{p}{p-1}$ and hence Lemma 6.5 and (33) yield

$$K_p(L_q(S;B)) \leq \left(5 \cdot \frac{p}{p-1} \cdot \frac{q}{q-1} \right)^{\frac{p-1}{p}} S_{p/(p-1)}(B^*) = \left(\frac{5pq}{(p-1)(q-1)} \right)^{1-1/p} K_p(B),$$

as desired. $\square$

We will also need the following observation: For $q \geq p$, recall that, by Hölder's inequality, $L_q(\Omega) \subset L_p(\Omega)$ for any probability space Ω and $\ell_p(A) \subset \ell_q(A)$ for any set A where both inclusions are contractive. Hence, for all $x, y \in B$, we have

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

and

$$\left(||x||^q + \frac{||y||^q}{K^q}\right)^{1/q} \leq \left(||x||^p + \frac{||y||^p}{K^p}\right)^{1/p}.$$

These two facts together yield that

$$K_q(B) \leq K_p(B) \quad \text{whenever} \quad q \geq p \tag{35}$$

It can be seen analogously that $S_q(B) \leq S_p(B)$ whenever $q \leq p$.

Recall from Section 1 that B is superreflexive if and only if it admits an equivalent norm which has modulus of convexity of power type p for some $p \geq 2$ (see Theorem 1.28). This motivates our interest in such norms and we shall prove metric Markov cotype for spaces with modulus of convexity of power type p.

Proposition 6.8. *Let B be a Banach space with modulus of convexity of power type $p \geq 2$. Then for every random vector U with values in B such that $E||U||^p < \infty$ we have*

$$||EU||^p + \frac{1}{(2^{p-1}-1)K_p(B)^p} E||U - EU||^p \leq E||U||^p. \tag{36}$$

Before we give the proof we observe that the existence of a constant such that (36) holds for every p-integrable random vector in fact gives an equivalent characterization of uniform convexity of power type p; this due to [Pis75]. It can be thought of as a generalization of the well-known variance formula for Hilbert space valued random vectors; namely that, in this case, equality holds in (36) with constant 1.

Proof of Proposition 6.8. Note that 36 follows from Jensen's inequality if $E\|U - EU\|^p = 0$, so we can restrict ourselves to the case $E\|U - EU\|^p > 0$. Let θ denote the supremum over all constants satisfying (36) for every p-integrable random vector with non-zero centralized pth moment. Observe that this supremum is finite and non-negative by Jensen's inequality. Fix an arbitrary $\varphi > \theta$. Then there exists a p-integrable random vector, U_0, such that $E\|U_0 - EU_0\|^p > 0$ and $\varphi E||U_0 - EU_0||^p > E||U_0||^p - ||EU_0||^p$. Applying inequality (31) pointwise to the vectors $U_0/2 + EU_0/2$ and $U_0/2 - EU_0/2$ we obtain

$$2\left\|\frac{U_0}{2} + \frac{EU_0}{2}\right\|^p + \frac{2}{K_p(B)^p}\left\|\frac{U_0}{2} - \frac{EU_0}{2}\right\|^p \leq ||U_0||^p + ||EU_0||^p.$$

By the way U_0 was chosen and θ defined, integrating this on either side yields

$$\begin{aligned}
\varphi E||U_0 - EU_0||^p &> E||U_0||^p - ||EU_0||^p \\
&\geq 2\left(E\left\|\frac{U_0}{2}+\frac{EU_0}{2}\right\|^p - \left\|E\left(\frac{U_0}{2}+\frac{EU_0}{2}\right)\right\|^p\right) + \frac{2}{K_p(B)^p}E\left\|\frac{U_0}{2}-\frac{EU_0}{2}\right\|^p \\
&\geq 2\theta E\left\|\frac{U_0}{2}+\frac{EU_0}{2} - E\left(\frac{U_0}{2}+\frac{EU_0}{2}\right)\right\|^p + \frac{2}{K_p(B)^p}E\left\|\frac{U_0}{2}-\frac{EU_0}{2}\right\|^p \\
&= 2\theta E\left\|\frac{U_0}{2}-\frac{EU_0}{2}\right\|^p + \frac{2}{K_p(B)^p}E\left\|\frac{U_0}{2}-\frac{EU_0}{2}\right\|^p \\
&= \left(\frac{\theta}{2^{p-1}}+\frac{1}{2^{p-1}K_p(B)^p}\right)E\left\|U_0 - EU_0\right\|^p.
\end{aligned}$$

This yields $\varphi > \theta/2^{p-1} + 1/(2^{p-1}K_p(B)^p)$ and letting $\varphi \searrow \theta$ we obtain $\theta \geq \frac{1}{(2^{p-1}-1)K_p(B)^p}$ yielding the desired. □

6.2 Vector valued martingales and metric Markov cotype

Let B be a Banach space with $K_p(B) < \infty$ for some $p \geq 2$. Furthermore, let $(M_k)_{k=0}^n$ be a sequence of p-integrable random vectors with values in B which is a martingale with respect to some filtration $\mathcal{F}_0 \subset \cdots \subset \mathcal{F}_n$. By reiterating the proof of Proposition 6.8 with the unconditional expectation replaced by the conditional we obtain

$$\begin{aligned}
E\left(\|M_n - M_0\|^p \mid \mathcal{F}_{n-1}\right) &\geq \left\|E\left(M_n - M_0 \mid \mathcal{F}_{n-1}\right)\right\|^p \\
&\quad + \frac{1}{(2^{p-1}-1)K_p(B)^p}E\left(\left\|M_n - M_0 - E\left(M_n - M_0 \mid \mathcal{F}_{n-1}\right)\right\|^p \mid \mathcal{F}_{n-1}\right) \\
&= \|M_{n-1} - M_0\|^p + \frac{1}{(2^{p-1}-1)K_p(B)^p}E\left(\|M_n - M_{n-1}\|^p \mid \mathcal{F}_{n-1}\right).
\end{aligned}$$

Integrating both sides of this inequality yields

$$E\left\|M_n - M_0\right\|^p \geq E\left\|M_{n-1} - M_0\right\|^p + \frac{1}{(2^{p-1}-1)K_p(B)^p}E\left\|M_n - M_{n-1}\right\|^p.$$

Iterating over $n-1, \cdots, 1$ gives the following inequality attributed to [Pis75]:

$$E\left\|M_n - M_0\right\|^p \geq \frac{1}{(2^{p-1}-1)K_p(B)^p}\sum_{k=1}^{n} E\left\|M_k - M_{k-1}\right\|^p. \tag{37}$$

We shall need the following extention of Pisier's inequality:

Lemma 6.9. *Let $p \geq 2$ and $q > 1$. Let B be a Banach space with $K_p(B) < \infty$ and $(M_k)_{k=0}^n$ a B-valued, q-integrable martingale with respect to the filtration $\mathcal{F}_0 \subset \cdots \subset \mathcal{F}_n$. If $q \geq p$ then*

$$E\left\|M_n - M_0\right\|^q \geq \frac{1}{(2^{q-1}-1)K_p(B)^q}\sum_{k=1}^{n} E\left\|M_k - M_{k-1}\right\|^q,$$

and if $q \le p$ then

$$E\left\|M_n - M_0\right\|^q \ge \frac{((1-1/p)(1-1/q))^{q(1-1/p)}}{5^{q(1-1/p)}(2K_p(B))^q n^{1-q/p}} \sum_{k=1}^{n} E\left\|M_k - M_{k-1}\right\|^q .$$

Note that the estimate also holds for $q = 1$, but in this case the statement is quite trivial.

Proof. The case $q \ge p$ is a consequence of (37) and (35) so assume $q \le p$. Then by corollary 6.7 we have

$$K_p(L_q(\Omega; B)) \le \left(\frac{5pq}{(p-1)(q-1)}\right)^{1-1/p} K_p(B) =: K,$$

where (Ω, μ) is the probability space on which the martingale is defined. By definition of $K_p(L_q(\mu, B))$ we then have

$$\begin{aligned}
&\left(E\left\|M_{n-1} - M_0 + \frac{M_n - M_{n-1}}{2}\right\|^q\right)^{p/q} + \frac{1}{(2K)^p}\left(E\left\|M_n - M_{n-1}\right\|^q\right)^{p/q} \\
&= \left\|M_{n-1} - M_0 + \frac{M_n - M_{n-1}}{2}\right\|_{L_q}^p + \frac{1}{K^p}\left\|\frac{M_n - M_{n-1}}{2}\right\|_{L_q}^p \\
&\le \frac{\left(E\left\|M_n - M_0\right\|^q\right)^{p/q} + \left(E\left\|M_{n-1} - M_0\right\|^q\right)^{p/q}}{2}.
\end{aligned} \tag{38}$$

Using the martingale property and conditional Jensen's inequality we obtain the following two estimates:

$$\begin{aligned}
E\left\|M_{n-1} - M_0\right\|^q &= E\left\|E\left(M_n - M_0 \mid \mathcal{F}_{n-1}\right)\right\|^q \\
&\le E\left(E\left(||M_n - M_0||^q \mid \mathcal{F}_{n-1}\right)\right) = E\left\|M_n - M_0\right\|^q ,
\end{aligned}$$

and

$$\begin{aligned}
E\left\|M_{n-1} - M_0 + \frac{M_n - M_{n-1}}{2}\right\|^q &= E\left(E\left(\left\|M_{n-1} - M_0 + \frac{M_n - M_{n-1}}{2}\right\|^q \mid \mathcal{F}_{n-1}\right)\right) \\
&\ge E\left\|E\left(M_{n-1} - M_0 + \frac{M_n - M_{n-1}}{2} \mid \mathcal{F}_{n-1}\right)\right\|^q = E\left\|M_{n-1} - M_0\right\|^q .
\end{aligned}$$

Inserting these in (38) yields

$$\left(E\left\|M_{n-1} - M_0\right\|^q\right)^{p/q} + \frac{1}{(2K)^p}\left(E\left\|M_n - M_{n-1}\right\|^q\right)^{p/q} \le \left(E\left\|M_n - M_0\right\|^q\right)^{p/q} .$$

Repeating the argument for the first term on the left iteratively, we obtain

$$\begin{aligned}
(2K)^p\left(E\left\|M_n - M_0\right\|^q\right)^{p/q} &\ge \sum_{k=1}^{n}\left(E\left\|M_k - M_{k-1}\right\|^q\right)^{p/q} \\
&\ge n\left(\frac{1}{n}\sum_{k=1}^{n} E\left\|M_k - M_{k-1}\right\|^q\right)^{p/q} \\
&= \frac{1}{n^{p/q-1}}\left(\sum_{k=1}^{n} E\left\|M_k - M_{k-1}\right\|^q\right)^{p/q} ,
\end{aligned}$$

where we used convexity of the function $t \mapsto t^{p/q}$. Raising both sides to the power $\frac{q}{p}$ yields the desired estimate. $\square$

Before we state the main theorem of this section, recall from Section 4, that for an $n \times n$ scalar valued matrix $A = (a_{ij})$ we define the linear operator $A \otimes I_B^n : L_p^n(B) \to L_p^n(B)$ by $(A \otimes I_B^n)f(i) = \sum_{j=1}^n a_{ij} f(j)$, for $f \in L_p^n(B)$ and $i \in [n]$.

