Delft University of Technology
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft Institute of Applied Mathematics

The Blum-Hanson Property

A thesis submitted to the
Delft Institute of Applied Mathematics
in partial fulfillment of the requirements

for the degree

MASTER OF SCIENCE
in
APPLIED MATHEMATICS

by

MARTIJN VAN OOSTERHOUT
Delft, the Netherlands
September 2009

MSc THESIS APPLIED MATHEMATICS

"The Blum-Hanson Property"

MARTIJN VAN OOSTERHOUT

Delft University of Technology

Daily supervisor

Dr. M.H.A. Haase

Responsible professor

Prof.dr. B. de Pagter

Other thesis committee members

Dr. C. Kraaikamp

Dr.ir. R.J. Fokkink

September 2009

Delft, the Netherlands

Contents

Preface

First, I would like to apologise to any non-mathematicians reading this. No matter how clearly I write, anyone without a basic introduction to functional analysis is not going to be able to make much sense of what I write. This is also going to include most bachelor students. Trust me when I tell you this is an interesting topic, but there is no way I'm going to be able to explain all the necessary theory and still be able to get to the point.

Ergodic theory is a relatively well established area of mathematics, but the application of general functional analysis tools to it is relatively recent. As such, some of the basic definitions presented here look somewhat different to those found in most books on ergodic theory. The Blum-Hanson property is studied here in a context that is probably far beyond what original authors expected.

After a brief introduction to Ergodic theory, Chapter 2 covers all the background and definitions needed to understand the later chapters. Chapter 3 covers the major positive results to date regarding the Blum-Hanson property. Chapter 4 introduces a new class of operators that are interesting for the results in the final chapter, the Generalised Shift Operators. Results for this new class of operators are then used to analyse a number of existing examples in Chapter 5. Here existing results are extended to cover more spaces and a new counter-example is introduced based on a suggestion by László Zsidó.

For completeness some referenced results are included in Appendix A. Additionally, for students who have only completed a basic course in functional analysis, some extra concepts are defined in Appendix B.

Chapter 1

An introduction to Ergodic theory

Ergodic theory is a branch of mathematics whose name gives no clue as to what it may be about. The exact origin of the word 'ergodic' is not really known, though the common version is that L. Boltzmann in 1885 used the word 'ergode' as a special kind of 'monode'. There are however other versions, as discussed by [Mat88].

Simply put, Ergodic theory is the study of dynamical systems and how they evolve over time. The canonical example is that of a set of gas particles contained inside a box. Here we know from experience that eventually the gas will spread throughout the box such that the energy and the particles are equally distributed. This is the second law of thermodynamics, but that doesn't really tell us why. We would like to be able to examine a dynamical system to determine what it will do in the long run. Hence the interest in formalising the problem mathematically.

The process of formalising is as follows: we start with the state space. In the example above it would be a set of vectors indicating both the position and the velocity of each gas particle. So if we have d gas particles, we have a state space of $\mathbb{R}^{6d}$ vectors (assuming 3D space). As mathematicians don't like being specific when it isn't necessary, the state space is usually indicated as Ω. The next step is to describe how the state evolves. Given the current state then using Newton's laws of mechanics and other physical laws we can define a function φ which takes the current state and returns the next one. Hence we have

$$\varphi : \Omega \to \Omega.$$

Now, assume we have a starting state ω_0 then the next state is $\varphi(\omega_0)$ followed by $\varphi^2(\omega_0)$. By repeatedly applying this function we produce a sequence of states

$$\mathrm{orb}_+(\omega_0) := \{\varphi^n(\omega_0) : n \in \mathbb{N}_0\}.$$

This sequence is called the *orbit* of ω_0. The goal of Ergodic Theory is to say something about what happens to these states as $n \to \infty$.

Now comes the first problem: φ is probably not a nice function, which is another way of saying it is not linear. Not being linear makes it quite hard to work with. Additionally, the given state space Ω is also quite unwieldy, given that in the real world we would never be able to measure the state accurately enough anyway. So we change our perspective: instead, we consider an observable

$$f : \Omega \to \mathbb{R}.$$

This could be, for example, the temperature at a given point in the given state. For each point in the space ($\mathbb{R}^{3d}$ is this case) there would be an observable for that point. These observables may be nonlinear, but they together form a vector space, since any linear combination of observables is also an observable. We could then look at how the function φ affects the set of observables. Looking at it this way has a huge advantage, since the induced mapping

$$T_\varphi := (f \mapsto f \circ \varphi)$$

is *linear*. This is nice as now the methods of functional analysis can be applied to the problem, with success. Instead of looking at the orbit of φ we look at the orbit of T_φ applied to the observable f. Even from a physical point of view it is clearly much more useful to be able to study how the temperature and pressure evolves than predicting the location of each individual particle. As a bonus it is also easier.

One slight snag is that we have a function which gives us the next state, but the real world is nowhere near that simple. No measuring device is good enough to calculate an instantaneous pressure or temperature, instead they measure averages over time. So instead of just following the orbit of T_φ, we consider the averages over time and how they behave, i.e.

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_\varphi^n f(\omega_0) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\varphi^n(\omega_0)) = \cdots ?$$

However, all this is for nought if we still need to determine the initial state ω_0. What we would really like is if the result of the above series was independent of the initial value. This is what Boltzmann suggested and he had an idea for what it would be, the now well-known phrase "time-mean equals space-mean". He suggested there would be some probability measure μ on Ω such that:

Hypothesis 1.1 (The Ergodic Hypothesis)**.** *For each initial state $\omega_0 \in \Omega$ and each (reasonable) observable $f : \Omega \to \mathbb{R}$, it is true that "time mean equals space mean", i.e.,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\varphi^n(\omega_0)) = \int_\Omega f \, d\mu$$

Going back to a physical interpretation of this, that the system would converge to some kind of average over all the observables over the state space.

A major step forward was the proof of the Mean Ergodic Theorem by von Neumann, which proved the following for Hilbert spaces.

Theorem 1.2 (The Mean Ergodic Theorem, von Neumann, 1931)**.** *Let $(\Omega, \Sigma, \mu; \varphi)$ be such that μ is φ-invariant and consider the induced operator $T := T_\varphi$. Then for each $f \in L^2(\Omega, \Sigma, \mu)$ the limit*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} T^j f$$

exists in the L^2-sense and is a fixed point of T.

Proof. [vN32]. □

Operator-theoretic Ergodic Theory

This is where the mathematicians took off. Once the problem was posed they began a study of what conditions on T_φ, Ω and f would guarantee the above result. While being originally based on a physical problem, it has since developed into a field of mathematics that has some startling interactions with other fields, such as number theory and group theory. Being operator-theoretic means that the exact origins of the definitions are irrelevant, only the properties of the spaces we apply them to. So, here we will be concentrating on the following specific cases:

- Ω will not be considered directly. Instead we will only concern ourselves with the vector space of measurements, i.e. real-valued functions on Ω. Specifically, we will primarily consider the standard spaces, ℓ^p and L^p with $1 \leqslant p \leqslant \infty$.

- T_φ (often indicated by just T) will be a linear operator that will be a contraction in general, though we will also consider the class of power-bounded operators.
- In the series above we will be looking at both weak and strong convergence, as well as convergence of sub-sequences.

Continuing the abstraction, the concept of ergodic theory was extended to so called topological dynamical systems.

Definition 1.3 (TDS)**.** The pair $(K;\varphi)$ is called a *topological dynamical system* (TDS) if K is a non-empty compact space and $\varphi : K \to K$ is continuous.

A TDS is simply what was described in the introduction but applied to an abstract compact space K. The extension to measure spaces is captured in the following definition.

Definition 1.4 (MDS)**.** The quadruple $(\Omega, \Sigma, \mu; \varphi)$ is called a *measure-preserving dynamical system* (MDS) if (Ω, Σ, μ) is a probability space, $\varphi : \Omega \to \Omega$ is measurable and μ is φ-invariant, i.e. $\mu(A) = \mu(\varphi^{-1}A)$ for all $A \in \Sigma$.

However, this is operator-theoretic ergodic theory and hence we do not trouble ourselves with the precise origins of the terminology. Instead, we see that the induced operator T_φ is a linear operator and try to apply the power of functional analysis to it. The ergodic hypothesis is captured in the following definition.

Definition 1.5 (Mean Ergodic, operator theoretic)**.** An operator T on a Banach space X is called mean ergodic if

$$\frac{1}{N}\sum_{n=1}^{N} T^n x \to Px$$

in the norm of X for some operator $P \in BL(X)$.

Definition 1.6 (Mean Ergodic Projection)**.** The operator $P \in BL(X)$ above, when it exists, is called the *mean ergodic projection* of T.

It can be shown that P is the projection onto the fixed space of T along $\overline{\mathrm{ran}(I-T)}$. On a Hilbert space this is the orthogonal projection onto $\mathrm{Fix}(T)$. Note that showing mean-ergodicity can be a task by itself, but fortunately, for many common spaces the results are simple.

Example 1.7. Every power bounded operator on a reflexive Banach space is mean-ergodic.

This is obviously a very large class of operators.

Example 1.8 (Dunford-Schwartz, 1956)**.** Let (Ω, Σ, μ) be a finite measure space. An operator $T \in BL(L^1(\Omega, \Sigma, \mu))$ is a Dunford-Schwartz operator (or complete contraction) if

$$\|T\|_1 \leqslant 1 \qquad\qquad \|T\|_\infty \leqslant 1.$$

If T is a Dunford-Schwartz operator, then T is mean-ergodic.

Proof. [DS56] □

Example 1.9. Let $a \in \mathbb{T}$ (the torus) and let ϕ_a be the rotation by a on $\mathbb{T}$. Then the induced operator T_{ϕ_a} on $C(\mathbb{T})$ is mean-ergodic.

Considering finer classifications of operators we have the following definitions.

Definition 1.10 (Strongly mixing, operator theoretic)**.** The operator T on a Banach space X is called strongly mixing if the sequence $(T^n)_n$ converges in the weak operator topology. Specifically, there exists an operator $P \in BL(X)$ such that

$$\langle T^n f, g\rangle \to \langle Pf, g\rangle$$

for all $f \in X, g \in X^*$.

Strongly mixing is a very strong condition. Any strongly mixing operator is automatically mean-ergodic and in fact the operator P is the mean-ergodic projection mentioned earlier. The origin of this term is that a strongly mixing dynamical system diffuses over time, that if we choose any part of our space, then over time all the "particles" in there will be distributed equally over the entire space. Between these two levels there is another form of mixing.

Definition 1.11 (Weakly mixing, operator theoretic)**.** An operator T on a Banach space X is called *weakly mixing* if there exists an increasing sequence of positive integers $(n_i)_i$ of density 1 such that $(T^{n_i})_i$ converges weakly. Specifically there exists an operator $P \in BL(X)$ such that

$$\langle T^{n_i} f, g\rangle \to \langle Pf, g\rangle$$

for all $f \in X, g \in X^*$.

The only change is shifting from simple convergence to the convergence of a subsequence of density 1. A small change but one with significant implications. It essentially means that the diffusion is perfect, except if we examine the measurements at "bad" times, where these "bad" times are of density 0, i.e. almost never. For a long time it was an open problem to find MDSs that were weakly mixing but not strongly mixing. These were found by Chacon in [Cha67].

We can summarise the above relationships as follows.

$$\boxed{\text{Strongly mixing}} \longrightarrow \boxed{\text{Weakly mixing}} \longrightarrow \boxed{\text{Mean-ergodic}}$$

We are not discussing ergodic theory in general, here we are specifically interested in the following result.

Theorem 1.12 (Blum-Hanson (1960))**.** *Let $(\Omega, \Sigma, \mu; \varphi)$ be an MDS with $T := T_\varphi$ its induced operator on $L^p(\Omega, \Sigma, \mu), 1 \leqslant p < \infty$. The MDS is strongly mixing if and only if for every increasing sequence of positive integers $(n_k)_k$ and for every $f \in L^p(\Omega, \Sigma, \mu)$ we have the strong convergence:*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} T^{n_j} f = \int_\Omega f \, d\mu \cdot \mathbf{1}$$

Proof. See [BH60]. □

This result opened a new area of research in ergodic theory. The goal of this thesis is to give an overview of the results in the area, in particular the recent results by Müller and Tomilov [MT07].

Chapter 2

Background

2.1 Power-bounded operators

Every bounded linear operator has a norm which gives an upper bound to the growth of the operator. It is defined as

$$\|T\| = \sup_{\|x\|=1} \|Tx\|.$$

This is the *induced norm* or *operator norm.* If $\|T\| \leqslant 1$ then T is called a contraction. Intuitively this means that it shrinks the input vector to something smaller.

Definition 2.1. An operator T on a Banach space X is called *power bounded* if the powers of the operator are uniformly bounded, that is

$$\sup_{n\in\mathbb{N}} \|T^n\| < C$$

for some constant C.

Obviously every contraction is power-bounded, so the interesting cases involving power-bounded operators focus on operators whose norms are greater than one.

Definition 2.2. Let T be a power-bounded operator on the Banach space X. Then one can define a new norm on X as

$$|||x||| := \sup_{n\geqslant 0} \|T^n x\|$$

for all $x \in X$.

Since we have the bounds $\|x\| \leqslant |||x||| \leqslant C\|x\|$ the new norm is equivalent to the old norm (i.e. it induces the same topology), however with respect to this new norm T is a contraction. It is worth noting that while the new norm preserves the topology, it does not preserve other properties. For example, after renorming a Hilbert space is no longer a Hilbert space. In general the properties of a space are divided into topological properties, which are preserved by the change of norm, and metric properties, which depend on the precise norm.

Example 2.3. As a simple example, consider the operator T defined on $\mathbb{R}^2$ by the matrices

$$T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

This operator first scales the axes and then rotates the space by θ. Intuitively, as the point gets rotated we would expect it to get pulled in and out but never escape or go to zero, except in the case where θ is a multiple of π. The actual behaviour is more complicated, depending on the exact value of θ.

Lemma 2.4. *The operator T as defined above is power-bounded whenever $|\cos\theta| < \frac{4}{5}$.*

Proof. By expansion we have that

$$T = \begin{bmatrix} 2\cos\theta & -\frac{1}{2}\sin\theta \\ 2\sin\theta & \frac{1}{2}\cos\theta \end{bmatrix}.$$

The eigenvalues are solutions of the characteristic equation

$$(2\cos\theta - \lambda)(\tfrac{1}{2}\cos\theta - \lambda) + \sin^2\theta = 1 - \tfrac{5}{2}\lambda\cos\theta + \lambda^2.$$

So we have

$$\lambda = \tfrac{5}{4}\cos\theta \pm \tfrac{1}{2}\sqrt{\tfrac{25}{4}\cos^2\theta - 4} = \tfrac{5}{4}\cos\theta \pm \sqrt{\tfrac{25}{16}\cos^2\theta - 1}.$$

We need to examine three cases:

- When $|\cos\theta| < \frac{4}{5}$ then the value under the square root is negative giving a pure imaginary number, so
$$|\lambda|^2 = \tfrac{25}{16}\cos^2\theta - \tfrac{25}{16}\cos^2\theta + 1 = 1$$
and T is power-bounded.
- When $|\cos\theta| > \frac{4}{5}$ the value under the square root is positive and for the larger of the two eigenvalues
$$|\lambda|^2 = \tfrac{25}{16}\cos^2\theta + \tfrac{25}{16}\cos^2\theta - 1 > 1$$
So we can decompose T as QDQ^{-1} where D is a diagonal matrix containing the eigenvalues. Now $T^n = QD^nQ^{-1}$ and as D^n is not bounded, T is not power-bounded.
- When $|\cos\theta| = \frac{4}{5}$ we have the eigenvalue 1 with multiplicity two. Now
$$T = \begin{bmatrix} \frac{8}{5} & -\frac{3}{10} \\ \frac{6}{5} & \frac{4}{10} \end{bmatrix}.$$
As this matrix has a deficient (or defective) eigenspace, it decomposes as
$$T = Q\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}Q^{-1}.$$
Now,
$$T^n = Q\begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}Q^{-1}$$
and T is again not power bounded.

□

As you can see, power-boundedness can be a very subtle property. It is also intimately related to the spectrum, in particular the spectral radius.

Definition 2.5 (Spectral radius)**.** The spectral radius of an operator T, denoted by $\rho(T)$, is defined as

$$\rho(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$$

where $\sigma(T)$ is the spectrum of T.

It is the spectral radius that reveals the underlying growth pattern. If the spectral radius is strictly less than one, the powers will converge to zero. If the spectral radius is greater than one the powers will be unbounded. Only when the spectral radius is exactly one can other behaviours appear.

The relationship with power-bounded operators is via Gelfand's Formula.

Theorem 2.6 (Gelfand's formula)**.** *Let T be an operator on a Banach space X. Then the spectral radius ($\rho(T)$) is given by*
$$\rho(T) = \lim_{n\to\infty} \|T^n\|^{1/n}.$$

Proof. [Gel41] □

This shows that any power-bounded operator has a spectral radius of 1, just like a contraction. There has been much research into the exact relationship between power-bounded operators and contractions.

2.2 Polynomially-bounded operators

Another class of functions we will consider are the polynomially-bounded operators. Consider the following definition.

Definition 2.7 (H^∞ norm)**.** Define the $\|\cdot\|_{H^\infty}$ norm on the space of complex polynomials as

$$\|q\|_{H^\infty} := \sup_{\substack{z\in\mathbb{C} \\ |z|=1}} |q(z)|.$$

Now we define the set of polynomials

$$\mathcal{P} = \{q \text{ is a polynomial} : \|q\|_{H^\infty} \leqslant 1\}$$

This is the set of all polynomials which are modulus less than one on the unit disc. This definition only looks at the boundary of the unit-disc, but as q is a polynomial and thus analytic, the maximum modulus principle applies.

Definition 2.8. The operator T is called polynomially bounded if there exists a constant C such that
$$\|q(T)\| \leqslant C \cdot \|q\|_{H^\infty}$$
for any polynomial q.

Example 2.9. As an example of an operator that is polynomially bounded, let $(\lambda_n)_n$ be a sequence of complex numbers such that $|\lambda_n| \leqslant 1$. Then the multiplication operator $M \in BL(\ell^p)$, defined by

$$Me_n := \lambda_n e_n,$$

is polynomially bounded. To see this note $q(M)e_n = q(\lambda_n)e_n$ where q is a polynomial and that $q(M)$ is still a multiplication operator. Hence $\|q(M)\| = \sup_n |q(\lambda_n)| \leqslant \|q\|_{H^\infty}$ which is bounded by the definition of the polynomial q.

The classical result in this area is the von Neumann inequality.

Theorem 2.10 (von Neumann inequality, 1951)**.** *Let T be a contraction on a Hilbert space. Then for any polynomial $q(z)$ we have*
$$\|q(T)\| \leqslant \|q\|_{H^\infty}.$$
In particular T is polynomially bounded by 1.

Proof. [vN51] □

In other words, for a fixed contraction T, the polynomial functional calculus map is itself a contraction.

The converse, whether all polynomially bounded operators are similar to a contraction, was for a very long time an open question but was finally laid to rest by G. Pisier in [Pis97]. The answer is no.

Definition 2.8 is only really useful in Hilbert spaces. This is because on general ℓ^p spaces only very special operators are polynomially-bounded.

