## SEMINAR OF

Proceedings, Universities of Malaga and Seville (Spain), September 2003-June 2004

$$
\lim _{|z| \rightarrow 1} \int_{\overline{\mathbb{D}}} \frac{\left(1-|z|^{q}\right)^{\beta-\alpha}}{|1-\bar{w} z|^{\beta}} d \mu(w)=0 .
$$



Daniel Girela Álvarez<br>Genaro López Acedo<br>Rafael Villa Caro<br>(editors)

# SEMINAR OF MATHEMATICAL ANALYSIS 

Proceedings, Universities of Malaga and Seville (Spain), September 2003-June 2004


Editorial Universidad de Sevilla

Serie: Ciencias

Correspondiente a la primera edición impresa de 2004

COMITÉ EDITORIAL:<br>José Beltrán Fortes<br>(Director de la Editorial Universidad de Sevilla)<br>Manuel Espejo y Lerdo de Tejada<br>Juan José Iglesias Rodríguez<br>Juan Jiménez-Castellanos Ballesteros<br>Isabel López Calderón<br>Juan Montero Delgado<br>Lourdes Munduate Jaca<br>Jaime Navarro Casas<br>$\mathrm{M}^{\mathrm{a}}$ del Pópulo Pablo-Romero Gil-Delgado<br>Adoración Rueda Rueda<br>Rosario Villegas Sánchez

Reservados todos los derechos. Ni la totalidad ni parte de este libro puede reproducirse o transmitirse por ningún procedimiento electrónico o mecánico, incluyendo fotocopia, grabación magnética o cualquier almacenamiento de información y sistemas de recuperación, sin permiso escrito de la Editorial Universidad de Sevilla.

EDICIÓN DIGITAL DE LA PRIMERA EDICIÓN IMPRESA DE 2004
© EDITORIAL UNIVERSIDAD DE SEVILLA 2016
Porvenir, 27 - 41013 Sevilla
Tlfs.: 954487 447; 954487 452; Fax: 954487443
Correo electrónico: eus4@us.es
Web: [http://www.editorial.us.es](http://www.editorial.us.es)
© DANIEL GIRELA ÁLVAREZ
GENARO LÓPEZ ACEDO
RAFAEL VILLA CARO (eds.) 2016
ISBNe: 978-84-472-2050-2
Digitalización y realización interactiva:
Santi García. santi@elmaquetador.es

## Contents

Preface ..... 9
A remark on Carleson measures from $H^{p}$ to $L^{q}(\mu)$ for $0<p<$$q<\infty$11O. Blasco
Subsets of classical Banach spaces failing the fixed point pro- perty ..... 21
P. N. Dowling
Periodic solutions for differential inclusions in Banach spaces ..... 31
Jesús García Falset
Free propagators and Feynman-Kac propagators ..... 47
Archil Gulisashvili
Survey on operators acting on distribution spaces of local regu- larity $\psi(t)$ ..... 65
Silvia I. Hartzstein
Universality in Orlicz spaces ..... 87
Francisco L. Hernández
Projection constants, isometric imbeddings and spherical de- signs ..... 99
Hermann König
Weighted norm inequalities for singular integral operators ..... 119
José María Martell
Weighted norm inequalities and extrapolation ..... 131
José María Martell
Littlewood-Paley-Stein theory for semigroups and its applica- tions to the characterization of Banach spaces ..... 151
Teresa Martínez
Fixed point theory the Picard operators technique ..... 175
Adrian Petruşel
Equilateral sets in finite-dimensional normed spaces ..... 195
Konrad J. SwanepoelAnalytic capacity and Calderón-Zygmund theory with non dou-bling measures239Xavier Tolsa
Discrete and continuous-time models of financial derivatives ..... 273
Hong Kun Xu

## Preface

This volume consists of the lecture notes of the Seminar on Mathematical Analysis corresponding to the period September 2003-June 2004.

The Seminar http://www.us.es/danamate/seminario/indice.htm is held at the Universities of Málaga and Seville, and it was conceived from the main idea of inviting relevant researchers from different fields of Mathematical Analysis.

This Seminar is possible thanks to the public announcement from the Junta de Andalucía for the promotion of research activities jointly organized by different research groups from those belonging to Plan Andaluz de Investigación. The participating groups are FQM-104 (main researcher: Dr. Arias de Reyna Martínez), FQM-127 (main researcher: Dr. Dominguez Benavides) from the University of Seville, and FQM-210 (main researcher: Dr. Girela Álvarez) from the University of Malaga.

The Seminar coordinator is Dr. Lopez Acedo. The organizing board is completed by Dr. Espínola García, Dr. García Vázquez, Dr. Girela Álvarez, Dra. Japón Pineda, Dr. Pérez Moreno and Dr. Villa Caro.

The Editors, Daniel Girela Álvarez<br>Genaro López Acedo<br>Rafael Villa Caro

# A remark on Carleson measures from $H^{p}$ to $L^{q}(\mu)$ for $0<p<q<\infty$ 

Oscar Blasco *


#### Abstract

In this note we shall investigate Carleson measures on the closure of the unit disc $\mathbb{D}$, i.e. finite, positive Borel measures $\mu$ for which the formal identity $J_{\mu}: H^{p} \rightarrow L^{q}(\mu)$ exists, for given values of $0<p<$ $q<\infty$, as a bounded operator from the Hardy space $H^{p}(\mathbb{D})$ into the Lebesgue space $L^{q}(\mu)$.


## 1 Introduction

These notes contain an extended version of the lecture I presented in October of 2003 in Málaga and they are part of the material in a joint paper with Hans Jarchow (see [2]).

We are going to work on the open unit disk $\mathbb{D}=\{z \in \mathbb{C}:|z|<1\}$ in the complex plane, its closure $\overline{\mathbb{D}}$ and the unit circle $\mathbb{T}=\partial \mathbb{D}$. In the sequel, $m$ will be the Haar measure on $\mathbb{T}$ (Borel algebra), so that $d m \equiv d t / 2 \pi$ and $d A(z)$ the normalized area measure. Given a Borel set $B \subseteq \mathbb{T}$, we shall often write $|B|$ instead of $m(B)$. The Lebesgue spaces $L^{p}(m)$ will also be denoted $L^{p}(\mathbb{T}), 0<p \leq \infty$. The canonical norm ( $p$-norm if $0<p<1$ ) on $L^{p}(\mathbb{T})$ is $\|\cdot\|_{p}$.

Let $\mathcal{I}$ be the collection of half-open intervals in $\mathbb{T}$ of the form $I=\left\{e^{i t}\right.$ : $\left.\theta_{1} \leq t<\theta_{2}\right\}$ where $0 \leq \theta_{1}<\theta_{2}<2 \pi$. With each $0 \neq z \in D$, we associate the interval $I(z) \in \mathcal{I}$ such that $|I(z)|=1-|z|$ and $z /|z|$ is the center of $I(z)$. Let $S(z)$ be the half-open Carleson box over $I(z)$ which has $z$ on its 'inner arc'; this inner arc and the boundary part 'to the right' are supposed to belong to $S(z)$. For convenience, let us also put $I(0)=\mathbb{T}, S(0)=\mathbb{D}$, and for any $I \in \mathcal{I}$ we write $S(I)$ the corresponding Carleson box $S\left(z_{I}\right)$ for $z_{I}=\left|z_{I}\right| \zeta_{I}$ where $\zeta_{I}$ is the center of $I$ and $1-\left|z_{I}\right|=|I|$. We shall write

[^0]$P_{z}(\zeta)=\frac{1-|z|^{2}}{|1-z \zeta|^{2}}$ for the Poisson kernel. Clearly one has there exists $C>0$ such that
\[

$$
\begin{equation*}
\frac{\chi_{I(z)}}{1-|z|^{2}} \leq C P_{z} . \tag{1}
\end{equation*}
$$

\]

Let $f: \mathbb{D} \rightarrow \mathbb{C}$ be analytic. For each $0<r<1, f_{r}: \overline{\mathbb{D}} \rightarrow \mathbb{C}: z \mapsto f(r z)$ is continuous, analytic on $\mathbb{D}$, and $M_{p}(f, r):=\left\|f_{r}\right\|_{p}<\infty$ for all $0<p \leq \infty$. The classical Hardy space $H^{p}(\mathbb{D})$ consists of all analytic functions $f: \mathbb{D} \rightarrow \mathbb{C}$ such that $\|f\|_{H^{p}}:=\sup _{r<1} M_{p}(f, r)$ is finite. Again, we get a Banach space if $1 \leq p<\infty$, and a $p$-Banach space if $0<p<1$. The usual Banach space of bounded analytic functions will be denoted by $H^{\infty}(\mathbb{D})$. If $0<q<p<\infty$, then $H^{\infty}(\mathbb{D}) \hookrightarrow H^{p}(\mathbb{D}) \hookrightarrow H^{q}(D)$ continuously with 'norm' one.

Recall that $f$ is in $H^{p}(\mathbb{D})$, for $0<p<\infty$, then $f^{*}\left(e^{i t}\right)=\lim _{r \rightarrow 1} f_{r}\left(e^{i t}\right)$ exists $m$-a.e. on $\mathbb{T}$ (Fatou's Theorem). Moreover, an element $f^{*}$ of $L^{p}(\mathbb{T})$ is generated in this way, and $f \mapsto f^{*}$ defines an isometric embedding $H^{p}(\mathbb{D}) \rightarrow$ $L^{p}(\mathbb{T})$. Its range is the closure $H^{p}(\mathbb{T})$ of the set of polynomials in $L^{p}(\mathbb{T})$. This leads to the identification of $H^{p}(\mathbb{D})$ and $H^{p}(\mathbb{T})$, and to the use of $H^{p}$ as a common symbol.

We shall be investigating Carleson measures on $\overline{\mathbb{D}}$, i.e. finite, positive Borel measures $\mu$ for which the formal identity $J_{\mu}: H^{p} \rightarrow L^{q}(\mu)$ exists, for given values of $0<p<q<\infty$, as a bounded operator from the Hardy space $H^{p}(\mathbb{D})$ into the Lebesgue space $L^{q}(\mu)$.

A characterization of measures on $\mathbb{D}$ for which $J_{\mu}$ is bounded for $p<q$ was obtained by P. Duren, using a modification of the argument given by L. Carleson in the case $p=q$.

Theorem 1. (see [5], page 163) Let $\mu$ be a finite measure on $\mathbb{D}$ and let $0<p<q<\infty$. Then $J_{\mu}: H^{p}(\mathbb{D}) \rightarrow L^{q}(\mu)$ is bounded if and only if

$$
\mu_{D}(S(z)) \leq C \cdot|I(z)|^{q / p} \quad \forall 0 \neq z \in \mathbb{D} .
$$

Examples of measures where $J_{\mu}: H^{p}(\mathbb{D}) \rightarrow L^{q}(\mu)$ is bounded for $p<q$ had appeared, for instance, in the result due to Hardy-Littlewood [8].

Theorem 2. (see [5], page 87) Let $0<p<q<\infty$. Then

$$
\begin{equation*}
\left(\int_{\mathbb{D}}(1-|z|)^{\frac{q}{p}-2}|f(z)|^{q} d A(z)\right)^{1 / q} \leq C\|f\|_{p} \tag{2}
\end{equation*}
$$

for all $f \in H^{p}(\mathbb{D})$.

Another example for such a Carleson measures is given by the embedding from $H^{1}$ into the Bergman space $B^{2}$

$$
\begin{equation*}
\left(\int_{\mathbb{D}}|f(z)|^{2} d A(z)\right)^{1 / 2} \leq C\|f\|_{1} . \tag{3}
\end{equation*}
$$

Let us remark that (3) can also be shown as a consequence of Hardy inequality (see [5], page 48).

We shall present here an alternative proof of Duren's theorem, which does not use ideas in Carleson approach. We shall see that it is actually equivalent to the Hardy and Littlewood result in Theorem 2.

When studying Carleson measures is, sometimes, important to consider not only measures on $\mathbb{D}$ but in $\overline{\mathbb{D}}$. For instance measures concentrated on $\mathbb{T}$, or measures coming from composition operators. We shall denote $\mu_{\mathbb{D}}$ and $\mu_{\mathbb{T}}$ the induced measures on $\mathbb{D}$ and $\mathbb{T}$.

Let $0<p, q<\infty$. A measure $\mu$ on $\overline{\mathbb{D}}$ is called a ( $p, q$ )-Carleson measure if $f \mapsto f$ defines a (linear, bounded) operator

$$
J_{\mu}: H^{p}(\mathbb{D}) \longrightarrow L^{q}(\mu)
$$

If $\mu$ is a $(p, q)$ - Carleson measure then $J_{\mu_{\mathbb{D}}}: H^{p}(\mathbb{D}) \rightarrow L^{q}\left(\mu_{\mathbb{D}}\right): f \mapsto f$ and $J_{\mu_{\mathbb{T}}}: H^{p}(\mathbb{T}) \rightarrow L^{q}\left(\mu_{\mathbb{T}}\right): f^{*} \mapsto f^{*}$ are well-defined operators.

We first observe that this notion only depends on the ratio $p / q$.
Lema 1. (see [2]) Let $\mu$ be a measure on $\overline{\mathbb{D}}$ and let $0<p, s, q<\infty$. Then $\mu$ is $(p, q)$-Carleson measure if and only if $\mu$ is $(s p, s q)$-Carleson measure.

This says that the case $p<q$ and be reduced to $p / q<1$. Our main theorem then establishes the following characterization.
Theorem 3. Let $\mu$ be a finite measure on $\overline{\mathbb{D}}$ and let $0<p<1$. Then the following statements are equivalent:
(i) $J_{\mu}: H^{p}(\mathbb{D}) \rightarrow L^{1}(\mu)$ is bounded if and only if $\mu_{\mathbb{T}}=0$ and

$$
\mu_{\mathbb{D}}(S(z)) \leq C \cdot|I(z)|^{1 / p} \quad \forall 0 \neq z \in \mathbb{D}
$$

(ii) There exists $C>0$ such that

$$
\int_{\mathbb{D}}(1-|z|)^{\frac{1}{p}-2}|f(z)| d A(z) \leq C\|f\|_{p}
$$

for all $f \in H^{p}(\mathbb{D})$.
Direct proofs of (i) and (ii) in Theorem 3 can be found in [5]. The proof of (i) follows the same steps as the one by $p=q$. The proof of (ii) uses factorization, but also it can be achieved by using interpolation (see [6]).

## 2 Proof of Theorem 3

We shall make use of a characterization of Carleson measures in terms of the Poisson kernel. The following lemma is a modification of Lemma 3.3 in [7], see also [1] for a proof.

Lema 2. (see [2]) Let $\mu$ be a finite measure on $\overline{\mathbb{D}}$ and let $0<\alpha<\beta$. Then

$$
\sup \left\{\mu_{\mathbb{D}}(S(z)), \mu_{\mathbb{T}}(I(z))\right\} \leq C \cdot|I(z)|^{\alpha} \quad \forall 0 \neq z \in \mathbb{D}
$$

if and only if

$$
\sup _{|z|<1} \int_{\overline{\mathbb{D}}} \frac{\left(1-|z|^{q}\right)^{\beta-\alpha}}{|1-\bar{w} z|^{\beta}} d \mu(w)<\infty .
$$

Proof of the theorem. (i) $\Rightarrow$ (ii) Consider $d \mu_{\mathbb{D}}(z)=(1-|z|)^{\frac{1}{p}-2} d A(z)$ and $d \mu_{\mathbb{T}}=0$. Clearly
$\mu_{\mathbb{D}}(S(z)) \approx(1-|z|) \int_{|z|}^{1}(1-r)^{\frac{1}{p}-2} d r=(1-|z|) \int_{0}^{1-|z|} s^{\frac{1}{p}-2} d s=\frac{p}{1-p}(1-|z|)^{\frac{1}{p}}$ Hence $\mu$ is a $(p, 1)$-Carleson measure.
(ii) $\Rightarrow$ (i) Let $\mu$ be a $(p, 1)$-Carleson measure. Take $z \in \mathbb{D}$ and $f(w)=$ $\frac{1}{(1-\bar{z} w)^{2 / p}}$. Hence $\|f\|_{p}=\frac{1}{\left(1-|z|^{2}\right)^{1 / p}}$ and the assumption gives that

$$
\int_{\overline{\mathbb{D}}} \frac{1}{|1-\bar{z} w|^{2 / p}} d \mu(w) \leq C \frac{1}{\left(1-|z|^{2}\right)^{1 / p}}
$$

Hence an application of Lemma 2 for $\alpha=1 / p$ and $\beta=2 / p$ shows that

$$
\max \left\{\mu_{\mathbb{D}}(S(z)), \mu_{\mathbb{T}}(I(z))\right\}, \leq C \cdot|I(z)|^{1 / p} \quad \forall 0 \neq z \in \mathbb{D} .
$$

Let us see that $\mu_{\mathbb{T}}=0$
Every open set $\Omega \subseteq \mathbb{T}$ is the union of countably many disjoint intervals $I(z)$ and $p<1$, we may conclude that $\mu_{\mathbb{T}}(\Omega)^{p} \leq C \cdot|\Omega|$. By regularity of these measures, we even get $\mu_{\mathbb{T}}(B)^{p} \leq C \cdot|B|$ for all Borel sets $B \subseteq \mathbb{T}$. In particular, $\mu_{\mathbb{T}} \ll m$ and so $d \mu_{\mathbb{T}}=F d m$ for some $F \in L^{1}(m)$. From Lebesgue differenciation theorem one gets $F(\zeta)=\lim _{|I| \rightarrow 0, \zeta \in I} \frac{1}{|I|} \int_{I} F d m \leq$ $\lim _{|I| \rightarrow 0, \zeta \in I}|I|^{1 / p-1}=0 m$-a.e. This gives $\mu_{\mathbb{T}}=0$.

Conversely, by Lemma 2, we assume that

$$
\sup _{|z|<1} \int_{\mathbb{D}} \frac{\left(1-|w|^{2}\right)^{2-1 / p}}{|1-\bar{w} z|^{2}} d \mu(z)<\infty .
$$

Writting $f(z)=\int_{\mathbb{D}} \frac{f(w)}{(1-\bar{w} z)^{2}} d A(w)$ one has

$$
\begin{aligned}
\int_{\overline{\mathbb{D}}}|f(z)| d \mu(z) & \leq \int_{\overline{\mathbb{D}}}\left(\int_{\mathbb{D}} \frac{|f(w)|}{|1-\bar{w} z|^{2}} d A(w)\right) d \mu(z) \\
& =\int_{\mathbb{D}}\left(\int_{\overline{\mathbb{D}}} \frac{d \mu(z)}{|1-\bar{w} z|^{2}}\right)|f(w)| d A(w) \\
& \leq C \int_{\mathbb{D}}|f(w)|(1-|w|)^{1 / p-2} d A(w) \\
& \leq C\|f\|_{p}
\end{aligned}
$$

## 3 Compactness of $(p, q)$-Carleson measures

We say that a measure $\mu$ on $\overline{\mathbb{D}}$ is a compact $(p, q)$-Carleson measure if the formal identity $J_{\mu}: H^{p}(\mathbb{D}) \rightarrow L^{q}(\mu)$ exists as a compact operator.