Theorem 6.10. *Let B be a Banach space with $K_p(B) < \infty$ for some $p \geq 2$. Then for every $n, t \in \mathbb{N}$, every $n \times n$ symmetric, stochastic matrix $A = (a_{ij})$ and every $x_1, \cdots, x_n \in B$ there exist $y_1, \cdots, y_n$ such that, for all $q > 1$, we have*

$$\max\left\{ \sum_{i=1}^n \|x_i - y_i\|^q , \left(\frac{((1-1/p)(1-1/q))^{1-1/p}}{16 \cdot 5^{1-1/p} K_p(B)} \right)^q t^{q/p} \sum_{i,j=1}^n a_{ij} \|y_i - y_j\|^q \right\} \leq \sum_{i,j}^n \mathscr{A}_t(A)_{ij} \|x_i - x_j\|^q .$$

Proof. The structure of the proof is as follows: First we find a suitable martingale, integrating which gives initial estimates using the above martingale inequalities. Then we choose our y_i's cleverly and obtain bounds on either side of these estimates to yield the desired inequality. Let us first find our martingale. To this end, define $f \in L_p^n(B)$ by $f(i) = x_i$ and for each $l \in [n]$ let $Z_0^l, Z_1^l, \cdots Z_t^l$ denote the Markov chain on $[n]$ which starts at l and has transition matrix A, i.e., $Z_0^l = l$ a.s. and $P\left(Z_m^l = j \mid Z_{m-1}^l = i\right) = a_{ij}$, for all $m \in [t]$ and $i, j \in [n]$. For $0 \leq m \leq t$ define $f_m = (A^{t-m} \otimes I_B^n) f \in L_p^n(B)$ and $M_m^l = f_m(Z_m^l)$, for $l \in [n]$ and $0 \leq m \leq t$. We claim that $(M_m^l)_{m=0}^t$ is a martingale with respect to the filtration generated by $Z_0^l, \cdots, Z_t^l$. Indeed, if we interpret $1/P(Z_{m-1}^l = i)$ as 0 whenever $(Z_{m-1}^l = i)$ is a null set and let $L := A \otimes I_B^n$, we have by example 1.55 that

$$\begin{aligned}
E\left(f(Z_m^l) \mid Z_{m-1}^l\right) &= \sum_{i=1}^n 1_{(Z_{m-1}^l = i)} \frac{1}{P(Z_{m-1}^l = i)} \int_{(Z_{m-1}^l = i)} f(Z_m^l) dP \\
&= \sum_{i=1}^n 1_{(Z_{m-1}^l = i)} \frac{1}{P(Z_{m-1}^l = i)} \sum_{j=1}^n x_j P(Z_m^l = j, Z_{m-1}^l = i) \\
&= \sum_{i=1}^n 1_{(Z_{m-1}^l = i)} \frac{1}{P(Z_{m-1}^l = i)} \sum_{j=1}^n x_j P(Z_m^l = j \mid Z_{m-1}^l = i) P(Z_{m-1}^l = i) \\
&= \sum_{i=1}^n 1_{(Z_{m-1}^l = i)} \sum_{j=1}^n x_j a_{ij} = \sum_{i=1}^n 1_{(Z_{m-1}^l = i)} (Lf)(i),
\end{aligned}$$

demonstrating that $E(f(Z_m^l) \mid Z_{m-1}^l) = Lf(Z_{m-1}^l)$. Now, using basic properties of Markov chains, we obtain

$$\begin{aligned}
E\left(M_m^l \mid Z_0^l, \cdots, Z_{m-1}^l\right) &= E\left(L^{t-m} f(Z_m^l) \mid Z_{m-1}^l\right) = L^{t-m} E\left(f(Z_m^l) \mid Z_{m-1}^l\right) \\
&= L^{t-m}(Lf(Z_{m-1}^l)) = L^{t-(m-1)}(Z_{m-1}^l) = M_{m-1}^l,
\end{aligned}$$

demonstrating the claim (we are here also using some well-known properties of conditional expectations for scalar valued random variables, which carry over to the vector valued case *verbatim*). Hence by Lemma 6.9 (the integrability assumptions are automatically satisfied) we have for each $l \in [n]$ that $KE\left\|M_t^l - M_0^l\right\|^q \geq \sum_{m=1}^t E\left\|M_m^l - M_{m-1}^l\right\|^q$, where

$$K = \begin{cases} \left(2^{q-1}-1\right) K_p(B)^q & \text{if } q \geq p \\ \frac{5^{q(1-1/p)}(2K_p(B))^q t^{1-q/p}}{((1-1/p)(1-1/q))^{q(1-1/p)}} & \text{if } q \leq p \end{cases}.$$

Now let $Z_0, \cdots, Z_t$ be the Markov chain on $[n]$ with transition matrix A where Z_0 is uniformly distributed on $[n]$. Then, using the formula defining M_m^l and averaging over $l \in [n]$, we obtain

$$KE\left\|f(Z_t) - (L^t f)(Z_0)\right\|^q \geq \sum_{m=1}^t E\left\|(L^{t-m} f)(Z_m) - (L^{t-m+1} f)(Z_{m-1})\right\|^q. \tag{39}$$

Using that A is the transition matrix and the Markov chain property the left hand side equals $K\sum_{i,j=1}^n \left\|f(i) - (L^t f)(j)\right\|^q P(Z_t = i, Z_0 = j) = K\sum_{i,j=1}^n \left\|f(i) - (L^t f)(j)\right\|^q \frac{(A^t)_{ij}}{n}$. A similar computation for the right hand side yields that (39) is equivalent to

$$K\sum_{i,j=1}^n (A^t)_{ij}\left\|f(i) - (L^t f)(j)\right\|^q \geq \sum_{m=1}^t \sum_{i,j=1}^n a_{ij}\left\|(L^{t-m} f)(i) - (L^{t-m+1} f)(j)\right\|^q. \tag{40}$$

Now we define $y_i = \frac{1}{t}\sum_{k=0}^{t-1}(L^k f)(i) = \frac{1}{t}\sum_{j=1}^n \sum_{k=0}^{t-1}(A^k)_{ij} x_j$, for $i \in [n]$. The rest of the proof is a matter of obtaining estimates of either side of (40) in terms of the x_i's and y_i's using convexity. First observe that since A^t is symmetric and stochastic we have

$$\frac{1}{t}\sum_{k=1}^t (L^k f)(i) = y_i + \frac{1}{t}(L^t f)(i) - \frac{1}{t}x_i = y_i - \frac{1}{t}\sum_{s=1}^n (A^t)_{is}(x_i - x_s).$$

This observation together with convexity of $x \mapsto \|x\|^q$ yields

$$\begin{aligned}
&\sum_{m=1}^t \sum_{i,j=1}^n a_{ij}\left\|(L^{t-m} f)(i) - (L^{t-m+1} f)(j)\right\|^q \\
&\geq t\sum_{i,j=1}^n a_{ij}\left\|\frac{1}{t}\sum_{m=1}^t \left((L^{t-m} f)(i) - (L^{t-m+1} f)(j)\right)\right\|^q \\
&= t\sum_{i,j=1}^n a_{ij}\left\|y_i - y_j + \frac{1}{t}\sum_{s=1}^n (A^t)_{js}(x_j - x_s)\right\|^q \\
&\geq \frac{t}{2^{q-1}}\sum_{i,j=1}^n a_{ij}\left\|y_i - y_j\right\|^q - \frac{1}{t^{q-1}}\sum_{i,j=1}^n a_{ij}\left\|\sum_{s=1}^n (A^t)_{js}(x_j - x_s)\right\|^q \\
&= \frac{t}{2^{q-1}}\sum_{i,j=1}^n a_{ij}\left\|y_i - y_j\right\|^q - \frac{1}{t^{q-1}}\sum_{j=1}^n \left\|\sum_{s=1}^n (A^t)_{js}(x_j - x_s)\right\|^q \\
&\geq \frac{t}{2^{q-1}}\sum_{i,j=1}^n a_{ij}\left\|y_i - y_j\right\|^q - \frac{1}{t^{q-1}}\sum_{j,s=1}^n (A^t)_{js}\left\|x_j - x_s\right\|^q.
\end{aligned} \tag{41}$$

The left hand side of 40 has the following estimate:

$$\begin{aligned}
&\sum_{i,j=1}^{n}(A^t)_{ij}\left\|f(i)-(L^tf)(j)\right\|^q=\sum_{i,j=1}^{n}(A^t)_{ij}\left\|\sum_{s=1}^{n}(A^t)_{js}(x_i-x_s)\right\|^q\\
&\le\sum_{i,j,s=1}^{n}(A^t)_{ij}(A^t)_{js}\left\|x_i-x_s\right\|^q\\
&\le 2^{q-1}\sum_{i,j,s=1}^{n}(A^t)_{ij}(A^t)_{js}(\left\|x_i-x_j\right\|^q+\left\|x_j-x_s\right\|^q)\\
&=2^q\sum_{i,j=1}^{n}(A^t)_{ij}\left\|x_i-x_j\right\|^q.
\end{aligned}\tag{42}$$