Example 2.11. For example, suppose S is the right-shift operator and suppose we have a complex sequence $(a_n)_{n=0}^{\infty}$ such that $\sum_{n=0}^{\infty}|a_n|$ diverges. Then suppose the sequence of polynomials $q_N(z) = \sum_{n=0}^{N} a_n z^n$ are such that $q_N(z) \in \mathcal{P}$. Then

$$\|q_N(S)\|_{\ell^1\to\ell^1} \geqslant \|q_N(S)e_0\|_{\ell^1\to\ell^1} = \sum_{n=0}^{N}|a_n|,$$

and so the right shift operator is not polynomially bounded on ℓ^1. The adjoint (the left shift) gives a similar example for ℓ^∞.

To see that such a sequence exists we need a result from Zygmund and Paley ([MP81]). Consider the series

$$f(x) = \sum_{n=0}^{\infty} c_n \varepsilon_n \mathrm{e}^{inx}$$

where $\sum_{n=0}^{\infty} c_n^2 < \infty$ and $(\varepsilon_n)_n$ is a Rademacher sequence, i.e. it is a sequence of independent random variables taking the values ± 1 with the same probability $\frac{1}{2}$. Then the above series converges uniformly almost surely if

$$\sum_{n=2}^{\infty} c_n^2 \log(n)^{1+\varepsilon} < \infty$$

for any $\varepsilon > 0$. Now use $c_n = \frac{1}{n}$ and find a bounded realisation and set $a_n = \varepsilon_n c_n$ and we have the required sequence of polynomials.

In [Zar05] it is shown that more generally contractions aren't polynomially bounded in general Banach spaces.

A suggested extension of the term polynomial boundedness to general ℓ^p (even L^p) is known as Matsaev's conjecture. Here the requirement is that for any polynomial $q(z)$ and contraction T we have

$$\|q(T)\|_{L^p\to L^p} \leqslant \|q(S)\|_{\ell^p\to\ell^p}.$$

In other words, the shift operator represents the worst case. For $p = 2$ it can be shown that this is the von Neumann inequality. For $p = 1$ and $p = \infty$ it is a simple calculation, Example 2.11 shows an extreme case. For all other p it is an open question.

The main question with respect to polynomially bounded operators is how they relate to power-bounded operators, since a polynomially bounded operator is clearly also power-bounded. The relationships between contractions, power-bounded and polynomially bounded operators can be described as follows.

$$\boxed{\text{Contraction}} \longrightarrow \boxed{\text{Polynomially-bounded}} \longrightarrow \boxed{\text{Power-bounded}}$$

2.3 The Blum-Hanson property

Returning to the primary topic of this thesis, the definition of the Blum-Hanson property is as follows.

Definition 2.12 (Blum-Hanson property, sequence)**.** A sequence $(x_n)_{n=1}^{\infty}$ in a Banach space X has the Blum-Hanson property if for each increasing sequence of positive integers $(k_n)_{n=1}^{\infty}$ the limit

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} x_{k_n}$$

exists in the norm topology.

An important thing to note is that a sequence can only have the Blum-Hanson property if it is bounded.

Lemma 2.13. *Suppose a sequence $(x_n)_n$ in a Banach space X has the Blum-Hanson property. Then it is bounded.*

Proof. Suppose the sequence is not bounded. Then we can inductively construct an increasing sequence of positive integers $(k_n)_n$ such that

$$\begin{aligned} \|x_{k_1}\| &\geqslant 1 && \text{for } n = 1 \\ \|x_{k_n}\| &\geqslant n^2 + \left\| \sum_{j=1}^{n-1} x_{k_j} \right\| && \text{for } n > 1. \end{aligned}$$

Then

$$\begin{aligned} \left\| \frac{1}{n}\sum_{j=1}^{n} x_{k_j} \right\| &= \frac{1}{n}\left\| x_{k_n} + \sum_{j=1}^{n-1} x_{k_j} \right\| \\ &\geqslant \frac{1}{n}\left| \|x_{k_n}\| - \left\| \sum_{j=1}^{n-1} x_{k_j} \right\| \right| \\ &\geqslant \frac{1}{n}\left| \left(n^2 + \left\| \sum_{j=1}^{n-1} x_{k_j} \right\| \right) - \left\| \sum_{j=1}^{n-1} x_{k_j} \right\| \right| \\ &= n \end{aligned}$$

and so the sequence $(x_n)_n$ does not have the Blum-Hanson property. $\square$

If the above limit exists for all subsequences $(k_n)_n$ then necessarily the limit must be independent of the subsequence. To see this consider the possibility of two different subsequences converging to different limits. Then one could construct a third series which does not converge at all, which is a contradiction.

The application to ergodic theory is whether the orbit of a point under the powers of an operator has the required property.

Definition 2.14 (Blum-Hanson property, operator)**.** Let X be a Banach space. The operator $T \in BL(X)$ has the Blum-Hanson property if for every increasing sequence of positive integers $(k_n)_{n=1}^{\infty}$ the limit

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} T^{k_n} x$$

exists in the norm topology for each $x \in X$.

The limit depends on x. We note:

- If an operator has the Blum-Hanson property for all $x \in X$ then as noted above, T must be a power-bounded operator by application of the principle of uniform boundedness.
- By using the sequence $k_n = n$ we have convergence to the mean-ergodic projection, so any operator with the Blum-Hanson property is mean-ergodic. By the equality of the limits every subsequence converges to this projection.

The origin of this term, as mentioned earlier, is from the paper by Blum-Hanson [BH60] where it was proved that the Blum-Hanson property is equivalent to the strong mixing property whenever T is an operator induced by a measure preserving transformation.

This lead to research into determining for what other classes of operators this equivalence could be shown. It became apparent that two kinds of equivalences can be shown. The first kind is the strong kind, where the equivalence is pointwise.

Definition 2.15 (Blum-Hanson equivalence, individual)**.** A mean ergodic operator $T \in BL(X)$ satisfies the individual (or strong) Blum-Hanson equivalence if for each $x \in X$ the following are equivalent.

- $T^n x \to Px$ weakly,
- $\frac{1}{N}\sum_{n=1}^{N} T^{k_n} x \to Px$ in norm for all increasing sequences of positive integers $(k_n)_n$,

where $P \in BL(X)$ is the mean ergodic projection of T.

The second kind is weaker, where the equivalence is only for the operator as a whole.

Definition 2.16 (Blum-Hanson equivalence, operator)**.** A mean ergodic operator $T \in BL(X)$ satisfies the operator (or weak) Blum-Hanson equivalence if the following are equivalent.

- $T^n \to P$ in the weak operator topology,
- $\frac{1}{N}\sum_{n=1}^{N} T^{k_n} \to P$ in the strong operator topology for all increasing sequences of positive integers $(k_n)_n$,

where $P \in BL(X)$ is the mean ergodic projection of T.

Clearly the strong equivalence implies the weak equivalence. The strong equivalence is much more useful as it does not require any properties of the operator as a whole, it only needs properties of an individual orbit.

As noted, the operator P in the above definition, if it exists, is simply the mean-ergodic projection (Definition 1.6) and is thus the projection onto the fixed space of T. Since for any $x \in \mathrm{Fix}(T)$ the above equivalences are trivially true, it is sufficient (that is, without loss of generality) to show that the equivalences are true for the case where $P = \mathbf{0}$, i.e. operators with a trivial fixed space.

When proving the Blum-Hanson equivalence, one implication is easy and actually true for all Banach spaces.

Lemma 2.17. *Let X be a Banach space. If the sequence $(x_n)_n \in X$ has the Blum-Hanson property then the sequence $(x_n)_n$ converges weakly.*

Proof. The Blum-Hanson property gives strong mean convergence along any subsequence, so in particular $\frac{1}{N}\sum_{n=1}^{N} x_n \to x$ strongly. Hence if $(x_n)_n$ is weakly Cauchy then $x_n \to x$ weakly.

If $(x_n)_n$ does not converge weakly, then there must exist a $y \in X^*$ such that $\langle x_n, y\rangle$ does not converge. This sequence is bounded so we can create two subsequences $(r_j)_j$ and $(s_j)_j$ such that $\lim_{j\to\infty}\langle x_{r_j}, y\rangle \neq \lim_{j\to\infty}\langle x_{s_j}, y\rangle$. Now we can create a new sequence $(k_j)_j$ taking 1 element from r_j, 2 from s_j, 4 from r_j, etc. In this sequence $\frac{1}{N}\sum_{n=1}^{N}\langle x_{k_j}, y\rangle$ diverges weakly and thus also strongly. Contradiction. $\square$

Corollary 2.18. *Let T be an operator on a Banach space X. If T has the Blum-Hanson property then $(T^n)_n$ converges in the weak operator topology.*

In contrast, the other implication is far from easy in the general case. There have been proofs for various specific spaces, but nothing general. There is however a partial converse.

Theorem 2.19. *Let T be an operator on the Banach space X. Then if T is weakly mixing then*

$$\frac{1}{N}\sum_{n=1}^{N} T^{k_n}x$$

converges in norm for all subsequences $(k_n)_n$ of positive lower density. If X^* is separable, the above holds "if and only if".

Proof. [Jon71] □

Since strongly mixing implies weakly mixing, any counterexample of the Blum-Hanson equivalence must use sequences of zero lower density.

Generalisations

In recent times there has been research into variations of the above property. They are beyond the scope of this thesis but brief descriptions are included here for reference.

Different subsequences. Instead of just considering all increasing subsequences of $\mathbb{N}$, one could look at the convergence requiring only sequences from a particular class. For example:

- the class $\mathcal{N}_+$ of all subsets of $\mathbb{N}$ having strictly positive lower density;
- the class $\mathcal{N}_{rd}$ of all relatively dense subsets of $\mathbb{N}$.

With this terminology we say that a sequence $(x_k)_k$ in a Banach space has the Blum-Hanson property with respect to $\mathcal{N}$ if the Cesàro means $\frac{1}{n}\sum_{j=1}^{n} x_{k_j}$ are convergent in the norm topology for all subsequences belonging to $\mathcal{N}$. Just as in the usual case, it can be shown for each of the classes above that if the Cesàro means converge then they must converge to the same limit. However, it is no longer true that the sequence must be bounded. See [Zsi07].

Weak convergence. Instead of looking only at the strong convergence of the averages, we could look at the weak convergence instead. The first results in this direction were by [JL76] where it was shown to be related to weak mixing in general and to the convergence along subsequences of positive lower density. This work has been extended by [Zsi07].

2.4 Similarity

Definition 2.20 (Similarity)**.** Two bounded linear operators $S : X \to X$ and $T : Y \to Y$ are called similar if there exists a bounded invertible operator $Q : X \to Y$ such that:

$$T = QSQ^{-1}$$

The way to think of this is that the operators do the same thing but in different bases. The Jordan decomposition theorem shows that every finite dimensional matrix is similar to an almost diagonal matrix. Similarity preserves many properties.

Lemma 2.21. *Let S and T be two similar operators. Then:*

1. *S is power-bounded if and only if T is power-bounded.*
2. *S is polynomially bounded if and only if T is polynomially bounded.*
3. *For any sequence of polynomials $(p_n)_n$, $p_n(S)$ converges weakly if and only if $p_n(T)$ converges weakly.*
4. *For any sequence of polynomials $(p_n)_n$, $p_n(S)$ converges strongly if and only if $p_n(T)$ converges strongly.*

   As a direct consequence:

   - *S is mean ergodic if and only if T is mean ergodic.*
   - *S has the Blum-Hanson property if and only if T has the Blum-Hanson property.*
   - *S is strongly mixing if and only if T is strongly mixing.*

Proof. As S and T are similar there exists a Q such that $S = QTQ^{-1}$. In each case we show the "if", the "only if" follows by reversing the roles of S and T.

1. Suppose S is power-bounded. Then
$$\sup_n \|T^n\| = \sup_n \|QS^nQ^{-1}\| \leqslant \|Q\|\|Q^{-1}\| \sup_n \|S^n\|$$
and T is power-bounded.
2. Suppose S is polynomially bounded. Let $p(z)$ be a polynomial, then:
$$\|p(T)\| = \|p(QSQ^{-1})\| = \|Qp(S)Q^{-1}\| \leqslant \|Q\|\|Q^{-1}\|\|p(S)\|$$
and T is polynomially bounded.
3. Suppose S converges weakly for the sequence polynomials $(p_n)_n$ to $\hat{S}$. Without loss of generality we can choose $\hat{S}$ to be zero. Choose an $\varepsilon > 0, x \in Y, y \in Y^*$ and find M such that $|\langle p_M(S)Q^{-1}x, Q^*y\rangle| < \varepsilon$. Then
$$|\langle p_M(T)x, y\rangle| = |\langle Qp_M(S)Q^{-1}x, y\rangle| = |\langle p_M(S)Q^{-1}x, Q^*y\rangle| \leqslant \varepsilon$$
and so T converges weakly to zero for the same sequence of polynomials.
4. Suppose S converges strongly for the sequence polynomials $(p_n)_n$. Choose an $\varepsilon > 0$ and find M and N such that $\|p_M(S) - p_N(S)\| < \varepsilon$. Then for any $x \in Y$:
$$\sup_{\|x\|=1} \|p_M(T)x - p_N(T)x\| = \|Q(p_M(S) - p_N(S))Q^{-1}x\| \leqslant \|Q\|\|Q^{-1}\|\varepsilon$$
and T converges strongly for the same sequence of polynomials.

□

One of the early results on polynomial boundedness relates to similarity.

Theorem 2.22. *Let T be a power-bounded operator on a Hilbert space H such that the inverse T^{-1} is also power-bounded, i.e. $\|T^n\| \leqslant C$ and $\|T^{-n}\| \leqslant C$ for some C for all $n \in \mathbb{N}$. Then there exists a self-adjoint operator Q such that*
$$\tfrac{1}{C} \leqslant \|Q\| \leqslant C$$
and QTQ^{-1} is a unitary transformation.

What this theorem proves is that under the stated conditions, the operator is actually similar to a unitary operator.

Proof. [SN47] Consider two elements $f, g \in H$. Then the sequence $(\langle T^n f, T^n g\rangle)_n$ is a bounded sequence and so we can define a linear functional $\phi(f, g) : H \times H \to \mathbb{C}$ as a Banach limit of the sequence as $n \to \infty$. By the properties of the Banach limit we have:

$$\phi(a_1 f_1 + a_2 f_2, b_1 g_1 + b_2 g_2) = a_1 \overline{b_1} \phi(f_1, g_1) + a_1 \overline{b_2} \phi(f_1, g_2) + a_2 \overline{b_1} \phi(f_2, g_1) + a_2 \overline{b_2} \phi(f_2, g_2)$$

which makes $\phi(f, g)$ a hermitian sesquilinear form, i.e. it is linear in f and antilinear in g. Furthermore the inequalities

$$\langle T^n f, T^n f\rangle = \|T^n f\|^2 \leqslant C^2 \|f\|^2$$

and

$$\|f\|^2 = \|T^{-n} T^n f\|^2 \leqslant \|T^{-n}\|^2 \|T^n f\|^2 = C^2 \langle T^n f, T^n f\rangle$$

show that, for all $n \in \mathbb{N}$

$$\frac{1}{C^2} \|f\|^2 \leqslant \langle T^n f, T^n f\rangle \leqslant C^2 \|f\|^2.$$

This uniformly bounds the terms of the Banach limit, hence

$$\frac{1}{C^2} \|f\|^2 \leqslant \phi(f, f) \leqslant C^2 \|f\|^2.$$

Using a standard result on bounded hermitian sesquilinear forms (Theorem A.5) there exists a linear operator $A \in BL(H)$ such that $\phi(f, g) = \langle Af, g\rangle$. Since $\phi(f, f) \in \mathbb{R}$ by another standard result (Lemma A.6) we have that A is hermitian.

Now, by the shift invariance of the Banach limit we have that

$$\phi(Tf, Tg) = \phi(f, g)$$

from which follows

$$\langle T^* ATf, g\rangle = \langle ATf, Tg\rangle = \phi(Tf, Tg) = \phi(f, g) = \langle Af, g\rangle$$

which means that

$$T^* AT = A.$$

Note that since $\phi(f, f) \geqslant 0$ we know that A is even positive, so let Q be the positive self-adjoint square root of A, i.e. $Q^2 = A$. From the above we have

$$\frac{1}{C} \leqslant \|Q\| \leqslant C \text{ and } \frac{1}{C} \leqslant \|Q^{-1}\| \leqslant C$$

Defining $U = QTQ^{-1}$ we see from

$$U^* U = Q^{-1}(T^* QQT)Q^{-1} = Q^{-1}(QQ)Q^{-1} = I$$

that U is isometric. As it admits an inverse $U^{-1} = QT^{-1}Q^{-1}$, it is unitary. This completes the proof. □

Corollary 2.23. *Let T be a linear operator on a Hilbert space such that T and T^{-1} are power-bounded. Then T is polynomially bounded.*

Proof. A consequence of the von Neumann inequality applied to the previous theorem. □

The conclusion is that the classes mentioned in the Lemma 2.21 are preserved by similarity. Proving the Blum-Hanson equivalence is the same as showing that the class of strongly mixing operators is the same as the class of operators with the Blum-Hanson property.

However, contractions do not remain contractions under similarity. For our purposes this is a small problem, but easily remedied by considering them as part of a larger class of operators which are similar to a contraction. The result of Theorem 2.22 is that all power-bounded operators whose inverse is also power-bounded are actually within the class of operators similar to a contraction.

With what we have so far we can create the following diagram for Hilbert spaces.

$$
\begin{tikzcd}
\boxed{\text{Blum-Hanson property}} \arrow[d] & \boxed{\text{Polynomially-bounded}} \arrow[dd] & \\
\boxed{\text{Strongly mixing}} \arrow[d] \arrow[dr] & & \boxed{\text{Similar to a contraction}} \arrow[ul] \arrow[dl] \\
\boxed{\text{Weakly mixing}} \arrow[d] & \boxed{\text{Power-bounded}} \arrow[dl] & \\
\boxed{\text{Mean-ergodic}} & &
\end{tikzcd}
$$

Apart from the implications relating to polynomially boundedness, the remainder is true for all reflexive Banach spaces. The question is: can this diagram be simplified, are any of these classes the same? And in particular for our purposes, are the Blum-Hanson property and strong mixing classes the same?

2.5 Similarity to a contraction

By considering the larger class of operators which are similar to a contraction we are left with the difficulty that showing an operator is similar to a contraction involves actually finding the similarity. What we would like is some properties of these operators we can use to simplify our proofs. There are some criteria such as Theorem 2.22 but they are quite special. It is usually simpler to prove something is not similar to a contraction by showing (for example) it is not polynomially-bounded, which by von Neumann's inequality implies it is not similar to a contraction.

One criterion we do have for Hilbert spaces is the following.

Theorem 2.24. *Let T be a contraction on a Hilbert space H. Define the space $H_0(T) = \{x : T^n x \to 0 \text{ weakly}\}$. Then $H_0(T) = H_0(T^*)$.*

Proof. [Fog63] □

Because similarity is in general not preserved by taking adjoints, we cannot use this directly. However, we have the following.