As the boundedness, the condition of compactness only depends on $p / q$.
Lema 3. (see[2]) Let $0<p, q, r<\infty$ be given and let $\mu$ be a measure on $\overline{\mathbb{D}}$. Then $\mu$ is a compact $(p, q)$-Carleson measure if and only if $\mu$ is a compact ( $p s, q s$ )-Carleson measure.

Lema 4. (see [2]) Let $\mu$ be a finite measure on $\overline{\mathbb{D}}$ and let $0<\alpha<\beta$. Then

$$
\lim _{|I| \rightarrow 0} \frac{\max \left\{\mu_{\mathbb{D}}(S(I)), \mu_{\mathbb{T}}(I)\right\}}{|I|^{\alpha}}=0
$$

if and only if

$$
\lim _{|z| \rightarrow 1} \int_{\overline{\mathbb{D}}} \frac{\left(1-|z|^{q}\right)^{\beta-\alpha}}{|1-\bar{w} z|^{\beta}} d \mu(w)=0
$$

We now present the proof of the formulation of compact embeddings.
Theorem 4. Let $0<p<q<\infty$ and $\mu$ a measure on $\overline{\mathbb{D}}$. Then $\mu$ is a compact $(p, q)$-Carleson measure if and only if $\mu_{\mathbb{T}}=0$ and

$$
\lim _{|I| \rightarrow 0} \frac{\mu_{\mathbb{D}}(S(I))}{|I|^{q / p}}=0 .
$$

Proof. Take an increasing sequence $r_{n}$ converging to 1. Put $f_{n}(w)=\frac{\left(1-r_{n}^{2}\right)^{\frac{1}{p}}}{\left(1-r_{n} w\right)^{\frac{2}{P}}}$. Hence $\left\|f_{n}\right\|_{p}=1$ for all $n \in \mathbb{N}$. By assumption there exists a subsequence
$f_{r_{n_{k}}}$ convergent in $L^{q}(\mu)$. Note that since the pointwise limit is zero then

$$
\lim _{k \rightarrow \infty} \int_{\mathbb{D}} \frac{\left(1-r_{n_{k}}^{2}\right)^{q / p}}{\left|1-r_{n_{k}} w\right|^{2 q / p}} d \mu(w)=0
$$

The proof of the implication is completed by invoking Lemma 4.
Conversely, assume $\mu_{\mathbb{T}}=0$ and $\lim _{|I| \rightarrow 0} \frac{\mu_{\mathbb{D}}(S(I))}{\mid I I^{q / p}}=0$.
Let us show that $J_{\mu_{\mathbb{D}}}: H^{p / q}(\mathbb{D}) \rightarrow L^{1}\left(\mu_{\mathbb{D}}\right)$ is compact (what is enough invoking Lemma 3).

Lemma 4 gives that for $\epsilon>0$ there exists $\delta>0$ such that, for $1-|z|<\delta$,

$$
\int_{\mathbb{D}} \frac{\left(1-|z|^{2}\right)^{2-q / p}}{|1-\bar{w} z|^{2}} d \mu(z)<\epsilon .
$$

Same argument as in Theorem 3 implies

$$
\begin{aligned}
\int_{\overline{\mathbb{D}}}|f(z)| d \mu(z) & \leq \int_{\overline{\mathbb{D}}} \int_{\mathbb{D}} \frac{|f(w)|}{|1-\bar{w} z|^{2}} d A(w) d \mu(z) \\
& =\int_{\mathbb{D}}\left(\int_{\overline{\mathbb{D}}} \frac{d \mu(z)}{|1-\bar{w} z|^{2}}\right)|f(w)| d A(w) \\
& \leq C \int_{|w| \leq 1-\delta}|f(w)|(1-|w|)^{q / p-2} d A(w) \\
& +C \epsilon \int_{|w|>1-\delta}|f(w)|(1-|w|)^{q / p-2} d A(w) \\
& \leq C \int_{|w| \leq 1-\delta}|f(w)|(1-|w|)^{q / p-2} d A(w)+C \epsilon\|f\|_{p}
\end{aligned}
$$

Let $\left(f_{n}\right)$ be a bounded sequence in $H^{p}(\mathbb{D})$. Then $\left(f_{n}\right)$ is relatively compact in $\mathcal{H}(\mathbb{D})$ and then there exists a subsequence convergent uniformly on compact sets. This and the previous estimates finish the proof.

Corollary 5. Suppose that $p \leq r<q$. Every $(p, q)$-Carleson measure is a compact ( $p, r$ )-Carleson measure.

Let $X$ and $Y$ be quasi-Banach spaces with separating duals. Recall that an operator $u: X \rightarrow Y$ is completely continuous if $\lim _{n}\left\|u x_{n}\right\|_{Y}=0$ holds for every weak null sequence ( $x_{n}$ ) in $X$.

Theorem 6. Let $0<p<1$ and $\mu$ on $\overline{\mathbb{D}}$ be ( $p, 1$ )-Carleson measure. Then $J_{\mu}: H^{p}(\mathbb{D}) \rightarrow L^{1}(\mu)$ is completely continuous.

Proof. Since $\mu_{\mathbb{T}}=0$ then $J_{\mu}$ is now the formal identity $H^{p}(\mathbb{D}) \rightarrow L^{1}\left(\mu_{\mathbb{D}}\right)$ : $f \mapsto f$ since $\mu=\mu_{\mathbb{D}}$. Let $\left(f_{n}\right)$ be a weak null sequence in $H^{p}(\mathbb{D})$. By continuity of point evaluations, $\lim _{n} f_{n}(z)=0 \forall z \in D$. Also, $\left(f_{n}\right)$ is uniformly integrable in $L^{1}\left(\mu_{\mathbb{D}}\right)$ : given $\varepsilon>0$ there is a $\delta>0$ such that if $B \subseteq \mathbb{D}$ is any Borel set with $\mu(B)<\delta$ then $\int_{B}\left|f_{n}\right| d \mu<\varepsilon$ for all $n$. But $f_{n} \rightarrow 0$ pointwise, so that Egorov's Theorem provides us with a Borel set $B \subseteq \mathbb{D}$ such that $\mu(B)<\delta$ and $\lim _{n} f_{n}(z)=0$ uniformly on $\mathbb{D} \backslash B$. Accordingly, there is an $n_{\varepsilon} \in \mathbb{N}$ such that $\int_{\mathbb{D} \backslash B}\left|f_{n}\right| d \mu<\varepsilon$ for $n \geq n_{\varepsilon}$. We conclude that $\left\|f_{n}\right\|_{L^{1}(\mu)}<2 \varepsilon$ for all $n \geq n_{\varepsilon}:\left(f_{n}\right)$ is a null sequence in the Banach space $L^{1}(\mu)$.

Theorem 7. Let $0<p<1$, and let $\mu$ be ( $p, 1$ )-Carleson measure on $\overline{\mathbb{D}}$. Then $J_{\mu}: H^{p}(\mathbb{D}) \rightarrow L^{1}(\mu)$ is a weakly compact operator if and only if $J_{\mu}$ is compact.

Proof. Assume $J_{\mu}$ is weakly compact. The compactness follows essentially by repeating an argument from the proof of Theorem 6. Let $\left(f_{n}\right)$ be a bounded sequence in $H^{p}(\mathbb{D})$. By Montel's Theorem, some subsequence of $\left(f_{n}\right)$ converges locally uniformly to some $f \in \mathcal{H}(\mathbb{D})$. By Fatou's Lemma, $f$ is in $H^{p}(\mathbb{D})$. Therefore it suffices to look at a bounded sequence $\left(f_{n}\right)$ in $H^{p}(\mathbb{D})$ which converges to zero pointwise. By hypothesis and since $\mu_{\mathbb{T}}=0$, $\left(f_{n}\right)$ is uniformly integrable in $L^{1}(\mu)=L^{1}\left(\mu_{\mathbb{D}}\right)$. But $f_{n} \rightarrow 0$ pointwise on $\mathbb{D}$. In combination with Egorov's Theorem this yields $\lim _{n}\left\|f_{n}\right\|_{L^{1}(\mu)}=0$.

## 4 Applications

We shall use the previous results to analyze embedding between Hardy and weighted Bergman spaces. Let $\rho:(0,1] \rightarrow[0, \infty)$ be an integrable function. Let us denote by $A^{p}(\rho)$ the space of analytic functions in the unit disc such that

$$
\int_{\mathbb{D}}|f(z)|^{p} \rho(1-|z|) d A(z)<\infty .
$$

The case $\rho(t)=t^{\alpha p-1}$ is usually denoted $A_{\alpha}^{p}$. The reader is referred to [1] for some results on these spaces.

Theorem 8. Let $0<p<q<\infty$ and let $\rho:(0,1] \rightarrow[0, \infty)$ be an integrable function. Then
(i) $H^{p}(\mathbb{D}) \subset A^{q}(\rho)$ if and only if

$$
\int_{0}^{s} \rho(t) d t \leq C s^{\frac{q-p}{p}}
$$

(ii) $H^{p}(\mathbb{D})$ is compactly contained into $A^{q}(\rho)$ if and only if

$$
\lim _{s \rightarrow 0} s^{\frac{p-q}{p}}\left(\int_{0}^{s} \rho(t) d t\right)=0 .
$$

In particular, $H^{p} \subset A_{\alpha}^{q}$ if and only if $\alpha \geq \frac{1}{p}-\frac{1}{q}$.
$H^{p}$ is compactly contained into $A_{\alpha}^{q}$ if and only if $\alpha>\frac{1}{p}-\frac{1}{q}$.
Proof. (i) Consider $d \mu(z)=\rho(1-|z|) d A(z)$. Using Theorem 1 we have that the condition for the embedding is that

$$
\int_{S(I)} \rho(1-|z|) d A(z)=|I| \int_{1-|I|}^{1} \rho(1-r) d r \leq C|I|^{q / p} .
$$

(ii) Same argument but applying Theorem 4.

## Acknowledgement

I would like to thank Daniel Girela, Genaro López, Tomás Domínguez Benavides and their groups of Málaga and Sevilla for their kind invitation to deliver lectures at both Universities and their hospitality during my visit.

## References

[1] O. Blasco, Operators on weighted Bergman spaces and applications, Duke Math. J., 66 , (1992), 443-467.
[2] O. Blasco, H. Jarchow A note of Carleson measures on Hardy spaces, To appear.
[3] J. Cima, A. Matheson Completely continuous composition operators. Trans. Amer. Math. Soc. 344 (1994), 849-856.
[4] J. Diestel, H. Jarchow, A. Tonge, Absolutely summing operators, Cambrigde Univ. Press, Cambrigde, 1995.
[5] P. Duren, $H^{p}$-spaces Academic Press, New York, 1970.
[6] T.M. Flett, On the rate of growth of mean values of holomorphic and harmonic functions. Proc. London Math. Soc. 20 (1970), 749-768.
[7] J.B. Garnett, Bounded analytic functions, Academic Press, New York, 1981.
[8] G.H. Hardy and J.E. Littlewood, Some properties of fractional integrals II. Math. Z. 34 (1932), 403-439.
[9] H. Hunziker, Kompositionsoperatoren auf klassischen Hardyräumen. Dissertation Universität Zürich (1989).
[10] H. Hunziker, H. Jarchow, Composition operators which improve integrability. Math. Nachr. 152 (1991), 83-99.
[11] H. Jarchow, Compactness properties of composition operators. Rend. Circ. Mat. Palermo, Suppl. 56 (1998), 91-97.
[12] D.H. Luecking, Embedding derivatives of Hardy spaces into Lebesgue spaces. Proc. London Math. Soc. 63 (1991), 565-619.
[13] J.V. Ryff, Subordinate $H^{p}$ functions. Duke Math. J. 33 (1966), 347354.
[14] D. Sarason, Weak compactness of holomorphic composition operators on $H^{1}$. Lecture Notes in Math. 1511 (1992), 75-79.
[15] I.V. Videnskii, On an analogue of Carleson measures. Soviet Math. Dokl. 37 (1988), 186-190.
[16] P. Wojtaszczyk, Banach spaces for analysts, Cambrigde Univ. Press, Cambridge, 1991.
[17] K. Zhu, Operators on Bergman spaces, Marcel Dekker, Inc., New York, 1990.
[18] A. Zygmund, Trigonometric series, Cambrigde Univ. Press, New York, 1959.

## Author's address:

O. Blasco

Departamento de Análisis Matemático
Universidad de Valencia 46100-Burjassot, Valencia, Spain. E-mail: oscar.blasco@uv.es

# Subsets of classical Banach spaces failing the fixed point property 

Patrick N. Dowling


#### Abstract

In this expository note, we consider the classical nonreflexive spaces $c_{0}, \ell^{1}, L^{1}[0,1]$ and $C[0,1]$ and try to identify subsets of these spaces that fail the fixed point property for nonexpansive mappings.


## 1 Introduction

A large portion of metric fixed point theory revolves around the following type of problem:
Let $K$ be a closed bounded convex non-empty subset of a Banach space $X$. Let $T: K \rightarrow K$ be a nonexpansive mapping; that is, $\|T x-T y\| \leq\|x-y\|$ for all $x, y \in K$. Does $T$ have a fixed point; that is, does there exist a point $x_{0} \in K$ so that $T x_{0}=x_{0}$ ?

If $K$ is a closed bounded convex non-empty subset of a Banach space $X$ with the property that every nonexpansive mapping $T: K \rightarrow K$ has a fixed point, then we say that $K$ has the fixed point property. A Banach space $X$ is said to have the fixed point property if every closed bounded convex non-empty subset of $X$ has the fixed point property. A Banach space $X$ is said to have the weak fixed point property if every weakly compact convex non-empty subset of $X$ has the fixed point property. It is well known that uniformly convex spaces and, more generally, reflexive Banach spaces with normal structure have the fixed point property [8]. Also well known is that Schur spaces and spaces with the uniform Kadec-Klee property have the weak fixed point property [8]. In particular, the space $\ell^{1}$, and the Hardy space $H^{1}$, have the weak fixed point property but they both fail to have the fixed point property.

The aim of this short note is to illustrate the difficulties that one encounters when one tries to determine whether or not a closed bounded convex non-empty subset of a Banach space has the fixed point property in the

[^1]classical Banach spaces $c_{0}, \ell^{1}, L^{1}[0,1]$ and $C[0,1]$, each equipped with their canonical norm. It is not our intent to be encyclopedic, but rather to focus on selected examples and short proofs that are key to understanding the fixed point property in these spaces.

## 2 The failure of the fixed point property in $\ell^{1}$

In $\ell^{1}$, the Banach space of absolutely summable sequences of real numbers, the standard example of a closed bounded convex non-empty subset is the following

Example 1. Let $\left(e_{n}\right)$ denote the canonical unit vector basis of $\ell^{1}$ and define

$$
K=\left\{\sum_{n=1}^{\infty} t_{n} e_{n}: t_{n} \geq 0 \text { for all } n \geq 1, \text { and } \sum_{n=1}^{\infty} t_{n}=1\right\}
$$

$K$ is clearly a closed bounded convex non-empty subset of $\ell^{1}$. Now define $T: K \rightarrow K$ by

$$
T\left(\sum_{n=1}^{\infty} t_{n} e_{n}\right)=\sum_{n=1}^{\infty} t_{n} e_{n+1}
$$

It is trivial to check that $T$ is a nonexpansive fixed point free mapping on $K$.
Counterbalancing this example is the fact that $\ell^{1}$, equipped with it's canonical norm, has the weak*-uniform Kadec-Klee property, and hence closed bounded convex non-empty which are weak* compact have the fixed point property - in particular, the closed unit ball of $\ell^{1}$ has the fixed point property [8]. Therefore a closed bounded convex non-empty subset of $\ell^{1}$ may fail the fixed point property, but it will be contained in a closed bounded convex superset (for example, a multiple of the unit ball of $\ell^{1}$ ) that has the fixed point property.

The most important feature to note about the set $K$ in Example 1, is that the sequence $\left(e_{n}\right)$, consisting of the unit vector basis elements of $\ell^{1}$, lies in $K$, and this sequence converges weak ${ }^{*}$ to 0 , which is not in $K$. Of course, this means that $K$ is not weak* (sequentially) compact. Roughly speaking, all closed bounded convex non-weak* compact subsets of $\ell^{1}$ exhibit the type of behavior seen in Example 1. To make this statement more specific, we need a theorem of Dowling, Lennard and Turett [3].

Theorem 2. Let $X$ be a Banach space with a norm $\|\cdot\|$, and let $K$ be a closed bounded convex subset of $X$. Let $\left(\varepsilon_{n}\right)$ be a null sequence in $(0,1)$. If $K$ contains a sequence $\left(x_{n}\right)$ such that

$$
\sum_{n=1}^{\infty}\left(1-\varepsilon_{n}\right)\left|t_{n}\right| \leq\left\|\sum_{n=1}^{\infty} t_{n} x_{n}\right\| \leq \sum_{n=1}^{\infty}\left(1+\varepsilon_{n}\right)\left|t_{n}\right|,
$$

for all $\left(t_{n}\right) \in \ell^{1}$, then $K$ contains a closed convex subset $C$ such that there is a nonexpansive affine mapping $T: C \rightarrow C$ which fails to have a fixed point in $C$.

Remark 3. The sequence $\left(x_{n}\right)$ in Theorem 2 is often refereed to as an asymptotically isometric $\ell^{1}$-basic sequence. The reason for this is because $\left(x_{n}\right)$ is equivalent to the unit vector basis of $\ell^{1}$ (in particular, $\left(x_{n}\right)$ is equivalent to the sequence ( $e_{n}$ ) in Example 1) with the equivalence constants approach 1 as $n \rightarrow \infty$.

It is proved in [2], that if $K$ is a closed bounded convex subset of $L^{1}[0,1]$ which is not weakly compact, then $K$ contains a sequence such that a translate of this sequence by a certain fixed element of $L^{1}[0,1]$ is a multiple of an asymptotically isometric $\ell^{1}$-basic sequence. Combining this with Theorem 2 , we obtain the following result.

Corollary 4. If $K$ is a closed bounded convex subset of $L^{1}[0,1]$ which is not weakly compact, then $K$ contains a closed convex non-empty subset $C$ such that there is a nonexpansive affine mapping $T: C \rightarrow C$ which fails to have a fixed point in $C$.

This corollary can now be used to characterize weak compactness of closed bounded convex subsets of $L^{1}[0,1]$ (equipped with it's canonical norm).

Corollary 5. Let $K$ be a closed bounded convex subset of $L^{1}[0,1]$. Then the following are equivalent;
(a) $K$ is weakly compact,
(b) Every closed convex subset of $K$ has the fixed point property for continuous affine self maps,
(c) Every closed convex subset of $K$ has the fixed point property for nonexpansive affine self maps.

Remark 6. The word affine cannot be dropped from Corollary 5 because of Alspach's example of a closed bounded convex non-empty subset of $L^{1}[0,1]$ which is weakly compact but fails the fixed point property for nonexpansive mappings (see [1]). The Banach space $\ell^{1}$ is isometric to a subspace of $L^{1}[0,1]$, so Corollary 5 can be used to characterize weak compactness in $\ell^{1}$. However, in $\ell^{1}$ compactness and weak compactness are equivalent, so no Alspach-like examples exist in $\ell^{1}$. As a result we can easily obtain the following result.

Corollary 7. A closed bounded convex non-empty subset $K$ of $\ell^{1}$ is compact if and only if every closed convex non-empty subset of $K$ has the fixed point property for nonexpansive self maps.