Using that $0=\sum_{i,j=1}^{n}\delta_{ij}\left\|x_i-x_j\right\|^q=\sum_{i,j=1}^{n}(I_n)_{ij}\left\|x_i-x_j\right\|^q$, where I_n is the $n\times n$ identity matrix, we have

$$\begin{aligned}
&\sum_{i,j=1}^{n}(A^t)_{ij}\left\|x_i-x_j\right\|^q=\sum_{i,j=1}^{n}\left(\frac{1}{t}\sum_{m=0}^{t-1}A^mA^{t-m}\right)_{ij}\left\|x_i-x_j\right\|^q\\
&\le\frac{2^{q-1}}{t}\sum_{i,j=1}^{n}\sum_{s=1}^{n}\sum_{m=0}^{t-1}(A^m)_{is}(A^{t-m})_{sj}\left(\left\|x_i-x_s\right\|^q+\left\|x_s-x_j\right\|^q\right)\\
&=2^{q-1}\sum_{i,s=1}^{n}\left(\frac{1}{t}\sum_{m=0}^{t-1}A^m\right)_{is}\left\|x_i-x_s\right\|^q+2^{q-1}\sum_{j,s=1}^{n}\left(\frac{1}{t}\sum_{m=0}^{t-1}A^{t-m}\right)_{sj}\left\|x_s-x_j\right\|^q\\
&=2^q\sum_{i,j=1}^{n}\mathscr{A}_t(A)_{ij}\left\|x_i-x_j\right\|^q+\frac{2^{q-1}}{t}\sum_{i,j=1}^{n}(A^t)_{ij}\left\|x_i-x_j\right\|^q.
\end{aligned}$$

If $t\ge 2^q$, this yields

$$\sum_{i,j=1}^{n}(A^t)_{ij}\left\|x_i-x_j\right\|^q\le 2^{q+1}\sum_{i,j=1}^{n}\mathscr{A}_t(A)_{ij}\left\|x_i-x_j\right\|^q,$$

and if $t\le 2^q$ we have the bound

$$\begin{aligned}
&\sum_{i,j=1}^{n}(A^t)_{ij}\left\|x_i-x_j\right\|^q=\sum_{i,j,r=1}^{n}a_{ir}(A^{t-1})_{rj}\left\|x_i-x_j\right\|^q\\
&\le 2^{q-1}\sum_{i,j,r=1}^{n}a_{ir}(A^{t-1})_{rj}\left(\left\|x_i-x_r\right\|^q+\left\|x_r-x_j\right\|^q\right)\\
&=2^{q-1}\sum_{i,r=1}^{n}a_{ir}\left\|x_i-x_r\right\|^q+2^{q-1}\sum_{j,r=1}^{n}(A^{t-1})_{rj}\left\|x_r-x_j\right\|^q\\
&\le 2^{q-1}t\sum_{i,j=1}^{n}\mathscr{A}_t(A)_{ij}\left\|x_i-x_j\right\|^q\\
&\le 2^{2q-1}\sum_{i,j=1}^{n}\mathscr{A}_t(A)_{ij}\left\|x_i-x_j\right\|^q.
\end{aligned}$$

In either case we have

$$\sum_{i,j=1}^{n} (A^t)_{ij} \|x_i - x_j\|^q \leq 4^q \sum_{i,j=1}^{n} \mathscr{A}_t(A)_{ij} \|x_i - x_j\|^q . \tag{43}$$

Finally, combining (41) and (42) into (40) we get

$$\frac{t}{2^{q-1}} \sum_{i,j=1}^{n} a_{ij} \|y_i - y_j\|^q - \frac{1}{t^{q-1}} \sum_{j,s=1}^{n} (A^t)_{js} \|x_j - x_s\|^q \leq K 2^q \sum_{i,j=1}^{n} (A^t)_{ij} \|x_i - x_j\|^q ,$$

which, after rearranging and applying (43), gives

$$t \sum_{i,j=1}^{n} a_{ij} \|y_i - y_j\|^q \leq 4^q K \sum_{i,j=1}^{n} (A^t)_{ij} \|x_i - x_j\|^q \leq 16^q K \sum_{i,j=1}^{n} \mathscr{A}_t(A) \|x_i - x_j\|^q .$$

Now, since

$$\frac{((1-1/p)(1-1/q))^{q(1-1/p)}}{5^{q(1-1/p)} 2^q K_p(B)^q t^{1-q/p}} \leq \frac{t^{q/p-1}}{(2^{q-1}-1) K_p(B)^q},$$

this yields

$$\left(\frac{((1-1/p)(1-1/q))^{1-1/p}}{16 \cdot 5^{1-1/p} K_p(B)} \right)^q t^{q/p} \sum_{i,j=1}^{n} a_{ij} \|y_i - y_j\|^q \leq \sum_{i,j=1}^{n} \mathscr{A}_t(A)_{ij} \|x_i - x_j\|^q .$$

The last thing to observe is that by definition of the y_i's we have

$$\sum_{i=1}^{n} \|x_i - y_i\|^q = \sum_{i=1}^{n} \left\| \frac{1}{t} \sum_{j=1}^{n} \sum_{k=0}^{t-1} (A^k)_{ij} (x_i - x_j) \right\|^q \leq \sum_{i,j=1}^{n} \mathscr{A}_t(A)_{ij} \|x_i - x_j\|^q .$$

This concludes the proof. □

Reducing to the case $q = 2$ we obtain the following corollary:

Corollary 6.11. *Let B be a Banach space with $K_p(B) < \infty$ for some $p \geq 2$. Then B has metric Markov cotype p with exponent 2 and we have the estimate*

$$C_p^{(2)}(B) \leq \sqrt{2} \frac{16 \cdot 5^{1-1/p} K_p(B)}{((1-1/p)/2)^{1-1/p}} \leq 320 K_p(B).$$

Proof of Theorem 5.8. Observe that metric Markov cotype p with exponent q is stable under isomorphism of normed spaces. Theorem 1.28 yields that there exists some $p \geq 2$ such that B is isomorphic to a uniformly convex Banach space of power type p. The theorem then follows from the above corollary. □

7 The Base Graph

In this section we shall describe the proof of Lemma 5.15 as given in [MN14]. It involves a beautiful blend of functional analytic results, discrete Fourier analysis as well as techniques from combinatorics. As this is a thesis in functional analysis, we will focus on the analytical parts of the proof, and state without proof the combinatorial parts. Throughout this section we will work with Banach spaces over the complex numbers only, but the results can be extended to real Banach spaces by a standard complexification strategy.

7.1 K-convex Banach spaces

We now introduce the notion of K-convex Banach spaces and prove an important theorem about such spaces which will be crucial in demonstrating the spectral properties of the base graphs from Lemma 5.15. For a Banach space B we denote by $L_p(\mathbb{F}_2^n; B)$ the Banach space of functions $f : \mathbb{F}_2^n \to B$ with the norm $\|f\|_p = \left(\frac{1}{2^n}\sum_{x\in\mathbb{F}_2^n}\|f(x)\|^p\right)^{1/p}$, where $\mathbb{F}_2$ denotes the field of order 2 (these are just Bochner spaces on the measurable space $(\mathbb{F}_2^n, \mathcal{P}(\mathbb{F}_2^n))$ equipped with the normalized counting measure, $\tau/2^n$, making it a probability space). Furthermore, for $A \subset [n]$, the *Walsh function* $W_A : \mathbb{F}_2^n \to \{\pm 1\}$ is given by $W_A(x) = (-1)^{\sum_{j\in A} x_j}$ (this definition is independent of whether the exponent is computed modulo 2 or in $\mathbb{N}$). Observe that W_A and W_B are independent whenever $A \neq B$, that W_A is a Rademacher variable whenever $A \neq \emptyset$ and that $W_\emptyset = 1$. Hence any $f : \mathbb{F}_2^n \to B$ has the unique expansion $f = \sum_{A\subset[n]} W_A \hat{f}(A)$, where

$$\hat{f}(A) := \frac{1}{2^n}\int_{\mathbb{F}_2^n} f W_a d\tau = \frac{1}{2^n}\sum_{y\in\mathbb{F}_2^n} f(y)W_A(y).$$

To see this in the scalar valued case it suffices to observe that the Walsh functions form an orthonormal basis for $L_2(\mathbb{F}_2^n)$ (note that, as sets, $L_p(\mathbb{F}_2^n; B)$ and $L_2(\mathbb{F}_2^n; B)$ are equal for any $p \geq 1$ and any Banach space B). Now the vector valued case follows by observing that for any $\psi \in B^*$ we have $\widehat{\psi \circ f} = \psi(\hat{f})$. For any function $\varphi : \mathbb{F}_2^n \to \mathbb{C}$ and $f : \mathbb{F}_2^n \to B$ we define their convolution in the usual way:

$$\varphi * f(x) := \frac{1}{2^n}\int_{\mathbb{F}_2^n} \varphi(x-y)W_A(y)d\tau(x) = \frac{1}{2^n}\sum_{y\in\mathbb{F}_2^n}\varphi(x-y)f(y) = \sum_{A\subset[n]}\hat{\varphi}(A)\hat{f}(A)W_A(x).$$

To see that $(\varphi * f)^\wedge(A) = \hat{\varphi}(A)\hat{f}(A)$ observe that, for any $x, y \in \mathbb{F}_2^n$ and $A \subset [n]$, we have $W_A(x+y) = W_A(x)W_A(y)$, and hence

$$\begin{aligned}
(\varphi * f)^\wedge(A) &= \frac{1}{2^n}\sum_{x\in\mathbb{F}_2^n}\frac{1}{2^n}\sum_{y\in\mathbb{F}_2^n}\varphi(x-y)f(y)W_A(x)\\
&= \frac{1}{2^n}\sum_{y\in\mathbb{F}_2^n}\frac{1}{2^n}\sum_{x\in\mathbb{F}_2^n}\varphi(x-y)W_A(x-y)f(y)W_A(y)\\
&= \frac{1}{2^n}\sum_{y\in\mathbb{F}_2^n}\hat{\varphi}(A)f(y)W_A(y) = \hat{\varphi}(A)\hat{f}(A).
\end{aligned}$$

Following the notation in [MN14] we define the *Rademacher projections* in the following way: for $m = 0, \cdots, n$ the m-level Rademacher projection $\mathbf{Rad}_m : L_p(\mathbb{F}_2^n; B) \to L_p(\mathbb{F}_2^n; B)$ is defined by $\mathbf{Rad}_m(f) := \sum_{|A|=m} \hat{f}(A) W_A$. Finally, for $z \in \mathbb{C}$, we define the operator $e^{z\Delta} : L_p(\mathbb{F}_2^n; B) \to L_p(\mathbb{F}_2^n; B)$ by

$$e^{z\Delta} f := \sum_{m=0}^{n} e^{zm} \mathbf{Rad}_m(f) = \sum_{A \subset [n]} e^{z|A|} \hat{f}(A) W_A = R_z * f,$$

where

$$R_z(x) := \prod_{j=1}^{n} (1 + e^z (-1)^{x_j}) = (1 - e^z)^{\|x\|_1} (1 + e^z)^{n - \|x\|_1}.$$