Lemma 2.25. *Let P be a contraction on a Hilbert space and so $T = SPS^{-1}$ is similar to a contraction. Then*

$$H_0(T) \cap H_0(T^*)^{\perp} = \{0\}.$$

Proof. [Fog64] First we have

$$H_0(T) = \{x : SP^nS^{-1}x \to 0 \text{ weakly}\} = \{x : S^{-1}x \in H_0(P)\} = S(H_0(P))$$

and
$$H_0(T^*)^\perp = [S^{*-1}(H_0(P^*))]^\perp = \{x : S^{-1}x \perp H_0(P^*)\} = S(H_0(P^*)^\perp).$$
However, from Theorem 2.24 we know $H_0(P^*) = H_0(P)$, so
$$H_0(T) \cap H_0(T^*)^\perp = \{0\}$$
as required. □

Similarity to contractions is reasonably well studied for Hilbert spaces, but for L^p spaces or Banach spaces there are very few results.

Chapter 3

Positive results for the Blum-Hanson property

As noted earlier the term "Blum-Hanson property" was introduced during the exploration of the strong mixing property. The strong mixing property is a strong property and thus it is useful to have a different way of describing it. However, it is still an area of research under which conditions the two are in fact related.

The equivalence of the strong-mixing condition and the Blum-Hanson property has been proved in the cases where T is:

- an isometry induced by a measure preserving transformation on $L^p, 1 \leqslant p < \infty$ [BH60];
- a Hilbert space contraction [JK71];
- a contraction on L^1 [AS72];
- a positive contraction on $L^p, 1 < p < \infty$ [AS75, Bel75];
- in a certain class of positive power bounded functions on L^1 [Mil80];
- certain power-bounded operators on Hilbert-spaces [Kan79];
- a contraction on $\ell^p, 1 \leqslant p < \infty$ [MT07].

Similarly, there have been counterexamples, where T is:

- an isometry on $C(K)$ where K is a compact Hausdorff space [AHR74];
- a power-bounded operator on ℓ^2 [MT07].

What follows is a selection of the positive results that have been found, with annotated proofs to give an idea of the effort one needs to go to prove the equivalence.

3.1 Blum-Hanson with a measure preserving transformation

The original Blum-Hanson proof [BH60] was the beginning of a whole area of research. It is the oldest and (almost) the easiest of all the proofs. It considers the simple case: operators induced by measure preserving transformations on a probability space. Because of this, all that needs to be shown is that it is true in a Hilbert space and the others follow by density arguments. Here only the L^2 case will be shown as this is the most important part of the proof.

The proof by Blum-Hanson uses the following lemma.

Lemma 3.1. *Let $\{C_{i,j} : i, j \geqslant 1\}$ be a bounded double sequence of numbers such that*

$$\lim_{|i-j|\to\infty} C_{i,j} = 0.$$

Then

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i,j=1}^{n} C_{i,j} = 0$$

Proof. Let $C = \sup_{i,j} |C_{i,j}|$ and fix $\varepsilon > 0$. Choose $M \in \mathbb{N}$ such that $|C_{i,j}| < \varepsilon$ for $|i-j| > M$. Then for $n > M$ we have

$$\begin{aligned} \frac{1}{n^2} \sum_{i,j=1}^{n} |C_{i,j}| &= \frac{1}{n^2} \sum_{\substack{|i-j|\leqslant M \\ i,j\leqslant n}} |C_{i,j}| + \underbrace{\frac{1}{n^2} \sum_{\substack{|i-j|>M \\ i,j\leqslant n}} |C_{i,j}|}_{<\varepsilon} \\ &\leqslant \frac{(2M+1)nC}{n^2} + \varepsilon \end{aligned}$$

We can now choose an $N > \frac{(2M+1)C}{\varepsilon}$ and then for $n > N$

$$\frac{1}{n^2} \sum_{i,j=1}^{n} |C_{i,j}| \leqslant 2\varepsilon.$$

And the result follows. □

The heart of the proof of Blum-Hanson is the following theorem. With this it is proved that the required result holds for simple functions on L^2. After this, the result can be extended to the general case on L^p using standard density arguments.

Note that the theorem here is presented in purely functional analytic terms. The original proof was done in terms of strongly mixing sets. Here the invertible measure preserving φ is left behind and we only look at the induced operator (which does not need to be invertible).

Theorem 3.2 (Blum-Hanson 1960)**.** *Let $(\Omega, \Sigma, \mu; \varphi)$ be a strongly mixing MDS with $T := T_\varphi$ its induced operator on $L^2(\Omega, \Sigma, \mu)$. Then for any increasing sequence of positive integers $(k_n)_n$ and any $f \in L^2(\Omega, \Sigma, \mu)$ we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} T^{k_i} f = Pf$$

strongly in the L^2-norm, where P is the weak limit of $(T^n)_n$.

Proof. [BH60] The first thing to note is due to the μ-invariance of T we have:

$$\langle T^m f, T^n g\rangle = \begin{cases} \langle T^{m-n} f, g\rangle & \text{when } m \geqslant n \\ \langle f, T^{n-m} g\rangle = \overline{\langle T^{n-m} g, f\rangle} & \text{when } m \leqslant n \end{cases}$$

Let P be the mean ergodic projection onto $\mathrm{Fix}(T)$, then $T^n \to P$ weakly. P is an orthogonal projection which means that $\langle Pf, g\rangle = \langle Pf, Pg\rangle$. As P is the projection onto the fixed space and commutes with T we have

$$\langle Pf, Pg\rangle = \langle Pf, T^n Pg\rangle = \langle Pf, PT^n g\rangle = \langle Pf, T^n g\rangle.$$

So

$$\begin{aligned}\left\|\frac{1}{n}\sum_{i=1}^{n}T^{k_i}f-Pf\right\|_2^2&=\left\langle\frac{1}{n}\sum_{i=1}^{n}T^{k_i}f-Pf,\frac{1}{n}\sum_{j=1}^{n}T^{k_j}f-Pf\right\rangle\\&=\left\langle\frac{1}{n}\sum_{i=1}^{n}T^{k_i}f,\frac{1}{n}\sum_{j=1}^{n}T^{k_j}f\right\rangle-\left\langle\frac{1}{n}\sum_{i=1}^{n}T^{k_i}f,Pf\right\rangle\\&\quad-\left\langle Pf,\frac{1}{n}\sum_{j=1}^{n}T^{k_j}f\right\rangle+\langle Pf,Pf\rangle\\&=\frac{1}{n^2}\left\langle\sum_{i=1}^{n}T^{k_i}f,\sum_{j=1}^{n}T^{k_j}f\right\rangle-\langle Pf,Pf\rangle\\&=\frac{1}{n^2}\left\langle\sum_{\substack{i,j=1\\i>j}}^{n}T^{k_i-k_j}f,f\right\rangle+\frac{1}{n^2}\overline{\left\langle\sum_{\substack{i,j=1\\i\leqslant j}}^{n}T^{k_j-k_i}f,f\right\rangle}-\langle Pf,Pf\rangle\end{aligned}$$

Now we define $C_{i,j}$ as

$$C_{i,j}=\begin{cases}\langle T^{k_i-k_j}f,f\rangle-\langle Pf,Pf\rangle & \text{when } i>j\\ \overline{\langle T^{k_j-k_i}f,f\rangle}-\langle Pf,Pf\rangle & \text{when } i\leqslant j\end{cases}$$

Clearly $\lim_{|i-j|\to\infty}C_{i,j}=0$ due to the strong mixing of T implying that $\langle T^nf,f\rangle\to\langle Pf,Pf\rangle$. The conjugation has no effect as the convergence is to a real number. So by Lemma 3.1 we have that $\frac{1}{n}\sum_{i=1}^{n}T^{k_i}\to P$ in the strong operator topology. $\square$

Note that this is an example of the individual Blum-Hanson equivalence, i.e. the equivalence holds for each orbit and does not rely on any of the properties of the operator along other orbits.

This proof by Blum-Hanson was simple but it already showed a fruitful way of looking at the problem. Namely, splitting the sum over the two sides of an inner product and then using the weak convergence to show that terms sufficiently far away from the diagonal become small enough. Then something similar to Lemma 3.1 finishes the job.

3.2 Blum-Hanson for Hilbert space contractions

The first improvement on Blum-Hanson's result came in [JK71] in 1971. Here it was shown that all Hilbert space contractions which weakly converge have the Blum-Hanson property. The proof is similar to that of Blum-Hanson, though actually shorter.

Theorem 3.3 (Jones-Kuftinec 1971)**.** *Let T be a Hilbert space contraction. Then for any $x\in H$, T^nx converges weakly if and only if $\frac{1}{n}\sum_{j=1}^{n}T^{k_j}x$ converges strongly for all strictly increasing sequences (k_j) of positive integers.*

Proof. Let $x_0:=\lim\frac{1}{n}\sum_{j=1}^{n}T^jx$ which always exists since T is mean-ergodic. If $T^nx\not\rightharpoonup x_0$ weakly then apply Lemma 2.17 and there exists a sequence $(k_j)_j$ such that $\frac{1}{n}\sum_{j=1}^{n}T^{k_j}x$ does not converge strongly.

Conversely, suppose $T^nx\to x_0$ weakly. Since $x_0=Px$ is a fixed point of T we can assume without loss of generality $x_0=0$. Since T is a contraction $\inf\|T^nx\|=\lim\|T^nx\|=\alpha$ exists. If $\alpha=0$ then we are done, as the Cesàro means of a convergent sequence converge. If $\alpha>0$ then we can assume without loss of generality that $\alpha=1$ by considering x/α.

Fix an $\varepsilon > 0$ and choose an integer M such that $\|T^M x\|^2 < 1+\varepsilon$. Because of the weak convergence to zero of $(T^k x)_k$ we can choose an integer K such that $|\langle T^M x, T^{M+k}x\rangle| < \varepsilon$ for all $k \geqslant K$. Then for $m \geqslant M$ and $n \geqslant K+m$ we have

$$\begin{aligned}
2\operatorname{Re}\langle T^m x, T^n x\rangle &= \|T^m x + T^n x\|^2 - \underbrace{\|T^m x\|^2}_{\geqslant 1} - \underbrace{\|T^n x\|^2}_{\geqslant 1} \\
&\leqslant \|T^{m-M}(T^M x + T^{n-m+M}x)\|^2 - 2 \\
&\leqslant \|T^M x + T^{n-m+M}x\|^2 - 2 \\
&\leqslant \underbrace{\|T^M x\|^2}_{<1+\varepsilon} + \underbrace{\|T^{n-m+M}x\|^2}_{<1+\varepsilon} + 2\underbrace{\operatorname{Re}\langle T^m x, T^{n-m+M}x\rangle}_{<\varepsilon} - 2 \\
&\leqslant 4\varepsilon
\end{aligned}$$

Now consider any increasing sequence of positive integers $(k_j)_{j=1}^{\infty}$, then we have

$$\begin{aligned}
\left\|\frac{1}{N}\sum_{j=1}^{N} T^{k_j}x\right\|^2 &= \operatorname{Re}\left\langle \frac{1}{N}\sum_{i=1}^{N} T^{k_i}x, \frac{1}{N}\sum_{j=1}^{N} T^{k_j}x\right\rangle \\
&= \frac{1}{N^2}\sum_{i,j=1}^{N} \operatorname{Re}\langle T^{k_i}x, T^{k_j}x\rangle \\
&= \frac{1}{N^2}\sum_{i=1}^{N} \operatorname{Re}\langle T^{k_i}x, T^{k_i}x\rangle + \frac{2}{N^2}\sum_{i<j}^{N} \operatorname{Re}\langle T^{k_i}x, T^{k_j}x\rangle \\
&= \frac{1}{N^2}\sum_{i=1}^{N} \underbrace{\|T^{k_i}x\|^2}_{\leqslant \|x\|^2} + \frac{2}{N^2}\sum_{\substack{i<M \\ i<j}}^{N} \underbrace{\operatorname{Re}\langle T^{k_i}x, T^{k_j}x\rangle}_{\leqslant \|x\|^2} \\
&\quad + \frac{2}{N^2}\sum_{\substack{i\geqslant M \\ i<j<K+i}}^{N} \underbrace{\operatorname{Re}\langle T^{k_i}x, T^{k_j}x\rangle}_{\leqslant \|x\|^2} + \frac{1}{N^2}\sum_{\substack{i\geqslant M \\ j\geqslant K+i}}^{N} \underbrace{2\operatorname{Re}\langle T^{k_i}x, T^{k_j}x\rangle}_{<4\varepsilon} \\
&\leqslant \frac{\|x\|^2}{N} + \frac{2(M-1)N\|x\|^2}{N^2} + \frac{2(K-1)N\|x\|^2}{N^2} + 4\varepsilon \\
&= \frac{(2M+2K-3)}{N}\|x\|^2 + 4\varepsilon
\end{aligned}$$

Here we used that $k_i \geqslant i$ and $k_j - k_i \geqslant j - i$. Now for $N > (2M+2K-3)\|x\|^2/\varepsilon$ we have that $\|\frac{1}{N}\sum_{j=1}^{N} T^{k_j}x\|^2 \leqslant 5\varepsilon$ and so $(T^n x)_n$ has the Blum-Hanson property. □

Note what was done in the last step: splitting the double sum over i and j and considering 4 different cases.

1. $i = j$.
2. $i < \min\{j, M\}$.
3. $M \leqslant i < j < K+i$.
4. $M \leqslant i$ and $j \geqslant K+i$.

These areas are depicted in Figure 3.1. Due to symmetry the proof only considers half the possibilities. The important thing to note is that of all these areas, only the last one grows faster than

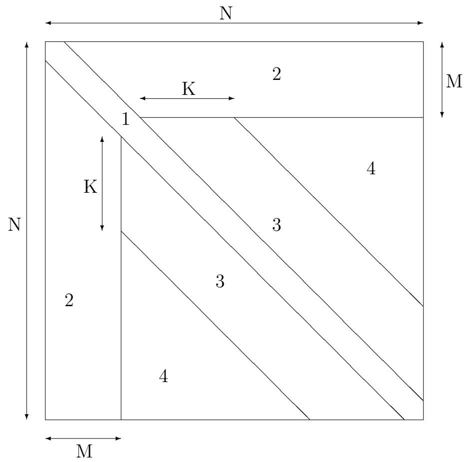


Figure 3.1: Division of cases for Theorem 3.3

linear with respect to N, but this is precisely the area that was proved could be made arbitrarily small. The conclusion is then straightforward.

It is worth noting that the proof of the positive L^p contraction case in [Bel75] actually runs along similar lines to this one, only there use is made of the identity

$$\|f\|_p^p = \langle f, \Phi(f) \rangle$$

where $\Phi(f)$ is the canonical duality mapping, which only works for positive f. Also, use is made of the following L^p inequality.

Definition 3.4 (L^p inequality)**.** Let $1 < p < \infty$. Let $f, g \in L_+^p(\Omega, \Sigma, \mu)$. Then for any $0 < \varepsilon < 1$ we have, with $\alpha = (p-1) + \frac{1}{p-1}$

$$\int_\Omega \Phi(f) \cdot g \, \mathrm{d}\mu \leqslant \varepsilon \|f\|_p^p + \varepsilon \|g\|_p^p + \frac{1}{\varepsilon^\alpha} \int_\Omega f \cdot \Phi(g) \, \mathrm{d}\mu$$

This inequality is used to work around the fact that the duality mapping is non-linear, but again it only works for positive functions. Together these make the proof very similar to the Hilbert space case, except that the result only applies to positive operators. For non-positive operators on general L^p spaces it is still an unsolved problem.

3.3 Blum-Hanson on L^1 contractions

That the Blum-Hanson property is true for all contractions on L^1 was proved by M. Akcoglu and L. Sucheston [AS72]. The proof is much more sophisticated than the two presented so far as we are no longer on a Hilbert space or even a reflexive space which means we no longer have many of the useful properties that helped before. A consequence of this is that the Blum-Hanson equivalence is only proved for the operator as a whole; the operator must converge weakly for all orbits to be able to show the Blum-Hanson property for any single orbit.

Theorem 3.5 (Blum-Hanson for L^1 contractions)**.** *Let T be a contraction on $L^1(\Omega, \Sigma, \mu)$. The following are equivalent:*

1. *$(T^n f)_n$ converges weakly in L^1 for all $f \in L^1(\Omega, \Sigma, \mu)$.*

2. $\frac{1}{n}\sum_{i=1}^{n} T^{k_i} f$ *converges strongly in* L^1 *for all* $f \in L^1(\Omega, \Sigma, \mu)$ *and for all increasing sequences* $(k_n)_{n=1}^{\infty}$ *of positive integers.*

First we state that without loss of generality we may assume the measure space is finite. To see this, note that the convergence of $(T^n f)_n$ is determined only by the set $\bigcup_n \{|T^n f| > 0\}$ which is sigma-finite.

Then consider the sigma-finite space $L^1(\Omega, \Sigma, \mu)$. Take a strictly positive $u \in L^1(\Omega, \Sigma, \mu)$ and define a new measure $\mathrm{d}\lambda = u^{-1} \cdot \mathrm{d}\mu$. We now have a new space $L^1(\Omega, \Sigma, \mu)$ which is finite and an operator $J : L^1(\Omega, \Sigma, \mu) \to L^1(\Omega, \Sigma, \lambda)$ between the two spaces defined as

$$Jf := f \cdot u.$$

This operator is an invertible isometry.

$$\|Jf\|_{L^1(\Omega,\Sigma,\lambda)} = \int_\Omega |Jf| \,\mathrm{d}\lambda = \int_\Omega |f| \,\mathrm{d}\mu = \|f\|_{L^1(\Omega,\Sigma,\mu)}$$

We also introduce the concept on a linear modulus.

Definition 3.6 (Linear modulus)**.** Let T be an operator on L^1. Define the operator $\tilde{T} \in BL(L^1_+)$ as

$$\tilde{T} f := \sup_{|g| \leqslant f} |Tg| \qquad f \in L^1_+, g \in L^1.$$

Then the linear modulus (denoted $|T|$) is a positive linear extension of $\tilde{T}$ to L^1 such that

- $|Tf| \leqslant |T||f|$ for each $f \in L^1$.
- $\| \, |T| \, \| \leqslant \|T\|$

Note that here $\leqslant$ is used to represent the partial order in the Banach lattice.

The existence of such an extension was proved in [CK64]. The linear modulus has the following additional properties.

(i) $|T^n f| \leqslant |T|^n |f|$ for each $f \in L^1$.

(ii) $|aT| = |a||T|$

(iii) $|T_1 + T_2| \leqslant |T_1| + |T_2|$

(iv) $T \geqslant 0$ implies $T = |T|$

The linear modulus is not necessarily unique. However, in [CK64] it is shown that that there exists a linear space of operators (which is dense in the strong operator topology on $BL(L^1)$) where the linear modulus is unique and can be explicitly calculated.

The relevant properties of the linear modulus which we will use are:

- the linear modulus does not have a greater norm, hence if T is a contraction, so is $|T|$.
- $|T|$ has the same invariant sets as T.

  Definition 3.7 (Invariant set)**.** Let T be a linear operator on $L^1(\Omega, \Sigma, \mu)$. A set $A \subset \Omega$ is an invariant set for T if for all functions $f \in L^1$ such that $\mathrm{supp}(f) \subseteq A$ we have $\mathrm{supp}(Tf) \subseteq A$.

The first lemma is rather technical. It essentially states that either the orbit of an arbitrary point goes strongly to zero, or the linear modulus has a fixed point.