In summary, if $K$ is a closed bounded convex non-empty subset of $\ell^{1}$, then $K$ has the fixed point property if it is weak* compact; if $K$ is not weak* compact, then while $K$ may or may not have the fixed point property, it will have a subset that fails the fixed point property.

## 3 The failure of the fixed point property in $c_{0}$

In $c_{0}$, the Banach space of sequences of real numbers converging to 0 , weakly compact convex non-empty sets have the fixed point property, by the celebrated result of Maurey [10]. In particular, if $\left(e_{n}\right)$ is the canonical unit vector basis of $c_{0}$, then, since $\left(e_{n}\right)$ is a weakly null sequence, the closed convex hull of the sequence ( $e_{n}$ ), $\overline{\operatorname{co}}\left\{e_{n}: n \geq 1\right\}$, is weakly compact and so has the fixed point property. On the other hand, let $\left(s_{n}\right)$ is the canonical summing basis of $c_{0}$; that is, $s_{n}=e_{1}+e_{2}+\cdots+e_{n}$. Note that if we denote the closed convex hull of $\left(s_{n}\right)$ by $K$, then

$$
K=\left\{\sum_{n=1}^{\infty} t_{n} e_{n}: 1=t_{1} \geq t_{2} \geq t_{3} \geq \cdots \geq 0\right\}
$$

One can easily see that the mapping $T: K \rightarrow K$ defined by

$$
T\left(\sum_{n=1}^{\infty} t_{n} e_{n}\right)=e_{1}+\sum_{n=1}^{\infty} t_{n} e_{n+1}
$$

is a nonexpansive fixed point free mapping on $K$. Moreover, if $K_{1}$ is any subset of $c_{0}$ that contains $K$, then $K_{1}$ also fails the fixed point property. To see this, do the following:

For all $u=\left(u_{1}, u_{2}, \cdots\right) \in c_{0}$, let $u^{*}=\left(u_{1}^{*}, u_{2}^{*}, \cdots\right)$ be the decreasing rearrangement of $u$. First we note that $u^{*} \in c_{0}$. Secondly, for each $n \in \mathbb{N}$ define $\tilde{u_{n}}=u_{n}^{*} \wedge 1$. Finally, define $T: c_{0} \rightarrow c_{0}$ by

$$
S(u)=\left(1, \tilde{u_{1}}, \tilde{u_{2}}, \tilde{u_{3}}, \ldots\right)=e_{1}+\sum_{n=1}^{\infty} \tilde{u_{n}} e_{n+1} .
$$

Note that $S(u) \in K$, for all $u \in c_{0}$, and the mapping $T$ defined above is equal to the mapping $S$, when restricted to $K$. Therefore, if $K_{1}$ is any subset of $c_{0}$ which contains $K$ as a subset, then the mapping $S$ maps $K_{1}$ into $K_{1}$. Since the range of $S$ is contained in $K$, any fixed point $S$ must be an element of $K$. However, $S=T$ on $K$, so a fixed point of $S$ must also be a fixed point of $T$, and since $T$ has no fixed points, neither does $S$. This example (and some variants of it) can be found in the paper of Llorens-Fuster and Sims [9].

In many respects, the behavior exhibited by the above example is canonical in $c_{0}$. To see this we first need to define the notion of an asymptotically isometric $c_{0}$-summing basic sequence.

Definition 8. A sequence $\left(w_{n}\right)$ in a Banach space $X$ is an asymptotically isometric $c_{0}$-summing basic sequence if there exists a null sequence $\left(\varepsilon_{n}\right)$ in $(0, \infty)$ such that

$$
\sup _{n}\left(\frac{1}{1+\varepsilon_{n}}\right)\left|\sum_{j=n}^{\infty} t_{j}\right| \leq\left\|\sum_{n=1}^{\infty} t_{n} w_{n}\right\| \leq \sup _{n}\left(1+\varepsilon_{n}\right)\left|\sum_{j=n}^{\infty} t_{j}\right|
$$

for all $\left(t_{n}\right) \in c_{00}$, the space of finitely non-zero sequences.
Remark 9. If, in the above definition, we let $\varepsilon_{n}=0$ for all $n$, then the sequence ( $w_{n}$ ) is behaving exactly like the summing basis $\left(s_{n}\right)$ in $c_{0}$. The concept of an asymptotically isometric $c_{0}$-summing basic sequence was introduced by Dowling, Lennard and Turett in [5]. In that paper (which is quite technical), the authors proved that if $K$ is a closed bounded convex non-empty subset of a Banach space $X$, and $K$ contains an asymptotically isometric $c_{0}$-summing basic sequence, then $K$ contains a closed bounded convex non-empty subset which fails the fixed point property. They also proved that every closed bounded convex non-empty subset of $c_{0}$ (equipped with its canonical norm) which is not weakly compact contains a sequence which is a multiple of an asymptotically isometric $c_{0}$-summing basic sequence. Consequently, if $K$ is a closed bounded convex non-weakly compact subset of $c_{0}$,
then $K$ contains a closed bounded convex non-empty subset that fails the fixed point property. Building on this in a subsequent paper [6], Dowling, Lennard and Turett proved that if $K$ is a closed bounded convex non-weakly compact subset of $c_{0}$, then $K$ fails the fixed point property. Therefore, by combining this result with the result of Maurey [10], we obtain the following complete characterization of the fixed point property in $c_{0}$.

Theorem 10. If $K$ is a closed bounded convex non-empty subset of $c_{0}$, equipped with its canonical norm, then $K$ has the fixed point property if and only if $K$ is weakly compact.

## 4 The failure of the fixed point property in $L^{1}[0,1]$

The well known example of Alspach [1] shows that the Banach space $L^{1}[0,1]$ fails to have the weak fixed point property. In fact, by building on Alspach's example, one can show that any subset of $L^{1}[0,1]$ with non-empty interior fails the fixed point property. Even more can be said: any subset of $L^{1}[0,1]$ that contains a non-trivial order interval fails the fixed point property. We will give a proof of this result in the special case of subsets of $L^{1}[0,1]$ containing the order interval determined by the identically 0 element and the identically 1 element (the set $C$ below). First, we need to recall some of the details of Alspach's construction. Let

$$
C=\left\{f \in L^{1}[0,1]: 0 \leq f(t) \leq 1, \text { for all } t \in[0,1]\right\}
$$

and define $T: C \rightarrow C$ by

$$
T f(t)= \begin{cases}\min \{2 f(2 t), 1\} & \text { for } 0 \leq t \leq \frac{1}{2} \\ \max \{2 f(2 t-1)-1,0\} & \text { for } \frac{1}{2}<t \leq 1\end{cases}
$$

for all $f \in C$.
Alspach showed that the mapping $T$ is an isometry on $C$ which has two fixed points; namely 0 and $\chi_{[0,1]}$. The mapping $T$ is an isometric self map of the closed convex subset $C_{0}=\left\{f \in C: \int_{[0,1]} f=1 / 2\right\}$ of $C$, and therefore $T$ has no fixed points in $C_{0}$ since $C_{0}$ contains neither of the points 0 nor $\chi_{[0,1]}$.

Alspach's example was modified by R. Sine [11], to produce a fixed point free nonexpansive mapping on all of $C$. This modification is achieved as follows; define $S: C \rightarrow C$ by $S(f)=\chi_{[0,1]}-f$, for all $f \in C$. The mapping $S$ is clearly an isometry of $C$ onto $C$. Thus the mapping $S T$ is a nonexpansive mapping on $C$. Sine proved that $S T$ is fixed point free on $C$. Using Sine's result we obtain the following result which appears in [4].

Theorem 11. Let $K$ be a closed bounded convex subset of $L^{1}[0,1]$ which contains the order interval $C=\left\{f \in L^{1}[0,1]: 0 \leq f(t) \leq 1\right.$, for all $t \in$ $[0,1]\}$. Then $K$ fails the fixed point property for nonexpansive mappings.

Proof. Define the mapping $R: K \rightarrow K$ by

$$
R f(t)=\min \{|f(t)|, 1\}, \text { for } 0 \leq t \leq 1, \text { for all } f \in K
$$

It is easily seen that $R$ is a nonexpansive mapping on $K$ and $R(f) \in C$ for all $f \in K$. Now define $U: K \rightarrow K$ by

$$
U(f)=S T(R(f)), \text { for all } f \in K
$$

The mapping $U$ is nonexpansive since all of the mappings, $R, S$, and $T$ are nonexpansive.

To show that $U$ is fixed point free, suppose that $f \in K$ is a fixed point of $U$, that is, $U(f)=f$. Since $f \in K, R(f) \in C$, and since $S T$ maps $C$ into $C, f=U(f)=S T(R(f)) \in C$. Note that the mapping $R$ restricted to $C$ is the identity on $C$. Therefore, $f=S T(R(f))=S T(f)$ and so $f$ is a fixed point of $S T$ in $C$. This contradicts Sine's result that $S T$ has no fixed point in $C$ [11]. This completes the proof.

Remark 12. One should note that even though we stated Theorem 11 for a closed bounded convex subset $K$ containing the order interval $\left\{f \in L^{1}[0,1]\right.$ : $0 \leq f(t) \leq 1$, for all $t \in[0,1]\}$, the proof only requires $K$ to contain the order interval and neither the closedness, boundedness nor convexity of $K$ is used.

The proof of the more general result that any subset of $L^{1}[0,1]$ that contains a non-trivial order interval fails the fixed point property, is not difficult but it is somewhat technical. An improvement of this result (with an even more technical proof) appears in a recent paper of Dowling, Lennard and Turett [7], where it is shown that every subset of $L^{1}[0,1]$ that contains the (non-trivial) intersection of an order interval and a hyperplane fails to have the fixed point property for nonexpansive mappings.

## 5 The failure of the fixed point property in $C[0,1]$

Since $C[0,1]$ contains every separable Banach space isometrically, it is not surprising that very little can be said about the fixed point property for a general closed bounded convex non-empty subset of $C[0,1]$. However, we can obtain a result that looks very similar to the result stated in Theorem 11.

Our main ingredient is a very slight modification of an example from the text of Goebel and Kirk [8, page 30]. Let

$$
C=\{f \in C[0,1]: 0 \leq f(t) \leq 1, \text { for all } t \in[0,1]\}
$$

and define $T: C \rightarrow C$ by

$$
T f(t)=\min \{1, \max \{0, f(t)+2 t-1\}\}, \text { for } 0 \leq t \leq 1,
$$

for all $f \in C$. It is trivial to show that $T$ is nonexpansive. Also, since $T(f)(t)>f(t)$ for some $t>1 / 2$ or $T(f)(t)<f(t)$ for some $t<1 / 2, T$ fails to have a fixed point.
Theorem 13. Let $K$ be a closed bounded convex subset of $C[0,1]$ which contains the set $C=\{f \in C[0,1]: 0 \leq f(t) \leq 1$, for all $t \in[0,1]\}$. Then $K$ fails the fixed point property for nonexpansive mappings.
Proof. Define the mapping $R: K \rightarrow K$ by

$$
R f(t)=\min \{|f(t)|, 1\}, \text { for } 0 \leq t \leq 1, \text { for all } f \in K
$$

It is easily seen that $R$ is a nonexpansive mapping on $K$ and $R(f) \in C$ for all $f \in K$. Now define $U: K \rightarrow K$ by

$$
U(f)=T(R(f)), \text { for all } f \in K
$$

The mapping $U$ is nonexpansive since the mappings $R$ and $T$ are nonexpansive.

To show that $U$ is fixed point free, suppose that $f \in K$ is a fixed point of $U$, that is, $U(f)=f$. Since $f \in K, R(f) \in C$, and since $T$ maps $C$ into $C, f=U(f)=T(R(f)) \in C$. Note that the mapping $R$ restricted to $C$ is the identity on $C$. Therefore, $f=T(R(f))=T(f)$ and so $f$ is a fixed point of $T$ in $C$. This contradicts the fact that $T$ has no fixed point in $C$ and so the proof is complete.

Remark 14. Again we note, just as we did after Theorem 11, neither the boundedness, closedness nor the convexity of $K$ is used in the proof of Theorem 13.

## Acknowledgement

This paper is based on a series of talks given at the Universidad de Sevilla during February, 2003. The author wishes to express his gratitude to Tomas Domínguez Benavides, María Japón and Genaro López for the invitation to visit the Universidad de Sevilla and present these results in their seminar, and for the wonderful hospitality he received during his stay.

## References

[1] D. Alspach, A fixed point free nonexpansive mapping, Proc. Amer. Math. Soc., 82 (1981), 423-424.
[2] P.N. Dowling, The fixed point property for subsets of $L_{1}[0,1]$, Contemporary Math. 232 (1999), 131-137.
[3] P.N. Dowling, C.J. Lennard and B. Turett, The fixed point property for subsets of some classical Banach spaces, Nonlinear Anal., 49 (2002), 141-145.
[4] P.N. Dowling, C.J. Lennard and B. Turett, Some more examples of subsets of $c_{0}$ and $L^{1}[0,1]$ failing the fixed point property, Contemporary Math., 328 (2003), 171-176.
[5] P. N. Dowling, C. J. Lennard, and B. Turett, Characterizations of weakly compact sets and new fixed point free maps in $c_{0}$, Studia Math., 154 (2003), no. 3, 277-293.
[6] P.N. Dowling, C.J. Lennard and B. Turett, Weak compactness is equivalent to the fixed point property in $c_{0}$, Proc. Amer. Math. Soc., 132 (2004), 1659-1666.
[7] P.N. Dowling, C.J. Lennard and B. Turett, New fixed point free nonexpansive maps on weakly compact, convex subsets of $L^{1}[0,1]$, preprint.
[8] Kazimierz Goebel and W.A. Kirk, Topics in metric fixed point theory, Cambridge University Press, Cambridge, 1990.
[9] E. Llorens-Fuster and B. Sims, The fixed point property in $c_{0}$, Canad. Math. Bull., 41 (1998), no. 4, 413-422.
[10] B. Maurey, Points fixes des contractions de certains faiblement compacts de Lscalarproduct1, Seminar on Functional Analysis, 1980-1981, École Polytech., Palaiseau, 1981, pp. Exp. No. VIII, 19.
[11] R. Sine, Remarks on an example of Alspach, Nonlinear Anal. and Appl., Marcel Dekker, (1981), 237-241.

## Author's address:

Patrick N. Dowling<br>Department of Mathematics and Statistics<br>Miami University,<br>Oxford, OH 45056<br>E-mail: dowlinpn@muohio.edu

# Peridic solutions for differential inclusions in Banach spaces 

Jesús García Falset *


#### Abstract

The purpose of this note is to prove the existence of periodic solutions for certain type of differential inclusions posed in Banach spaces with the fixed point property for nonexpansive mappings (FPP for short). In particular, we obtain that the equation $$
\left.u_{t}-\Delta u+|u|^{s-1} u=f \quad \text { in }\right] 0, \infty[\times \Omega,
$$ where $s>1, \Omega$ is a bounded open domain of $\mathbb{R}^{n}$ with smooth boundary and there exists $1<p<\infty$ such that $f(t, x)$ is $w$-periodic (in $t$ ) and $f \in L^{p}(] 0, w[\times \Omega)$ admits a weak $w$-periodic solution.


## 1 Introduction

In this paper we show that if $X$ is a real Banach space with the FPP and $A \subset X \times X$ is an $m$-accretive operator on $X$ with $\overline{D(A)}$ convex, then the existence of a bounded integral solution on $\mathbb{R}^{+}$of the differential inclusion

$$
\begin{equation*}
u^{\prime}(t)+A u(t) \ni f(t) \tag{1}
\end{equation*}
$$

where $f$ is $w$-periodic and $f \in L_{l o c}^{1}(0, \infty, X)$, is equivalent to the existence of a $w$-periodic integral solution of such inclusion (some results of this type were known see [2], [7],[12] and [13]). Thus, if one wants to know if Problem 1 admits a $w$-periodic integral solution it will be enough to study if it has a bounded integral solution. In this sense, we deal with a kind of accretive operators (see Definition 5 below ) for which it is possible to know when Problem 1 has a bounded integral solution, this fact allows us to extend to reflexive Banach spaces with the FPP the well known result of periodic

[^2]solutions for maximal monotonous operators on a Hilbert space (see [7]). Moreover, using these results we obtain the existence of periodic solutions of the problem
\[

\left\{$$
\begin{array}{l}
\left.u_{t}-\Delta u+|u|^{\alpha-1} u=f, \text { on }\right] 0, \infty[\times \Omega, \\
u(t, x)=0, \text { on }[0, \infty[\times \partial \Omega,
\end{array}
$$\right.
\]

where $\Omega$ is a bounded open domain in $\mathbb{R}^{n}$ with smooth boundary $\partial \Omega, f(t, x)$ is a given $L^{p}$-function on $] 0, w[\times \Omega, 1<p<\infty$, periodic in $t$ with period $w$, and $\alpha>1$.

## 2 Preliminaries

Throughout this note we assume that $X$ is a real Banach space, denote by $B_{X}$ its closed unit ball, by $X^{\star}$ the dual space of $X$, and by $2^{X}$ the collection of subsets of $X$. A mapping $A: X \rightarrow 2^{X}$ will be called an operator on $X$. The domain of $A$ is denoted by $D(A)$ and its range by $R(A)$. We define the normalized duality mapping by

$$
J(x):=\left\{j \in X^{\star}:\langle x, j\rangle=\|x\|^{2},\|j\|=\|x\|\right\} .
$$

Let $\langle y, x\rangle_{+}:=\max \{\langle y, j\rangle: j \in J(x)\}$.
An operator $A$ on $X$ is accretive if and only if $\langle u-v, x-y\rangle_{+} \geq 0$ for all $x, y \in D(A)$ and for each $u \in A(x), v \in A(y)$. If in addition $R(I+\lambda A)$ is precisely $X$ for all $\lambda>0$, then $A$ is said to be $m$ - accretive. Accretive operators were introduced by F.E. Browder [8] and T. Kato [14] independently (we refer the reader to [3],[6], [9] for background material on accretivity).

Let $X$ be a Banach space, let $A: D(A) \rightarrow 2^{X}$ be an $m$-accretive operator and $f \in L_{\text {loc }}^{1}(0, \infty, X)$. If we consider the following initial value problem:

$$
\left\{\begin{array}{l}
u^{\prime}(t)+A(u(t)) \ni f(t)  \tag{2}\\
u(0)=x_{0}
\end{array}\right.
$$

We say that a continuous function $u:[0,+\infty[\rightarrow X$ is a integral solution of (2) if $u(0)=x_{0}$ and moreover the inequality

$$
\|u(t)-x\|^{2}-\|u(s)-x\|^{2} \leq 2 \int_{s}^{t}\langle f(\tau)-y, u(\tau)-x\rangle_{+} d \tau
$$

holds whenever $0 \leq s \leq t$, and $(x, y) \in A$.