The fact that $\widehat{R_z}(A) = e^{z|A|}$ is seen by a simple computation in the case where $n = 1$. For $n > 1$ we proceed by induction. First consider the case where $A \subset [n]$ does not contain n. Then we can consider A as a subset of $[n-1]$ and hence

$$\begin{aligned}
\widehat{R_z}(A) &= \sum_{x \in \mathbb{F}_2^n} \left(\frac{1-e^z}{2}\right)^{\|x\|_1} \left(\frac{1+e^z}{2}\right)^{n-\|x\|_1} W_A(x) \\
&= \sum_{x \in \mathbb{F}_2^{n-1}} \left(\frac{1-e^z}{2}\right)^{\|x\|_1} \left(\frac{1+e^z}{2}\right)^{n-\|x\|_1} W_A(x) \\
&+ \sum_{x \in \mathbb{F}_2^{n-1}} \left(\frac{1-e^z}{2}\right)^{\|x\|_1+1} \left(\frac{1+e^z}{2}\right)^{n-\|x\|_1-1} W_A(x) \\
&= \left(\frac{1+e^z}{2} + \frac{1-e^z}{2}\right) \sum_{x \in \mathbb{F}_2^{n-1}} \left(\frac{1-e^z}{2}\right)^{\|x\|_1} \left(\frac{1+e^z}{2}\right)^{n-1-\|x\|_1} W_A(x) \\
&= e^{z|A|},
\end{aligned}$$

by the induction hypothesis. If $n \in A$ we can consider $A \setminus \{n\}$ as a subset of $[n-1]$ and write $\sum_{j \in A} x_j = \sum_{j \in A \setminus \{n\}} x_j + x_n$. Then we obtain by a computation analogous to the one above that

$$\widehat{R_z}(A) = \left(\frac{1+e^z}{2} - \frac{1-e^z}{2}\right) e^{z|A \setminus \{n\}|} = e^z e^{z(|A|-1)} = e^{z|A|},$$

demonstrating the claim. Using the fact that $e^{z\Delta} f = R_z * f$ we get for every $x \in \mathbb{F}_2^n$ that

$$e^{z\Delta} f(x) = \sum_{y \in \mathbb{F}_2^n} \left(\frac{1-e^z}{2}\right)^{\|x-y\|_1} \left(\frac{1+e^z}{2}\right)^{n-\|x-y\|_1} f(y), \tag{44}$$

so in particular

$$e^{z\Delta} \delta_y(x) = \left(\frac{1-e^z}{2}\right)^{\|x-y\|_1} \left(\frac{1+e^z}{2}\right)^{n-\|x-y\|_1}. \tag{45}$$

Finally, for $k \in \mathbb{N}$, we let $L_p^{\geq k}(\mathbb{F}_2^n; B)$ denote the subspace of functions f such that $\hat{f}(A) = 0$ whenever $|A| < k$. Observe that $L_p^{\geq}(\mathbb{F}_2^n; B)$ is invariant under the Rademacher projections and hence under $e^{z\Delta}$. K-convexity was introduced by Maurey and Pisier (*Studia math*, 1976):

Definition 7.1. Let B be a Banach space. The K-convexity constant of B is defined by $K(B) = \sup_n \|\mathbf{Rad}_1\|_{L_2(\mathbb{F}_2^n;B)\to L_2(\mathbb{F}_2^n;B)}$ and we say that B is K-convex if $K(B) < \infty$.

It is relatively straightforward to see that Hilbert spaces are K-convex. Indeed, let H be a Hilbert space and take $f \in L_2(\mathbb{F}_2^n; H)$. Write $f = \sum_{A\subset[n]} \hat{f}(A) W_A$ with $\hat{f}(A) \in H$. Then

$$\begin{aligned}
\|\mathbf{Rad}_1(f)\|_{L_2(\mathbb{F}_2^n;H)}^2 &= \frac{1}{2^n} \sum_{x\in\mathbb{F}_2^n} \left\langle \sum_{i=1}^n \hat{f}(\{i\}) W_{\{i\}}(x), \sum_{i=1}^n \hat{f}(\{i\}) W_{\{i\}}(x) \right\rangle \\
&= \sum_{i,j=1}^n \langle \hat{f}(\{i\}), \hat{f}(\{j\})\rangle \frac{1}{2^n} \sum_{x\in\mathbb{F}_2^n} W_{\{i\}}(x) W_{\{j\}}(x) \\
&= \sum_{i=1}^n \langle \hat{f}(\{i\}), \hat{f}(\{i\})\rangle = \sum_{i=1}^n \left\|\hat{f}(\{i\})\right\|_H^2,
\end{aligned}$$

using that the Walsh functions form a system of mutually independent Rademacher variables. It can be seen in the same way that $\|f\|_{L_2(\mathbb{F}_2^n;B)}^2 = \sum_{A\subset[n]} \left\|\hat{f}(\{A\})\right\|_H^2$, demonstrating that $\mathbf{Rad}_1$ is a contraction on $L_2(\mathbb{F}_2^n; B)$, for all $n \in \mathbb{N}$. This proves that H is K-convex with $K(H) \leq 1$. In fact all L_p spaces are K-convex and more generally we have:

Remark 7.2 (**Superreflexive Banach spaces are K-convex**)**.** It is a deep theorem due to Pisier [Pis82] that B is K-convex if and only if there exists $n_0 \in N$ and $\varepsilon_0 > 0$ such that for every injective, linear map $T : \ell_1^{n_0} \to B$ we have $\|T\|_{\ell_1^{n_0}\to B} \left\|T^{-1}\right\|_{T(\ell_1^{n_0})\to\ell_1^{n_0}} \geq 1 + \varepsilon_0$. Using this characterization of K-convexity we can see that superreflexivity implies K-convexity. Indeed, if B is not K-convex then for every $n \in \mathbb{N}$ and every $\varepsilon > 0$ we could find an injective, linear map $T : \ell_1^n \to B$ such that $\|T\|_{\ell_1^n\to B} \left\|T^{-1}\right\|_{T(\ell_1^n)\to\ell_1^n} < 1 + \varepsilon$. This implies that ℓ_1 is finitely representable in B (the proof of this claim is almost analogous to the argument in Example 1.30). Since ℓ_1 is not reflexive, this yields that B cannot be superreflexive.

The expansion properties of the graphs from Lemma 5.15 hinge on the following theorem:

Theorem 7.3. *Let $K, p \in (1, \infty)$ be arbitrary. Then there exist $\alpha = \alpha(K, p) \in (0, 1)$, $b = b(K, p) \in (2, \infty)$ and $c = c(K, p) \in (2, \infty)$ such that for every K-convex Banach space B with $K(B) \leq K$, every $k, n \in \mathbb{N}$ and every $t \in (0, \infty)$ we have the estimate*

$$\left\|e^{-t\Delta}\right\|_{L_p^{\geq k}(\mathbb{F}_2^n;B)\to L_p^{\geq k}(\mathbb{F}_2^n;B)} \leq ce^{-\alpha k \min\{t,t^b\}}.$$

The proof of the above theorem requires the following deep result about K-convex Banach spaces due to Pisier [Pis82], which we state without proof:

Theorem 7.4. *Let $K, p \in (1, \infty)$ be arbitrary. Then there exist $\phi = \phi(K, p) \in (0, \pi/4)$ and $M = M(K, p) \in (2, \infty)$ such that whenever B is a K-convex Banach space with $K(B) \leq K$ we have the estimate*

$$\left\|e^{-z\Delta}\right\|_{L_p(\mathbb{F}_2^n;B)\to L_p(\mathbb{F}_2^n;B)} \leq M, \tag{46}$$

for every $n \in \mathbb{N}$ and every $z \in \mathbb{C}$ satisfying $|\arg(z)| \leq \phi$.

We now introduce the notion of *harmonic measure*. This piece of theory draws on several more or less classical results from complex analysis, proofs of which can be found in Rudin's legendary text book on real and complex analysis [Rud87].

Definition 7.5. Let Ω be an open subset of the complex plane, which we may identify with $\mathbb{R}^2$. For a C^2 function $u : \Omega \to \mathbb{C}$ we define its *Laplacian*

$$\Delta u := \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} : \Omega \to \mathbb{C}.$$

We say that u is *harmonic* if $\Delta u = 0$ on all of Ω.

It follows directly from this definition (since the Laplacian of a real valued function is real valued) that a function is harmonic if and only if both its real and imaginary parts are harmonic. Hence it follows from the Cauchy-Riemann differential equations that any holomorphic function is harmonic. There is a local converse to this fact, namely:

Theorem 7.6. *[Rud87, Th. 11.9] Suppose Ω is simply connected and $u : \Omega \to \mathbb{R}$ is harmonic. Then there exists a holomorphic function $f : \Omega \to \mathbb{C}$ such that $u = \Re(f)$. In particular u has partial derivatives of all orders.*

The following theorem provides a solution to the so-called *Dirichlet boundary problem*: given a continuous function $f : T \to \mathbb{C}$, where $T \subset \mathbb{C}$ denotes the unit circle, find a continuous function $u : \overline{D} \to \mathbb{C}$ which is harmonic on D and such that $u|_T = f$ (here $D \subset \mathbb{C}$ denotes the open unit disc).

Theorem 7.7. *[Rud87, Sec. 11.5] Let $f : T \to \mathbb{C}$ be continuous. Then there exists a continuous function $Hf : \overline{D} \to \mathbb{C}$ which is harmonic on D and such that $(Hf)|_T = f$. This function is given by*

$$Hf(re^{i\theta}) := \begin{cases} f(e^{i\theta}) & \text{if } r = 1, \\ \frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta - t) f(e^{it})dt & \text{if } 0 \leq r < 1 \end{cases},$$

where

$$P_r(t) := \sum_{-\infty}^{\infty} r^{|n|}e^{int}, \qquad 0 \leq r < 1,\ t \in \mathbb{R}$$

is the Poisson kernel*. The integral in the definition of Hf is known as the* Poisson integral *of f. Conversely, if $u : \overline{D} \to \mathbb{C}$ is continuous and harmonic on D, then $H(u|_T) = u$.*

Remark 7.8. Note that, by definition of the Poisson kernel, $Hf(z)$ is independent of the choice of argument of $z \in \overline{D}$, i.e., $Hf(re^{i\theta}) = Hf(re^{i(\theta+2\pi)})$ so that Hf is well defined. Moreover, since $P_0(t) = 1$, for all $t \in \mathbb{R}$, we have $Hf(0) = \int_T f(z)\frac{dz}{2\pi}$.