Lemma 3.8. *Let T be a contraction on $L^1(\Omega, \Sigma, \mu)$. Suppose $f \in L^1$ and that $(T^n f)_n$ converges weakly in L^1 as $n \to \infty$. Then either $\lim_{n\to\infty}\|T^n f\| = 0$ or there exists a non-zero $p \in L^1_+(\Omega, \Sigma, \mu)$ such that $|T|p = p$.*

Proof. Since T is a contraction, $\alpha := \lim_{n\to\infty}\|T^n f\|$ exists. If $\alpha = 0$ then the lemma is proved, so assume $\alpha > 0$. Now define a positive linear functional $\pi : L^\infty \to \mathbb{C}$ as

$$\pi(h) := L_{n\to\infty} \int_\Omega |T^n f| \cdot h \, \mathrm{d}\mu$$

where L is a Banach limit. This is a bounded linear functional on L^∞ and hence a bounded finitely additive μ-continuous measure. This functional is non-zero because $\pi(\mathbf{1}) = \alpha > 0$.

We claim that π is a countably additive measure. To see this define the measures

$$\pi_n(A) := \int_A T^n f \, \mathrm{d}\mu \qquad A \in \Sigma.$$

By the weak convergence of $(T^n f)_n$ we have that $\lim_{n\to\infty} \pi_n(A)$ exists. Since $T^n f \in L^1$ we have that π_n are all μ-continuous measures.

By the Vitali-Hahn-Saks theorem (Theorem A.3) these measures are uniformly μ-continuous. Hence by Lemma A.4 the measures

$$|\pi_n|(A) := \int_A |T^n f| \, \mathrm{d}\mu \qquad A \in \Sigma$$

are also uniformly μ-continuous, meaning that for $A_i \searrow \emptyset$

$$\lim_{i\to\infty} \sup_n |\pi_n|(A_i) = 0.$$

Since the sup dominates the Banach limit we have

$$\lim_{i\to\infty} \pi(\mathbf{1}_{A_i}) = \lim_{i\to\infty} L_{n\to\infty}|\pi_n|(A_i) \leqslant \lim_{i\to\infty} \sup_n |\pi_n|(A_i) = 0,$$

and hence π is countably additive.

By the Radon-Nikodym theorem any countably additive measure can be identified by a function in L^1. Since the measure is positive there must exist a function $p \in L^1_+$ such that $\pi(h) = \int p \cdot h \, \mathrm{d}\mu$ for all $h \in L^\infty$. We know p is non-zero because $\pi(\mathbf{1}) = \alpha$. Now, for all $h \in L^\infty$ we have (noting $\pi(h) = \int p \cdot h \, \mathrm{d}\mu = L_{n\to\infty}(\int |T^n f| \cdot h \, \mathrm{d}\mu)$)

$$\begin{aligned}
\int_\Omega |T|p \cdot h \, \mathrm{d}\mu &= \int_\Omega p \cdot |T|^* \, h \, \mathrm{d}\mu \\
&= L_{n\to\infty} \int_\Omega |T^n f| \cdot |T|^* h \, \mathrm{d}\mu \\
&= L_{n\to\infty} \int_\Omega |T| \, |T^n f| \cdot h \, \mathrm{d}\mu \\
&\geqslant L_{n\to\infty} \int_\Omega |T^{n+1} f| \cdot h \, \mathrm{d}\mu \\
&= \int_\Omega p \cdot h \, \mathrm{d}\mu
\end{aligned}$$

This implies $|T|p \geqslant p$ but since $|T|$ is a contraction, we must have $|T|p = p$. □

The second lemma builds on the first showing that the separation shown in the previous lemma can actually be used to split the space Ω in two: on one space all orbits go to zero and on the other space there is a fixed point with a maximal support.

Lemma 3.9. *A contraction T on $L^1(\Omega, \Sigma, \mu)$ decomposes the space Ω into two disjoint sets F and G where $F \cup G = \Omega$ such that:*

1. *If $f \in L^1$ and $(T^n f)_n$ converges weakly then $(T^n f)_n$ converges strongly to zero on F, i.e.*

$$\lim_{n\to\infty} \int_F |T^n f| \, d\mu = 0$$

2. *There exists an $r \in L^1_+$ such that $|T|r = r$ and the support of r is G.*

Proof. Let $Q = \{q : q \in L^1_+, |T|q = q\}$, the set of positive fixed points of $|T|$.

- If Q is empty then we set $r = \mathbf{0}$ and G is empty.
- Otherwise consider the set $\tilde{Q} = \{\mathbf{1}_{\mathrm{supp}(q)} : q \in Q\}$. This is a $\vee$-stable set bounded by 1 and so by Theorem A.2 has a supremum $\tilde{q}$ and we can find a sequence $(\tilde{q}_n) \subset \tilde{Q}$ such that $\sup_n \tilde{q}_n = \tilde{q}$. Mapping this back to Q gives us a countable sequence in $(q_n)_n \in Q$.

  Without loss of generality we may assume that $\sup_n \|q_n\|$ is uniformly bounded. Now let $r = \sum_{n=0}^{\infty} q_n \cdot 2^{-n}$ and so $|T|r = r$ and $G = \mathrm{supp}(r) = \bigcup \mathrm{supp}(q_n) = \mathrm{supp}(\tilde{q})$.

In any case we have an $r \in Q$ such that $\mathrm{supp}(q) \subseteq \mathrm{supp}(r)$ for all $q \in Q$.

G is an invariant set, i.e. $TL^1(G, \mu) \subseteq L^1(G, \mu)$. To see this consider an $f \in L^1_+(G, \mu)$ and let $C > 0$. Now let $f_1 = Cr \wedge f$ and $f_2 = f - f_1$. Choose an $\varepsilon > 0$ and find a C such that $\|f_2\| < \varepsilon$. Now $f_1 \leqslant Cr$ and so $|T|f_1 \leqslant Cr$ and hence $|T|f_1 \in L^1(G, \mu)$. Since ε was arbitrary, $|T|f \in L^1(G, \mu)$ and so also $Tf \in L^1(G, \mu)$, as $|T|$ and T have the same invariant sets.

Now define a contraction $R : L^1(F, \mu) \to L^1(F, \mu)$ as $R := T|_F$. Clearly $R^n = T^n|_F$ and if $T^n f$ converges weakly then so does $R^n f$. Hence by Lemma 3.8 either $\lim_{n\to\infty} \|R^n f\| = 0$ or there is a non-zero $p \in L^1_+(F, \mu)$ such that $|R|p = p$. But $\mathrm{supp}(p)$ must be in G and therefore cannot be in F, so the second case is impossible. ☐

The final lemma actually shows the Blum-Hanson equivalence for Dunford-Schwartz operators, but only for the operator as a whole, not for individual orbits. Because we know the operator is bounded in both the L^1 and the L^∞ norm, we can use the previously proved Hilbert space case for L^2 and density arguments to move to other L^p spaces.

Lemma 3.10. *Let (Ω, Σ, μ) be a finite measure space and let T be an operator that is a contraction in both L^1 and L^∞. Then the following are equivalent:*

1. *There exists a number $1 \leqslant p_0 < \infty$ such that if $f \in L^{p_0}$ then $(T^n f)_n$ converges weakly in L^{p_0}.*
2. *For all $1 \leqslant p < \infty$ and $f \in L^p$ then $(T^n f)_n$ converges weakly in L^p.*
3. *For all $1 \leqslant p < \infty$ and $f \in L^p$ then $\frac{1}{N}\sum_{n=1}^{N} T^{k_n} f$ converges strongly in L^p for every increasing sequence of positive integers $(k_n)_{n=1}^{\infty}$.*

Proof. (1)→(2): Since the measure space is finite, L^∞ is a subset of each $L^p, 1 \leqslant p \leqslant \infty$. Fix an $\varepsilon > 0$ and a $1 \leqslant p \leqslant \infty$. Let q be the adjoint index. Let $f \in L^p, g \in L^q$ be arbitrary. Then $f \in L^\infty$ and $g \in L^\infty$. As L^∞ is also a dense subset of L^{p_0} we can find an $f' \in L^{p_0}, g' \in L^{q_0}$ such that $\|f - f'\|_{p_0} < \varepsilon$ and $\|g - g'\|_{q_0} < \varepsilon$. Now

$$\begin{aligned}
|\langle T^n f, g\rangle - \langle T^n f', g'\rangle| &\leqslant |\langle T^n (f - f'), g\rangle| + |\langle T^n f', g - g'\rangle| \\
&\leqslant \|T^n\|_{p_0} \underbrace{\|f - f'\|_{p_0}}_{<\varepsilon} \|g\|_{q_0} + \|T^n\|_{p_0} \|f'\|_{p_0} \underbrace{\|g - g'\|_{q_0}}_{<\varepsilon} \\
&\leqslant \varepsilon(\|g\|_{q_0} + \|f'\|_{p_0})
\end{aligned}$$

Since we chose the ε, whenever $(T^n)_n$ converges weakly in L^{p_0}, it converges weakly for all $L^p, 1 \leqslant p < \infty$, completing (1)→(2).

(2)→(3): Let p, $1 \leqslant p < \infty$ be fixed. Note that it is sufficient to prove the strong convergence of $\frac{1}{n}\sum_{i=1}^n T^{k_i} f$ for f in a dense subset of L^p. As noted above, because we are on a finite measure space L^∞ is dense in L^p. But if $f \in L^\infty$ then $f \in L^2$. Then the sequence $f_n = \frac{1}{n}\sum_{i=1}^n T^{k_i} f$ converges strongly in L^2 by Theorem 3.3 (Blum-Hanson equivalence for Hilbert spaces). Hence we have the Cauchy sequence

$$\lim_{n,m\to\infty} \|f_n - f_m\|_2 = 0.$$

Noting that (for $m > n$)

$$\begin{aligned}\|f_n - f_m\|_\infty &\leqslant \|f_n\|_\infty + \|f_m\|_\infty \\ &\leqslant 2\|f\|_\infty\end{aligned}$$

we have

$$\||f_n - f_m|^{p-1}\|_2 \leqslant 2\|f\|_\infty^{p-1}\|\mathbf{1}\|_2 = 2\|f\|_\infty^{p-1}\sqrt{\mu(\Omega)}.$$

Then the Cauchy-Schwartz inequality shows

$$\begin{aligned}\|f_n - f_m\|_p^p &= \langle |f_n - f_m|, |f_n - f_m|^{p-1}\rangle \\ &\leqslant \|f_n - f_m\|_2 \cdot \||f_n - f_m|^{p-1}\| \\ &\leqslant \|f_n - f_m\|_2 \cdot 2\|f\|_\infty^{p-1}\sqrt{\mu(\Omega)}\end{aligned}$$

so that $(f_n)_n$ is also a Cauchy sequence in L^p. Hence $\frac{1}{n}\sum_{i=1}^n T^{k_i} f$ converges strongly in L^p for each f.

(3)→(1): Apply Lemma 2.17. □

Now the completion of the main proof: showing Blum-Hanson equivalence for arbitrary contractions in L^1. The proof proceeds in two steps. First we show that the equivalence applies on G by using the previous lemma. Finally it is shown that the addition of functions on F do not affect the result. Note, the equivalence is only for the operator as a whole, not for each orbit.

Proof of Theorem 3.5. (2)→(1): Apply Lemma 2.17.

(1)→(2): Let $F \cup G = \Omega$ be the decomposition as given in Lemma 3.9 and let $r \in L^1_+$ be the function with support G such that $|T|r = r$. Note $r \geqslant 0$. If $r = \mathbf{0}$ then G is empty and the result follows directly from Lemma 3.9. So we may assume that $r \neq \mathbf{0}$.

Now consider the finite measure λ defined as $\mathrm{d}\lambda = r\,\mathrm{d}\mu$, finite because r is integrable. Then $\varphi \in L^1(\Omega, \Sigma, \lambda)$ if and only if $r\varphi \in L^1(\Omega, \Sigma, \mu)$. Define the operator $S := \frac{1}{r}Tr$. From

$$\int_\Omega |S\varphi|\,\mathrm{d}\lambda = \int_\Omega |Tr\varphi|\,\mathrm{d}\mu \leqslant \int_\Omega |r\varphi|\,\mathrm{d}\mu = \int_\Omega |\varphi|\,\mathrm{d}\lambda$$

we see that S is a contraction on $L^1(\Omega, \Sigma, \lambda)$. Let $f \in L^\infty(\Omega, \Sigma, \lambda)$ be an arbitrary bounded function. Then

$$|Sf| = \frac{1}{r}|T(rf)| \leqslant \frac{1}{r}|T|(r|f|) \leqslant \frac{1}{r}|T|(r\|f\|_\infty \cdot \mathbf{1}) = \|f\|_\infty \frac{1}{r}|T|r = \|f\|_\infty \mathbf{1}$$

So $\|Sf\|_\infty \leqslant \|f\|_\infty$ and we see that S is also a contraction on $L^\infty(\Omega, \Sigma, \lambda)$. Hence S is a Dunford-Schwartz operator.

Now, for each $h \in L^\infty(\Omega, \Sigma, \lambda)$ we have that

$$\int_\Omega S^n\varphi \cdot \overline{h}\,\mathrm{d}\lambda = \int_\Omega T^n r\varphi \cdot \overline{h}\,\mathrm{d}\mu$$

so if $(T^n)_n$ converges weakly, so does $(S^n)_n$. Since λ is a finite measure, Lemma 3.10 applies and $\frac{1}{n}\sum_{i=1}^n S^{k_i}\varphi$ converges strongly in $L^1(G,\Sigma,\lambda)$ for each $\varphi \in L^1(G,\Sigma,\lambda)$. This shows that

$$\frac{r}{n}\sum_{i=1}^n S^{k_i}\varphi = \frac{1}{n}\sum_{i=1}^n T^{k_i} r\varphi$$

is a Cauchy sequence in $L^1(G,\Sigma,\mu)$ and hence also in $L^1(\Omega,\Sigma,\mu)$. Since any $f \in L^1(G,\Sigma,\mu)$ can be written as $r\varphi$ with $\varphi \in L^1(G,\Sigma,\lambda)$ this proves (1)→(2) for $f \in L^1(G,\Sigma,\mu)$.

To prove (1)→(2) for general $f \in L^1(\Omega,\Sigma,\mu)$ it is sufficient to show that given an $\varepsilon > 0$ there is a Cauchy sequence $g_n \in L^1(\Omega,\Sigma,\mu)$ such that

$$\limsup_{n\to\infty}\left\|\frac{1}{n}\sum_{i=1}^n T^{k_i}f - g_n\right\| < \varepsilon$$

So choose an $f \in L^1(\Omega,\Sigma,\mu)$ and the above $\varepsilon > 0$ and choose an N such that

$$\int_F |T^{k_N}f|\,\mathrm{d}\mu < \varepsilon$$

which is possible because on F the functions go strongly to zero.

Let $g := \mathbf{1}_G T^{k_N}f$ and $g_n := \frac{1}{n}\sum_{i=1}^n T^{k_{N+i}-k_N}g$. Note that all the g_n are in G and so they form a Cauchy sequence in $L^1(G,\Sigma,\mu)$. So

$$\begin{aligned} g_n &= \frac{1}{n}\sum_{i=1}^n T^{k_{N+i}-k_N}(\mathbf{1}_G T^{k_N}f) \\ &= \frac{1}{n}\sum_{i=1}^n T^{k_{N+i}-k_N}(T^{k_N}f - \mathbf{1}_F T^{k_N}f) \\ &= \frac{1}{n}\sum_{i=1}^n T^{k_{N+i}}f - \frac{1}{n}\sum_{i=1}^n T^{k_{N+i}-k_N}\underbrace{(\mathbf{1}_F T^{k_N}f)}_{\|\cdot\|<\varepsilon} \end{aligned}$$

Since the first N terms of the sum do not affect the limit this shows that $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n T^{k_i}f = \lim_{n\to\infty} g_n$ which completes the proof of (1)→(2). □

As you can see, the proof here is very different from the Hilbert space case. As with all the other positive results so far, it relies on special inequalities for the space it is operating on.

3.4 Blum-Hanson on ℓ^p contractions

The most recent development has been the proof of the individual Blum-Hanson equivalence for general ℓ^p contractions, $1 \leqslant p < \infty$ by Müller-Tomilov in [MT07]. Although it does require special properties of the ℓ^p spaces, it is actually quite short and fairly straightforward.

Theorem 3.11. *(Müller-Tomilov 2007) Let $1 \leqslant p < \infty$, $T \in BL(\ell^p)$ be a contraction and $x \in \ell^p$. Suppose that the sequence $(T^n x)_n$ is weakly convergent and $(n_i)_i$ is an increasing sequence of positive integers. Then the limit*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N T^{n_i}x$$

exists in the norm topology.

Proof. Without loss of generality we may assume $T^n x \to 0$ weakly. For $p = 1$ the statement is obvious since for ℓ^1 the weak convergence of sequences is the same as convergence in the norm topology (this is the Schur property, for more details see [Meg98, Chapter 2]).

Let $1 < p < \infty$ and $(e_j)_j$ be the standard basis vectors for ℓ^p. Let P_r be the canonical projection onto $\operatorname{span}\{e_1, \dots, e_r\}$.

Since T is a contraction the limit $\lim_{n\to\infty} \|T^n x\|$ exists. Since the statement is clear if the limit is zero, without loss of generality we may assume that the limit is equal to 1.

Choose a $\delta > 0$. Find a positive integer t such that $t^{-(1-\frac{1}{p})} < \delta/2$. Now there exists an $\varepsilon \in (0, 1)$ such that

$$((1+\varepsilon)^p + 2^p s)^{1/p} < 2(s+1)^{1/p} - (s+1)\varepsilon$$

for $s = 1, \dots, t-1$. To see why this is true, note that substituting $\varepsilon = 0$ gives the inequality $1 + 2^p s < 2^p(s+1)$ which is true for all $s \in \mathbb{N}$.

Now, find a k such that $\|T^k x\| < 1+\varepsilon$. Because the elements of a member of ℓ^p eventually become small we can find an r such that $\|(I - P_r)T^k x\| < \varepsilon$. Finally, as on a finite dimensional space weak convergence is equivalent to pointwise convergence we can find a d such that $\|P_r T^{k+j} x\| < \varepsilon$ for all $j \geqslant d$.

Note this is the point were this proof becomes specific for ℓ^p. While it is possible to create a sequence of compact projections on a general separable L^p space which converge to the identity, it is not possible to also have the ranges of P_r and $I - P_r$ live on disjoint sets, which we will need below.

We now use induction over $1 \leqslant s \leqslant t$ to show that

$$\left\|\sum_{i=1}^{s} T^{m_i} x\right\| \leqslant 2s^{1/p}$$

whenever $k \leqslant m_1 < m_2 < \dots < m_s$, $s \leqslant t$ and $m_{i+1} - m_i \geqslant d$ for all i.