It is well known that the Problem 2 has a unique integral solution for each $x_{0} \in \overline{D(A)}$. This concept of solution was introduced by Bénilan in [4] and perhaps one of the most important results involving integral solution is:

Theorem 1. Let $X$ be a Banach space and let $A$ be m-accretive operator on $X$. Then:
(a) For every $x \in \overline{D(A)}$, the initial-value problem $u^{\prime}+A u \ni f, u(0)=x$, has a unique integral solution $u:\left[0,+\infty\left[\rightarrow \overline{D(A)}\right.\right.$ whenever $f \in L_{L o c}^{1}(0, \infty, X)$.
(b) If $u$ and $v$ are integral solutions of, respectively, $u^{\prime}+A u \ni f$ and $v^{\prime}+A v \ni g$ on $[0, \infty[$, then

$$
\|u(t)-v(t)\|^{2}-\|u(s)-v(s)\|^{2} \leq 2 \int_{s}^{t}\langle f(\tau)-g(\tau), u(\tau)-v(\tau)\rangle_{+} d \tau
$$

for $0 \leq s \leq t$.
An integral solution of Problem 2 cannot be interpreted as a solution of Cauchy problem in a pointwise sense, they are not strong solutions. However, every strong solution is an integral solution and moreover, under certain additional assumptions one may obtain more regularity of integral solutions.

We denote by $B V_{l o c}(0, \infty, X)$ the subspace formed by those functions in $L_{l o c}^{1}(0, \infty, X)$ which are of bounded variation on each compact subset of $[0, \infty[$. From Theorem 2.2 page 131 and Remark 1 page 133 of [3] we have:

Theorem 2. Let $X$ be a Banach space with Radon-Nikodym property, $f \in$ $B V_{l o c}(0, \infty, X)$. If $u$ is an integral solution of Problem 2 and $x_{0} \in D(A)$. Then $u$ is a strong solution of this problem.

Finally, if $X$ is a real Banach space, a self-mapping $T$ of a nonempty subset $C$ of $X$ is said to be a nonexpansive mapping if $\|T x-T y\| \leq\|x-y\|$, for all $x, y \in C$.

We say that $X$ has the weak fixed point property (WFPP for short)(resp. fixed point property (FPP in short)) if for each nonempty weakly compact convex (resp. bounded, closed, convex) set $K \subset X$ and each nonexpansive mapping $T: K \rightarrow K$ there exists an element $x \in K$ such that $x=T x$. Of course, both conditions coincide when the Banach space $X$ is reflexive. It is well known that under nice conditions of geometric type on the norm of $X$, the WFPP or the FPP can be guaranteed. Among many others, uniform convexity, uniform smoothness and normal structure along with the reflexivity are examples of such conditions (see [10], [11] and references within).

## 3 The main result

We are going to present a result that gives us the relationship between the existence of bounded integral solution and the existence of a $w$-periodic integral solution.

Theorem 3. Let $X$ be a Banach space with the WFPP. If $A: D(A) \rightarrow 2^{X}$ is an m-accretive operator such that $\overline{D(A)}$ is convex and locally weakly compact and $f \in L_{\text {Loc }}^{1}(0, \infty, X)$ is $w$-periodic. Then the initial value problem

$$
u^{\prime}(t)+A u(t) \ni f(t)
$$

has a w-periodic integral solution if and only if it has a bounded integral solution.

Proof. The necessary condition is obvious since a continuous periodic function is bounded. Conversely we consider $x \in \overline{D(A)}$, then by Theorem 1(a) the problem

$$
\begin{equation*}
u^{\prime}(t)+A u(t) \ni f(t) \quad u(0)=x \tag{3}
\end{equation*}
$$

has a unique integral solution $u:[0,+\infty[\rightarrow \overline{D(A)}$.
The structure of the proof follows several steps.
First step. Let us show that $v:[0,+\infty[\rightarrow \overline{D(A)}$ defined by $v(t)=$ $u(w+t)$ is the unique integral solution of:

$$
\begin{equation*}
v^{\prime}(t)+A v(t) \ni f(t) \quad v(0)=u(w) \tag{4}
\end{equation*}
$$

Indeed, given $0 \leq s \leq t$ and $\left(x^{\prime}, y\right) \in A$, by the definition of integral solution, we have to study if

$$
\begin{equation*}
\left\|v(t)-x^{\prime}\right\|^{2}-\left\|v(s)-x^{\prime}\right\|^{2} \leq 2 \int_{s}^{t}\left\langle f(\tau)-y, v(\tau)-x^{\prime}\right\rangle_{+} d \tau \tag{5}
\end{equation*}
$$

Notice that since $u$ is the integral solution of (3) and $v(t)=u(t+w)$, we have

$$
\begin{aligned}
\left\|u(w+t)-x^{\prime}\right\|^{2} & -\left\|v(s+w)-x^{\prime}\right\|^{2} \leq 2 \int_{s+w}^{t+w}\left\langle f(\tau)-y, u(\tau)-x^{\prime}\right\rangle_{+} d \tau= \\
& =2 \int_{s}^{t}\left\langle f(w+\tau)-y, u(w+\tau)-x^{\prime}\right\rangle_{+} d \tau
\end{aligned}
$$

Now we may use that $f$ is $w$-periodic and thus we derive that

$$
\left\|v(t)-x^{\prime}\right\|^{2}-\left\|v(s)-x^{\prime}\right\|^{2} \leq 2 \int_{s}^{t}\left\langle f(\tau)-y, v(\tau)-x^{\prime}\right\rangle_{+} d \tau
$$

Which means that $v$ is the integral solution of (4).
Second. We consider the Poincaré mapping $P: \overline{D(A)} \rightarrow \overline{D(A)}$ defined by $P x=u_{x}(w)$ where $x \in \overline{D(A)}$ and $u_{x}$ is the integral solution of (3). The mapping $P$ is well defined from the uniqueness of integral solutions for the Problem 3.

Notice that $P$ is nonexpansive. Let $x, y \in \overline{D(A)}$, then by Theorem 1 (b)

$$
\begin{aligned}
\|P x-P y\|^{2} & =\left\|u_{x}(w)-u_{y}(w)\right\|^{2} \\
& \leq\left\|u_{x}(0)-u_{y}(0)\right\|^{2}+2 \int_{0}^{t}\left\langle f(\tau)-f(\tau), u_{x}(\tau)-u_{y}(\tau)\right\rangle_{+} d \tau
\end{aligned}
$$

And therefore

$$
\|P x-P y\| \leq\|x-y\| .
$$

Third. Let us see that there exists a nonempty convex weakly compact subset $C$ which is $P$-invariant.

By hypothesis we know that Problem 1 has a bounded solution $u_{x_{0}}$. Consider now the sequence $\left(P^{n}\left(x_{0}\right)\right)$, by the first step of the proof for each $n \in \mathbb{N}$ it is clear that $P^{n}\left(x_{0}\right)=u_{x_{0}}(n w)$, which means that the sequence $\left\{P^{n}\left(x_{0}\right)\right\}$ is bounded. This allows us to say that

$$
R:=\limsup _{n \rightarrow \infty}\left\|P^{n}\left(x_{0}\right)-x_{0}\right\|<\infty
$$

Thus, if we take the subset

$$
C:=\left\{x \in \overline{D(A)} \cap B\left(x_{0}, 2 R\right): \limsup _{n \rightarrow \infty}\left\|P^{n}\left(x_{0}\right)-x\right\| \leq R\right\} .
$$

Clearly, $C$ is convex and weakly compact.
Let us see that $C$ is also $P$-invariant. Indeed, if $y \in C$ then

$$
\limsup _{n \rightarrow \infty}\left\|P^{n}\left(x_{0}\right)-P y\right\| \leq \limsup _{n \rightarrow \infty}\left\|P^{n-1}\left(x_{0}\right)-y\right\| \leq R,
$$

and moreover,

$$
\left\|P y-x_{0}\right\| \leq \limsup _{n \rightarrow \infty}\left\|P^{n}\left(x_{0}\right)-x_{0}\right\|+\limsup _{n \rightarrow \infty}\left\|P^{n}\left(x_{0}\right)-P y\right\| \leq 2 R .
$$

So, since $X$ enjoys the WFPP, $P$ has a fixed point on $C$. Then there exists an integral solution $u$ of the problem $u^{\prime}+A u \ni f$ which satisfies $u(0)=u(w)$ and again by the first step $u$ is $w$-periodic.

Corollary 4. (i). Let $X$ be a Banach space with the WFPP. If $A \subset X \times X$ is an $m$-accretive operator such that $\overline{D(A)}$ is convex and weakly compact and $f \in L_{L o c}^{1}(0, \infty, X)$ is w-periodic . Then the problem $u^{\prime}(t)+A u(t) \ni f(f)$ has an integral w-periodic solution.
(ii). Let $X$ be a Banach space with the FPP. If $A \subset X \times X$ is an $m$-accretive operator such that $\overline{D(A)}$ is convex and $f \in L_{L o c}^{1}(0, \infty, X)$ is $w$ periodic. Then the problem $u^{\prime}(t)+A u(t) \ni f(f)$ has an integral $w$-periodic solution if and only if it has a bounded integral solution.
(iii). Let $X$ be a Banach space with the FPP. If $A \subset X \times X$ is an $m$-accretive operator such that $\overline{D(A)}$ is convex and bounded. Then the problem $u^{\prime}(t)+A u(t) \ni f(f)$ has an integral w-periodic solution whenever $f \in L_{\text {Loc }}^{1}(0, \infty, X)$ and it is $w$-periodic.

## 4 Boundedness of solutions

Let $X$ be a Banach space, and let $A \subset X \times X$ be an $m$-accretive operator on $X$ such that $\overline{D(A)}$ is convex and $0 \in R(A)$. If we consider the initial value problem (2) with $f \in L^{1}(0, \infty, X)$, then the integral solutions of Problem 2 are bounded (see [15]). However, it fails when $f \in L_{l o c}^{1}(0, \infty, X)$.

In this section we are going to introduce a kind of accretive operators for which it will be possible to know when the initial value Problem 2 admits bounded integral solutions.

Definition 5. Let $X$ be a Banach space, an accretive operator $A \subset X \times X$ is said to be coercive if there exist $\gamma:\left[0,+\infty\left[\rightarrow\left[0,+\infty\left[\right.\right.\right.\right.$ with $\lim _{t \rightarrow \infty} \gamma(t)=\infty$ such that for each $(x, y) \in A$ we have that $\langle y, x\rangle_{+} \geq \gamma(\|x\|)\|x\|$.

Example 6. Let $H$ be a Hilbert space. Given $\alpha, \beta \geq 0$ we define the following operator:

$$
\begin{aligned}
& A: H \longrightarrow 2^{H} \\
& x \quad \mapsto \quad A(x)= \begin{cases}x\|x\|^{\alpha}+\beta \frac{x}{\|x\|}, & x \neq 0 \\
\beta B_{X}, & x=0 .\end{cases}
\end{aligned}
$$

Let us show that $A$ is an $m$-accretive and coercive operator on $H$.
If we take $\gamma:\left[0, \infty\left[\rightarrow\left[0, \infty\left[\right.\right.\right.\right.$ such that $\gamma(t)=t^{\alpha+1}+\beta$ it is easy to see that $\langle y, x\rangle_{+} \geq \gamma(\|x\|)\|x\|$.

To see that $A$ is accretive, it suffices to consider the argument developed in [13]. Thus the only thing that we have to prove is that $R(I+A)=X$. Indeed,

If $\|z\| \leq \beta$ by the definition of $A$ it is clear that $z \in 0+A 0$.

Then we may assume that $\|z\|>\beta$. In this case, define the function:

$$
\begin{array}{rll}
f:[0,+\infty[ & \longrightarrow & \mathbb{R} \\
t & \mapsto & f(t)=t\left(1+t^{\alpha}\right)
\end{array}
$$

Since $f$ is continuous there exists a unique $\left.t_{0} \in\right] 0, \infty\left[\right.$ such that $f\left(t_{0}\right)=$ $\|z\|-\beta$.

Now it is enough to take $x=\frac{t_{0}}{t_{0}+t_{0}^{\alpha+1}+\beta} z$ to see that $z=x+A x$.
The following result is interesting in order to have a method for obtaining coercive $m$-accretive operators.

Proposition 7. Let $X$ be a Banach space if $A \subset X \times X$ is an m-accretive operator with $0 \in A 0$. Then, for each $\alpha>0$ the new operator $B_{\alpha}:=I+\alpha A$ is coercive.

Proof. Consider $(x, y) \in B_{\alpha}$, then by definition there exists $u \in A x$ such that $y=x+\alpha u$, therefore

$$
\langle y, x\rangle_{+}=\|x\|^{2}+\alpha\langle u, x\rangle_{+} .
$$

Since by hypothesis $0 \in A 0$ and $A$ is accretive, it is clear that $\langle u-0, x-0\rangle_{+} \geq$ 0 . Consequently, it is enough to obtain the result to take $\gamma(t)=t$.

Example 8. Let $X$ be a Banach space and consider $B: X \rightarrow X$ defined by $B x=x+\frac{x}{1+\|x\|}$, then to see that $A$ is an $m$-accretive and coercive operator it is enough to apply the above proposition, since it is not difficult to see that the operator $A x=\frac{x}{1+\|x\|}$ is $m$-accretive on $X$.
Remark 9. An important concept of solution of the following initial value problem

$$
\begin{equation*}
u^{\prime}(t)+A u(t) \ni f(t) \quad u(0)=x_{0} \quad 0 \leq t \leq T \tag{6}
\end{equation*}
$$

can be found in page 134 of [3].
Definition 10. A function $u \in \mathcal{C}([0, T], X)$ is called a weak solution of the initial value problem (6) if there are sequences $\left(u_{n}\right) \subset W^{1, \infty}(0, T, X)$ and $\left(f_{n}\right) \subset L^{1}(0, T, X)$ such that
(a) $\frac{d}{d t} u_{n}(t)+A u_{n}(t) \ni f_{n}(t)$ a.e. $\left.t \in\right] 0, T[, n=1,2, \ldots$
(b) $\lim _{n \rightarrow \infty} u_{n}(t)=u(t)$ uniformly on $[0, T]$.
(c) $u(0)=x_{0}$ and $\lim _{n \rightarrow \infty} f_{n}=f$ in $L^{1}(0, T, X)$.

It is easy to see that if $u$ is a weak solution then it is an integral solution. Moreover, in page 134 of [3] we can find an important result involving this kind of solutions:

Corollary 11. Let $X$ be reflexive and let $A \subset X \times X$ be an m-accretive operator on $X$. Then for each $x_{0} \in \overline{D(A)}$, the initial value problem (6) has a unique weak solution.

As a consequence of the above Remark and Theorem 3 we obtain the following result.

Corollary 12. Let $X$ be a reflexive Banach space with the F.P.P.. If $A$ : $D(A) \rightarrow 2^{X}$ is an m-accretive and coercive operator with $\overline{D(A)}$ convex and $f \in L_{L o c}^{1}(0, \infty, X)$ is w-periodic. Then the initial value problem

$$
\begin{equation*}
u^{\prime}(t)+A u(t) \ni f(t) \tag{7}
\end{equation*}
$$

has a w-periodic integral solution.
Proof. By Theorem 3 it will be enough to see that the integral solutions of (7) are bounded. To see this, consider $x_{0} \in \overline{D(A)}$ and we take $u_{x_{0}}$ the unique integral solution of the problem

$$
\left\{\begin{array}{l}
u^{\prime}(t)+A(u(t)) \ni f(t) t \in[0, \infty[ \\
u(0)=x_{0} .
\end{array}\right.
$$

From Corollary 11 the initial value problem

$$
\left\{\begin{array}{l}
u^{\prime}(t)+A(u(t)) \ni f(t) \quad t \in[0, w] \\
u(0)=x_{0}
\end{array}\right.
$$

has a unique weak solution and therefore this weak solution will be $u_{x_{0}}$. This means that there exist sequences $\left(u_{n}\right)$ and $\left(f_{n}\right)$ as in Definition 10.

Since $f \in L_{L o c}^{1}(0, \infty, X)$, we know that $M:=\int_{0}^{w}\|f(t)\| d t<\infty$.
On the other hand, Since $A$ is coercive there exists $R>0$ such that $\langle y, x\rangle_{+}>\frac{M}{w}\|x\|$ whenever $(x, y) \in A$ and $\|x\| \geq R$.

Let us show that for every $n \in \mathbb{N}$ and for each $x \in D(A)$ the inequality

$$
\left\|u_{x_{0}}(n w)\right\| \leq K:=\max \left\{\left\|u_{x_{0}}(0)\right\|, R+M+|A x| w+2\|x\|\right\}
$$

holds, where $|A x|:=\inf \{\|u\|: u \in A x\}$.
Given $\epsilon>0$ there exists $n_{0} \in \mathbb{N}$ such that

$$
\int_{0}^{w}\left\|f_{n_{0}}(t)\right\| d t \leq M+\epsilon \text { and }\left\|u_{n_{0}}(t)-u_{x_{0}}(t)\right\| \leq \epsilon, \quad \forall t \in[0, w]
$$

If $\left\|u_{n_{0}}(t)\right\|>R$ for every $t \in[0, w]$, since

$$
\left.f_{n_{0}}(t)-u_{n_{0}}^{\prime}(t) \in A u_{n_{0}}(t), \text { a.e. } t \in\right] 0, w[,
$$

we have that

$$
\left.\left\langle f_{n_{0}}(t)-u_{n_{0}}^{\prime}(t), u_{n_{0}}(t)\right\rangle_{+} \geq \frac{M}{w}\left\|u_{n_{0}}(t)\right\| \text { a.e. } t \in\right] 0, w[.
$$

Therefore, there exists $j(t) \in J u_{n_{0}}(t)$ such that

$$
\left.\left\langle f_{n_{0}}(t)-u_{n_{0}}^{\prime}(t), j(t)\right\rangle \geq \frac{M}{w}\left\|u_{n_{0}}(t)\right\| \text { a.e. } t \in\right] 0, w[.
$$

Now using the differentiation rule of Kato, we have

$$
\left\|f_{n_{0}}(t)\right\|\left\|u_{n_{0}}(t)\right\| \geq\left\langle f_{n_{0}}(t), j(t)\right\rangle \geq\left\|u_{n_{0}}(t)\right\| \frac{d}{d t}\left\|u_{n_{0}}(t)\right\|+\frac{M}{w}\left\|u_{n_{0}}(t)\right\|
$$

a.e. $t \in] 0, w[$, therefore

$$
M+\epsilon \geq\left\|u_{n_{0}}(w)\right\|-\left\|u_{n_{0}}(0)\right\|+M
$$

Otherwise, let $\left.t_{0} \in\right] 0, w\left[\right.$ such that $\left\|u_{n_{0}}\left(t_{0}\right)\right\| \leq R$, if $t \geq t_{0}$ since for each $(x, y) \in A$ we have

$$
\left\|u_{n_{0}}(t)-x\right\|^{2} \leq\left\|u_{n_{0}}\left(t_{0}\right)-x\right\|^{2}+2 \int_{t_{0}}^{t}\left\langle f_{n_{0}}(\tau)-y, u_{n_{0}}(\tau)-x\right\rangle_{+} d \tau
$$

thus by a variant of Gronwal's lemma we derive that

$$
\left\|u_{n_{0}}(t)-x\right\| \leq R+\|x\|+\int_{t_{0}}^{t}\left\|f_{n_{0}}(\tau)-y\right\| d \tau \leq R+M+\epsilon+\|x\|+\|y\| w
$$

Note that the above argument implies that $\left\|u_{x_{0}}(w)\right\| \leq K$.
Suppose now that this has been proved for

$$
\left\{u_{x_{0}}(0), u_{x_{0}}(w), u_{x_{0}}(2 w), \ldots, u_{x_{0}}((n-1) w)\right\} .
$$

Consider now the following initial value problem

$$
\left\{\begin{array}{l}
u^{\prime}(t)+A(u(t)) \ni f(t) \quad t \in[0, n w] \\
u(0)=x_{0} .
\end{array}\right.
$$

It is clear that $u_{x_{0}}$ is the unique weak solution of the above problem, then we may argue as before to obtain that there exist sequences $\left(v_{n}\right)$ and $\left(g_{n}\right)$ as in Definition 10 and thus, given $\epsilon>0$ there exists $n_{0} \in \mathbb{N}$ such that

$$
\int_{(n-1) w}^{n w}\left\|g_{n_{0}}(t)\right\| d t \leq M+\epsilon
$$

(Notice that this is a consequence of the $w$-periodicity of $f$.) Moreover,

$$
\left\|v_{n_{0}}(t)-u_{x_{0}}(t)\right\| \leq \epsilon \quad \forall t \in[0, n w] .
$$

If $\left\|v_{n_{0}}(t)\right\| \geq R$ for all $t \in[(n-1) w, n w]$, then it is easy to see that

$$
\left\|u_{x_{0}}(n w)\right\| \leq\left\|u_{x_{0}}((n-1) w)\right\| \leq K .
$$

Otherwise, let $\left.s_{0} \in\right](n-1) w, n w\left[\right.$ such that $\left\|v_{n_{0}}\left(s_{0}\right)\right\| \leq R$, if $s \geq s_{0}$ we can write,

$$
\left\|v_{n_{0}}(s)-x\right\| \leq\left\|v_{n_{0}}\left(s_{0}\right)-x\right\|+\int_{s_{0}}^{s}\left\|g_{n_{0}}(\tau)-y\right\| d \tau
$$

And then

$$
\left\|v_{n_{0}}(s)-x\right\| \leq\left\|v_{n_{0}}\left(s_{0}\right)-x\right\|+\int_{(n-1) w}^{n w}\left\|g_{n_{0}}(\tau)-y\right\| d \tau \leq R+M+\epsilon+\|x\|+\|y\| w
$$

Consequently

$$
\left\|u_{x_{0}}(n w)\right\| \leq K .
$$

This means that the sequence $\left\{u_{x_{0}}(n w)\right\}$ is bounded.
Finally, we will show that the function $u_{x_{0}}($.$) is also bounded. Otherwise$ we may find $\left.t_{0} \in\right] m w,(m+1) w\left[\right.$ satisfying $\left\|u_{x_{0}}\left(t_{0}\right)\right\|>K+2 M$.