We need one more important theorem from complex analysis before we introduce harmonic measure, namely the Riemann mapping theorem. Recall that a *conformal map* between two open regions $\Omega_1, \Omega_2 \subset \mathbb{C}$ is a holomorphic bijection $\psi : \Omega_1 \to \Omega_2$ (this automatically implies that ψ^{-1} is also holomorphic and that $\psi'(z) \neq 0$, for all $z \in \Omega_1$). If such a conformal map exists we say that Ω_1 and Ω_2 are conformally equivalent. The following theorem is known as *The Riemann mapping theorem*:

Theorem 7.9. *[Rud87, Th.s 14.8 & 14.19] If $\Omega \subsetneq \mathbb{C}$ is open and simply connected then Ω is conformally equivalent to D. Moreover, for every $z_0 \in \Omega$ there exists a conformal equivalence which carries z_0 to the origin. Finally, if Ω is convex and bounded then $\overline{\Omega}$ and $\overline{D}$ are homeomorphic and any conformal equivalence between Ω and D extends to a homeomorphism between $\overline{\Omega}$ and $\overline{D}$.*

Now suppose Ω is an open, bounded, convex, simply connected region of the complex plane, and let $\psi : \Omega \to D$ be a conformal equivalence which extends to a homeomorphism. Then $\partial\Omega$ is compact and for $f \in C(\partial\Omega)$ we can view $f \circ \psi^{-1}$ as an element of $C(T)$, so it has a solution to the Dirichlet boundary problem, $H(f \circ \psi^{-1})$. The composition of a harmonic function with a conformal map is clearly harmonic, so $H_\psi f := H(f \circ \psi^{-1}) \circ \psi : \Omega \to \mathbb{C}$ is harmonic and extends to a continuous map on $\overline{\Omega}$ which is equal to f on $\partial\Omega$. Let $z_0 = \psi^{-1}(0)$ and let $\mu_{z_0} := \frac{dz}{2\pi} \circ \psi^{-1}$, the image measure of $\frac{dz}{2\pi}$ (i.e., normalized Lebesgue measure on T) under ψ^{-1}, which is a probability measure. As a consequence of the above remark we have

$$\begin{aligned} H_\psi f(z_0) = H(f \circ \psi^{-1})(0) &= \frac{1}{2\pi} \int_T H(f \circ \psi^{-1})(z) dz \\ = \int_T f \circ \psi^{-1} \frac{dz}{2\pi} &= \int_{\partial\Omega} f \left(\frac{dz}{2\pi} \circ \psi^{-1} \right) = \int_{\partial\Omega} f d\mu_{z_0}. \end{aligned}$$

By the Riesz representation theorem for the dual of $C(K)$, where K is a compact space, μ_{z_0} is unique with this property. Hence, it does not depend on ψ, only on z_0, i.e., on the fact that ψ carries z_0 to the origin. In particular, for any Borel set $E \subset \delta\Omega$ we have $\mu_{z_0}(E) = \int_{\psi(E)} dz/2\pi$ where $\psi : \Omega \to D$ is any conformal map with $\psi(z_0) = 0$, which exists by the Riemann mapping theorem. We summarize this in the following definition:

Definition 7.10. Let Ω be an open, bounded, convex, simply connected region of the complex plane. For $z_0 \in \Omega$, the unique Borel probability measure μ_{z_0} on $\partial\Omega$ such that every continuous $u : \overline{\Omega} \to \mathbb{C}$ which is harmonic on Ω satisfies

$$u(z_0) = \int_{\partial\Omega} u d\mu_{z_0}.$$

is called the *harmonic measure* for z_0. It is given by the image measure of normalized Lebesgue measure on T under the inverse of any conformal equivalence between Ω and D which carries z_0 to the origin.

Now, let $K, p \in (1, \infty)$ and let $\phi \in (0, \pi/4)$ be as in Theorem 7.4. Pick $a > 0$ such that the vertical line segment connecting $a + i\pi$ with $a - i\pi$ is contained in the unbounded pizza slice section $\{z \in \mathbb{C} \mid |\arg(z)| \le \phi\} \subset \mathbb{C}$. Specifically, we can put $a = \pi/\tan\phi$, which is the smallest a satisfying this. Define V_ϕ to be the bounded pizza slice section:

$$V_\phi := \left\{ z \in \mathbb{C} \mid |z| \le 2\sqrt{a^2 + \pi^2} \text{ and } |\arg(z)| \le \phi \right\},$$

which contains the line segment connecting $a + i\pi$ with $a - i\pi$ and that connecting the origin with $2a$. This section is bounded, convex and simply connected so any interior point of V_ϕ admits a harmonic measure on its boundary, which is the disjoint union of the two sets

$$V_\phi^0 = \{x \pm ix \tan\phi \mid 0 \le x < 2a\} \quad \text{and} \quad V_\phi^1 = \left\{ (2\sqrt{a^2 + \pi^2}) e^{i\theta} \mid |\theta| \le \phi \right\}.$$

Lemma 7.11. *Let $\phi \in (0, \pi/4)$, $a > 0$ and V_ϕ be as above and put $r := 2\sqrt{a^2 + \pi^2}$. For $t \in (0, 2a) \subset \mathrm{Int}(V_\phi)$ put $\theta_t := \mu_t(V_\phi^1)$. Then*

$$\theta_t \geq \frac{1}{2}\left(\frac{t}{r}\right)^{\pi/2\phi}. \tag{47}$$

Moreover, for every $\varepsilon \in (0,1)$ there exists a bounded, continuous function $\Psi_\varepsilon^t : V_\phi \to \mathbb{C}$ which is holomorphic (and hence harmonic) on the interior of V_ϕ such that

- $\Psi_\varepsilon^t(t) = 1$,
- $|\Psi_\varepsilon^t(z)| = \varepsilon$ *for every* $z \in V_\phi^0$,
- $|\Psi_\varepsilon^t(z)| = \frac{1}{\varepsilon^{(1-\theta_t)/\theta_t}}$ *for every* $z \in V_\phi^1$.

The proof of the first part of Lemma 7.11 (the content of which is found in [MN14, Lemma 5.7]) writes up an explicit conformal equivalence between V and D which takes t to the origin and computes the normalized arc length of the image of V_ϕ^1 under this map. The second part (which corresponds to [MN14, Lemma 5.8]) is essentially attributed to Pisier. Though they are not at all trivial we shall omit the proofs.

Proof of Theorem 7.3. For $K, p \in (1, \infty)$ take $\phi \in (0, \pi/4)$ and $M \in (2, \infty)$ according to Theorem 7.4. Let $a = \pi/\tan(\phi)$ as above and let B be a K-convex Banach space with $K(B) \leq K$. Let $t \in (0, \infty)$ and suppose first that $t \geq 2a$. Now, for $m = 0, \cdots, n$ we have for any $f \in L_p(\mathbb{F}_2^n; B)$ and $x \in \mathbb{F}_2^n$ that the function

$$s \mapsto e^{ims} e^{-(a+is)\Delta} f(x)$$

from $[-\pi, \pi]$ to B is Bochner integrable and by Example 1.49 we get

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ims} e^{-(a+is)\Delta} f(x) ds &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ims} \sum_{k=0}^{n} e^{-(a+is)k} \mathbf{Rad}_k(f)(x) ds \\ &= \sum_{k=0}^{n} e^{-ak} \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{is(m-k)} ds \mathbf{Rad}_k(f)(x) = e^{-am}\mathbf{Rad}_m(f)(x).\end{aligned}$$

By a slight abuse of notation we will write this as

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ims} e^{-(a+is)\Delta} ds = e^{-am}\mathbf{Rad}_m.$$

By Theorem 7.4 we have $\left\|e^{-(a+is)\Delta}\right\|_{L_p(\mathbb{F}_2^n;B)\to L_p(\mathbb{F}_2^n;B)} \leq M$ for every $s \in [-\pi, \pi]$ and hence, in the above notation,

$$\|\mathbf{Rad}_m\|_{L_p(\mathbb{F}_2^n;B)\to L_p(\mathbb{F}_2^n;B)} \leq \frac{e^{ma}}{2\pi}\int_{-\pi}^{\pi} \left\|e^{-(a+is)\Delta}\right\|_{L_p(\mathbb{F}_2^n;B)\to L_p(\mathbb{F}_2^n;B)} ds \leq M e^{ma}.$$

This implies, for any $k \leq n$ and any $z \in \mathbb{C}$ with $\Re(z) \geq 2a$ (so in particular for t), that

$$\begin{aligned}\left\|e^{-z\Delta}\right\|_{L_p^{\geq k}(\mathbb{F}_2^n;B)\to L_p^{\geq k}(\mathbb{F}_2^n;B)} &= \left\|\sum_{m=k}^{n} e^{-zm}\mathbf{Rad}_m\right\|_{L_p^{\geq k}(\mathbb{F}_2^n;B)\to L_p^{\geq k}(\mathbb{F}_2^n;B)} \\ &\leq \sum_{m=k}^{n} e^{-\Re(z)m} M e^{ma} \leq M\sum_{m=k}^{n} e^{-m\Re(z)/2} \leq M e^{-k\Re(z)/2}\sum_{m=0}^{\infty}(e^{-\Re(z)/2})^m \\ &= \frac{Me^{-k\Re(z)/2}}{1-e^{-\Re(z)/2}} \leq \frac{Me^{-k\Re(z)/2}}{1-e^{-a}} \leq \frac{Me^{-ka}}{1-e^{-a}},\end{aligned} \tag{48}$$

using that $a \leq \Re(z)/2$ by assumption. This demonstrates that for $t \geq 2a$ the conclusion of the Theorem holds with $\alpha(K,p) = 1/2$, $c(K,p) = M/(1-e^{-a})$ and any $b(K,p) > 2$.