For $s = 1$ the statement is obviously true. Now suppose $1 \leqslant s < t$ and $m_1, \dots, m_{s+1}$ satisfy the required conditions. Now

$$\begin{aligned}
\left\|\sum_{i=1}^{s+1} T^{m_i} x\right\| &\leqslant \left\|T^{m_1-k} \sum_{i=1}^{s+1} T^{m_i-m_1+k} x\right\| \\
&\leqslant \left\|\sum_{i=1}^{s+1} T^{m_i-m_1+k} x\right\| \\
&= \left\|(P_r + (I-P_r))(T^k x + \sum_{i=2}^{s+1} T^{m_i-m_1+k} x)\right\| \\
&\leqslant \left\|P_r T^k x + (I-P_r)\sum_{i=2}^{s+1} T^{m_i-m_1+k} x\right\| + \underbrace{\|(I-P_r)T^k x\|}_{<\varepsilon} + \underbrace{\left\|P_r \sum_{i=2}^{s+1} T^{m_i-m_1+k} x\right\|}_{<s\varepsilon} \\
&\leqslant \left(\underbrace{\|P_r T^k x\|^p}_{<1+\varepsilon} + \underbrace{\left\|(I-P_r)\sum_{i=2}^{s+1} T^{m_i-m_1+k} x\right\|^p}_{<2s^{1/p} \text{ by induction}}\right)^{\frac{1}{p}} + (s+1)\varepsilon \qquad (3.1) \\
&\leqslant ((1+\varepsilon)^p + 2^p s)^{1/p} + (s+1)\varepsilon \\
&\leqslant 2(s+1)^{1/p}
\end{aligned}$$

proves the statement for $s+1$ and completes the induction.

Note Equation 3.1, where we explicitly require that $\|P_r x + (I-P_r)y\|^p \leqslant \|P_r x\|^p + \|(I-P_r)y\|^p$.

What we have shown so far is that the growth of the sums of up to t elements of T^{m_i} can be bounded at $O(t^{1/p})$ if the sequence is sparse enough.

Now let $(n_i)_i$ be an arbitrary increasing sequence of positive integers and choose an $N > k+dt$. Write $N = k+mt+r$ where $1 \leqslant r \leqslant t$ and $m \geqslant d$ is a positive integer. Then we can take the first $k+r$ elements as fixed and divide the remaining elements into m subsequences of length t. In each of these subsequences the distance between each successive element is greater than d as $n_j - n_i \geqslant j-i$, so we can apply the results of the above induction.

$$\begin{aligned}\left\|\sum_{i=1}^{N} T^{n_i}x\right\| &\leqslant \left\|\sum_{i=1}^{k+r} T^{n_i}x\right\| + \sum_{s=1}^{m}\left\|\sum_{i=0}^{t-1} T^{n_{k+r+s+im}}x\right\| \\ &\leqslant (k+r)\|x\| + m\cdot 2t^{1/p}\end{aligned}$$

Using the bounds we have on the subsequences we can bound the Cesàro sums.

$$\begin{aligned}\frac{1}{N}\left\|\sum_{i=1}^{N} T^{n_i}x\right\| &\leqslant \frac{(k+r)\|x\|}{N} + \frac{m\cdot 2t^{1/p}}{N} \\ &\leqslant \frac{(k+r)\|x\|}{N} + \frac{m\cdot 2t^{1/p}}{mt} \\ &= \frac{(k+r)\|x\|}{N} + 2t^{1/p-1}\end{aligned}$$

We send $N \to \infty$ so the initial terms vanish.

$$\limsup_{N\to\infty} \frac{1}{N}\left\|\sum_{i=1}^{N} T^{n_i}x\right\| \leqslant 2t^{1/p-1} < \delta.$$

And since $\delta > 0$ was arbitrary we have

$$\lim_{N\to\infty} \frac{1}{N}\left\|\sum_{i=1}^{N} T^{n_i}x\right\| = 0.$$

□

The proof for ℓ^p is interesting in that it does not require very many special properties of the ℓ^p, in fact just the one noted above. The rest of the proof just uses properties of norms which are true for all Banach spaces and finally proves an individual Blum-Hanson equivalence.

This completes the presentation of the positive results for the Blum-Hanson property. The remaining results are examples and counterexamples produced over the years to help determine what further results may or may not be possible. First however we take a detour to prove some general results that will help in the analysis of the later examples.

Chapter 4

Generalised shift operators

Before proceeding to the examples and counterexamples, it is useful to have a general framework in which to study some of them. Many of the examples given are of a very specific form, which makes them easy to construct and analyse. This chapter will introduce some useful lemmas which will make our work in the next chapter much easier.

One of the most often studied operators in operator theory is the shift operator. There are two basic types, the left shift and the right shift, traditionally indicated by L and R. For this thesis we will work with a generalisation of these operators.

First we need to introduce something from another branch of mathematics, graph theory. There graphs are represented by a triplet $G = (V, E, w)$. Here V is the set of vertices, E is the set of edges (v, v') and $w : E \to \mathbb{R}$ is the weight function, which indicates the weight assigned to each edge. Here we use the convention $w(v, v') = 0$ for an edge that does not exist in the graph.

Note that the set of edges E is in fact redundant, since we could define the edges in terms of the weight function.

$$E := \{(v, v') : w(v, v') \neq 0\}.$$

Nevertheless, it is sometimes useful to be able to refer to the set of edges to simplify explanations.

We are particularly interested in directed weighted acyclic graphs.

- directed - each edge has a direction.
- weighted - each edge has a weight.
- acyclic - the graph has no loops

To see how an operator might be defined by a graph, note we can construct a real vector space using the vertices to represent the basis vectors. Using the weight function we can define a mapping from a basis vector to a linear combination of basis vectors. Formally:

$$Tv := \sum_{v' \in V} w(v, v')v'.$$

However, this sum is not necessarily well defined if there are infinitely many edges entering or leaving a node. So we will require the following assumption.

Remark 4.1. For the purposes of this thesis we will assume that there are only a finite number of edges leaving a cell and a finite number of edges entering a cell.

With this assumption we can use this graph to define an operator $T : c_{00}(V) \to c_{00}(V)$. Here we use $c_{00}(V)$ to represent the sequences with finitely many non-zero entries. Now the operator T can be defined on $c_{00}(V)$ as follows.

$$Tx := \sum_{v \in V} x_v \sum_{(v,v') \in E} w(v,v')v'$$

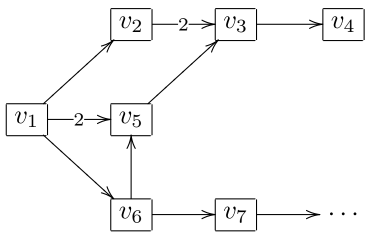
This is useful basis, as $c_{00}(V)$ is dense in $\ell^p(V), 1 \leqslant p < \infty$.

Definition 4.2 (Generalised shift operator)**.** A generalised shift operator $T : c_{00}(V) \to c_{00}(V)$ is an operator represented by a directed weighted acyclic graph $G = (V, E, w)$ and is defined as the linear extension of

$$Tv := \sum_{v' \in V} w(v,v')v'.$$

Most compact operators on ℓ^p could be represented as a graph, though only rarely as an acyclic graph. The assumption in Remark 4.1 reduces the scope even more. Nevertheless, this relatively small class of operators is very interesting class as we shall see.

Example 4.3. Consider the following graph:



These diagrams will appear regularly in this thesis, so there are a few conventions (from graph theory) to pay attention to:

- The boxes will be referred to as *cells* or *nodes* and the joins between them as *arcs* or *edges*.
- Where no weight is given (as is the case with most arcs) the weight is assumed to be 1.
- Cell v_1 is referred to as a *source* and cell v_4 is referred to as a *sink*.
- The use of $(\cdots)$ indicates a (possibly infinite) number of cells attached in sequence, all with weight 1. When the number of elements is infinite we may refer to it as an *infinite source* or *infinite sink*, as appropriate.
- Cells v_2, v_5 and v_6 are referred to as the *successors* of v_1.
- Cells v_2 and v_5 are the *predecessors* of v_3.

For convenience we define two operators [1]

- $\text{pred} : V \to \mathscr{P}(V)$
$$\text{pred}(v') := \{v : (v,v') \in E\}$$
- $\text{succ} : V \to \mathscr{P}(V)$
$$\text{succ}(v) := \{v' : (v,v') \in E\}$$

which define the predecessors and successors of a node. We can usefully extend these notions to sets of vertices.

[1] $\mathscr{P}(V)$ is the power set of V.

- $\operatorname{pred} : \mathscr{P}(V) \to \mathscr{P}(V)$
$$\operatorname{pred}(W) := \bigcup_{v \in W} \operatorname{pred}(v)$$
- $\operatorname{succ} : \mathscr{P}(V) \to \mathscr{P}(V)$
$$\operatorname{succ}(W) := \bigcup_{v \in W} \operatorname{succ}(v)$$

Remark 4.1 guarantees that $\operatorname{pred}(v)$ and $\operatorname{succ}(v)$ are finite sets for all $v \in V$. The sets also depend on the operator T. Where there are multiple operators used, the appropriate one will be indicated by a subscript, for example pred_T.

Additionally, the fact that the graph has no cycles means that $\operatorname{succ}^n(v)$ cannot contain v for any $n \geqslant 1$. This is used often in the proofs in this chapter.

So Example 4.3 now represents the operator $T : c_{00}(V) \to c_{00}(V)$ defined as:

$$\begin{array}{lll} Tv_0 := v_1 + 2v_4 + v_5 & Tv_1 := 2v_2 & Tv_3 := 0 \\ Tv_4 := v_3 & Tv_5 := v_4 + v_6 & Tv_n := v_{n+1} \quad (n = 6 \dots \infty) \end{array}$$

As we will see the properties of this operator are strongly related to the structure of the graph and only weakly related to the space on which it operates. Indeed, much of this thesis is studying the various properties of T and the range of spaces on which it is well defined and how that relates to the various properties we are interested in.

Lemma 4.4. *Under the assumption in Remark 4.1 that only finitely many edges leave any node, the operator $T : c_{00}(V) \to c_{00}(V)$ is well defined.*

Proof. For any $v_i \in V$ we know that Tv_i has only finitely many non-zero entries since a node in a graph may have only finitely many edges leaving it. Since an element $x \in c_{00}(V)$ has only finitely many non-zero elements it follows that $Tx \in c_{00}(V)$ for all x. □

It is also worth noting that if all the entries of $w(v, v')$ are positive, then the resulting operator T is also positive.

4.1 The adjoint

Something that will be very useful in the exploration of these operators is the adjoint of the operator T.

Lemma 4.5. *For a given graph $G = (V, E, w)$ with its associated operator $T : c_{00}(V) \to c_{00}(V)$, there is an adjoint operator $T^* : c_{00}(V) \to c_{00}(V)$ associated with a graph $G^* = (V, E^*, w^*)$ where $E^* = \{(i, j) : (j, i) \in E\}$ and $w^*(v, v') = w(v', v)$.*

Proof. First we note that from the above definition we have that

$$\operatorname{pred}_T(v) = \operatorname{succ}_{T^*}(v) \text{ and } \operatorname{succ}_T(v) = \operatorname{pred}_{T^*}(v)$$

hence the graph satisfies the assumption in Remark 4.1 and so by Lemma 4.4 the operator T^* is a well-defined operator on $c_{00}(V)$.

Now define a formal inner product:

$$\langle x, y \rangle = \sum_{v \in V} x_v y_v, \qquad x, y \in c_{00}(V)$$

Since both x and y have only finitely many non-zero entries, the above sum is well defined. So

$$\begin{aligned}
\langle Tx, y\rangle &= \sum_{v\in V}(Tx)_v y_v \\
&= \sum_{v\in V}\left(\sum_{(v',v)\in E} x_{v'}w(v',v)\right) y_v \\
&= \sum_{v\in V}\sum_{(v,v')\in E^*} x_{v'}w^*(v,v')y_v \\
&= \sum_{v'\in V} x_{v'}\left(\sum_{(v,v')\in E^*} y_v w^*(v,v')\right) \\
&= \langle x, T^*y\rangle
\end{aligned}$$

which shows that T^* is a adjoint of T. □

Simply put, the adjoint is the operator associated with the same graph, but with the arrows in the opposite direction and the same weights.

4.2 The norm

We now have an operator T defined on the algebra $c_{00}(V)$ but the question arises: does T also define a bounded operator on ℓ^p?

Theorem 4.6. *Any graph $G = (V, E, w)$ which meets the conditions*

$$\sup_{j\in V}\sum_{i\in V}|w(i,j)| \leqslant M \qquad\qquad \sup_{i\in V}\sum_{j\in V}|w(i,j)| \leqslant M$$

defines a bounded operator from $T : \ell^p(V) \to \ell^p(V)$ for all $1 \leqslant p \leqslant \infty$ and $\|T\|_p \leqslant M$.

Proof. Consider the 1-norm:

$$\begin{aligned}
\|T\|_{1\to 1} &= \sup_{\|x\|_1=1}\sum_j\left|\sum_i x_i w(i,j)\right| \leqslant \sup_{\|x\|_1=1}\sum_j\sum_i |x_i|\,|w(i,j)| \\
&= \sup_{\|x\|_1=1}\sum_i |x_i|\sum_j |w(i,j)| = \sup_i\sum_j |w(i,j)| \\
&\leqslant M
\end{aligned}$$

Now, consider the ∞-norm:

$$\begin{aligned}
\|T\|_{\infty\to\infty} &= \sup_{\|x\|_\infty=1}\sup_j\left|\sum_i x_i w(i,j)\right| \leqslant \sup_{\|x\|_\infty=1}\sup_j\sum_i |x_i|\,|w(i,j)| \\
&= \sup_j\sum_i |w(i,j)| \leqslant M
\end{aligned}$$

Now by Riesz-Thorin we see that T defines a bounded operator from $T : \ell^p(V) \to \ell^p(V)$ for all $1 \leqslant p \leqslant \infty$. □

Translating the result back to the original graph, the maximum absolute row sum is merely the maximum absolute sum of the weights *leaving* a cell. The maximum absolute column sum is the maximum absolute sum of the weights *entering* a cell. In particular:

Corollary 4.7.

$$\|T\|_1 = \sup_{v \in V} \sum_{v' \in \operatorname{succ}(v)} |w(v, v')| = \sup_{v \in V} \|Tv\|_1$$

and

$$\|T\|_\infty = \sup_{v' \in V} \sum_{v \in \operatorname{pred}(v)} |w(v, v')| = \sup_{v \in V} \|Tv\|_\infty.$$

So for Example 4.3 we can clearly see that $\|T\|_1 = 4$ (determined by e_0) and $\|T\|_\infty = 3$ (determined by e_2 and e_4).

Corollary 4.8. *For a graph $G = (V, E, w)$ with associated operator T we have:*

$$\|T\|_1 = \|T^*\|_\infty \text{ and } \|T\|_\infty = \|T^*\|_1$$

Proof. The adjoint swaps the roles of $\operatorname{pred}(v)$ and $\operatorname{succ}(v)$. The result follows. □

These are useful theorems, which will be used regularly during the rest of this thesis. In some cases we would like to be able to determine the exact norm, rather than just determining that it is bounded. For that we can use the following lemma.

Lemma 4.9. *For any graph $G = (V, E, w)$ with associated operator $T : c_{00}(V) \to c_{00}(V)$ we can partition V minimally in two separate ways into the sets $\{I_k\}_k, \{J_k\}_k, \quad I_k, J_k \subseteq V$ such that:* [2]

$$\bigsqcup I_k = \bigsqcup J_k = V$$

$$\operatorname{succ} I_0 = \operatorname{pred} J_0 = \emptyset$$

$$\operatorname{succ} I_k = J_k, \qquad \operatorname{pred} J_k = I_k, \qquad \forall k > 0$$

The partition is minimal in the sense that it is not possible to make a finer partition satisfying these properties. Let P_{I_k} be the projection onto the space $c_{00}(I_k)$. Then

$$T = \bigoplus_k (TP_{I_k}) \text{ and } T^* = \bigoplus_k (T^* P_{J_k})$$

Proof. Consider the relation between two elements of V defined as

$$a \sim b \iff \operatorname{succ}(a) \cap \operatorname{succ}(b) \neq \emptyset \text{ or } \operatorname{succ}(a) = \operatorname{succ}(b) = \emptyset,$$

that is to say a and b are related if they share a successor or both have no successors. This relation is reflexive and symmetric. The transitive closure of this relation defines a partition of V which is the minimal partition meeting the necessary requirements ([Wec92]). Let $\{I_k\}_k$ be this partition, where I_0 is the part where $\operatorname{succ}(I_0) = \emptyset$. For all $\{I_k\}_{k>0}$ we have that $\operatorname{pred}(\operatorname{succ}(I_k)) = I_k$.

By using the equivalence

$$a \approx b \iff \operatorname{pred}(a) \cap \operatorname{pred}(b) \neq \emptyset \text{ or } \operatorname{pred}(a) = \operatorname{pred}(b) = \emptyset,$$

we obtain similarly a partition $\{J_k\}_k$.

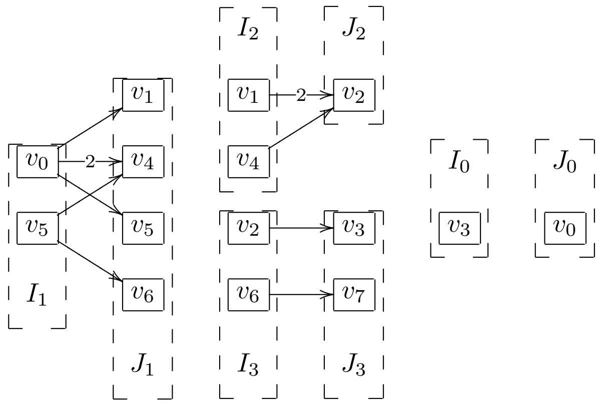
To show that the $\{I_k\}_{k>0}$ and $\{J_k\}_{k>0}$ can be mapped one-to-one, consider two elements $a, b \in I_k$ such that $a \sim b$. Then the set $C = \operatorname{succ}(a) \cap \operatorname{succ}(b)$ is non-empty. Choose a $c' \in C$. Now for each $a' \in \operatorname{succ}(a)$ and $b' \in \operatorname{succ}(b)$ we have that $a' \approx c' \approx b'$. Hence $\operatorname{succ}(I_k)$ is transitively closed under $\approx$ and hence must be a subset of some J_k. Similarly the $\operatorname{pred}(J_k)$ can be mapped onto some subset of I_k and hence there is a one-to-one mapping between the $\{I_k\}_{k>0}$ and $\{J_k\}_{k>0}$.

It remains to show that these sets can be used to split the operator T into disjoint operators. But this is clear, as $Tc_{00}(I_k) = c_{00}(J_k)$ by definition and the spaces $c_{00}(J_k)$ are disjoint. □

[2] $\bigsqcup$ is the exclusive (or disjoint) union

Once the operator is partitioned this way we can see that $c_{00}(I_0)$ is the kernel of T, the set of sinks and $c_{00}(J_0)$ the kernel of T^*, the set of sources. For the adjoint operator the $\{I_k\}_k$ and $\{J_k\}_k$ are simply swapped.

Using the above we can split the original operator into the direct sum of simpler operators. To see how this works, here is what we get when this process is applied to the Example 4.3.



Plus infinitely more trivial cases for $\{v_n\}_{n=7}^{\infty}$. Here you can clearly see that the original operator T has been split into a number of simpler operators. All the sources are in J_0 and all the sinks in I_0, as consistent with their definition.

Now we can prove the following theorem.

Lemma 4.10. *Given a decomposition as described in Lemma 4.9 the norm can be calculated as:*

$$\|T\|_p = \sup_k \|TP_{I_k}\|_p$$

Proof. Since T maps the I_k into the J_k which are all disjoint, the result is straightforward. □

The above proof provides the necessary information to calculate the 2-norm of an operator exactly, not just an upper bound because the decomposition produces operators which are simpler, often finite dimensional operators where the 2-norm is easy to calculate.