As a consequence of the above argument we may assume $\left\|u_{x_{0}}(t)\right\|>R$ for all $t \in\left[m w, t_{0}\right]$, however in this case applying the differentiation rule of Kato, as above, we arrive to

$$
\|u(t)\| \leq\|u(m w)\|+M
$$

which is a contradiction.
Corollary 13. Let $X$ be a reflexive Banach space with the $F P P$ and let $A \subset X \times X$ be an $m$-accretive and coercive operator on $X$ with $\overline{D(A)}$ convex. Then given $w>0$ and $f \in L^{1}(0, w, X)$ the initial value problem

$$
\begin{equation*}
u^{\prime}(t)+A u(t) \ni f(t), \quad t \in[0, w] \tag{8}
\end{equation*}
$$

has a weak solution $u$ such that $u(0)=u(w)$.

Proof. Consider $\tilde{f}$ the $w$-periodic extension of $f$ to $\mathbb{R}^{+}$. It is clear that Corollary 12 yields a $w$-periodic integral solution $u$ of the problem

$$
u^{\prime}(t)+A u(t) \ni \tilde{f}(t), \quad t \in[0, \infty[.
$$

Therefore by Remark $9, u$ is a weak solution of the initial value problem

$$
u^{\prime}(t)+A u(t) \ni f(t), \quad t \in[0, w]
$$

which satisfies that $u(0)=u(w)$.
Corollary 14 (see [7]). Let $H$ be a real Hilbert space. If $A \subset H \times H$ is maximal monotone, coercive and $\left.f \in L^{1}[0, w], H\right)$, then Problem 8 has an weak solution $u$ such that $u(0)=u(w)$.

Example 15. Let $H$ be a Hilbert space and $A$ the operator of Example (6) and let $f \in L_{l o c}^{1}(0, \infty, H)$ be a w-periodic function. Then the differential inclusion $u^{\prime}(t)+A u(t) \ni f(t) \quad t \geq 0$, has a w-periodic integral solution.

If $f \in B V_{\text {loc }}(0, \infty, H)$, by Theorem 2, it admits a w-periodic strong solution (see example of [13]).

Example 16. Let $X$ be a reflexive Banach space with the FPP. Then the differential equation $u^{\prime}(t)+u(t)+\frac{u(t)}{1+\|u(t)\|}=f(t), t \geq 0$, has a $w$-periodic integral solution whenever $f \in L_{L o c}^{1}(0, \infty, X)$ and moreover it is w-periodic.

If $f \in B V_{\text {loc }}(0, \infty, X)$, by Theorem 2, it admits a w-periodic strong solution.

## 5 Application

In the last decades, periodic partial differential equations have been the subject of extensive study (see for example [16] and references within). This section is concerned with the time periodic solutions of the problem

$$
\left\{\begin{array}{l}
\left.u_{t}-\Delta u+|u|^{\alpha-1} u=f, \text { on }\right] 0, \infty[\times \Omega,  \tag{9}\\
u(t, x)=0, \text { on }[0, \infty[\times \partial \Omega, \\
u(0, x)=u_{0}(x), \text { on } \Omega
\end{array}\right.
$$

where $\alpha>1, \Omega$ is a bounded open domain in $\mathbb{R}^{n}$ with smooth boundary $\partial \Omega$, $f(t, x)$ is a given $L^{p}$-function on $] 0, w[\times \Omega, 1<p<\infty$, periodic in $t$ with period $w$.

Given $\alpha>1$, let $\beta_{\alpha}$ be the following maximal monotone subset of $\mathbb{R} \times \mathbb{R}$

$$
\begin{aligned}
\beta_{\alpha}: & \mathbb{R} \\
t & \longrightarrow \mathbb{R} \\
& \beta_{\alpha}(t)=|t|^{\alpha} \operatorname{sign}_{0}(t),
\end{aligned}
$$

where, as usual, we denote by $\operatorname{sign} n_{0}$ the following function:

$$
\begin{aligned}
\operatorname{sign}_{0}: & \mathbb{R}
\end{aligned} \quad \longrightarrow \mathbb{R} .
$$

Given $1<p<\infty$, let $\tilde{\beta_{\alpha}} \subset L^{p}(\Omega) \times L^{p}(\Omega)$ be the operator defined by

$$
D\left(\tilde{\beta}_{\alpha}\right)=\left\{u \in L^{p}(\Omega) ; \exists v \in L^{p}(\Omega), \text { such that } v(x) \in \beta_{\alpha}(u(x)) \text { a. e. on } \Omega\right\}
$$

$$
\tilde{\beta_{\alpha}}(u)=\left\{v \in L^{p}(\Omega) ; \quad v(x) \in \beta_{\alpha}(u(x)) \text { a. e. on } \Omega\right\}
$$

From [5] it is clear that $\tilde{\beta_{\alpha}}$ is an $m$ accretive operator on $L^{p}(\Omega)$.
On the other hand, let $B \subset L^{p}(\Omega) \times L^{p}(\Omega)$ be the operator defined by

$$
B u=-\Delta u \quad \forall u \in D(B)
$$

where $D(B)=W_{0}^{1, p}(\Omega) \cap W^{2, p}(\Omega)$. The argument developed in [3] show that if we define the operator $A:=B+\tilde{\beta}_{\alpha}$, where $D(A)=D(B) \cap D\left(\tilde{\beta_{\alpha}}\right)$, then $A$ is a $m$-accretive operator on $L^{p}(\Omega)$.

Then the problem (9) may be rewritten as

$$
\left\{\begin{array}{l}
u^{\prime}(t)+A u(t)=f(t), \quad 0<t<\infty, \\
u(0)=u_{0} .
\end{array}\right.
$$

where $u($.$) is regarded as a function from \left[0, \infty\left[\right.\right.$ to $L^{p}(\Omega)$.
Proposition 17. Let $A \subset L^{p}(\Omega) \times L^{p}(\Omega)$ be the above $m$-accretive operator. Then $A$ is coercive in $L^{p}(\Omega)$.

Proof. First, we will prove that for each $u \in D(A)$ the inequality

$$
\begin{equation*}
\|u\|_{p} \leq\|u\|_{p+\alpha-1}(\mu(\Omega))^{\frac{\alpha-1}{p(p+\alpha-1)}} \tag{10}
\end{equation*}
$$

holds. Indeed, since $u \in D(A)$, then $u \in D\left(\tilde{\beta_{\alpha}}\right)$ and therefore $v=|u|^{\alpha} \operatorname{sign}_{0}(u) \in$ $L^{p}(\Omega)$ which means that $|u|^{p \alpha} \in L^{1}(\Omega)$.

Now, since $p>1, \alpha>1$ and $\Omega$ is bounded we have that $u \in L^{p+\alpha-1}(\Omega)$, and hence $u^{p} \in L^{\frac{p+\alpha-1}{p}}(\Omega)$, thus using the Hölder's inequality we obtain that

$$
\int_{\Omega}|u|^{p} \leq\left(\int_{\Omega}\left(|u|^{p}\right)^{\frac{\alpha+p-1}{p}}\right)^{\frac{p}{\alpha+p-1}}(\mu(\Omega))^{\frac{\alpha-1}{\alpha+p-1}}
$$

and this implies the inequality (10).
Let us show that $A$ is also coercive. Indeed, since $0=B 0$, and $0=\tilde{\beta_{\alpha}}(0)$, given $u \in D(A)$ we have

$$
\langle A u, u\rangle_{+}=\langle A u-A 0, u-0\rangle_{+} \geq 0 .
$$

The fact that $L^{p}(\Omega)$ is a uniformly smooth Banach space yields

$$
\langle A u, u\rangle_{+}=\langle B u-0, u-0\rangle+\left\langle\tilde{\beta_{\alpha}}(u)-0, u-0\right\rangle
$$

Moreover the normalized duality mapping on $L^{p}(\Omega)$ has the form

$$
J(u)=\|u\|_{p}^{2-p}|u|^{p-1} \operatorname{sign}_{0}(u)
$$

Then if we call $M=(\mu(\Omega))^{\frac{1-\alpha}{p(p+\alpha-1)}}$, we deduce, from the inequality (10), that $\int_{\Omega}|u|^{p+\alpha-1} \geq\|u\|_{p}^{p+\alpha-1} M$, and consequently, since $B$ is an accretive operator, we have that

$$
\langle A u, u\rangle_{+} \geq\left\langle\tilde{\beta_{\alpha}}(u), u\right\rangle=\|u\|_{p}^{2-p} \int_{\Omega}|u|^{p+\alpha-1} \geq\|u\|_{p}^{2-p}\|u\|_{p}^{p+\alpha-1} M=M\|u\|_{p}^{\alpha+1} .
$$

Thus, if we define $\gamma(t)=M t^{\alpha}$ we derive that

$$
\langle A u, u\rangle_{+} \geq \gamma\left(\|u\|_{p}\right)\|u\|_{p}
$$

which means that $A$ is coercive.
Since $L^{p}(\Omega)$ is a uniformly convex Banach space and $A \subset L^{p}(\Omega) \times L^{p}(\Omega)$ is $m$-accretive then $\overline{D(A)}$ is convex, which allows us to obtain along with Proposition 17 and Corollary 12 the following result:

Theorem 18. let $\Omega$ be a bounded and open domain in $\mathbb{R}^{n}$ with smooth boundary $\partial \Omega$, and let $A \subset L^{p}(\Omega) \times L^{p}(\Omega)$ with $1<p<\infty$, the above $m$ accretive operator. Then the initial value problem

$$
u^{\prime}(t)+A u(t)=f(t), \quad t \in[0, \infty[
$$

has a w-periodic integral solution whenever $f \in L_{l o c}^{1}\left(0, \infty, L^{p}(\Omega)\right)$ and it is $w$-periodic.

At this point, we would like to remark that following [1] and [5] it is possible to obtain the same conclusion as above if in Equation (9) we replace $\Delta u$ by $p$-Laplacian operator of $u$. Thus this allows us to compare the above theorem with the problem solved in [16].

### 5.1 Remarks.

Checking carefully the proof of Corollary 12 we may relax its hypothesis in the following sense:

Corollary 19. Let $X$ be a reflexive Banach space with the FPP, if $A \subset$ $X \times X$ is $m$-accretive with $\overline{D(A)}$ convex, $f \in L_{l o c}^{1}(0, \infty, X)$ is w-periodic and $M=\int_{0}^{w}\|f(t)\| d t$. Then the initial value problem

$$
u^{\prime}(t)+A u(t) \ni f(t)
$$

has a w-periodic integral solution whenever there exists $R>0$ such that if $(x, y) \in A$ with $\|x\|>R$, then $\langle y, x\rangle_{+}>\frac{M}{w}\|x\|$.
Example 20. Let $X$ be a reflexive Banach space with the FPP. Then the differential equation $u^{\prime}(t)+\frac{u(t)}{1+\|u(t)\|}=f(t) t \geq 0$, has an integral $w$-periodic solution whenever $f \in L_{\text {Loc }}^{1}(0, \infty, X)$ and moreover it is $w$-periodic and $\frac{1}{w} \int_{0}^{w}\|f(t)\| d t<1$.

As a particular case of this fact, the same conclusion holds when

$$
\sup \{\|f(t)\|: t \geq 0\}<1
$$

(see [12]).
By using the same argument of [13] it is easy to see that Theorem 2 of [13] can be generalized in the following way:
Theorem 21 ([13]). Let $X$ be a reflexive Banach space with the FPP. Let $F: \mathbb{R}^{+} \times X \rightarrow C V(X)$ be upper semi continuous, bounding such that for all $(t, x, y) \in \mathbb{R}^{+} \times X \times X,\langle F(t, x)-F(t, y), x-y\rangle_{+} \geq 0$, and moreover $F(t+w, x)=F(t, x),(w>0)$. Then the initial value problem

$$
\begin{equation*}
x^{\prime}(t)+F(t, x(t)) \ni 0 \tag{11}
\end{equation*}
$$

has a w-periodic strong solution if and only if it has a bounded strong solution on $[0, \infty[$.

Where $C V(X)$ denotes the collection non-empty compact convex subsets of $X$.
$F$ is bounding if it maps bounded subsets of $X$ into bounded subsets.
$F$ is upper semi-continuous in $X$ if for every $x_{0} \in X$ and every open set $G$ with $F x_{0} \subset G$ there exists a neighborhood $U$ of $x_{0}$ such that $F x \subset G$ for all $x \in U$.

## References

[1] Andreu F., Mazón J.M., Segura de León, Toledo J. Quasi-linear elliptic and parabolic equations in $L^{1}$ with nonlinear boundary conditions, Adv. Math. Sci. Appl. 7, (1997), 183-213.
[2] Baillon J., Haraux A., Comportement à l'infinie pour les équations d'évolution avec forcing périodique, Arch. Rat. Mec. Anal. 67 (1977), 101-109.
[3] V. Barbu, Nonlinear semigroups and differential equations in Banach spaces, Noordhoff International Publishing (1976).
[4] Ph. Bénilan, Equations d'évolution dans un espace de Banach quelconque et applications, Thesis Orsay, (1972).
[5] Ph. Bénilan, M.G. Crandall, Completely accretive operators, Semigroups theory and evolution equations Ph. Clément, E. Mitidieri and Ben de Payter (Edts) LN in Pure Applied Math. 135 Marcel Dekker, Inc. (1991).
[6] Ph. Bénilan, M.G. Crandall, A. Pazy, Evolution equations governed by accretive opertors. Forthcoming.
[7] H. Brezis, Operateurs Maximaux Monotones et semigroupes de contractions dans les espaces de Hilbert, Noth-Holland (1973).
[8] F. E. Browder, Nonlinear accretive operators in Banach spaces, Bull. Amer. Math. Soc. 73 (1967), 470-476.
[9] I. Cioranescu, Geometry of Banach spaces, duality mappings and nonlinear problems, Kluwer Academic Publishers (1990).
[10] K. Goebel, W.A. Kirk, Topics in metric fixed point theory Cambridge Univ. Pres. (1990).
[11] Handbook of Metric Fixed Point Theory, Edited by W.A. Kirk and B. Sims, Kluwer Academic Publishers (2001).
[12] E. Hanebaly, Un théorème du point fixe et solutions périodiques d'equations différentielles V-dissipatives, Fixed point theory and applications, Edit. par J.B.Baillon, M.A. Thera (1991), 221-230.
[13] E. Hanebaly, B. Marzouki, Periodic and almost periodic solutions for multi-valued differential equations in Banach spaces, Electronic J. of Differential Equations 2000 (2000) No 24, 1-16.
[14] T. Kato, Nonlinear semigroups and evolution equations, J. Math. Soc. Japan 19 (1967), 508-520.
[15] K. Kobayashi, I. Miyadera, On the asymptotic behavior of almost-orbits of nonlinear contraction semigroups in Banach spaces, Nonlinear Analysis 6 (1982), 349-365.
[16] Wang Yi-fu, Periodic solutions of Porous medium equations with weakly nonlinear sources, Northeast. Math. J. 16(4) (2000), 475-483.

## Author's address:

J. García Falset<br>Departamento de Análisis Matemático<br>Universitat de Valencia<br>Dr. Moliner 50<br>46100, Burjassot, Valencia, Spain<br>E-mail: garciaf@uv.es

# Free propagators and Feynman-Kac propagators 

Archil Gulisashvili


#### Abstract

In this note, backward free propagators associated with transition probability densities and Feynman-Kac propagators corresponding to time-dependent families of measures are studied. We are interested in the following problem: Determine in what form the properties of backward free propagators are inherited by backward Feynman-Kac propagators. The inheritance problem is studied in the present note in the case of the boundedness in $L^{r}$ and the boundedness in various spaces of continuous functions.


## 1 Introduction

Propagators are two-parametric generalizations of semigroups. The term "propagator" is not standard. Various other names were used to label these objects (evolution families, time-inhomogeneous semigroups, etc.) In the present note, we survey some of the results obtained in [7-11]. We begin with a general definition of a propagator. Let $B$ be a Banach space, and denote by $L(B, B)$ the space of all bounded linear operators on $B$.

Definition 1. A forward propagator on a Banach space $B$ is a two-parametric family, $S=\{S(t, \tau) \in L(B, B): 0 \leq \tau \leq t<\infty\}$, such that $S(t, \lambda) S(\lambda, \tau)=$ $S(t, \tau)$ for all $\tau \leq \lambda \leq t$, and $S(t, t)=I$ for all $0 \leq t<\infty$.

A backward propagator on $B$ is a two-parametric family of operators, $Q=\{Q(\tau, t) \in L(B, B): 0 \leq \tau \leq t<\infty\}$, such that $Q(\tau, t)=Q(\tau, \lambda) Q(\lambda, t)$ for all $\tau \leq \lambda \leq t$, and $Q(t, t)=I$ for all $0 \leq t<\infty$.

If such a family is defined only for $0 \leq \tau \leq t \leq T$ where $T>0$, then we will assume that the propagator properties hold for $0 \leq \tau \leq t \leq T$.