Now suppose that $t < 2a$, and let $\varepsilon \in (0,1)$. Take θ_t and Ψ_ε^t as in Lemma 7.11 and observe that there are probability measures, μ_t^0 and μ_t^1, on V_ϕ^0 and V_ϕ^1, respectively, such that $\mu_t(E) = (1-\theta_t)\mu_t^0(E\cap V_\phi^0) + \theta_t\mu_t^1(E\cap V_\phi^1)$, for all Borel sets $E \subset V_\phi$. Indeed, we just define $\mu_t^0(E) = \mu_t(E)/\mu_t(V_0) = \mu_t(E)/(1-\theta_t)$ and similarly for μ_t^1. Now, since e^{-mz} is holomorphic for every $m = 0, \cdots, n$ we obtain

$$\begin{aligned}e^{-t\Delta} = \Psi_\varepsilon^t(t)e^{-t\Delta} &= \int_{\partial V_\phi}\Psi_\varepsilon^t(z)e^{-z\Delta}d\mu_t(z) \\ &= (1-\theta_t)\int_{V_\phi^0}\Psi_\varepsilon^t(z)e^{-z\Delta}\mu_0(z) + \theta_t\int_{V_\phi^1}\Psi_\varepsilon^t(z)e^{-z\Delta}d\mu_1(z).\end{aligned}$$

Hence, since any $z \in V_\phi^1$ has $\Re(z) \geq 2a$ and any $z \in V_\phi^0$ has $\arg(z) = \phi$, we get by (48), Lemma 7.11 and Theorem 7.4 that

$$\begin{aligned}\left\|e^{-t\Delta}\right\|_{L_p^{\geq k}(\mathbb{F}_2^n;B)\to L_p^{\geq k}(\mathbb{F}_2^n;B)} &\leq (1-\theta_t)\varepsilon M + \frac{\theta_t}{\varepsilon^{(1-\theta_t)/\theta_t}}\cdot\frac{Me^{-ka}}{1-e^{-a}} \\ &\leq \varepsilon M + \frac{1}{\varepsilon^{2(r/t)^{\pi/2\phi}-1}}\cdot\frac{Me^{-ka}}{1-e^{-a}},\end{aligned}$$

using that $1/\theta_t \leq 2(r/t)^{\pi/2\phi}$. Since this estimate is valid for every $\varepsilon \in (0,1)$ we can put $\varepsilon = \exp(-\frac{1}{2}(t/r)^{\pi/2\phi}ka)$ to obtain

$$\begin{aligned}\left\|e^{-t\Delta}\right\|_{L_p^{\geq k}(\mathbb{F}_2^n;B)\to L_p^{\geq k}(\mathbb{F}_2^n;B)} &\leq M\exp\left(-\frac{1}{2}\left(\frac{t}{r}\right)^{\pi/2\phi}ka\right) + \frac{M}{1-e^{-a}}\cdot\exp\left(-\frac{1}{2}\left(\frac{t}{r}\right)^{\pi/2\phi}ka\right) \\ &\leq \left(M + \frac{M}{1-e^{-a}}\right)\exp\left(-\frac{a}{2r^{\pi/2\phi}}kt^{\pi/2\phi}\right).\end{aligned}$$

So taking $c(K,p) = M + M/(1-e^{-a})$, $\alpha(K,p) = \min\left\{1/2, a/(2r^{\pi/2\phi})\right\}$ and $b(K,p) = \pi/2\phi$, the conclusion of the theorem is valid for any $t > 0$. $\square$

7.2 Construction of the base graphs

Let $t \in (0,\infty)$ and $n \in \mathbb{N}$ and define $\tau_t = \frac{1-e^{-t}}{2}$ and $\sigma_t^n = \tau_t^{4\tau_t n}(1-\tau_t)^{(1-4\tau_t)n}$. Furthermore let $e_t^n : \{0,\cdots,n\} \to \mathbb{N}_0$ denote the function given by

$$e_t^n(k) = \left\lfloor\frac{\tau_t^k(1-\tau_t)^{n-k}}{\sigma_t^n}\right\rfloor.$$

The following lemma constitutes the afforementioned combinatorial parts of the proof of Lemma 5.15. Its proof can be found in [MN14, Lemma 7.2].

Lemma 7.12. *Fix $t \in (0, 1/4)$, $p \geq 1$ and an integer $n \geq 2^{13}$ such that*

$$\tau_t \geq \sqrt{\frac{p \log(18n)}{18n}}.$$

Let $G_t^n = (V_t^n, E_t^n)$ be the graph whose vertex set is $\mathbb{F}_2^n$ and where every $x, y \in \mathbb{F}_2^n$ are joined by $e_t^n(\|x-y\|_1)$ edges. Then G_t^n is d_t^n-regular, where $\frac{1}{3\sigma_t^n} \leq d_t^n \leq \frac{1}{\sigma_t^n}$. and for every metric space (X, d_X) and every $f, g : \mathbb{F}_2^n \to X$ we have

$$\begin{aligned} \frac{1}{3|E_t^n|} \sum_{(x,y)\in E_t^n} d_X(f(x), g(y))^p &\leq \frac{1}{2^n} \sum_{x,y\in\mathbb{F}_2^n} e^{-t\Delta}\delta_x(y) d_X(f(x), g(y))^p \\ &\leq \frac{3}{|E_t^n|} \sum_{(x,y)\in E_t^n} d_X(f(x), g(y))^p. \end{aligned} \tag{49}$$

In what follows, for each $n \in \mathbb{N}$, let $V_n \subset \mathbb{F}_2^n$ be a 'good linear code', i.e., a linear subspace such that $D_n := \dim(V_n) \geq n/10$ and the *Hamming weight* $k_n := \min_{x \in V_n \setminus \{0\}} \|x\|_1 \geq n/10$. There are many sources for the existence of linear codes with these properties, we refer to [Ost13, Sec. 4]. As usual, we let $V_n^\perp = \{x \in \mathbb{F}_2^n \mid \sum_{i=1}^n x_i y_i = 0 \ \forall y \in V_n\}$. Let $\pi : \mathbb{F}_2^n \to \mathbb{F}_2^n / V_n^\perp$ be the quotient map and for $f \in L_p(\mathbb{F}_2^n / V_n^\perp; B)$ define $\pi f \in L_p(\mathbb{F}_2^n; B)$ by $\pi f(x) = f(\pi(x))$. Then πf is constant on cosets of $V_n^\perp$, i.e., $\pi f(x+y) = \pi f(x)$ for every $x \in \mathbb{F}_2^n$ and $y \in V_n^\perp$, and clearly any function which satisfies this arises in this way. Furthermore $\|f\|_{L_p(\mathbb{F}_2^n/V_n^\perp;B)} = \|\pi f\|_{L_p(\mathbb{F}_2^n;B)}$ and $f \in L_p(\mathbb{F}_2^n/V_n^\perp; B)$ is mean-zero if and only if $\pi f \in L_p(\mathbb{F}_2^n; B)$ is mean-zero.

Lemma 7.13. *Let $L_p(\mathbb{F}_2^n/V_n^\perp; B)_0 \subset L_p(\mathbb{F}_2^n/V_n^\perp; B)$ denote the subspace of mean-zero functions. Then $\pi f \in L_p^{\geq k_n}(\mathbb{F}_2^n; B)$, for every $f \in L_p(\mathbb{F}_2^n/V_n^\perp; B)_0$.*

Proof. Let $A \subset [n]$ satisfy $|A| < k_n$, and assume first that A is non-empty. Then, if we let $1_A \in \mathbb{F}_2^n$ denote the vector whose i'th coordinate is 1 if and only if $i \in A$, we have by definition of k_n that $1_A \notin V_n = (V_n^\perp)^\perp$ (this equality is seen analogously to the result for real or complex, finite dimensional vector spaces). Hence there is a vector $y \in V_n^\perp$ such that $\sum_{i=1}^n y_i (1_A)_i = \sum_{i\in A} y_i = 1 \mod 2$. Hence $W_A(y) = -1$ and then, since πf is constant on cosets of $V_n^\perp$ we get

$$\begin{aligned} (\pi f)^\wedge(A) &= \frac{1}{2^n} \sum_{x\in\mathbb{F}_2^n} \pi f(x) W_A(x) = \frac{1}{2^n} \sum_{x\in\mathbb{F}_2^n} \pi f(x+y) W_A(x) \\ &= \frac{1}{2^n} \sum_{z\in\mathbb{F}_2^n} \pi f(z) W_A(z+y) = \frac{1}{2^n} \sum_{z\in F_2^n} \pi f(z) W_A(z) W_A(y) = -(\pi f)^\wedge(A), \end{aligned}$$

which implies $(\pi f)^\wedge(A) = 0$. If $A = \emptyset$ then $(\pi f)^\wedge(A) = 0$ by the mean-zero assumption. □

Using formula (44) for $e^{z\Delta}$ it is easily seen that if f is constant on cosets then so is $e^{z\Delta} f$, and hence for any $f \in L_p(\mathbb{F}_2^n/V_n^\perp; B)$ we can view $(e^{z\Delta}\pi) f$ as an element of $L_p(\mathbb{F}_2^n/V_n^\perp; B)$ as well with $\left\|(e^{z\Delta}\pi) f\right\|_{L_p(\mathbb{F}_2^n;B)} = \left\|(e^{z\Delta}\pi) f\right\|_{L_p(\mathbb{F}_2^n/V_n^\perp;B)}$.