We now have easy criteria to determine if a particular operator T determined by a graph is bounded in the given space, just by looking at the graph. We will see that we can prove similar criteria for the other types of boundedness.

4.3 Power-boundedness

The first thing to note is that if T is a generalised shift operator, then T^n is also one. From the point of view of the graph it just means you take n steps instead of 1. The most important lesson from Corollary 4.7 is that to determine the ℓ^1 norm or the ℓ^∞ norm we only need to look at individual basis vectors. In particular we have:

Lemma 4.11. *For a graph $G = (V, E, w)$ with associated operator T we have*

$$\sup_{n\in\mathbb{N}} \|T^n\|_1 = \sup_{v\in V} \sup_{n\in\mathbb{N}} \|T^n v\|_1$$

$$\sup_{n\in\mathbb{N}} \|T^n\|_\infty = \sup_{v\in V} \sup_{n\in\mathbb{N}} \|T^{*n} v\|_1$$

Proof. The first directly follows from the application of Corollary 4.7 to the graph associated with the operator T^n. The second follows with the application of Corollary 4.8. ☐

However, it is usually less work than that.

Lemma 4.12. *Consider a graph $G = (V, E, w)$ with associated operator T. Suppose T is positive ($w \geqslant 0$). Let $\lvert\!\lvert\!\lvert \cdot \rvert\!\rvert\!\rvert$ be the power-bounded norm from Definition 2.2 generated by T. Define the sets $K_1, K_2, K_3 \subseteq V$ as:*

$$K_1 := \{v \in V : \exists v' \in V, 0 < |w(v', v)| < 1\}$$

$$K_2 := \{v \in V : \mathrm{pred}(v) = \emptyset\}$$

and let $K_3 \subseteq V$ be a set such that

$$\bigcup_{n \in \mathbb{N}} \mathrm{succ}^n(K_3) = V.$$

Then

$$\sup_n \|T^n\|_1 = \sup_{v \in K_1 \cup K_2 \cup K_3} \lvert\!\lvert\!\lvert v \rvert\!\rvert\!\rvert_1$$

Here we have that:

- K_1 is the set of nodes where a predecessor node has a weight less than one.
- K_2 is the set of sources.
- K_3 is not a fixed set of nodes like K_1 and K_2. This set may be chosen at will and constitutes a representative subset of V for the purposes of this lemma. Clearly you would like to choose this set to be as small as possible.

Proof. Suppose it is not true. Let $m \in V$ be a vertex not in $K_1 \cup K_2 \cup K_3$. Consider the set of predecessors of m. This set cannot be empty because $m \notin K_2$. For each $v' \in \mathrm{pred}(m)$ we have

$$Tv' = \sum_{v \in V} w(v', v)v \geqslant w(v', m)m$$

in the order of ℓ^p because $w(v', v) \geqslant 0$. Since T is positive it is order preserving so we have

$$T^{n+1}v' = T^n \sum_{v \in V} w(v', v)v \geqslant w(v', m)T^n m$$

and finally by the monotonicity of the norm we have

$$\|T^{n+1}v'\| \geqslant \|T^n \sum_{v \in V} w(v', v)v\| \geqslant |w(v', m)|\|T^n m\|.$$

We know $|w(v', m)| \geqslant 1$ because $m \notin K_1$, so $\lvert\!\lvert\!\lvert v' \rvert\!\rvert\!\rvert \geqslant \lvert\!\lvert\!\lvert m \rvert\!\rvert\!\rvert$.

So far we have shown that either $m \in K_1 \cup K_2$ or all the predecessors of m have a greater than or equal power-bounded norm. Hence we can define the function $S : \mathscr{P}(V) \to \mathscr{P}(V)$ as

$$S(W) := W \cup \{v : v \in \mathrm{succ}(W), v \notin K_1 \cup K_2\}$$

so that for any $W \subseteq V$

$$\sup_{v \in W} \lvert\!\lvert\!\lvert v \rvert\!\rvert\!\rvert = \sup_{v \in S(W)} \lvert\!\lvert\!\lvert v \rvert\!\rvert\!\rvert.$$

Note that

$$W \cup \mathrm{succ}(W) \subseteq S(W) \cup K_1 \cup K_2$$

and hence if $(K_1 \cup K_2) \subseteq W$ then

$$(W \cup \operatorname{succ}(W)) \subseteq S(W).$$

Hence for an $n \in \mathbb{N}$ with $W = K_1 \cup K_2 \cup K_3$ we have

$$\bigcup_{j=0}^{n} \operatorname{succ}^j(W) \subseteq S^n(W).$$

By the assumption on K_3 the left-hand side approaches V as $n \to \infty$ and so

$$\sup_{v \in K_1 \cup K_2 \cup K_3} |||v||| = \sup_{v \in S^j(K_1 \cup K_2 \cup K_3)} |||v||| = \sup_{v \in V} |||v|||.$$

□

4.4 Examples

The above lemma is exceptionally useful as for most of the operators we will consider the above sets are simple or even empty. To make the uses of the above lemma clearer, here follow some examples.

Example 4.13. Consider the following operator T on the space $V = \{e_j\}_{j=-\infty}^{+\infty}$.

$$Te_j = \begin{cases} e_{j+1} & j \neq 0 \\ \frac{1}{4} e_{j+1} & j = 0 \end{cases}$$

$$\cdots \longrightarrow \boxed{e_{-1}} \longrightarrow \boxed{e_0} \xrightarrow{\frac{1}{4}} \boxed{e_1} \longrightarrow \cdots$$

This is a somewhat trivial example but it shows some important principles. In accordance with Lemma 4.12 we determine the three sets.

$$\begin{aligned} K_1 &= \{e_1\} \\ K_2 &= \emptyset \\ K_3 &= \{e_k : k < -1\} \end{aligned}$$

Definition of the first two is clear, but the third seems odd. The requirement is that it contains enough elements such that for all $x \in V$ there exists an n such that $x \in \operatorname{succ}^n(K_3)$. If K_3 were a finite set then we would be able to find a n such that e_n is not in $\operatorname{succ}^n(K_3)$. So K_3 is defined to contain all the basis vectors in the infinite source.

However, the presence of an infinite number of elements does not pose any difficulty. This is because the original operator T restricted to K_3 is a partial isometry. To put it in functional analytic terms, for any $j < -2$

$$\sup_n \|T^n e_j\| = \max\{\sup_{n \leqslant -j-2} \underbrace{\|e_{j+n}\|}_{=1}, \sup_n \|T^n e_{-1}\|\}.$$

So even though there are an infinite number of elements, the supremum of the power-bounded norm for the elements of K_3 is simply the maximum of 1 and the power-bounded norm of *any* element of the set. Thus for the purposes of calculation it is sufficient to take one element as a representative for the whole set and the same norm will apply to all of them. Hence

$$|||T|||_1 = \max\{1, |||e_{-1}|||, |||e_0|||\} = 1.$$

Now we consider the power-bounded norm in ℓ^∞ by considering the operator T^*. As shown earlier, the adjoint simply reverses the direction of the arrows. Note that there is a self-similarity between T and T^* by the mapping

$$e_j \mapsto e_{1-j}$$

and hence $\|T\|_\infty = \|T^*\|_1 = \|T\|_1 = 1$. The use of similarity to simplify calculations is common. Finally the Riesz-Thorin theorem proves that $\|T\|_p \leqslant 1$ for $1 < p < \infty$.

Example 4.14. Consider the following operator T on the space $V = \{e_j\}_{j=0}^{+\infty} \cup \{f_j\}_{j=0}^{+\infty}$.

$$\begin{aligned} Te_j &= e_{j+1} + \begin{cases} f_j & j = 2 \text{ or } j = 4 \\ 0 & \text{otherwise} \end{cases} \\ Tf_j &= f_{j-1} \qquad (j = 0 \dots \infty) \end{aligned}$$

$$\begin{tikzcd} \boxed{e_0} \arrow[r] & \boxed{e_1} \arrow[r] & \boxed{e_2} \arrow[r] \arrow[d] & \boxed{e_3} \arrow[r] & \boxed{e_4} \arrow[r] \arrow[d] & \boxed{e_5} \arrow[r] & \cdots \\ \boxed{f_0} & \boxed{f_1} \arrow[l] & \boxed{f_2} \arrow[l] & \boxed{f_3} \arrow[l] & \boxed{f_4} \arrow[l] & \boxed{f_5} \arrow[l] & \cdots \arrow[l] \end{tikzcd}$$

In accordance with Lemma 4.12 we determine the three sets.

$$\begin{aligned} K_1 &= \emptyset \\ K_2 &= \{e_0\} \\ K_3 &= \{e_0\} \cup \{f_k : k \geqslant 5\} \end{aligned}$$

As in the previous example we see that

$$\sup_n \|T^n\|_1 = \max\{\|e_0\|_1, \|f_5\|_1\}.$$

Clearly $\|f_5\|_1 = 1$. To calculate $\|e_0\|_1$ we need to examine carefully what happens to the powers of T.

$$\begin{aligned} T^2 e_0 &= e_2 \\ T^3 e_0 &= e_3 + f_2 \\ T^4 e_0 &= e_4 + f_1 \\ T^5 e_0 &= e_5 + f_0 + f_4 \\ T^6 e_0 &= e_6 + f_3 \\ T^7 e_0 &= e_7 + f_2 \end{aligned}$$

What is happening is that there is always precisely one element of the $\{e_j\}_j$ present and a varying number of the $\{f_j\}_j$ which appear and disappear depending on the value of n. Using δ notation we can describe this succinctly.

$$\|T^n e_0\| = 1 + \delta_{3 \leqslant n \leqslant 5} + \delta_{5 \leqslant n \leqslant 9}$$

In general, if there is a link between e_j and f_j then we get the term

$$\delta_{j+1 \leqslant n \leqslant 2j+1}.$$

The examples in the next chapter make extensive use of this construction. Continuing with our current example gives us

$$\|e_0\|_1 \leqslant 3$$

which means

$$\sup_n \|T^n\|_1 \leqslant 3.$$

Since T is similar to T^* we also have

$$\sup_n \|T^n\|_\infty \leqslant 3$$

and so

$$\sup_n \|T^n\|_p \leqslant 3$$

for all $1 \leqslant p \leqslant \infty$ by Riesz-Thorin.

4.5 Weak convergence

For the operators we have created above we also wish to know whether they have weakly converging powers, that is, whether they are strongly mixing. What is important here is a form of recurrence. Consider any two elements $x, y \in V$ and consider the set

$$R_{x,y} := \{n \in \mathbb{N} : y \in \operatorname{succ}^n(x)\}.$$

What this is referring to is if we start at the vertex x and then for each n check if the vertex y is n steps away. What we wish to know is if for each pair (x, y) this set is finite. If so, we have the following.

Lemma 4.15. *Consider a graph $G = (V, E, w)$ with associated operator T. Suppose that T is power-bounded by M and that $R_{x,y}$ is a finite set for each $x, y \in V$. Then the sequence $(T^n)_n$ is weakly convergent to zero in all ℓ^p for $1 < p < \infty$.*

Proof. From the assumption we know that for any two basis vectors v, v' there exists an N such that

$$\langle T^n v, v' \rangle = 0 \qquad \forall n \geqslant N.$$

Fix a p and choose an arbitrary $x \in \ell^p, y \in \ell^q$ where q is the adjoint index. Now choose an $\varepsilon > 0$ and find two elements $x', y' \in c_{00}(V)$ such that

$$\|x - x'\|_p \leqslant \varepsilon \text{ and } \|y - y'\|_q \leqslant \varepsilon$$

which can be done because sequences with finitely many non-zero elements are dense in ℓ^p and ℓ^q. Now since x' and y' have finitely many non-zero entries there exists an N such that $\langle T^n x', y' \rangle = 0$ for all $n > N$. So for these n

$$\begin{aligned}
\langle T^n x, y \rangle &= \langle T^n x, y - y' \rangle + \langle T^n (x - x'), y' \rangle + \langle T^n x', y' \rangle \\
&\leqslant M\|x\|_p \varepsilon + M\|y'\|_q \varepsilon \\
&= M\varepsilon(\|x\|_p + \|y'\|_q)
\end{aligned}$$

so $(T^n)_n$ is weakly convergent. ☐

Note that ℓ^1 and ℓ^∞ are excluded as $c_{00}(V)$ is not dense in ℓ^∞. For ℓ^1 weak convergence is the same as norm convergence.

Chapter 5

Examples and counterexamples

Since actual proofs of the Blum-Hanson equivalence have been somewhat slow in coming, a lot of work has been focused on finding examples and counterexamples, operators with specific combinations of properties. In this way a clearer understanding is reached as to what may or may not be possible.

5.1 Foguel's counterexample

In [SN47] Sz. Nagy showed that every power-bounded operator with a power-bounded inverse is similar to a contraction (see Theorem 2.22). Twelve years later ([SN59]) it was shown that every *compact* power-bounded operator is similar to a contraction. It was conjectured that every power-bounded operator is similar to a contraction.

This was quickly shown to be incorrect. [Fog64] introduced a class of operators which were shown to be power-bounded but not similar to a contraction. It proved a useful idea since later Lebow in [Leb68] showed these operators were also not polynomially bounded. In [Pis97] a very similar class of operators was introduced which were shown to be polynomially bounded but also not similar to a contraction.

Definition 5.1 (Foguel's counterexample 1964)**.** Define the Hilbert space H with the basis vectors $\{e_n\}_{n=0}^{\infty}, \{f_n\}_{n=0}^{\infty}$. Let $(r_j)_j$ be a sequence of increasing natural numbers with Hadamard gaps, i.e. $r_{j+1}/r_j > q > 1$. Define the operator $T : H \to H$ as

$$\begin{aligned} Te_0 &= 0 \\ Te_i &= e_{i-1} \\ Tf_i &= f_{i+1} \qquad i \notin \{r_j\} \\ Tf_i &= f_{i+1} + e_i \qquad i \in \{r_j\} \end{aligned}$$

Here is the graph representation using the sequence $r_j = 3^j$, which is the version introduced by Halmos [Hal64] and what most people referred to later on.

$$\begin{tikzcd}
\boxed{e_0} & \boxed{e_1} \arrow[l] & \boxed{e_2} \arrow[l] & \boxed{e_3} \arrow[l] & \cdots \arrow[l] & \boxed{e_9} \arrow[l] & \cdots \arrow[l] \\
\boxed{f_0} \arrow[r] & \boxed{f_1} \arrow[r] & \boxed{f_2} \arrow[r] & \boxed{f_3} \arrow[r] \arrow[u] & \cdots \arrow[r] & \boxed{f_9} \arrow[r] \arrow[u] & \cdots
\end{tikzcd}$$

We, however, will consider the operator T mapping $c_{00} \to c_{00}$ and its unique extensions to $\ell^p, 1 \leqslant p \leqslant \infty$. Additionally we will be considering cases where (r_j) is a general increasing sequence of positive integers. This operator is clearly a generalised shift operator as described in the previous chapter. As such it is a bounded operator. Foguel proved it power-bounded but it is instructive to examine exactly why it is. By application of Theorem 4.6 we see that $\|T\|_p \leqslant 2$ for all $1 \leqslant p \leqslant \infty$.

Lemma 5.2. *Foguel's operator T, with (r_j) a sequence with Hadamard gaps, is power-bounded on $\ell^p, 1 \leqslant p \leqslant \infty$.*

Proof. Determining the sets for Lemma 4.12 we have that $K_1 = K_3 = \{f_0\}$ and $K_2 = \emptyset$. Note also there is a self-similarity between T and T^* (as in Example 4.13). So

$$\sup_n \|T^n\|_\infty = \sup_n \|T^{*n}\|_1 = \sup_n \|T^n\|_1 = \|f_0\|_1$$

Since T is a positive operator the 1-norm of $T^n(f_0)$ is simply the sum of the basis vectors. Hence

$$\|T^n f_0\|_1 = 1 + \operatorname{card}\{k \in \{r_j\} : k < n < 2k + 2\}$$

where card is the number of elements in the set. By the definition of $(r_j)_j$ this is uniformly bounded. □

The way to view the above formula is to see the initial vector travel to the right along the sequence of f_j occasionally setting one of the e_j. These e_j travel to the left and are eventually annihilated. The further one proceeds the longer the values in the e_j persist, so the sequence must become "sparse" if the norm is to be bounded. We prove the result for all ℓ^p instead of just Hilbert spaces as Foguel and Halmos did.

Lemma 5.3. *Foguel's operator T based on the sequence $(r_k)_k$ is power-bounded in ℓ^p if there exists an M such that*

$$\sup_{k \in \mathbb{N}} \operatorname{card}\{k \in \{r_j\} : k < n < 2k + 2\} < M$$

Proof. This follows directly from the calculation of the norm above:

$$\|T^n f_0\|_\infty = \|T^n f_0\|_1 = 1 + \operatorname{card}\{k \in \{r_j\} : k < n < 2k + 2\} \leqslant 1 + M.$$

By Riesz-Thorin T is power-bounded for all ℓ^p. □

We now have an explicit calculation for the norm and we know that for sequences with Hadamard gaps the norm is bounded. However, there are non-Hadamard sequences which also lead to a power-bounded operator. Similarly there are sequences of zero upper density which lead to a non-power-bounded operator.

Example 5.4. Consider the sequence

$$r_j = \lceil j^{1+\varepsilon} \rceil$$

for some $\varepsilon > 0$. Then we have

$$\overline{d}(r_j) = \limsup_{j \to \infty} \frac{j}{r_j} \leqslant \limsup_{j \to \infty} j^{-\varepsilon} = 0.$$

However, for any $n \in \mathbb{N}$

$$\begin{aligned}
\operatorname{card}\{k \in \{r_j\} : k < n < 2k + 2\} &= \operatorname{card}\{k \in \{r_j\} : \tfrac{n}{2} + 1 < k < n\} \\
&= \operatorname{card}\{k = \lceil j^{1+\varepsilon} \rceil : j \in \mathbb{N}, \tfrac{n}{2} + 1 < k < n\} \\
&= \operatorname{card}\{j^{1+\varepsilon} \in (k-1, k] : j \in \mathbb{N}, \tfrac{n}{2} + 1 < k < n\} \\
&= \operatorname{card}\{j^{1+\varepsilon} \in (\tfrac{n}{2}, n-1] : j \in \mathbb{N}\} \\
&= \operatorname{card}\{(1+\varepsilon)\ln(j) \in (\ln(\tfrac{n}{2}), \ln(n-1)] : j \in \mathbb{N}\}
\end{aligned}$$

The interval on the right-hand side remains approximately the same size ($\ln 2$), while the $\ln(j)$ come closer together as $j \to \infty$, making the count unbounded. In any case, the result in Lemma 5.3 is stronger than Foguel's proof, which required Hadamard gaps.

The other property of interest for this operator is that it is not similar to a contraction.

Lemma 5.5. *Foguel's operator T on a Hilbert space using any sequence $(r_j)_j$ is not similar to a contraction.*

Proof. Recall from Theorem 2.24 the definition $H_0(T) = \{x : T^n x \to 0 \text{ weakly}\}$.

As $Te_0 = 0$, clearly $e_0 \in H_0(T)$.

Now, $T^{2r_j+1} f_0 = e_0 + f_{2r_j+1}$ and so $T^{2r_j+1} f_0 \to e_0$ weakly. Consider an element $z \in H_0(T^*)$. Then

$$\begin{aligned}\langle e_0, z\rangle &= \lim_{j\to\infty} \langle T^{2r_j+1} f_0, z\rangle \\ &= \lim_{j\to\infty} \langle f_0, T^{*2r_j+1} z\rangle \\ &= 0\end{aligned}$$

and so $e_0 \in H_0(T^*)^\perp$.