The assumptions in the definition of a forward propagator are called the flow conditions, while backward propagators satisfy the backward flow conditions. In this note, we will study backward propagators. This choice is not discriminating, since there is a simple link between propagators and
backward propagators. In the case of a finite time-interval $[0, T]$, forward and backward propagators are connected by time reversal $\eta(t)=T-t$. For instance, if $S(t, \tau)$ with $0 \leq \tau \leq t \leq T$ is a propagator on $B$, then $Q(t, \tau)=S(\eta(t), \eta(\tau))$ is a backward propagator on $B$.

We will say that a backward propagator $Q$ is strongly continuous if for every $x \in B$, the function $(\tau, t) \rightarrow Q(\tau, t) x$ is continuous. A backward propagator $Q$ is called uniformly bounded if

$$
\|Q(\tau, t)\|_{B \rightarrow B} \leq M
$$

where $M$ does not depend on $\tau$ and $t$. If for every compact subset $K$ of the set $0 \leq \tau \leq t<\infty$ we have

$$
\|Q(\tau, t)\|_{B \rightarrow B} \leq M_{K}
$$

for all $(\tau, t) \in K$, then $Q$ will be called locally uniformly bounded. A backward propagator $Q$ is called separately strongly continuous if for every fixed $t$ and $x \in B$, the function $\tau \rightarrow Q(\tau, t) x$ is continuous on $[0, t]$, and for every fixed $\tau$ and $x \in B$, the function $t \rightarrow Q(\tau, t) x$ is continuous on $[\tau, T]$ (if $T=\infty$, then we consider the interval $[t, \infty)$ instead of the interval $[t, T]$ ). Similar definitions can be given in the forward case.

The following assertion obtained in [11] shows that the joint strong continuity and the separate strong continuity of propagators are equivalent in the presence of the local uniform boundedness:

Theorem 2. For a backward propagator $Q$ on $B$, the following are equivalent:
(i) The strong continuity.
(ii) The strong separate continuity and the uniform local boundedness.

The same result holds for forward propagators.
A simplest example of a propagator can be obtained from a semigroup $S_{t}$ on a Banach space $B$. Consider the family of operators given by $\left\{S_{t-\tau}\right\}$. Then it is simultaneously a forward and a backward propagator on $B$.

## 2 Backward propagators generated by transition functions

Various examples of backward propagators arise in the theory of Markov processes. Let $E$ be a locally compact second countable Hausdorff topological space. It is known that $E$ is $\sigma$-compact and metrisable. We will fix a
metric $\rho: E \times E \rightarrow[0, \infty)$ generating the topology of $E$, and denote by $\mathcal{E}$ the $\sigma$-algebra of Borel subsets of $E$.

Let $P(r, x ; s, A)$, where $0 \leq r<s<\infty, x \in E$, and $A \in \mathcal{E}$, be a transition probability function. This means that the following conditions hold:

1. For fixed $r, s$, and $A, P$ is a nonnegative Borel measurable function on $E$.
2. For fixed $r, s$, and $x, P$ is a Borel measure on $\mathcal{E}$.
3. $P(r, x ; s, E)=1$ for all $r, s$, and $x$.
4. $P(r, x ; s, A)=\int_{E} P(r, x ; u, d y) P(u, y ; s, A)$ for all $r<u<s$, and $A$.

Given a transition probability function $P$, we can define a family of contraction operators on the space $L_{\mathcal{E}}^{\infty}$ of bounded Borel functions on $E$ by

$$
\left\{\begin{array}{l}
Y(\tau, t) f(x)=\int_{E} f(y) P(\tau, x ; t, d y), \quad 0 \leq \tau<t<\infty \\
Y(t, t) f(x)=f(x), 0 \leq t<\infty
\end{array}\right.
$$

for all $x \in E$ and $f \in L_{\mathcal{E}}^{\infty}$. This family will be called the free backward propagator associated with $P$.

Let us fix a non-negative Borel measure $m$ on $(E, \mathcal{E})$ (the reference measure). We will write $d x$ instead of $m(d x)$, and will always assume that $0<m(A)<\infty$ for any compact subset $A$ of $E$ having nonempty interior. By $L^{r}$ with $1 \leq r \leq \infty$ will be denoted Lebesgue spaces with respect to the measure $m$. The space of all Borel functions from $L^{r}$ will be denoted by $L_{\mathcal{E}}^{r}$. If $P$ is a transition probability function, then we will say that $P$ possesses density $p$, if there exists a nonnegative function $p(r, x ; s, y)$ such that

$$
P(r, x ; s, A)=\int_{A} p(r, x ; s, y) d y
$$

for all $A \in \mathcal{E}$. In this case, the free backward propagator $Y$ is defined on the space $L^{\infty}$ by

$$
\left\{\begin{array}{l}
Y(t, \tau) f(x)=\int_{E} f(y) p(\tau, x ; t, y) d y, \quad 0 \leq \tau<t<\infty \\
Y(t, t) f(x)=f(x), 0 \leq t<\infty
\end{array}\right.
$$

A rich source of transition probability densities is the theory of second order non-divergence or divergence form parabolic partial differential equations on $R^{n}$. If there exists a fundamental solution for such an equation,
then it can be used as a transition density. Numerous results concerning the existence of fundamental solutions in the case of parabolic equations with time-dependent coefficients can be found in $[4,5,12,14]$ (see also the references in these papers). Next we will give some details.

Let us consider the following equation:

$$
\begin{equation*}
\frac{\partial u}{\partial \tau}+L u=0 \tag{1}
\end{equation*}
$$

In $1, L$ stands for the differential operator given by

$$
\begin{equation*}
L=\sum_{i, j=1}^{n} a_{i j}(\tau, x) \frac{\partial^{2}}{\partial x_{i} \partial x_{j}}+\sum_{i=1}^{n} b_{i}(\tau, x) \frac{\partial}{\partial x_{i}} \tag{2}
\end{equation*}
$$

(non-divergence form), or by

$$
\begin{equation*}
L u=\sum_{i, j=1}^{n} \frac{\partial}{\partial x_{i}}\left[a_{i j}(\tau, x) \frac{\partial}{\partial x_{j}}\right]+\sum_{i=1}^{n} b_{i}(\tau, x) \frac{\partial u}{\partial x_{i}} \tag{3}
\end{equation*}
$$

(divergence form). Let us also consider the final value problem,

$$
\left\{\begin{array}{l}
\frac{\partial u}{\partial \tau}+L u=0, \quad 0 \leq \tau<t \leq T  \tag{4}\\
u(t)=f
\end{array}\right.
$$

for equation 1. Solutions to problem 4 with $L$ in the divergence form are understood in the weak sense. The following results are known:

Non-divergence form. Let $L$ be as in formula 2, and assume that

1. The functions $a_{i j}$ and $b_{i}$ are bounded and measurable on $[0, T] \times R^{n}$;
2. There exists a constant $\gamma>0$ such that for all $(\tau, x) \in[0, T] \times R^{n}$ and any collection of real numbers $\lambda_{1}, \cdots, \lambda_{n}$,

$$
\sum_{i, j=1}^{n} a_{i j}(\tau, x) \lambda_{i} \lambda_{j} \geq \gamma \sum_{i=1}^{n} \lambda_{i}^{2}
$$

3. There exists a constant $\delta$ with $0<\delta \leq 1$ such that

$$
\begin{aligned}
& \sum_{i, j=1}^{n}\left|a_{i j}\left(\tau_{1}, x_{1}\right)-a_{i j}\left(\tau_{2}, x_{2}\right)\right|+\sum_{i=1}^{n}\left|b_{i}\left(\tau_{1}, x_{1}\right)-b_{i}\left(\tau_{2}, x_{2}\right)\right| \leq C\left(\left|x_{1}-x_{2}\right|^{\delta}+\left|\tau_{1}-\tau_{2}\right|^{\delta}\right) \\
& \text { for all }\left(\tau_{1}, x_{1}\right),\left(\tau_{2}, x_{2}\right) \in[0, T] \times R^{n} .
\end{aligned}
$$

Then there exists a unique fundamental solution $p(\tau, x ; t, y)$ of equation 1 . The function $p$ satisfies the following conditions: it is jointly continuous, strictly positive,

$$
\begin{equation*}
p(\tau, x ; t, y) \leq M(t-\tau)^{-\frac{n}{2}} \exp \left\{-\frac{\alpha|x-y|^{2}}{t-\tau}\right\} \tag{5}
\end{equation*}
$$

and

$$
\begin{equation*}
\left|\frac{\partial p}{\partial x_{i}}(\tau, x ; t, y)\right| \leq M(t-\tau)^{-\frac{n+1}{2}} \exp \left\{-\frac{\alpha|x-y|^{2}}{t-\tau}\right\} \tag{6}
\end{equation*}
$$

For $f \in C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ and $t>0$, the function

$$
u(\tau, x)=\int_{R^{n}} f(y) p(\tau, x ; t, y) d y
$$

is in $C_{b}^{1,2}\left([0, t] \times R^{n}\right)$ and satisfies 4 (see e.g., $\left.[4,5,12]\right)$.
Divergence form. Let $L$ be as in 3, and suppose that conditions 1 and 2 hold for $a_{i j}$ and $b_{i}$. Then there exists a unique fundamental solution $p(\tau, x ; t, y)$ of equation 1 satisfying estimate 5 (more information on fundamental solutions in the divergence case can be found in [14]).

An important special example of a transition probability density is the Gaussian density,

$$
p_{t}(z)=(2 \pi t)^{-\frac{n}{2}} \exp \left\{-\frac{|z|^{2}}{2 t}\right\}
$$

on $R^{n}$. In this case, the free backward propagator is related to the heat semigroup,

$$
\begin{equation*}
S_{t} f(x)=f \star p_{t}(x) \tag{7}
\end{equation*}
$$

by the standard formula

$$
\begin{equation*}
Y(\tau, t) f=S_{t-\tau} f \tag{8}
\end{equation*}
$$

Our next goal is to define and study Feynman-Kac propagators. They are connected with perturbations of free propagators by time-dependent functions or measures. We will need some definitions and results from the theory of Markov processes.

Let $\Omega=E^{[0, \infty)}$ denote the space of all paths in $E$ equipped with the cylindrical $\sigma$-algebra $\mathcal{F}$, and let $P$ be a transition probability function. Then there exists a non-terminating non-homogeneous Markov pro$\operatorname{cess}\left(X_{t}, \mathcal{F}_{t}^{\tau}, P_{\tau, x}\right),(\tau, t) \in D_{\infty}$, on $(\Omega, \mathcal{F})$ with the phase space $(E, \mathcal{E})$. Here $X_{t}(\omega)=\omega(t), \mathcal{F}_{t}^{\tau}=\sigma\left(X_{s}: \tau \leq s \leq t\right)$, and $P_{\tau, x}$ with $0 \leq \tau \leq T$ and $x \in E$
is a measure on $\mathcal{F}$ such that

$$
\begin{array}{r}
P_{\tau, x}\left(\omega\left(t_{1}\right) \in A_{1} ; \cdots ; \omega\left(t_{k-1}\right) \in A_{k-1} ; \omega\left(t_{k}\right) \in A_{k}\right) \\
=\int_{A_{1}} P\left(\tau, x ; t_{1}, d x_{1}\right) \int_{A_{2}} P\left(t_{1}, x_{1} ; t_{2}, d x_{2}\right) \cdots \int_{A_{k}} P\left(t_{k-1}, x_{k-1} ; t_{k}, d x_{k}\right)
\end{array}
$$

for all $\tau<t_{1}<t_{2}<\cdots<t_{k} \leq T$ and $A_{i} \in \mathcal{E}$ for $1 \leq i \leq k$. We will also use non-homogeneous Markov processes defined on a general probability space $(\Omega, \mathcal{F})$ (see $[3,6,22])$.

The Markov property of the process $X$ follows from the definitions and can be formulated as follows: For all $0 \leq \tau \leq s \leq t \leq T, x \in E$, and $g \in L_{\mathcal{E}}^{\infty}$,

$$
Y(t, s) g\left(X_{s}\right)=E_{\tau, x}\left(g\left(X_{t}\right) \mid \mathcal{F}_{s}^{\tau}\right)
$$

$P_{\tau, x}$ a.s. We will restrict ourselves to progressively measurable processes. A process $X$ is called $\mathcal{F}_{t}^{\tau}$-progressively measurable, or simply progressively measurable, if for every $\tau$ and $t$ with $0 \leq \tau<t \leq T$, the mapping $(s, \omega) \rightarrow$ $\omega(s)$ of $[\tau, t] \times \Omega$ into $E$ is $\sigma\left(\mathcal{B}_{[\tau, t]} \times \mathcal{F}_{t}^{\tau}\right) / \mathcal{E}$-measurable. It is known that every left- or right-continuous process is progressively measurable (see [6, 23] for more information concerning progressive measurability of stochastic processes). We will denote by $\mathcal{M}$ the class of all transition probability functions $P$ such that there exists a progressively measurable process $X$ corresponding to $P$. If $P \in \mathcal{M}$, then we will always choose a progressively measurable process $X$ to represent $P$.

It is known that the condition

$$
\lim _{t-s \rightarrow 0+} P_{\tau, x}\left(\rho\left(X_{s}, X_{t}\right)>\epsilon\right)=0
$$

for all $0 \leq \tau<s<t \leq T, x \in E$, and $\epsilon>0$, or equivalently, the condition

$$
\begin{equation*}
\lim _{t-s \rightarrow 0+} \int_{E} P\left(s, y ; t, G_{\epsilon}(y)\right) P(\tau, x ; s, d y)=0 \tag{9}
\end{equation*}
$$

where $\epsilon>0, x, y \in E, G_{\epsilon}(y)=\{x \in E: \rho(x, y)>\epsilon\}$, and $0 \leq \tau<s<$ $t \leq T$, guarantees that $P \in \mathcal{M}$ (this follows from Theorem 4 in Chapter 1.6 in [6]). The equivalent conditions mentioned above are called the stochastic continuity conditions for the process $X$. It is also known that the uniform stochastic continuity condition for the transition function $P$,

$$
\lim _{t-s \rightarrow 0+} \sup _{y \in E} P\left(s, y ; t, G_{\epsilon}(y)\right)=0
$$

for all $\epsilon>0$, is stronger than condition 9 , and implies the existence of a process $X$ corresponding to $P$ and such that $X$ is right-continuous and possesses left-hand limits (see [6], Theorem 2 on p. 75).

The sample paths of progressively measurable processes are Borel measurable functions. Hence, if $X$ is progressively measurable, then the functional

$$
\begin{equation*}
A_{V}(\tau, t)=\int_{\tau}^{t} V\left(s, X_{s}\right) d s \tag{10}
\end{equation*}
$$

is defined for appropriate Borel functions $V$ on $[0, T] \times E$. Moreover, for all $0 \leq \tau \leq t \leq T$, the random variable $A_{V}(\tau, t)$ is $\mathcal{F}_{t}^{\tau}$-measurable.

Definition 3. Let $P \in \mathcal{M}$, and let $V$ be a Borel function on $[0, T] \times E$. We will call the family of linear operators,

$$
Y_{V}(\tau, t) g(x)=E_{\tau, x} g\left(X_{t}\right) \exp \left\{-\int_{\tau}^{t} V\left(s, X_{s}\right) d s\right\}, \quad 0 \leq \tau \leq t \leq T
$$

the backward Feynman-Kac propagator associated with $P$ and $V$.
For a time-dependent measure $\mu$, the backward propagator $Y_{\mu}$ will be defined in Section 5.

## 3 Non-autonomous classes of functions and measures

Let $P$ be a transition probability function from the class $\mathcal{M}$, and assume that $V$ is a Borel function on the set $[0, T] \times E$, where $T>0$ is a fixed number, and $\mu$ is a family $\{\mu(t): 0 \leq t \leq T\}$ of signed measures on $(E, \mathcal{E})$. Consider the following potential operators:

$$
\begin{equation*}
N_{V}(\tau, t, x)=\int_{\tau}^{t} Y(\tau, s) V(s)(x) d s, \quad(\tau, t) \in D_{T}, \quad x \in E \tag{11}
\end{equation*}
$$

and

$$
\begin{equation*}
N(\mu)(\tau, t, x)=\int_{\tau}^{t} Y(\tau, s) \mu(s)(x) d s, \quad(\tau, t) \in D_{T}, \quad x \in E \tag{12}
\end{equation*}
$$

It is assumed in 11 and 12 that the integrals on the right-hand side make sense. In 12 , we also assume that $P$ possesses density $p$.
Definition 4. Let $P \in \mathcal{M}$. Then we say that $V$ belongs to the class $\hat{\mathcal{P}}_{f}^{*}$, provided that

$$
\sup _{(\tau, t) \in D_{T}} \sup _{x \in E} N(|V|)(\tau, t, x)<\infty
$$

Let $V \in \hat{\mathcal{P}}_{f}^{*}$. Then we say that $V$ belongs to the class $\mathcal{P}_{f}^{*}$, provided that

$$
\lim _{t-\tau \rightarrow 0+} \sup _{x \in E} N(|V|)(t, \tau, x)=0 .
$$

Suppose that $P \in \mathcal{M}$ possesses density $p$. Then we say that $\mu$ belongs to the class $\hat{\mathcal{P}}_{m}^{*}$, provided that

$$
\sup _{\tau:(\tau, t) \in D_{T}} \sup _{x \in E} N(|\mu|)(\tau, t, x)<\infty .
$$

If $\mu \in \hat{\mathcal{P}}_{m}^{*}$, then we say that $\mu$ belongs to the class $\mathcal{P}_{m}^{*}$, provided that

$$
\lim _{t \rightarrow \tau \rightarrow 0+} \sup _{x \in E} N(|\mu|)(\tau, t, x)=0
$$

The classes defined above were studied in [11]. In the case of the heat semigroup, these classes were introduced in $[8,9]$.

The following non-standard approximation was considered in [11]:
Definition 5. Let $P \in \mathcal{M}, V \in \mathcal{P}_{f}^{*}$, and $V_{k} \in \mathcal{P}_{f}^{*}$ for all $k \geq 1$. Then we will say that the sequence $V_{k} \zeta$-approaches $V$ provided that

$$
\lim _{k \rightarrow \infty} \sup _{(\tau, t) \in D_{T}} \sup _{x \in E}\left|N\left(V-V_{k}\right)(\tau, t, x)\right| \rightarrow 0,
$$

and

$$
\lim _{t-\tau \rightarrow 0+} \sup _{k \geq 1} \sup _{x \in E} N\left(\left|V_{k}\right|\right)(\tau, t, x)=0 .
$$

If $P \in \mathcal{M}$ possesses density $p$, then we will say that a sequence $\mu_{k} \in \mathcal{P}_{m}^{*}$, $k \geq 1, \zeta$-approaches $\mu \in \mathcal{P}_{f}^{*}$ provided that

$$
\lim _{k \rightarrow \infty} \sup _{(\tau, t) \in D_{T}} \sup _{x \in E}\left|N\left(\mu-\mu_{k}\right)(\tau, t, x)\right| \rightarrow 0,
$$

and

$$
\lim _{t-\tau \rightarrow 0+} \sup _{k \geq 1} \sup _{x \in E} N\left(\left|\mu_{k}\right|\right)(\tau, t, x)=0 .
$$

The next results from [11] explain why this type of approximation is useful in the theory of Feynman-Kac propagators.