Lemma 7.14. *Let $K, p > 1$. Then there exist $n(K,p) \in \mathbb{N}$, $\delta(K,p) \in (0,1)$, and a sequence of connected, regular graphs $(H_n(K,p))_{n \geq n(K,p)}$ satisfying the following:*
For each $n \geq n(K,p)$, $H_n(K,p)$ has $m_n := |\mathbb{F}_2^n / V_n^\perp| = 2^{D_n}$ vertices and degree of regularity $d_n(K,p)$ satisfying $d_n(K,p) \leq e^{(\log m_n)^{1-\delta(K,p)}}$. Moreover, for each K-convex Banach space, B, with $K(B) \leq K$ we have the uniform bound

$$\sup_{n \geq n(K,p)} \gamma_+ \left(H_n(K,p), \|\cdot\|_B^p \right) \leq 9^{p+1}.$$

Proof. Let α, b and c denote the constants (depending on K and p) according to Theorem 7.3. Let $n \in \mathbb{N}$ and put $t_n = \left(\frac{\log(2c)}{k_n \alpha} \right)^{1/b}$. Then, since $k_n \to \infty$ as $n \to \infty$, we have $t_n \to 0$ as $n \to \infty$. Hence there is $n(K,p) \in \mathbb{N}$ such that, for every $n \geq n(K,p)$, the assumptions of Lemma 7.12 are satisfied for t_n, p and n. Furthermore, there is $\delta = \delta(K,p) \in (0,1)$ such that

$$\frac{1}{\tau_{t_n}^{8n\tau_{t_n}}} \leq e^{(\log m_n)^{1-\delta}},$$

whenever $n \geq n(K,p)$ (recall $m_n = 2^{D_n}$ and $k_n, D_n \geq n/10$). Let $n \geq n(K,p)$ and put $t = t_n$. Then the graph G_t^n from Lemma 7.12 is d_t^n-regular on 2^n vertices and

$$d_t^n \leq \frac{1}{\sigma_t^n} \leq \frac{1}{\tau_t^{8n\tau_t}} \leq e^{(\log m_n)^{1-\delta}},$$

where the second estimate follows from the definition of σ_t^n in the beginning of this subsection. Now define $H_n = H_n(K,p)$ to be the graph whose vertex set it $\mathbb{F}_2^n / V_n^\perp$ and where the number of edges between two vertices $S, T \in \mathbb{F}_2^n / V_n^\perp$ is the number of edges in G_t^n with one end point in S and one in T divided by $|V_n^\perp|$. So if $E_n = E_n(K,p)$ denotes the edge set of H_n, we have for $S = x + V_n^\perp$ and $T = y + V_n^\perp$, $x, y \in \mathbb{F}_2^n$, that

$$\begin{aligned} E_n(S,T) &= \frac{1}{|V_n^\perp|} \sum_{u,v \in V_n^\perp} e_t^n(\|x + u - (y + v)\|_1) \\ &= \frac{1}{|V_n^\perp|} \sum_{u,v \in V_n^\perp} e_t^n(\|x - y + (u - v)\|_1) \\ &= \sum_{u \in V_n^\perp} e_t^n(\|x - y + u\|_1) = \sum_{u \in T} e_t^n(\|x - u\|_1). \end{aligned}$$

Hence, the degree of $S = x + V_n^\perp$ is $\sum_{T \in \mathbb{F}_2^n / V_n^\perp} \sum_{u \in T} e_t^n(\|x - u\|_1) = d_t^n$, so H_n is d_t^n-regular. Furthermore, by definition of D_n, the number of vertices of H_n is 2^{D_n}. Let B be a K-convex Banach space with $K(B) \leq K$. Then, for $f \in L_p(\mathbb{F}_2^n / V_n^\perp; B)_0 \setminus \{0\}$, we have by Lemma 7.13 that $\pi f \in L_p^{\geq k_n}(\mathbb{F}_2^n; B)$ and hence by Theorem 7.3 we have

$$\frac{\left\| (e^{-t\Delta} \pi) f \right\|_{L_p(\mathbb{F}_2^n / V_n^\perp; B)}}{\|f\|_{L_p(\mathbb{F}_2^n / V_n^\perp; B)}} \leq c e^{-\alpha k_n \min\{t, t^b\}} = \frac{1}{2}, \tag{50}$$

where the last equality follows from the definition of $t = t_n$ and the fact that $t < 1/4$ and $b > 2$, whence $t^b \le t$. Now let $Q = (q_{S,T})_{S,T\in\mathbb{F}_2^n/V_n^\perp}$ be the $m_n \times m_n$ matrix given by

$$q_{S,T} = \frac{1}{|V_n^\perp|}\sum_{u\in S, v\in T} \tau_t^{\|u-v\|_1}(1-\tau_t)^{n-\|u-v\|_1}.$$

This matrix is clearly symmetric and since, for fixed $S \in \mathbb{F}_2^n/V_n^\perp$, we have

$$\begin{aligned}
\sum_{T\in\mathbb{F}_2^n/V_n^\perp} q_{S,T} &= \frac{1}{|V_n^\perp|}\sum_{u\in S}\sum_{T\in\mathbb{F}_2^n/V_n^\perp}\sum_{v\in T}\tau_t^{\|u-v\|_1}(1-\tau_t)^{n-\|u-v\|_1} \\
&= \sum_{x\in\mathbb{F}_2^n}\tau_t^{\|x\|_1}(1-\tau_t)^{n-\|x\|_1} \\
&= \sum_{k=0}^{n}\sum_{\substack{x\in\mathbb{F}_2^n \\ \|x\|_1=k}}\tau_t^k(1-\tau_t)^{n-k} \\
&= \sum_{k=0}^{n}\binom{n}{k}\tau_t^k(1-\tau_t)^{n-k} = 1,
\end{aligned}$$

we see that Q is also stochastic. Furthermore, for every $f \in L_p(\mathbb{F}_2^n/V_n^\perp; B)$, we have

$$\begin{aligned}
(Q\otimes I_B^{m_n})f(S) &= \sum_{T\in F_2^n/V_n^\perp} q_{S,T}f(T) \\
&= \sum_{T\in F_2^n/V_n^\perp}\frac{1}{|V_n^\perp|}\sum_{u\in S,v\in T}\tau_t^{\|u-v\|_1}(1-\tau_t)^{n-\|u-v\|_1}f(T) \\
&= \frac{1}{|V_n^\perp|}\sum_{u\in S}\sum_{T\in F_2^n/V_n^\perp}\sum_{v\in T}\tau_t^{\|u-v\|_1}(1-\tau_t)^{n-\|u-v\|_1}\pi f(v) \\
&= \frac{1}{|V_n^\perp|}\sum_{u\in S}(e^{-t\Delta}\pi)f(u) = (e^{-t\Delta}\pi)f(S),
\end{aligned}$$

where by $(e^{-t\Delta}\pi)f(S)$ we mean the (constant) value $(e^{-t\Delta}\pi)f$ takes on the coset S. This demonstrates that $Q\otimes I_B^{m_n} = e^{-t\Delta}\pi$, and hence by (50) we have $\lambda_B^{(p)}(Q) \ge 1/2$. Moreover, Lemma 4.4 yields

$$\gamma_+(Q, \|\cdot\|^p) \le \left(1 + \frac{4}{\lambda_B^{(p)}(Q)}\right)^p \le 9^p.$$

Now, let $f, g : \mathbb{F}_2^n/V_n^\perp \to B$ be arbitrary. Then by definition of the absolute spectral gap, formula (45) and Lemma 7.12 we have

$$\begin{aligned}
&\frac{1}{|\mathbb{F}_2^n/V_n^\perp|^2}\sum_{S,T\in\mathbb{F}_2^n/V_n^\perp}\|f(S)-g(T)\|_B^p\\
&\le \frac{9^p}{|\mathbb{F}_2^n/V_n^\perp|}\sum_{S,T\in\mathbb{F}_2/V_n^\perp} q_{S,T}\,\|f(S)-g(T)\|_B^p\\
&= \frac{9^p}{2^{D_n}}\sum_{S,T\in\mathbb{F}_2^n/V_n^\perp}\frac{1}{|V_n^\perp|}\sum_{v\in T,u\in S}\tau_t^{\|u-v\|_1}(1-\tau_t)^{n-\|u-v\|_1}\,\|f(S)-g(T)\|_B^p\\
&= \frac{9^p}{2^n}\sum_{x,y\in F_2^n} e^{-t\Delta}\delta_x(y)\,\|\pi f(x)-\pi g(y)\|_B^p\\
&\le \frac{3\cdot 9^p}{|E_t^n|}\sum_{(x,y)\in E_t^n}\|\pi f(x)-\pi g(y)\|_B^p\\
&= \frac{3\cdot 9^p\cdot 2^{D_n}}{2^n d_t^n}\sum_{S,T\in\mathbb{F}_2^n/V_n^\perp}\frac{1}{|V_n^\perp|}\sum_{u\in S,v\in T}E_t^n(u,v)\,\|f(S)-g(T)\|_B^p\\
&= \frac{3\cdot 9^p}{|\mathbb{F}_2^n/V_n^\perp|d_t^n}\sum_{(S,T)\in E(H_n)}\|f(S)-g(T)\|_B^p\,,
\end{aligned}$$

from which it follows that $\gamma_+(H_n,\|\cdot\|_B^p)\le 3\cdot 9^p\le 9^{p+1}$ as desired. □

As a consequence of Lemma 7.14 we obtain the following generalization of Lemma 5.15:

Corollary 7.15. *For every $\delta\in(0,1)$ and $p>1$ there exist $n_0^p(\delta)\in\mathbb{N}$ and a sequence of graphs $(H_n^p(\delta))_{n\ge n_0^p(\delta)}$ such that $H_n^p(\delta)$ is $d_n^p(\delta)$-regular on m_n vetices with $d_n^p(\delta)\le e^{(\log m_n)^{1-\delta}}$ and $2^{n/10}\le m_n\le 2^n$. If B is a Banach space with $K(B)<\infty$, we have $\gamma_+(H_n^p(\delta),\|\cdot\|_B^p)<\infty$, for all $n\ge n_0^p(\delta)$, and there exists $\delta_0^p(B)\in(0,1)$ such that $\gamma_+(H_n^p(\delta),\|\cdot\|_B^p)\le 9^{p+1}$, for all $\delta\in(0,\delta_0^p(B)]$ and $n\ge n_0^p(\delta)$.*