Hence by Lemma 2.25 T is not similar to a contraction. □

So we see that similarity to a contraction is independent of whether the operator is power-bounded or not, but this proof only covers Hilbert spaces. As noted earlier, similarity to a contraction is not a well studied area for general ℓ^p spaces.

Finally we show that the Blum-Hanson property is in fact not dependent on the sequence (r_j) but is intrinsic to the structure of the operator.

Theorem 5.6. *Let T be Foguel's operator with any sequence $(r_j)_j$. Then T does not have the Blum-Hanson property.*

Proof. Consider again the sequence $(\langle T^n f_0, e_0\rangle)_n$. This sequence will be 1 whenever $n = 2r_j+1$ for some j and 0 otherwise. In particular it will be zero for all even n. So $(T^n)_n$ does not converge in the weak operator topology. Applying Lemma 2.17 shows that T does not have the Blum-Hanson property. □

This makes Foguel's operator an example of a power-bounded operator which is not similar to a contraction and which also does not have the Blum-Hanson property. As $(T^n)_n$ does not converge in the weak operator topology, it is not counterexample to the Blum-Hanson equivalence, but it laid the groundwork for the following counterexample. It was however the first example of a power-bounded operator which is not polynomially bounded.

Theorem 5.7. *Let T be Foguel's operator on ℓ^2 with any sequence $(r_j)_j$ with Hadamard gaps. Then T is not polynomially bounded.*

Proof. [Leb68] □

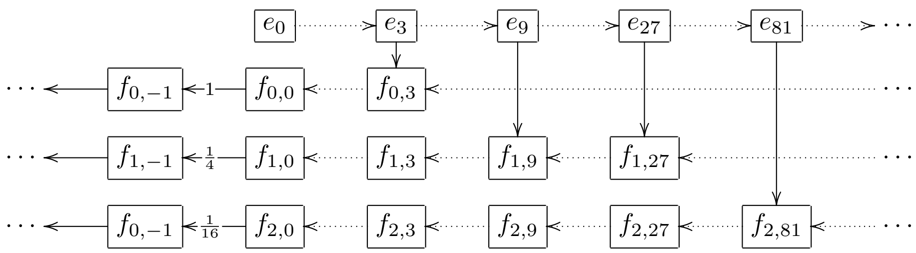
5.2 Müller-Tomilov's example

In [MT07] Müller-Tomilov introduced an operator that in many ways was an extension of the operator defined by Foguel. The goal was to make an operator that was weakly convergent to zero while still being power-bounded and still not possessing the Blum-Hanson property, thus making it a counterexample for the Blum-Hanson equivalence for power-bounded operators on $\ell^p, 1 \leqslant p < \infty$. The definition is as follows.

Definition 5.8 (Müller-Tomilov's example 2007)**.** Let H be the Hilbert space with an orthonormal basis formed by the vectors $\{e_i\}(i \geqslant 0)$ and $\{f_{i,j}\}(i \geqslant 0, j \in \mathbb{Z})$. Define the function $r : \mathbb{N} \to \mathbb{N}$ by $r(k) = \lfloor \log_2 k \rfloor + 1$, so $r(k) = s$ whenever $2^{r-1} \leqslant k < 2^r (k \geqslant 1, s \geqslant 1)$. Then define $T \in B(H)$ by:

$$
\begin{array}{ll}
Tf_{i,j} = f_{i,j-1} & i \geqslant 1, j \neq 0 \\
Tf_{i,0} = 4^{-i} f_{i,-1} & i \geqslant 1 \\
Te_j = e_{j+1} & (j \notin 3^k : k = 1, 2, \ldots) \\
Te_{3^k} = e_{3^k+1} + f_{r(k),3^k} & (k = 1, 2, \ldots)
\end{array}
$$

Note the use of the 3^k as in Foguel's operator; it plays much the same role here. The $r(k)$ also serves a specific purpose, as will be seen later. The associated graph is as follows:



There are a number of features of this operator not present in Foguel's operator. Firstly, there are no sinks, which means that values are not annihilated as in Foguel's operator. Instead the sequences of f are carefully scaled so they stay away from zero but do not cause any other problems. These sequences do not materially affect the properties we are interested in here.

Secondly, the arcs coming from the $(e_j)_j$ are at intervals of 3^k similar to Foguel's operator, but instead of all linking to one chain they link to a sequence to chains of $(f_{j,k})_k$. To each of the $(f_{j,k})_k$ there is never an infinite number of links to a single chain of f_j (unbounded yes, but not infinitely many). The importance of this will become clear later on.

Stepping away from just Hilbert spaces, considering the above operator on c_{00} with its unique extensions to ℓ^p we note that this is clearly a generalised shift operator.

Lemma 5.9. *The Müller-Tomilov operator T is power-bounded on all $\ell^p, 1 \leqslant p \leqslant \infty$.*

Proof. Applying the definitions from Lemma 4.12 to T we consider the sets

$$
\begin{aligned}
K_1 &= \{f_{n,-1} : n = 1, 2, \ldots\} \\
K_2 &= \{e_0\} \\
K_3 &= \{e_0\} \cup \{f_{n,k} : k > 3^{2^k}\}
\end{aligned}
$$

Now because T restricted to each subspace $\{f_{n,k}\}_{k\in\mathbb{Z}}$ is almost an isometry we have (as in Example 4.13)

$$\begin{aligned}
\sup_{v\in K_1} \|v\|_1 &= 1 \\
\sup_{v\in K_2} \|v\|_1 &= \|e_0\|_1 \\
\sup_{v\in K_3} \|v\|_1 &= \max\{1, \|e_0\|\}_1
\end{aligned}$$

hence we only need to explicitly calculate $\|e_0\|$. As the operator is positive we can simply sum the elements.

$$\begin{aligned}
\|T^n e_0\|_1 &= 1 + \underbrace{\sum_{i\in\mathbb{N}} \delta_{\{3^i<n\leqslant 2\cdot 3^i\}}}_{\leqslant 1} + \sum_{i\in\mathbb{N}} 4^{-r(i)} \underbrace{\delta_{\{2\cdot 3^i<n\}}}_{\leqslant 1} \\
&\leqslant 2 + \sum_{s=1}^{\infty} 4^{-s}\cdot \operatorname{card}\{k\in\mathbb{N} : r(s)=k\} \\
&\leqslant 2 + \sum_{s=1}^{\infty} 4^{-s}2^s \leqslant \tfrac{5}{2}
\end{aligned}$$

Hence

$$\sup_n \|T^n\|_1 = \sup_{v\in K_1\cup K_2\cup K_3} \|v\| \leqslant \tfrac{5}{2}.$$

Similarly $\sup_n \|T^n\|_\infty = \sup_n \|T^{*n}\|_1$. Hence we need to consider the sets

$$\begin{aligned}
K_1 &= \{f_{n,0} : n = 1,2,\ldots\} \\
K_2 &= \emptyset \\
K_3 &= \{e_k : k > 1\} \cup \{f_{n,k} : k < -1\}
\end{aligned}$$

Similar to above (as in Example 4.13)

$$\sup_{v\in K_3} \|v\|_\infty \leqslant \sup_{v\in K_1\cup\{e_1\}} \|v\|_\infty$$

hence we only need to calculate explicitly for K_1, as $\|e_1\|_\infty = 1$. Note how if we restrict the operator to $\{e_j\}_j \cup \{f_{i,j}\}_{j\geqslant 0}$ we get an operator that is similar to itself and so

$$\begin{aligned}
\|f_{i,0}\|_\infty &= \|e_0\|_1 \\
&\leqslant \tfrac{5}{2}
\end{aligned}$$

Hence

$$\sup_n \|T\|_\infty = \sup_{v\in K_1\cup K_2\cup K_3} \|v\| \leqslant \tfrac{5}{2}$$

and by Riesz-Thorin T is power-bounded on all $\ell^p, 1 \leqslant p \leqslant \infty$. □

Lemma 5.10. *The Müller-Tomilov operator T has weakly convergent powers, i.e. $(\langle T^n f, g\rangle)_n$ converges for all $f \in \ell^p, g \in \ell^q, 1 < p < \infty$.*

Proof. We want to use Lemma 4.15. We need to consider a number of possibilities:

$$\operatorname{card}\{R_{x,y}\} \leqslant \begin{cases} 1 & x \in \{e_k\}_k, y \in \{e_k\}_k \\ 0 & x \in \{f_{i,k}\}_k, y \in \{e_k\}_k \\ 2^i & x \in \{e_k\}_k, y \in \{f_{i,k}\}_k \\ 0 & x \in \{f_{i,k}\}_k, y \in \{f_{i,k}\}_k \end{cases}$$

Since in all cases $R_{x,y}$ is finite, $(T^n)_n$ converges weakly in $\ell^p, 1 < p < \infty$. □

Lemma 5.11. *The Müller-Tomilov operator T does not have the Blum-Hanson property.*

Proof. Consider the sequence $(k_n)_n$ defined by $k_n = 2 \cdot 3^n + 1$. Then

$$T^{k_n} e_0 = e_{k_n} + f_{r(n),0} + \sum_{j=1}^{n-1} 4^{-r(j)} f_{r(j),2\cdot 3^j - 2\cdot 3^{k_n}}$$

And thus for each $s \in \mathbb{N}$:

$$\frac{1}{2^s - 1} \left\| \sum_{n=1}^{2^s-1} T^{k_n} e_0 \right\| \geqslant \frac{1}{2^s - 1} \left| \sum_{n=1}^{2^s-1} \langle T^{k_n} e_0, f_{s,0} \rangle \right| = \frac{2^{s-1}}{2^s - 1} \geqslant \frac{1}{2}$$

Hence the sequence $\frac{1}{N} \|\sum_{n=1}^{N} T^{k_n} e_0\|$ does not converge to zero as $N \to \infty$. Since the sequence $\frac{1}{N} \sum_{n=1}^{N} T^{k_n} e_0$ does converge weakly to zero it does not converge strongly in the norm topology. □

Summarising the effects of the construction on the properties.

- Power-boundedness is primarily dependent on the 3^k, but as in Foguel's operator, any Hadamard sequence will do.
- Not having the Blum-Hanson property is controlled by $r(k)$ growing slow enough. The logarithm to any base would be sufficient, whereas linear growth would not have worked.
- Weak-convergence (strong-mixing) is inherent to the structure itself, as long as $r(k)$ is not constant, in which case it reduces to Foguel's operator.

Whether this operator is polynomially bounded or not has not yet been determined.

5.3 A new example

One of effects of studying these generalised shift operators is that after a while you can create new operators with specified characteristics. So, when László Zsidó asked the following

> Is there for every $\varepsilon > 0$ a Müller-Tomilov type counterexample with $\sup_n \|T^n\| < 1 + \varepsilon$ or is there some $\lambda > 1$ such that the strong mixing property implies the Blum-Hanson property for the orbits of any linear operator T on a Hilbert space with $\sup_n \|T^n\| < \lambda$?

it was clear to me that you could make a operator above with an arbitrarily small ε. After understanding how the power-boundedness of the Müller-Tomilov operator is calculated (Lemma 5.9) I quickly came up with the following.

Example 5.12. Let H be the Hilbert space with an orthonormal basis formed by the vector $\{e_i\}(i \geqslant 0)$ and $\{f_{i,j}\}(i \geqslant 0, j \geqslant 0)$. Define the function $r : \mathbb{N} \to \mathbb{N}$ by $r(k) = \lfloor \log_2 k \rfloor + 1$, so $r(k) = s$ whenever $2^{r-1} \leqslant k < 2^r (k \geqslant 1, s \geqslant 1)$. Choose an $\varepsilon > 0$. Then define $T \in B(H)$ by:

$$
\begin{aligned}
Tf_{i,j} &= f_{i,j-1} && i \geqslant 1, j \neq 0 \\
Te_j &= e_{j+1} && (j \notin 3^k : k = 1, 2, \dots) \\
Te_{3^k} &= e_{3^k+1} + \varepsilon f_{r(k),3^k} && (k = 1, 2, \dots)
\end{aligned}
$$

The graph formulation is as follows.

$$
\begin{tikzcd}
\boxed{e_0} \arrow[r, dotted] & \boxed{e_3} \arrow[r, dotted] \arrow[d, "\varepsilon"] & \boxed{e_9} \arrow[r, dotted] \arrow[dd, "\varepsilon"] & \boxed{e_{27}} \arrow[r, dotted] \arrow[dd, "\varepsilon"] & \boxed{e_{81}} \arrow[r, dotted] \arrow[ddd, "\varepsilon"] & \cdots \\
\boxed{f_{0,0}} & \boxed{f_{0,3}} \arrow[l, dotted] & & & & \cdots \arrow[llll, dotted] \\
\boxed{f_{1,0}} & \boxed{f_{1,3}} \arrow[l, dotted] & \boxed{f_{1,9}} \arrow[l, dotted] & \boxed{f_{1,27}} \arrow[l, dotted] & & \cdots \arrow[ll, dotted] \\
\boxed{f_{2,0}} & \boxed{f_{2,3}} \arrow[l, dotted] & \boxed{f_{2,9}} \arrow[l, dotted] & \boxed{f_{2,27}} \arrow[l, dotted] & \boxed{f_{2,81}} \arrow[l, dotted] & \cdots \arrow[l, dotted]
\end{tikzcd}
$$

As you can see this operator is very similar to the previous one. The differences are:

- The lack of infinite sinks on the left hand side. As noted earlier, these are not relevant to the properties of the operator we are interested in.
- The down arrows are scaled by the ε chosen in the definition.

Now to show that this operator has the required properties. The proofs are essentially simplified versions of those for the Müller-Tomilov operator.

Lemma 5.13. *The operator T is power-bounded by $1+\varepsilon$ on all $\ell^p, 1 \leqslant p \leqslant \infty$.*

Proof. Applying the definitions from Lemma 4.12 to T we consider the sets

$$
\begin{aligned}
K_1 &= \emptyset \\
K_2 &= \{e_0\} \\
K_3 &= \{e_0\} \cup \{f_{n,k} : k > 3^{2^k}\}
\end{aligned}
$$

Now because T restricted to each subspace $\{f_{n,k}\}_{k\in\mathbb{Z}}$ is an isometry we have (as in Example 4.13)

$$
\begin{aligned}
\sup_{v\in K_2} |||v|||_1 &= |||e_0|||_1 \\
\sup_{v\in K_3} |||v|||_1 &= \max\{1, |||e_0|||\}_1
\end{aligned}
$$

hence we only need to explicitly calculate $|||e_0|||$. As the operator is positive we can simply sum the elements.

$$
\begin{aligned}
\|T^n e_0\|_1 &= 1 + \varepsilon \underbrace{\sum_{i\in\mathbb{N}} \delta_{\{3^i < n \leqslant 2\cdot 3^i\}}}_{\leqslant 1} \\
&\leqslant 1 + \varepsilon
\end{aligned}
$$

Hence

$$
\sup_n \|T^n\|_1 = \sup_{v\in K_1\cup K_2\cup K_3} |||v||| \leqslant 1 + \varepsilon.
$$

Similarly $\sup_n \|T^n\|_\infty = \sup_n \|T^{*n}\|_1$. Hence we need to consider the sets

$$\begin{aligned} K_1 &= \emptyset \\ K_2 &= \emptyset \\ K_3 &= \{e_k : k > 1\} \cup \{f_{n,0} : n \in \mathbb{N}\} \end{aligned}$$

Similar to above (as in Example 4.13)

$$\sup_{v \in K_3} \|v\|_\infty = \max\{1, \sup_{v \in \{f_{n,0} : n \in \mathbb{N}\}} \|v\|_\infty\}$$

hence we only need to calculate explicitly for $\{f_{n,0} : n \in \mathbb{N}\}$, as $\|e_1\|_\infty = 1$. Note how if we restrict the operator to $\{e_j\}_j \cup \{f_{i,j}\}_j$ we get an operator that is similar to itself and so

$$\begin{aligned} \|f_{i,0}\|_\infty &= \|e_0\|_1 \\ &\leqslant 1 + \varepsilon \end{aligned}$$

Hence

$$\sup_n \|T\|_\infty = \sup_{v \in K_1 \cup K_2 \cup K_3} \|v\| \leqslant 1 + \varepsilon$$

and so by Riesz-Thorin T is power-bounded by $1 + \varepsilon$ on all $\ell^p, 1 \leqslant p \leqslant \infty$. □

Lemma 5.14. *The operator T has weakly convergent powers, i.e. $(\langle T^n f, g\rangle)_n$ converges for all $f \in \ell^p, g \in \ell^q, 1 < p < \infty$.*

Proof. This proof is precisely the same as in the Müller-Tomilov case, see Lemma 5.10. □

Lemma 5.15. *The operator T does not have the Blum-Hanson property.*

This proof is practically identical to Lemma 5.11, just with an ε thrown in.

Proof. Consider the sequence $(k_n)_n$ defined by $k_n = 2 \cdot 3^n + 1$. Then

$$T^{k_n} e_0 = e_{k_n} + \varepsilon f_{r(n),0}$$

And thus for each $s \in \mathbb{N}$:

$$\frac{1}{2^s - 1} \left\| \sum_{n=1}^{2^s - 1} T^{k_n} e_0 \right\| \geqslant \frac{1}{2^s - 1} \left| \sum_{n=1}^{2^s - 1} \langle T^{k_n} e_0, f_{s,0} \rangle \right| = \varepsilon \frac{2^{s-1}}{2^s - 1} \geqslant \frac{\varepsilon}{2}$$

Hence the sequence $\frac{1}{N} \|\sum_{n=1}^N T^{k_n} e_0\|$ does not converge to zero as $N \to \infty$. Since the sequence $\frac{1}{N} \sum_{n=1}^N T^{k_n} e_0$ does converge weakly to zero it does not converge strongly in the norm topology. □

The above shows that the question by László was easily answered and in fact in an even more general form than he wanted: we have shown the requested operators exist in all $\ell^p, 1 < p < \infty$, not just Hilbert spaces.

5.4 The $C(K)$ counterexample

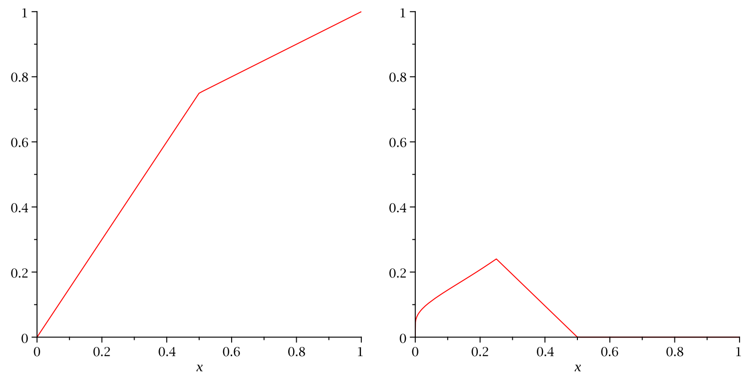
The first counterexample to the idea that all contractions on a Banach space would have the Blum-Hanson equivalence was given in [AHR74].