Theorem 6. Let $P \in \mathcal{M}$, and let $V \in \mathcal{P}_{f}^{*}$ and $V_{k} \in \mathcal{P}_{f}^{*}$ be such that $V_{k}$ $\zeta$-approaches $V$. Then

$$
\lim _{k \rightarrow \infty} \sup _{(\tau, t) \in D_{T}}\left\|Y_{V}(\tau, t)-Y_{V_{k}}(\tau, t)\right\|_{L_{\varepsilon}^{\infty} \rightarrow L_{\mathcal{\varepsilon}}^{\infty}}=0 .
$$

Suppose that $P \in \mathcal{M}$ possesses density $p$, and let $\mu \in \mathcal{P}_{m}^{*}$ and $\mu_{k} \in \mathcal{P}_{m}^{*}$ be such that $\mu_{k} \zeta$-approaches $\mu$. Then

$$
\lim _{k \rightarrow \infty} \sup _{(\tau, t) \in D_{T}}\left\|Y_{\mu}(\tau, t)-Y_{\mu_{k}}(\tau, t)\right\|_{\infty \rightarrow \infty}=0 .
$$

Theorem 7. (a) Let $P \in \mathcal{M}$ and $V \in \mathcal{P}_{f}^{*}$. For $k \geq 1,0 \leq \tau \leq T$, and $x \in E$, put

$$
g_{k}(\tau, x)=k N(V)\left(\tau, \min \left(\tau+\frac{1}{k}, T\right), x\right) .
$$

Then the following conditions hold:

$$
\begin{equation*}
g_{k} \in \mathcal{P}_{f}^{*} \tag{13}
\end{equation*}
$$

for all $k \geq 1$;

$$
\begin{equation*}
\lim _{k \rightarrow \infty} \sup _{(\tau, t) \in D_{T}} \sup _{x \in E}\left|N\left(V-g_{k}\right)(\tau, t, x)\right|=0 ; \tag{14}
\end{equation*}
$$

and

$$
\begin{equation*}
\lim _{t \rightarrow \tau \rightarrow 0+} \sup _{k \geq 1} \sup _{x \in E} N\left(\left|g_{k}\right|\right)(\tau, t, x)=0 . \tag{15}
\end{equation*}
$$

(b) Suppose that $P \in \mathcal{M}$ possesses density $p$, and let $\mu \in \mathcal{P}_{m}^{*}$. For $k \geq 1$, $0 \leq \tau \leq T$, and $x \in E$, put

$$
\begin{equation*}
g_{k}(\tau, x)=k N(V)\left(\tau, \min \left(\tau+\frac{1}{k}, T\right), x\right) . \tag{16}
\end{equation*}
$$

Then conditions 13-15 in part (a) of Lemma 1 hold with $\mu$ instead of $V$.
It is clear from theorems 2 and 3 that imposing various restrictions on the potentials $N(V)$ and $N(\mu)$, we can get new interesting classes of functions and measures. The following notation will be used in the sequel: the symbol $B C$ will stand for the space of all bounded continuous functions on $E$ equipped with the norm

$$
\|f\|_{C}=\sup _{x \in E}|f(x)|,
$$

by $C_{\infty}$ will be denoted the space of all continuous functions on $E$ vanishing at infinity, and by $B U C$ will be denoted the space of bounded uniformly continuous functions on $E$. The following classes were introduced in [11]:

Definition 8. We define the function classes $\mathcal{P}_{f, c}^{*}$ and $\mathcal{P}_{f, u}^{*}$ as follows:

$$
\begin{gathered}
V \in \mathcal{P}_{f, c}^{*} \Longleftrightarrow V \in \mathcal{P}_{f}^{*} \text { and } N(V)(\tau, t, \cdot) \in B C \text { for all }(\tau, t) \in D_{T} \\
V \in \mathcal{P}_{f, u}^{*} \Longleftrightarrow V \in \mathcal{P}_{f}^{*} \text { and } N(V)(\tau, t, \cdot) \in B U C \text { for all }(\tau, t) \in D_{T}
\end{gathered}
$$

Definition 9. We define the class $\mathcal{D}_{f, c}^{*}$ as follows: A function $V \in \mathcal{P}_{f}^{*}$ belongs to this class if there exists exists a sequence $V_{k} \in \mathcal{P}_{f}^{*}$ such that $V_{k}(t, \cdot) \in B C$ for all $k \geq 1$ and $0 \leq t \leq T$, and $V_{k} \zeta$-approaches $V$. The class $\mathcal{D}_{f, u}^{*}$ is defined similarly. Here we require the condition $V_{k}(t, \cdot) \in B U C$ for all $k \geq 1$ and $0 \leq t \leq T$, and $V_{k} \zeta$-approaches $V$.

Remark 10. If $P$ possesses density $p$, then the classes of time-dependent measures $\mathcal{P}_{m, c}^{*}, \mathcal{P}_{m, u}^{*}, \mathcal{D}_{m, c}^{*}$, and $\mathcal{D}_{m, u}^{*}$ can be defined similarly.

The next lemma describes the relations between the classes in definitions 4 and 6.

Lemma 11. The following assertions hold:

1. $\mathcal{P}_{f, c}^{*} \subset \mathcal{D}_{f, c}^{*}$ and $\mathcal{P}_{m, c}^{*} \subset \mathcal{D}_{m, c}^{*}$.
2. $\mathcal{P}_{f, u}^{*} \subset \mathcal{D}_{f, u}^{*}$ and $\mathcal{P}_{m, u}^{*} \subset \mathcal{D}_{m, u}^{*}$.
3. If $V \in \mathcal{P}_{f}^{*}$, and there exists a sequence $V_{k} \in \mathcal{P}_{f, c}^{*}$ such that $V_{k} \zeta$ approaches $V$, then $V \in \mathcal{P}_{f, c}^{*}$. Similarly, if $\mu \in \mathcal{P}_{m}^{*}$, and there exists a sequence $V_{k} \in \mathcal{P}_{f, c}^{*}$ such that $V_{k} \zeta$-approaches $V$, then $\mu \in \mathcal{P}_{m, c}^{*}$.
4. If $V \in \mathcal{P}_{f}^{*}$, and there exists a sequence $V_{k} \in \mathcal{P}_{f, u}^{*}$ such that $V_{k} \zeta$ approaches $V$, then $V \in \mathcal{P}_{f, u}^{*}$. Similarly, if $\mu \in \mathcal{P}_{m}^{*}$, and there exists a sequence $V_{k} \in \mathcal{P}_{f, u}^{*}$ such that $V_{k} \zeta$-approaches $V$, then $\mu \in \mathcal{P}_{m, u}^{*}$.

The proof of this lemma can be found in [11].

## 4 The functional $A_{\mu}$ and the backward propagator $A_{\mu}$

Now we are ready to define the Feynman-Kac propagator $Y_{\mu}$ for a timedependent measure $\mu$. First we should determine what functional replaces functional 10 in the Feynman-Kac formula for $Y_{\mu}$ with $\mu \in \mathcal{P}_{m}^{*}$. It is natural to try to $\zeta$-approximate $\mu$ by a sequence of functions $g_{k} \in \mathcal{P}_{f}^{*}$, and prove that the corresponding sequence of functionals $A_{g_{k}}$ converges in an appropriate sense. Then the limit of the sequence $A_{g_{k}}$ can be a good candidate to replace $A_{V}$. It was shown in [11] that $A_{\mu}$ exists and satisfies the following condition:

$$
\begin{equation*}
\lim _{k \rightarrow \infty} \sup _{\tau: 0 \leq \tau \leq T} \sup _{x \in E} E_{\tau, x} \sup _{t: \tau \leq t \leq T}\left|A_{\mu}(\tau, t)-A_{g_{k}}(\tau, t)\right|^{n}=0 \tag{17}
\end{equation*}
$$

for every $n \geq 1$. In 17 , the functions $g_{k}$ are defined by 16 . It can be shown that the functional $A_{\mu}$ is unique up to equivalence (see [11]). The sufficient conditions for the existence of $A_{\mu}$ are: $P \in \mathcal{M}, P$ possesses density $p$, and $\mu \in \mathcal{P}_{m}^{*}$.

Definition 12. Let $P \in \mathcal{M}$ be a transition probability function possessing density $p$, and let $\mu \in \mathcal{P}_{m}^{*}$. Then the family of operators,

$$
Y_{\mu}(\tau, t) g(x)=E_{\tau, x} g\left(X_{t}\right) \exp \left\{-A_{\mu}(\tau, t)\right\}, \quad(\tau, t) \in D_{T},
$$

is called the backward Feynman-Kac propagator associated with $P$ and $\mu$.

## 5 The inheritance problem for Feynman-Kac propagators

The free backward propagator $Y$ associated with a transition probability function $P$ and the backward Feynman-Kac propagators $Y_{V}$ and $Y_{\mu}$ are closely connected. Various properties of $Y$ are inherited by $Y_{V}$ and $Y_{\mu}$ (see the results below). However, it is interesting to mention that some simple properties are not inherited. Among such properties is the $L^{1}$-boundedness. This was shown in [9] for forward propagators and the space $R^{3}$. Next we will describe how to construct such a counterexample in the case of backward Feynman-Kac propagators and the space $R^{n}$ with $n \geq 2$. Let the free backward propagator $Y$ be given by 7 with $S_{t}$ defined by 8 , and consider the following function on the space $[0,1] \times R^{n}$ :

$$
V(t, x)=-\frac{1}{\sqrt{1-t} \ln \frac{e}{1-t}|x|} .
$$

It follows from Lemma 7 in [9] that

$$
\begin{equation*}
V \in \mathcal{P}_{f}^{*} \Leftrightarrow \lim _{t \rightarrow 0+} \sup _{h: 0 \leq h \leq 1-t} \int_{0}^{t} \frac{1}{\sqrt{s} \sqrt{1-s-h} \ln \frac{e}{1-s-h}} d s=0 . \tag{18}
\end{equation*}
$$

Lemma 7 was established in [9] in the case $\mathrm{n}=3$, the general case is similar. We have

$$
\begin{aligned}
\sup _{h: 0 \leq h \leq 1-t} \int_{0}^{t} \frac{1}{\sqrt{s} \sqrt{1-s-h} \ln \frac{e}{1-s-h}} d s & \leq \int_{0}^{t} \frac{1}{\sqrt{s} \sqrt{t-s} \ln \frac{e}{t-s}} d s \\
& \leq \frac{\sqrt{2}}{\sqrt{t} \ln \frac{2 e}{t}} \int_{0}^{\frac{t}{2}} \frac{1}{\sqrt{s}} d s \\
& +\frac{\sqrt{2}}{\sqrt{t} \ln \frac{2 e}{t}} \int_{\frac{t}{2}}^{t} \frac{1}{\sqrt{t-s}} d s \\
& =\frac{2 \sqrt{2}}{\sqrt{t} \ln \frac{2 e}{t}} \int_{0}^{\frac{t}{2}} \frac{1}{\sqrt{s}} d s \rightarrow 0
\end{aligned}
$$

as $t \rightarrow 0+$. Therefore, it follows from 18 that $V \in \mathcal{P}_{f}^{*}$.
Now suppose that $Y_{V}(\tau, 1) \in L\left(L^{1}, L^{1}\right)$ for some $\tau 0 \leq \tau<1$. Then $Y_{V}^{*}(\tau, 1) 1 \in L^{\infty}$, and hence, the function given by

$$
x \rightarrow E_{x} \exp \left\{\int_{0}^{1-\tau} \frac{1}{\sqrt{s} \ln \frac{e}{s}\left|B_{s}\right|} d s\right\}
$$

where $B_{s}$ is a standard Brownian motion in $R^{n}$, belongs to the space $L^{\infty}$. It follows that the function

$$
\phi_{\tau}(x)=E_{x} \int_{0}^{1-\tau} \frac{1}{\sqrt{s} \ln \frac{e}{s}\left|B_{s}\right|} d s
$$

also belongs to $L^{\infty}$. We have

$$
\begin{aligned}
\phi_{\tau}(x) & =\int_{0}^{1-\tau} \frac{1}{\sqrt{s} \ln \frac{e}{s}} d s \int_{R^{n}}|y|^{-1}(2 \pi s)^{-\frac{n}{2}} \exp \left\{-\frac{|x-y|^{2}}{2 s}\right\} d y \\
& \geq c \int_{0}^{1-\tau} \frac{1}{\sqrt{s} \ln \frac{e}{s}}(2 \pi s)^{-\frac{n}{2}} d s \int_{y:|x-y| \leq \sqrt{s}}|y|^{-1} d y \\
& \geq c \int_{0}^{1-\tau} \frac{1}{\sqrt{s} \ln \frac{e}{s}}(|x|+\sqrt{s})^{-1} .
\end{aligned}
$$

This provides a contradiction to the boundedness of the function $\phi$.
Hence, we conclude that $V \in \mathcal{P}_{f}^{*}$, but $Y_{V}(\tau, 1) \notin L\left(L^{1}, L^{1}\right)$.
The next results were obtained in [11]. The first theorems concern the $L^{s}$-boundedness and the ( $L^{s}-L^{q}$ )-smoothing property:

Theorem 13. (a) Let $P \in \mathcal{M}$. Then for any $V \in \mathcal{P}_{f}^{*}, Y_{V}$ is a backward propagator on $L_{\mathcal{E}}^{\infty}$.
(b) Suppose that $P \in \mathcal{M}$ possesses the density $p$, and let $V \in \mathcal{P}_{f}^{*}$. Then $Y_{V}$ is a backward propagator on $L^{\infty}$.
(c) Suppose that $P \in \mathcal{M}$ possesses the density $p$, and let $\mu \in \mathcal{P}_{m}^{*}$. Then $Y_{\mu}$ is a backward propagator on $L^{\infty}$.

Theorem 14. Let $1<s<\infty$ and $1 \leq r<s$. Then the following are true: (a) Let $P \in \mathcal{M}$ and $V \in \mathcal{P}_{f}^{*}$. Suppose that the free backward propagator $Y$ satisfies $Y(\tau, t) \in L\left(L_{\mathcal{E}}^{r}, L_{\mathcal{E}}^{r}\right)$ for all $(\tau, t) \in D_{T}$. Then $Y_{V}$ is a backward propagator on $L_{\mathcal{E}}^{s}$. If, in addition, $Y$ is uniformly bounded on $L_{\mathcal{E}}^{r}$ and strongly continuous on $L_{\mathcal{E}}^{s}$, then $Y_{V}$ is a strongly continuous backward propagator on $L_{\mathcal{E}}^{s}$.
(b) If $P \in \mathcal{M}$ possesses the density $p$, and if $Y(\tau, t) \in L\left(L^{r}, L^{r}\right)$ for all $(\tau, t) \in D_{T}$, then $Y_{V}$ is a backward propagator on $L^{s}$. If, in addition, $Y$ is uniformly bounded on $L^{r}$ and strongly continuous on $L^{s}$, then $Y_{V}$ is a strongly continuous backward propagator on $L^{s}$.
(c) Suppose that $P \in \mathcal{M}$ possesses the density $p$ and let $\mu \in \mathcal{P}_{m}^{*}$. If $Y(\tau, t) \in$ $L\left(L^{r}, L^{r}\right)$ for all $0 \leq \tau<t \leq T$, then $Y_{\mu}$ is a backward propagator on $L^{s}$. If in addition, $Y$ is uniformly bounded on $L^{r}$ and strongly continuous on $L^{s}$, then $Y_{\mu}$ is a strongly continuous backward propagator on $L^{s}$.

Theorem 15. Let $1<s<q \leq \infty$ and $1 \leq r<s$. Then the following are true:
(a) Let $P \in \mathcal{M}$ and $V \in \mathcal{P}_{f}^{*}$. Suppose that $Y(\tau, t) \in L\left(L_{\mathcal{E}}^{r}, L_{\mathcal{E}}^{\frac{r q}{s}}\right)$ for all $0 \leq \tau<t \leq T$. Then $Y_{V}(\tau, t) \in L\left(L_{\mathcal{E}}^{s}, L_{\mathcal{E}}^{q}\right)$ for all $0 \leq \tau<t \leq T$.
(b) If $P \in \mathcal{M}$ possesses the density $p, V \in \mathcal{P}_{f}^{*}$, and $Y(\tau, t) \in L\left(L^{r}, L^{\frac{r q}{p}}\right)$ for all $0 \leq \tau<t \leq T$, then $Y_{V}(\tau, t) \in L\left(L^{s}, L^{q}\right)$.
(c) If $P \in \mathcal{M}$ possesses the density $p, \mu \in \mathcal{P}_{m}^{*}$, and $Y(\tau, t) \in L\left(L^{r}, L^{\frac{r q}{p}}\right)$ for all $0 \leq \tau<t \leq T$, then $Y_{\mu}(\tau, t) \in L\left(L^{s}, L^{q}\right)$.

The next group of results concerns the inheritance of the boundedness in the spaces of continuous functions on $E$.

Definition 16. A backward $B C$-propagator is called a backward Feller propagator. A backward $C_{\infty}$-propagator is called a backward Feller-Dynkin propagator. If a backward $L_{\mathcal{E}}^{\infty}$-propagator $Q$ is such that $Q(\tau, t) \in L\left(L_{\mathcal{E}}^{\infty}, B C\right)$ for all $0 \leq \tau<t \leq T$, then it is said that $Q$ satisfies the strong Feller condition. If a backward $L_{\mathcal{E}}^{\infty}$-propagator $Q$ is such that $Q(\tau, t) \in L\left(L_{\mathcal{E}}^{\infty}, B U C\right)$ for all $0 \leq \tau<t \leq T$, then it is said that $Q$ satisfies the strong $B U C$-condition.

Remark 17. If $Q$ is a backward $L^{\infty}$-propagator, then we may replace the space $L_{\mathcal{E}}^{\infty}$ by the space $L^{\infty}$ in the definition of the strong Feller and the strong $B U C$-condition.

Theorem 18. Let $P \in \mathcal{M}$ and $V \in \mathcal{P}_{f}^{*}$. Then the following assertions hold: (a) If $Y$ satisfies the strong Feller condition, then $Y_{V}$ also satisfies the same condition.
(b) If $Y$ satisfies the strong $B U C$-condition, then $Y_{V}$ also satisfies the same condition.

Theorem 19. Let $P \in \mathcal{M}, V \in \mathcal{P}_{f}^{*}$, and suppose that $Y$ satisfies the strong Feller condition. Then the following assertions hold:
(a) If $Y$ is a backward Feller-Dynkin propagator, then $Y_{V}$ also has the same
property.
(b) If $Y$ is a strongly continuous backward Feller-Dynkin propagator, then $Y_{V}$ also has the same property.

Theorem 20. Let $P \in \mathcal{M}, V \in \mathcal{P}_{f}^{*}$, and suppose that $Y$ satisfies the strong $B U C$-condition. Then the following assertion holds:
If $Y$ is a strongly continuous backward BUC-propagator, then $Y_{V}$ also has the same property.