Proof. First observe that by the way $\delta(K,p)$ is defined in Lemma 7.14, for any fixed $p>1$ we may assume that it decreases continuously to 0 with K. Let $\delta\in(0,1)$ and $p>1$ be given. First assume that $\delta>\delta(2,p)$. Then it follows from Proposition 4.2 that any connected regular graph which is not bipartite has finite absolute spectral gap with respect to $\|\cdot\|_B^p$. Hence we can let $n_0^p(\delta)$ be the smallest integer such that $e^{(\log m_n)^{1-\delta}}\ge 3$ and, for each $n\ge n_0^p(\delta)$, let $H_n^p(\delta)$ be the m_n-cycle (with m_n as defined in Lemma 7.14) with self loops. These graphs satisfy the bounds and finiteness assumptions per default. In case $\delta\le\delta(2,p)$ put $K_\delta^p=\sup\{K\ge 2\mid \delta(K,p)\ge\delta\}$ which is finite since $\delta(2,p)\ge\delta$ by assumption and because of the monotonicity assumptions above. Also, as $\delta(K,p)>0$, for all $K,p>1$, it is clear that $\lim_{\delta\to 0^+}K_\delta^p=\infty$. Now set $n_0^p(\delta)=n(K_\delta^p,p)$ and, for each $n\ge n_0(\delta)$, put $H_n^p(\delta)=H_n(K_\delta^p,p)$. This defines $H_n^p(\delta)$ for every $\delta\in(0,1)$ and $p>1$, satisfying the degree and vertex set bounds and finiteness of absolute spectral gaps. For a K-convex Banach space B we put $\delta_0^p(B)=\inf\left\{\delta\in(0,\delta(2,p)]\mid K_\delta^p<K(B)\right\}$ where we interpret this infimum as being equal to $\delta(2,p)$ if the set is empty. Note that $\delta_0^p(B)>0$ since $K_\delta^p\to\infty$ as $\delta\to 0^+$. Then for any $0<\delta\le\delta_0^p(B)$ we have $K_\delta^p\ge K(B)$, and thus for such δ and $n\ge n_0^p(\delta)$ we have $H_n^p(\delta)=H_n(K_\delta^p,p)$. Hence $\gamma_+(H_n^p(\delta),\|\cdot\|^p)\le 9^{p+1}$ by Lemma 7.14, concluding the proof. □

Final remarks

We saw in Section 3 that if a group has property (T) then its Schreier coset graphs with respect to a finite Kazhdan (and hence generating) set with Kazhdan constant k have lower bounds on their expansion properties in terms of k. More recently, spectral conditions of graphs associated to a group Γ have been proven to imply property (T) for Γ. Namely, let Γ be a group with finite, generating set S. We then define the *link graph* of S, denoted by $L(S)$, to be the graph whose vertex set is S, and where $(s,t) \in S \times S$ is an edge if and only if $s^{-1}t \in S$. The following theorem, due to [Żuk03], gives a spectral criterion for property (T):

Theorem. *Let Γ be a group with finite, symmetric generating set S not containing the identity. If $\lambda(L(S)) > 1/2$ then Γ has property (T) with Kazhdan pair $(S, \sqrt{2(2-\lambda(L(S))^{-1})})$.*

The estimate $\lambda(L(S)) > 1/2$ gives an upper bound on the constant in the Poincaré inequality for functions on the link graph with values in Hilbert space. This observation was used in [LS17] to prove spectral criteria for Banach space versions of property (T). It can be shown that a group Γ has property (T) if and only if every affine isometric action of Γ on a Hilbert space has a fixed point. This characterization, due to Delorme and Guichardet, is why property (T) is also referred to as a fixed point property for group actions, and is used to generalize property (T) to Banach spaces: if B is a Banach space, then Γ has property (F_B) if every affine isometric action of Γ on B has a fixed point. This definition is due to Bader, Furman, Gelander and Monod, who introduced it as well as other generalizations of property (T), and systematically studied them (*Acta matematica*, 2007). The aforementioned work by [LS17] introduces Poincaré inequalities as criteria for property (F_B):

Theorem. *Let S be a finite, symmetric generating set for Γ, not containing the identity and let B be a Banach space. Suppose there exists $p \in (1,\infty)$ and a constant $C \in (0,1)$ such that*

$$\inf_{x \in B} \|f - x\|_{\ell_p(L(S);B)} \leq C \, \|\nabla f\|_{L_p(E(L(S));B)},$$

for every $f : S \to B$, where $\nabla f : E(L(S)) \to B$ is given by $\nabla f(s,t) = f(s) - f(t)$. Then Γ has property (F_B).

They proceed to give spectral criteria, in terms of the operator norm of the adjacency matrix associated to the link graph, for obtaining Poincaré inequalities as in the above theorem, and hence for a group to have property (F_B). These criteria are then applied to certain random groups, to prove that, with high probability, such groups have property (F_B), for all uniformly curved Banach spaces B. Having studied constructions of expander sequences in this thesis, I would be very interested in moving on to study such spectral criteria for fixed point properties and further applications to random groups.

References

[AM85] N. Alon and V. D. Milman. "λ_1, isoperimetric inequalities for graphs, and superconcentrators". In: *Journal of Combinatorial Theory* 38(1) (1985), pp. 73–88.

[BCL94] K. Ball, E. A. Carlen, and E. H. Lieb. "Sharp uniform convexity and smoothness inequalities for trace norms". In: *Inventiones mathematica* 115 (1994), pp. 463–482.

[Bea82] Bernard Beauzamy. *Introduction to Banach Spaces and their Geometry*. Vol. 86. Notas de Matemática. North-Holland Publishing Company, 1982.

[BL00] Yoav Benyamini and Joram Lindenstrauss. *Geometric Nonlinear Functional Analysis - Volume 1*. Vol. 48. Colloquium publications. American Mathematical Society, 2000.

[BL76] Jöran Bergh and Jörgen Löfström. *Interpolation Spaces*. Vol. 223. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1976.

[BO07] Nathaniel P. Brown and Narutaka Ozawa. *C*-algebras and Finite-Dimensional Approximations*. American Mathematical Society, 2007.

[Cla36] James A. Clarkson. "Uniformly convex spaces". In: *Trans. Amer. Math* 40 (1936), pp. 396–414.

[Enf72] Per Enflo. "Banach spaces which can be given an equivalent uniformly convex norm". In: *Israel Journal of Mathematics* 13 (1972), pp. 281–288.

[Fol99] Gerald B. Folland. *Real Analysis, Modern Techniques and Their Applications*. Second Edition. Pure and Applied Mathematics. John Wiley & Sons, Inc., 1999.

[FRR95] Steven C Ferry, Andrew Ranicki, and Jonathan M. Rosenberg, eds. *Novikov conjectures, index theorems and rigidity. Vol. 1*. Vol. 226. London Mathematical Society lecture note series. Cambridge University Press, 1995.

[Gro77] Jonathan L. Gross. "Every Connected Regular Graph of Even Degree is a Schreier Coset Graph". In: *Journal of Combinatorial Theory (B)* 22 (1977), pp. 227–232.

[Hyt+16] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis. *Analysis in Banach Spaces - Volume I: Martingales and Littlewood-Paley Theory*. Vol. 63. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics. Springer Verlag, 2016.

[KY06] Gennadi Kasparov and Guoliang Yu. "The coarse geometric Novikov conjecture and uniform convexity". In: *Advances in Mathematics* 206 (1) (2006), pp. 1–56.

[Lin63] Joram Lindenstrauss. "On the Modulus of smoothness and Divergent Series in Banach Spaces". In: *Michigan Math. J.* 10 (1963), pp. 241–252.

[LS17] Tim de Laat and Mikael de la Salle. "Banach space actions and L^2-spectral gap". In: *arXiv e-prints*, arXiv:1705.03296 (2017).

[Mat97] Jiří Matoušek. "On embedding Expanders into ℓ_p spaces". In: *Israel Journal of Mathematics* 102 (1997), pp. 189–197.

[Maz29] S. Mazur. "Une remarque sur l'homéomorphie des champs fonctionnels". In: *Studia Mathematica* 1.1 (1929), pp. 83–85.

[Mei08] John Meier. *Groups, Graphs and Trees - An Introduction to the Geometry of Infinite Groups*. Vol. 73. London Mathematical Society Student Texts. Cambridge Univeristy Press, 2008.

[Mim18] Masato Mimura. “Sphere equivalence, Banach expanders, and extrapolation”. In: *Int. Math. Res. Notices* 2015 (2018), pp. 4392–4420. arXiv: `1310.4737 [math.GR]`.

[MN14] Manor Mendel and Assaf Naor. “Nonlinear Spectral Calculus and Super-Expanders”. In: *Publ.math.IHES* 119 (2014), pp. 1–95. DOI: `https://doi.org/10.1007/s10240-013-0053-2`.

[Nao12] Assaf Naor. “An introduction to the Ribe program”. In: *Jpn. J. Math.* 7.2 (2012), pp. 167–233.

[NY12] Piotr W. Nowak and Guoliang Yu. *Large Scale Geometry*. EMS Textbooks in Mathematics. European Mathematical Society, 2012.

[Osa14] Damian Osajda. “Small cancellation labellings of some infinite graphs and applications”. In: *arXiv e-prints*, arXiv:1406.5015 (2014).

[Osa18] Damian Osajda. “Residually finite non-exact groups”. In: *Geom. Funct. Anal.* 28 (2018), pp. 509–517. DOI: `https://doi.org/10.1007/s00039-018-0432-4`.

[Ost13] Mikhail I. Ostrovskii. *Metric Embeddings - Bilipschitz and Coarse Embeddings into Banach Spaces*. Vol. 49. Studies in Mathematics. De Gruyter, 2013.

[Pin73] Mark S. Pinsker. “On the complexity of a concentrator”. In: *7th International Teletraffic Conference*. 1973.

[Pis10] Gilles Pisier. “Complex Interpolation between Hilbert, Banach and Operator Spaces”. In: *Memoirs Amer. Math* 208.978 (2010). arXiv: `0802.0476 [math.FA]`.

[Pis75] Gilles Pisier. “Martingales with Values in Uniformly Convex Spaces”. In: *Israel J. Math.* 20 (1975), pp. 326–250.

[Pis82] Gilles Pisier. “Holomorphic semigroups and the geometry of Banach spaces”. In: *Ann. of Math.* 2nd ser. 115.2 (1982), pp. 375–392.

[Rud87] Walter Rudin. *Real and Complex Analysis*. 3rd edition. McGraw-Hill International editions. McGraw-Hill Book Company, 1987.

[RVW02] Omar Reingold, Salil Vadhan, and Avi Wigderson. “Entropy waves, the zig-zag graph product, and new constant-degree expanders”. In: *An. of Math.* Second Series 155.1 (2002), pp. 157–187.

[Sha99] Yehuda Shalom. “Bounded Generation and Kazhdan’s Property (T)”. In: *Publ. math. IHES* 90 (1999), pp. 145–168.

[Tit26] Edward Charles Titchmarsh. “Reciprocal formulae involving series and integrals”. In: *Math. Z.* 25 (1926), pp. 321–381.

[Zhu93] Kehe Zhu. *An Introduction to Operator Algebras*. Vol. 9. Studies in Advanced Mathematics. CRC press, 1993.

[Żuk03] Andrzej Żuk. “Property (T) and Kazhdan constants for discrete groups”. In: *Geometric & Functional Analysis GAFA*. 3rd ser. 13 (2003), pp. 643–670.