Example 5.16 ($C(K)$ counterexample)**.** Let $K := [0, 1)$ be the unit interval with the circle topology; this is clearly a compact space. Now we define the Banach space $X := C(K^2)$, the continuous functions on the torus. Define two functions $\phi, \alpha : K \to K$ as

$$\phi(x) := \begin{cases} \frac{3}{2}x & 0 \leqslant x < \frac{1}{2} \\ \frac{1}{2} + \frac{1}{2}x & \frac{1}{2} \leqslant x < 1 \end{cases}$$

$$\alpha(x) := \begin{cases} -\frac{A}{\ln x} & 0 \leqslant x \leqslant \frac{1}{4} \\ \frac{4A(\frac{1}{2}-x)}{\ln 4} & \frac{1}{4} < x \leqslant \frac{1}{2} \\ 0 & \frac{1}{2} < x < 1 \end{cases}.$$

Here A is a positive constant such that $\alpha(\frac{1}{4}) < \frac{1}{4}$, for example $A = \frac{1}{3}$. See Figure 5.1 for a visualisation of these functions.



(a) $\phi(x)$ (b) $\alpha(x)$

Figure 5.1: Graphs of functions for $C(K)$ counterexample.

By definition these functions are both continuous. Using these two functions we can define our transition function $\tau : X \to X$ as

$$\tau(x, y) := (\phi(x) + \alpha(y), \phi(y))$$

and so it is also continuous. Note that the first term is modulo 1 as we are on the torus. So now we define our operator $T : X \to X$ as the operator induced by τ, i.e.

$$(Tf)(x) := f(\tau(x)).$$

Note that T is an contraction on $C(K^2)$.

Lemma 5.17. *The operator T is strongly mixing, i.e. $(T^n f)_n$ converges weakly to $\mathbf{1} \cdot f(0, 0)$.*

Proof. Given that T is a contraction and hence T^n is a contraction, it is sufficient to show that $((T^n f)(x))_n$ converges pointwise. Hence it is sufficient to show that $\tau(x) \to (0, 0)$ for all $x \in K$.

Let the starting point be (x, y) and let (x_n, y_n) be the n-th point in the orbit. So $(x_0, y_0) = (x, y)$. Note that $\phi(x) > x$ so the sequence $(y_n)_n$ is a monotonically increasing sequence converging to 1, which equals 0 on the torus. Find an N such that $y_n > \frac{1}{2}$ for all $n > N$. Because $\alpha(y_n) = 0$ for all $n > N$ the $(x_n)_{n>N}$ also form a monotonically increasing sequence.

Hence $(x_n, y_n) \to (0, 0)$ and $(T^n)_n$ converges weakly. □

Lemma 5.18. *The operator T does not have the Blum-Hanson property.*

Proof. What we need to do is demonstrate that there exists an $f \in C(K^2)$ and a sequence of positive integers $(m_i)_i$ such that

$$\frac{1}{N}\sum_{i=1}^{N} T^{m_i} f$$

does not converge in the norm of $C(K^2)$.

Choose an integer $M \in \mathbb{N}$. Consider the point $x = (0, \frac{1}{4} \cdot \frac{2}{3}^M)$ so that $\phi^M(y_0) = \frac{1}{4}$. Now consider the sequence

$$\delta_n = x_n - x_{n-1} = (\phi(x_{n-1}) - x_{n-1}) + \alpha(y_{n-1}).$$

This sequence is the amount the x_n changes after each iteration. So

$$x_n = \sum_{i=1}^{n} \delta_n \bmod 1.$$

Define a function $r : K^2 \to \mathbb{N}_0$ as

$$r(x) := \operatorname{card}\{x_n : \tfrac{1}{8} < x_n < \tfrac{7}{8}\}$$

which is the number of times the orbit of x is in the given strip. Note that for $n \leqslant M$ we have that $0 < \delta_n < \frac{1}{2}$ so if $\sum_{i=1}^{n} \delta_n > N$ for some integer N we see that $r(x) > N$. By calculation

$$\begin{aligned}\sum_{n=1}^{M} \delta_n &\geqslant \sum_{n=0}^{M-1} \alpha(y_n) \\ &= \sum_{n=0}^{M-1} \frac{A}{\ln 4 + (M-n)\ln\frac{3}{2}} \\ &= \sum_{n=1}^{M} \frac{A}{\ln 4 + n\ln\frac{3}{2}}\end{aligned}$$

which can be made arbitrarily large by choosing an sufficiently large M. Hence we can define a sequence $(t_n)_n \in K^2$ such that $r(t_n) > n$.

To complete the counterexample choose any positive continuous function $f \in C(K^2)$ such that $f(0, 0) = 0$ and $f(x, y) > 1$ for all $\frac{1}{8} < x < \frac{7}{8}$ and $y \in K$. By Lemma 5.17 we have that $T^n f \to \mathbf{0}$ weakly.

Now construct an increasing sequence of positive integers $(m_i)_{i=0}^{\infty}$ and a sequence of points $(x_r)_{r=1}^{\infty}$ recursively as follows. Let $m_0 = 0$. If the first 2^k elements are fixed, let $x_k = t_{m_{2^k-1}+2^k}$. We know that $r(x_k) > m_{2^k-1} + 2^k$ so we can fill in the next 2^k elements with indexes such that m_i remains an ascending sequence and the orbit of x_k falls inside the strip $((\frac{1}{8}, \frac{7}{8}), [0, 1))$.

Now for each k we have that the orbit of x_k is such that $f(\tau^{m_i}(x)) > 1$ for more than half the sequence $(m_i)_{i=0}^{2^r-1}$ and so for all k

$$\left\| \frac{1}{2^k}\sum_{i=0}^{2^k} (T^{m_i} f)(x_k) \right\| = \left\| \frac{1}{2^k}\sum_{i=0}^{2^k} f(\tau^{m_i}(x_k)) \right\| > \frac{1}{2}.$$

So along this sequence $(m_i)_i$ the sums do not converge uniformly to zero and hence T does not have the Blum-Hanson property. $\square$

Chapter 6

Conclusion

The conclusion is that while the Blum-Hanson equivalence is possible amongst contractions, although the case for general L^p is not yet done, the Müller-Tomilov operator means that for power-bounded operators on general $L^p, 1 < p < \infty$ it is impossible to have the Blum-Hanson equivalence. Since every power-bounded operator can be used to define an equivalent norm such that the operator is a contraction, this makes contractions on L^p spaces very special. As shown, contractions under the sup-norm on $C(K)$ also do not have the required property. The deciding factor is clearly not the topology and is instead a metric property. However, precisely what it is about L^p spaces that allows it to work is as yet unknown.

The remaining hole to be filled is general (non-positive) L^p contractions, either a proof or a counterexample. However, not a lot of progress seems to have been made in this area; they are amongst the most difficult spaces to analyse.

Bibliography

[AHR74] M. A. Akcoglu, J. P. Huneke, and H. Rost, *A counter example to the Blum Hanson theorem in general spaces*, Pacific J. Math. **50** (1974), 305–308. MR MR0350455 (50 #2947)

[AS72] M. A. Akcoglu and L. Sucheston, *On operator convergence in Hilbert space and in Lebesgue space*, Period. Math. Hungar. **2** (1972), 235–244, Collection of articles dedicated to the memory of Alfréd Rényi, I. MR MR0326433 (48 #4777)

[AS75] ———, *Weak convergence of positive contractions implies strong convergence of averages*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **32** (1975), 139–145. MR MR0377559 (51 #13730)

[Bel75] Alexandra Bellow, *An L^p-inequality with application to ergodic theory*, Houston J. Math. **1** (1975), no. 1, 153–159. MR MR0385070 (52 #5940)

[BH60] J. R. Blum and D. L. Hanson, *On the mean ergodic theorem for subsequences*, Bull. Amer. Math. Soc. **66** (1960), 308–311. MR MR0118803 (22 #9572)

[Car05] N. L. Carothers, *A short course on Banach space theory*, London Mathematical Society Student Texts, vol. 64, Cambridge University Press, Cambridge, 2005. MR MR2124948 (2005k:46001)

[Cha67] R. V. Chacon, *A geometric construction of measure preserving transformations*, Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 2, Univ. California Press, Berkeley, Calif., 1967, pp. 335–360. MR MR0212158 (35 #3033)

[CK64] R. V. Chacon and U. Krengel, *Linear modulus of a linear operator*, Proc. Amer. Math. Soc. **15** (1964), 553–559. MR MR0164244 (29 #1543)

[Con90] John B. Conway, *A course in functional analysis*, second ed., Graduate Texts in Mathematics, vol. 96, Springer-Verlag, New York, 1990. MR MR1070713 (91e:46001)

[DS56] Nelson Dunford and J. T. Schwartz, *Convergence almost everywhere of operator averages*, J. Rational Mech. Anal. **5** (1956), 129–178. MR MR0077090 (17,987g)

[DS58] ———, *Linear Operators. I. General Theory*, With the assistance of W. G. Bade and R. G. Bartle. Pure and Applied Mathematics, Vol. 7, Interscience Publishers, Inc., New York, 1958. MR MR0117523 (22 #8302)

[Fog63] S. R. Foguel, *Powers of a contraction in Hilbert space*, Pacific J. Math. **13** (1963), 551–562. MR MR0163170 (29 #473)

[Fog64] ———, *A counterexample to a problem of Sz.-Nagy*, Proc. Amer. Math. Soc. **15** (1964), 788–790. MR MR0165362 (29 #2646)

[Gel41] I. Gelfand, *Normierte Ringe*, Rec. Math. [Mat. Sbornik] N. S. **9 (51)** (1941), 3–24. MR MR0004726 (3,51f)

[Hal64] P. R. Halmos, *On Foguel's answer to Nagy's question*, Proc. Amer. Math. Soc. **15** (1964), 791–793. MR MR0165363 (29 #2647)

[JK71] Lee K. Jones and Velimir Kuftinec, *A note on the Blum-Hanson theorem*, Proc. Amer. Math. Soc. **30** (1971), 202–203. MR MR0281023 (43 #6742)

[JL76] Lee K. Jones and Michael Lin, *Ergodic theorems of weak mixing type*, Proc. Amer. Math. Soc. **57** (1976), no. 1, 50–52. MR MR0405134 (53 #8929)

[Jon71] Lee K. Jones, *A mean ergodic theorem for weakly mixing operators*, Advances in Math. **7** (1971), 211–216 (1971). MR MR0285690 (44 #2908)

[Kan79] Charn Huen Kan, *On Fong and Sucheston's mixing property of operators in a Hilbert space*, Acta Sci. Math. (Szeged) **41** (1979), no. 3-4, 317–325. MR MR555424 (81a:47010)

[Leb68] A. Lebow, *A power-bounded operator that is not polynomially bounded*, Michigan Math. J. **15** (1968), 397–399. MR MR0236753 (38 #5047)

[Mat88] Martin Mathieu, *On the origin of the notion "ergodic theory"*, Exposition. Math. **6** (1988), no. 4, 373–377. MR MR964761 (89i:01041)

[Meg98] Robert E. Megginson, *An introduction to Banach space theory*, Graduate Texts in Mathematics, vol. 183, Springer-Verlag, New York, 1998. MR MR1650235 (99k:46002)

[Mil80] Annie Millet, *Sur le théorème en moyenne d'Akcoglu-Sucheston*, Math. Z. **172** (1980), no. 3, 213–237. MR MR581440 (81h:47007)

[MN91] P. Meyer-Nieberg, *Banach lattices*, Universitext, Springer-Verlag, Berlin, 1991. MR MR1128093 (93f:46025)

[MP81] Michael B. Marcus and Gilles Pisier, *Random Fourier series with applications to harmonic analysis*, Annals of Mathematics Studies, vol. 101, Princeton University Press, Princeton, N.J., 1981. MR MR630532 (83b:60031)

[MT07] Vladimir Müller and Yuri Tomilov, *Quasisimilarity of power bounded operators and Blum-Hanson property*, J. Funct. Anal. **246** (2007), no. 2, 385–399. MR MR2321048 (2009a:47003)

[Pis97] Gilles Pisier, *A polynomially bounded operator on Hilbert space which is not similar to a contraction*, J. Amer. Math. Soc. **10** (1997), no. 2, 351–369. MR MR1415321 (97f:47002)

[Sch74] Helmut H. Schaefer, *Banach lattices and positive operators*, Springer-Verlag, New York, 1974, Die Grundlehren der mathematischen Wissenschaften, Band 215. MR MR0423039 (54 #11023)

[SN47] Béla Sz.-Nagy, *On uniformly bounded linear transformations in Hilbert space*, Acta Univ. Szeged. Sect. Sci. Math. **11** (1947), 152–157. MR MR0022309 (9,191b)

[SN59] ______, *Completely continuous operators with uniformly bounded iterates*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **4** (1959), 89–93. MR MR0108722 (21 #7436)

[vN32] Johann von Neumann, *Proof of the Quasi-Ergodic Hypothesis*, Proceedings of the National Academy of Sciences of the United States of America **18** (1932), no. 1, 70–82.

[vN51] ______, *Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes*, Math. Nachr. **4** (1951), 258–281. MR MR0043386 (13,254a)

[Wec92] Wolfgang Wechler, *Universal algebra for computer scientists*, EATCS Monographs on Theoretical Computer Science, vol. 25, Springer-Verlag, Berlin, 1992. MR MR1177406 (94a:08001)

[Zar05] Mohamed Zarrabi, *On polynomially bounded operators acting on a Banach space*, J. Funct. Anal. **225** (2005), no. 1, 147–166. MR MR2149921 (2006b:47007)

[Zsi07] László Zsidó, *Weak mixing properties of vector sequences*, The extended field of operator theory, Oper. Theory Adv. Appl., vol. 171, Birkhäuser, Basel, 2007, pp. 361–388. MR MR2308570 (2008i:47026)

Appendix A

External theorems

This appendix collects a number of the unusual or lesser known theorems referred to in the text.

Banach lattices

Definition A.1 ($\vee$-stable)**.** A set F is called $\vee$-stable if for any $f, g \in F$ we have that

$$f \vee g = \sup\{f, g\} \in F$$

Theorem A.2. *Let (Ω, Σ, μ) be a measure space. Let $F \subseteq L^1_+(\Omega, \Sigma, \mu)$ be a uniformly bounded $\vee$-stable set. Then $f := \sup F$ exists in the Banach lattice $L^1(\Omega; \mathbb{R})$ and there exists an increasing sequence $(f_n)_n \in F$ such that $\sup_n f_n = f$ and $\|f - f_n\| \to 0$. In particular $f \in F$ if F is closed.*

Proof. [MN91] □

Weak compactness of L^1

Theorem A.3 (Vitali-Hahn-Saks)**.** *Suppose $(\lambda_n)_n$ is a sequence of μ-continuous signed measures on a σ-algebra Σ such that*

$$\lambda(A) := \lim_{n\to\infty} \lambda_n(A)$$

exists in $\mathbb{R}$ for each $A \in \Sigma$. Then the measures $(\lambda_n)_n$ are uniformly μ-continuous and λ is a signed measure on Σ.

Proof. [Car05, Corollary 13.7] and [DS58] □

This result is related to the more general result that certain subsets of L^1 are weakly compact. More information on this can be found in [Car05, Chapter 13].

Lemma A.4. *If the sequence of countably additive measures $(\lambda_n)_n$ is uniformly μ-continuous, then the sequence $(|\lambda_n|)_n$ is uniformly μ-continuous.*

Proof. Fix $\varepsilon > 0$. Then there exists a $\delta > 0$ such that

$$\mu(A) < \delta \implies \lambda_n(A) < \varepsilon$$

for all n. Now, using the identity

$$|\lambda_n|(A) \leqslant 4 \sup_{B\subseteq A} |\lambda_n(B)|$$

we see that $|\lambda_n|(A) < 4\varepsilon$ and so the sequence $(|\lambda_n|)_n$ is uniformly μ-continuous. □

Hilbert space results

Theorem A.5. *([Con90, Theorem 2.2]) If $\phi : H \times H \to \mathbb{C}$ is a bounded sesquilinear form with bound M, then there exist unique operators $A, B \in BL(H)$ such that*

$$\phi(f, g) = \langle Af, g\rangle = \langle f, Bg\rangle$$

for all $f, g \in H$ and $\|A\| \leqslant M, \|B\| \leqslant M$.

Lemma A.6. *([Con90, Proposition 2.12]) If H is a complex Hilbert space and $A \in BL(H)$, then A is hermitian if and only if $\langle Ah, h\rangle \in \mathbb{R}$ for all $h \in H$.*

Appendix B

Notes for functional analysis students

To make this thesis a little more readable for people who have only had a basic course in functional analysis, here is the explanation of some terms.

Hausdorff space A Hausdorff space (also known a separated or T_2 space) is a space where any two points can be separated by open sets. Almost all spaces discussed in analysis are Hausdorff. In particular any metric space is Hausdorff.

Weak convergence A sequence $(x_n)_n$ in a Banach space X is said to be weakly convergent if $\langle x_n, y\rangle$ converges for each $y \in X^*$. For example, the sequence of basis vectors for an ℓ^p space, $1 < p < \infty$ is weakly convergent.

A sequence of operators $(T_n)_n$ on a Banach space X is said to be weakly convergent to $P \in BL(X)$ if

$$\langle T_n x, y\rangle \to \langle Px, y\rangle$$

for all $x \in X, y \in X^*$. For example, the powers of the right-shift operator on $\ell^p, 1 < p < \infty$ are weakly convergent, even though it is an isometry.

Strong convergence implies weak convergence, but not vice-versa. More information can be found in [Con90, Chapter 4].

Density For a subset of the positive integers $(n_i)_i \subseteq \mathbb{N}$ we define the upper density as

$$\overline{d}(n_i) = \limsup_{i\to\infty} \frac{\mathrm{card}\{n_j : n_j < i\}}{i}$$

and the lower density as

$$\underline{d}(n_i) = \liminf_{i\to\infty} \frac{\mathrm{card}\{n_j : n_j < i\}}{i}.$$

If these two values are the same then we just say density.

For example, the even numbers have density $\frac{1}{2}$, the primes have density 0 and $\mathbb{N}$ has density 1.

Banach limits A Banach limit is a generalisation of ordinary limits. Define the bounded linear functional $L : c \to \mathbb{C}$ as

$$L(x) := \lim_{n\to\infty} x_n.$$

This is a bounded linear functional that maps sequences to their limit. This functional may be linearly extended to a functional $L : \ell^\infty \to \mathbb{C}$ to give a generalised limit for all bounded sequences. For a glimpse of how this might work, consider how Cesàro means and Abel means are extensions of L, but neither exist for all bounded sequences.

Note the extension is not uniquely defined, so we always speak of *a* Banach limit. The properties you can use are

- L is shift invariant.
- L is positive.
- L maps the constant sequence $(1)_n$ to 1.
- L is a bounded linear functional.

More information about Banach limits can be found in [Con90, Section 3.7].

Banach Lattice A Banach lattice is, among many other things, a Banach space which is also a partially ordered set $(X, \leqslant)$ such that

$$x \vee y := \sup\{x, y\} \text{ and } x \wedge y := \inf\{x, y\}$$

exist for all $x, y \in X$. Similarly we can talk about the supremum or infimum of a set, if they exist. A lattice is called complete is every non-empty subset has a supremum and an infimum. A subset $D \subseteq X$ of a lattice is called $\vee$-stable ($\wedge$-stable) if $a, b \in D$ implies $a \vee b \in D$ ($a \wedge b \in D$).

For a more complete description see [Sch74].