Theorem 21. Let $P \in \mathcal{M}$. Then the following assertions hold: (a) Suppose that $Y$ is a backward strong Feller propagator. Suppose also that $Y$ is continuous on $B C$ in the topology of uniform convergence on compact subsets of $E$. Let $V \in \mathcal{P}_{f}^{*}$. Then for every $t \in(0, T]$ and any $g \in L_{\mathcal{E}}^{\infty}$, the function $(\tau, x) \rightarrow Y_{V}(\tau, t) g(x)$ is continuous on the set $[0, t) \times R^{n}$.
(b) Suppose that $Y$ is a strongly continuous backward BUC-propagator possessing the strong BUC-property. Let $V \in \mathcal{P}_{f}^{*}$. Then for every $t \in(0, T]$ and any $g \in L_{\mathcal{E}}^{\infty}$, the function $(\tau, x) \rightarrow Y_{\mu}(\tau, t) g(x)$ is continuous on the set $[0, t) \times R^{n}$.

If we assume that $V$ and $\mu$ belong to the classes in definitions 5 and 6 , we get stronger results.

Theorem 22. Let $P \in \mathcal{M}$ and $V \in \mathcal{D}_{f, c}^{*}$. Then the following assertions hold:
(a) If $Y$ is a backward Feller propagator, then $Y_{V}$ has the same property.
(b) If $Y$ is a backward Feller-Dynkin propagator, then $Y_{V}$ has the same property. If, in addition, $Y$ is strongly continuous on $C_{\infty}$, then $Y_{V}$ is also strongly continuous on $C_{\infty}$.

Theorem 23. Let $P \in \mathcal{M}$ and $V \in \mathcal{D}_{f, u}^{*}$. Then if $Y$ is a backward BUCpropagator, then $Y_{V}$ has the same property. If, in addition, $Y$ is strongly continuous on $B U C$, then $Y_{V}$ is also strongly continuous on BUC.

Remark 24. Theorem 10 (Theorem 11) holds for a time-dependent measure $\mu \in \mathcal{D}_{m, c}\left(\mu \in \mathcal{D}_{m, u}\right)$, provided that $P \in \mathcal{M}$ possesses density $p$.

## 6 Applications

Let us go back to the case of free backward propagators associated with fundamental solutions of equation 1 with $L$ in non-divergence form. Let $p$ be such a fundamental solution. It is not difficult to prove, using estimates 5 and 6 , that the backward free propagator $Y$ associated with the density $p$ is
( $L^{s}-L^{q}$ )-smoothing for all $1 \leq s \leq q \leq \infty$, and possesses the strong Feller and the strong $B U C$-property. It follows that $Y$ is a backward $B C$ - and $B U C$-propagator. Moreover, $Y$ is a backward $C_{\infty}$-propagator (use estimate 5).

The following simple lemma is useful in problems concerning the approximation by backward propagators:

Lemma 25. For every function $f \in B U C$ and $\epsilon>0$, we have

$$
\|f-Y(\tau, t) f\|_{C} \leq \sup _{\rho(x, y) \leq \epsilon}|f(x)-f(y)|+2\|f\|_{C} \sup _{x \in E} \int_{y: \rho(x, y)>\epsilon} p(\tau, x ; t, y) d y
$$

Using this lemma and Theorem 1, one gets the following assertion (see [11]):

Lemma 26. Let p be a fundamental function for equation 1 in non-divergence form. Then the free backward propagator $Y$ associated with $p$ is strongly continuous on $C_{\infty}$ and BUC.

Since the backward free propagator in the example above is $\left(L^{s}-L^{q}\right)$ smoothing for all $1 \leq s \leq q \leq \infty$, possesses the strong Feller and the strong $B U C$-property, and is a strongly continuous propagator on $C_{\infty}$ and $B U C$, the perturbed backward propagator $Y_{\mu}$ with $\mu \in \mathcal{P}_{m}^{*}$ satisfies $Y_{\mu}(\tau, t) \in$ $L\left(L^{s}, L^{q}\right)$ for all $(\tau, t) \in D_{T}$ and $1<s \leq q \leq \infty$, possesses the strong Feller and the strong $B U C$-property, and is a strongly continuous propagator on $C_{\infty}$ and $B U C$ (see theorems 3-8).

As in the non-divergence case, the backward free propagator $Y$ in the divergence case is such that $Y(\tau, t) \in L\left(L^{s}, L^{q}\right)$ for all $(\tau, t) \in D_{T}$ and $1 \leq s \leq q \leq \infty$ (this follows from the Gaussian estimate). The strong Feller property also holds for $Y$ (this follows from the Gaussian estimate and the continuity of $p$ ). Moreover, $Y$ is strongly continuous on $C_{\infty}$ (this fact can be obtained from the strong Feller property, the Gaussian estimate, Lemma 3, and the ideas in the proof of part (b) of Theorem 7.6 in [11]). Hence, using the theorems in Section 5, we see that the perturbed backward propagator $Y_{\mu}$ with $\mu \in \mathcal{P}_{m}^{*}$ satisfies $Y_{\mu} \in L\left(L^{s}, L^{q}\right)$ for all $(\tau, t) \in D_{T}$ and $1<s \leq q \leq \infty$, possesses the strong Feller property, and is a strongly continuous backward propagator on $C_{\infty}$.

## 7 Acknowledgements

It is my pleasure to thank Professor Genaro López for inviting me to give a mini-course on propagators at the Universidad de Sevilla in January, 2004.

Part of this work was done during my stay at Centre de Recerca Matematica (CRM) in Bellaterra, Spain, in November 2003 - April 2004 (Beca de profesores e investigadores extranjeros en régimen de año sabático del Ministerio de Educación, Cultura y Deporte de España, referencia SAB2002-0066). I am very grateful to Professor López, entire Faculty of Mathematics at the Universidad de Sevilla, and the staff of CRM for their wonderful hospitality.

## References

[1] M. Aizenman and B. Simon, Brownian motion and Harnack's inequality for Schrödinger operators, Comm. Pure Appl. Math. 35 (1982), 209-271.
[2] M. Demuth and J. A. van Casteren, Stochastic Spectral Theory for Selfadjoint Feller Operators. A functional integration approach, Birkhäuser Verlag, Basel, 2000.
[3] E. B. Dynkin, Théorie des processus markoviens, Dunod, Paris, 1963.
[4] S. D. Eidelman, Parabolic equations, Partial Differential Equations VI, Elliptic and Parabolic Operators, Yu. V. Egorov and M. A. Shubin (Eds.), Springer-Verlag, Berlin, 1994, p. 203-325.
[5] A. Friedman, Partial Differential Equations of Parabolic Type, Prentice Hall, Englewood Cliffs, NJ, 1964.
[6] I. I. Gihman and A. V. Skorokhod, The Theory of Stochastic Processes II, Springer-Verlag, Berlin, 1975.
[7] A. Gulisashvili, Classes of time-dependent measures and the behavior of Feynman-Kac propagators, C. R. Acad. Sci. Paris, Ser.I 334 (2002), 1-5.
[8] A. Gulisashvili, On the heat equation with a time-dependent singular potential, J. Funct. Anal. 194 (2002), 17-52.
[9] A. Gulisashvili, Nonautonomous Kato classes of measures and Feynman-Kac propagators, to be published in Trans. Amer. Math. Soc.
[10] A. Gulisashvili, Nonautonomous Kato classes and the behavior of backward Feynman-Kac propagators, in Analyse stochastique et théorie du potentiel, Saint Priest de Gimel (2002), Association Laplace-Gauss, Paris, 2003.
[11] A. Gulisashvili, Markov processes, classes of time-dependent measures, and Feynman-Kac propagators, submitted for publication.
[12] A. M. Il'in, A. S. Kalashnikov, and O. A. Oleinik, Linear equations of the second order of parabolic type, Russian Math. Surveyes 17 (1962), 1-143.
[13] G. W. Johnson and M. Lapidus, The Feynman Integral and Feynman's Operational Calculus, Clarendon Press, Oxford, 2000.
[14] V. Liskevich and Y. Semenov, Estimates for fundamental solutions of second-order parabolic equations, J. London Math. Soc. 62 (2000), 521-543.
[15] M. Nagasawa, Stochastic Processes in Quantum Physics, Birkhäuser, Basel, 2000.
[16] E. M. Ouhabaz, P. Stollmann, K.-Th. Sturm, and J. Voigt, The Feller property for absorption semigroups, J. Funct. Anal. 138 (1996), 351378.
[17] Qi Zhang, On a parabolic equation with a singular lower order term, Transactions Amer. Math. Soc. 348 (1996), 2811-2844.
[18] Qi Zhang, On a parabolic equation with a singular lower order term, Part 2: The Gaussian bounds, Indiana Univ. Math. J. 46 (1997), 989-1020.
[19] F. Räbiger, A. Rhandi, R. Schnaubelt, and J. Voigt, Non-autonomous Miyadera perturbation, Differential Integral Equations 13 (2000), 341368.1
[20] R. Schnaubelt and J. Voigt, The non-autonomous Kato class, Arch. Math. 72 (1999), 454-460.
[21] B. Simon, Schrödinger semigroups, Bull. Amer. Math. Soc. 7 (1982), 445-526.
[22] D. W. Stroock and S. R. S. Varadhan, Multidimensional Diffusion Processes, Springer, Berlin, 1997.
[23] J. Yeh, Martingales and Stochastic Analysis, World Scientific, Singapore, 1995.

## Author's address:

Archil Gulisashvili<br>Department of Mathematics<br>Ohio University<br>Athens, OH 45701 USA<br>E-mail: guli@math.ohiou.edu

# Survey on operators acting on distribution spaces of local regularity $\psi(t)$ 

Silvia I. Hartzstein *


#### Abstract

This is a survey of Besov and Triebel-Lizorkin spaces of distributions with local regularity measured by a function $\psi$, and integrability determined by the index $p$ and $q$, and the operators -generalizations of the classical fractional integral and derivative operators of order $\alpha$ acting between spaces of the same scale but of different local regularity. The underlying geometry is that of homogeneous type spaces and the moduli of continuity $\psi$ belong to a larger class than the potentials $t^{\alpha}$.


## 1 Introduction

The Littlewood-Paley theory provides a unifying approach to the study of most of the classical functional spaces on $\mathbb{R}^{n}$ as, for example, Lebesgue spaces, Hardy spaces, Sobolev spaces, different kind of Lipschitz spaces and $B M O$, together with their traces on subspaces. By means of this theory all these spaces can be characterized through the action of an appropriate family of operators. (For an insight on these topics, see, for example [5], [12], [13], [14] and [15]).

From this unifying approach arise the homogeneous Besov spaces, $\dot{B}_{p}^{\alpha, q}\left(R^{n}\right)$, and Triebel-Lizorkin spaces, $\dot{F}_{p}^{\alpha, q}\left(R^{n}\right)$, whose local regularity is determined by the potential $t^{\alpha}$ and their integrability by the index $p$ and $q$.

More precisely, let $\varphi$ be a function with the following properties:

$$
\begin{align*}
& \varphi \in \mathcal{S}\left(R^{n}\right)  \tag{1}\\
& \operatorname{supp} \hat{\varphi} \subset\left\{\xi \in R^{n}: 1 / 2 \leq|\xi| \leq 2\right\}  \tag{2}\\
& |\hat{\varphi}(\xi)| \geq c>0 \text { if } 3 / 5 \leq|\xi| \leq 5 / 3 \tag{3}
\end{align*}
$$

[^3]Set $\varphi_{t}=t^{-n} \varphi\left(t^{-1} x\right)$, for $t>0$. The Besov space $\dot{B}_{p}^{\alpha, q}\left(R^{n}\right), \alpha \in R, 1 \leq$ $p, q \leq \infty$ is the family of all $f \in \mathcal{S}^{\prime} / \mathcal{P}$, tempered distributions modulo polynomials, such that

$$
\begin{array}{ll}
\|f\|_{\dot{B}_{p}^{\alpha, q}}=\left(\int_{0}^{\infty}\left(t^{-\alpha}\left\|\varphi_{t} * f\right\|_{p}\right)^{q} \frac{d t}{t}\right)^{\frac{1}{q}}<\infty, & 0<p \leq \infty, 0<q<\infty \\
\|f\|_{\dot{B}_{p}^{\alpha, \infty}}=\sup _{0<t<\infty}\left(t^{-\alpha}\left\|\varphi_{t} * f\right\|_{p}\right)<\infty, & 0<p \leq \infty,
\end{array}
$$

and interchanging the order of the norms in the above definitions we obtain the Triebel-Lizorkin space $\dot{F}_{p}^{\alpha, q}\left(R^{n}\right), \alpha \in R, 1 \leq p, q \leq \infty, p \neq \infty$,

$$
\|f\|_{\dot{F}_{p}^{\alpha, q}}=\left\|\left(\int_{0}^{\infty}\left(t^{-\alpha}\left|\varphi_{t} * f\right|\right)^{q} \frac{d t}{t}\right)^{\frac{1}{q}}\right\|_{L^{p}}<\infty
$$

There is a formula, due to Calderón, closely related to the study of these spaces. For example, it shows that the definitions of these norms are independent of the choice of $\varphi$ satisfying (1), (2) and (3). Moreover, it allows the characterization of all the spaces we mentioned at the beginning as special cases of the Besov and Triebel-Lizorkin spaces.

There are many versions of this formula, one of its continuous variants stated as follows:

## Calderón's Reproducing Formula:

Assume that $\varphi$ satisfies (1), (2) and (3). Then there exists a function $\psi$ satisfying (1) and (2) such that for $f \in \mathcal{S}^{\prime} / \mathcal{P}$,

$$
f(x)=\int_{0}^{\infty}\left(\psi_{t} * \varphi_{t} * f\right)(x) \frac{d t}{t}
$$

where the integral converges in the distribution sense.
Some examples of Littlewood-Paley characterizations obtained by using the Calderón reproducing formula are the following.
Lebesgue and Hardy spaces satisfy $L^{p} \simeq \dot{F}_{p}^{0,2}, 1<p<\infty$ and, in general, $H^{p} \simeq \dot{F}_{p}^{0,2}, 0<p<\infty$.
For $0<\alpha<n$, the Lipschitz space $\dot{\Lambda}^{\alpha}$ of complex functions $f$ (modulo constants), such that there exists a constant $C$

$$
|f(x)-f(y)| \leq C|x-y|^{\alpha},
$$

for all $x$ and $y \in R^{n}$, satisfies $\dot{\Lambda}_{\alpha} \simeq \dot{B}_{\infty}^{\alpha, \infty}$.
The Sobolev space $\dot{L}_{k}^{p}$ of tempered distributions (modulus polynomials of
order at most $k-1$ ) such that all their (weak) derivatives of order $k$ are in $L^{p}$ satisfies $\dot{L}_{k}^{p} \simeq \dot{F}_{p}^{k, 2}, 1<p<\infty$.
Finally, if we consider the fractional derivative operator defined through the Fourier transform by $\widehat{D^{\alpha} f}(\xi)=C_{\alpha}|\xi|^{\alpha} \widehat{f}(\xi)$, then the fractional Sobolev space $\dot{L}_{\alpha}^{p}$ of tempered distributions such that $D^{\alpha} f$ is in $L^{p}$, satisfies $\dot{L}_{\alpha}^{p} \simeq$ $\dot{F}_{p}^{\alpha, 2}, 1<p<\infty$.

In the more general setting of spaces of homogeneous type there are neither convolutions nor Fourier transform. Nevertheless, since an appropriate family of operators, $\left\{Q_{t}\right\}_{t>0}$, can be constructed in this context, namely, a collection of operators whose kernels satisfy certain size, smoothness and moment conditions and the property $\int_{0}^{\infty} Q_{t} \frac{d t}{t}=I$ on $L^{2}$, Han and Sawyer in [8], first, (considering a countable family $\left\{D_{k}\right\}_{k \in Z}$ satisfying the same features), and, next, Deng and Han in [3] proved reproduction formulae which allowed them to define Besov and Triebel-Lizorkin spaces on spaces of homogeneous type, to develop Littlewood- Paley characterizations and to obtain atomic decompositions of them.

In the context of $\mathbb{R}^{n}$ the fractional integral operator of order $\alpha, 0<\alpha<n$ is defined by

$$
I_{\alpha} f(x)=\int_{R^{n}} \frac{f(y)}{|x-y|^{n-\alpha}} d y
$$

It is very useful to know the action of $I_{\alpha}$ on spaces of smooth functions when studying regularity of classical solutions of differential equations uniformly elliptical, as is the case of the Laplace equation. The classical result on the behavior of the operator on Lipschitz spaces of functions is that it improves smoothness, mapping $\dot{\Lambda}_{\beta}$ on $\dot{\Lambda}_{\beta+\alpha / n}$, if $\beta+\alpha / n<1$.

Nicely, a similar behavior of the fractional integral operator is noticed when studying its action in a more general context, that is, on the family of Besov and Triebel-Lizorkin spaces on $R^{n}$. Moreover, $I_{\alpha}$ is an isomorphism between $\dot{B}_{p}^{\beta, q}$ and $\dot{B}_{p}^{\alpha+\beta, q}$ and between $\dot{F}_{p}^{\beta, q}$ and $\dot{F}_{p}^{\alpha+\beta, q}$ and its inverse is the fractional derivative operator $D_{\alpha}$, (see [5]). The Fourier transform is the tool used to prove these results. An important application of the previous facts is the identification, that we mentioned before, between $\dot{L}_{\alpha}^{p}\left(R^{n}\right)$ and $\dot{F}_{p}^{\alpha, 2}\left(R^{n}\right), \alpha>0,1<p<\infty$.

In the context of spaces of homogeneous spaces Gatto, Segovia and Vági in [6] extended the result on Lipschitz spaces by defining appropriate quasimetrics, $\delta_{\alpha}$, in terms of an approximation to the identity.

We are interested in considering a larger class of Besov and TriebelLizorkin spaces of distributions, on spaces of homogeneous type. For this class, defined in [9], local regularity is determined by more general 'moduli
of continuity' than the potentials $t^{\alpha}$. Those are growth functions $\psi(t)$ as, for instance, $t^{\beta} \log (1+t)$ or $\max \left(t^{\alpha}, t^{\beta}\right)$. In the range of these spaces can be characterized, for example, the Lipschitz space $\dot{\Lambda}^{\psi}$ of all complex functions such that there exists a constant $C$ satisfying

$$
|f(x)-f(y)| \leq C \psi(\delta(x, y))
$$

where $\psi$ belongs to a class (defined later in this work) of non-negative and quasi-increasing function such that $\lim _{t \rightarrow 0^{+}} \psi(t)=0$ and $\lim _{t \rightarrow \infty} \psi(t)=\infty$.

In order to find the 'natural' isomorphisms mapping one space of local regularity $\psi_{1}$, onto another of regularity $\psi_{2}$ the author and B. Viviani defined, in [10], the Integral and Derivative operators of 'functional order' $\phi$, $I_{\phi}$ and $D_{\phi}$. This definitions are similar in spirit to those of fractional integral and derivative operators given in [6].

The purpose of this work is to make a survey of the definitions and results obtained by the author and B. Viviani in studying the action of Integral and Derivative operators of functional order $\phi$ on $\dot{\Lambda}^{\psi}, \dot{B}_{p}^{\psi, q}$ and $\dot{F}_{p}^{\psi, q}$. We avoid showing those proofs whose technical difficulties make them too long to include here, referring the reader to the corresponding papers for them.

## 2 Previous definitions and known results

### 2.1 Growth functions

Let first define the class of functions, 'moduli of continuity', measuring local regularity of the distribution spaces and also the order of our operators. (For more details on these class see, [16], [17]). In this section by $\phi(t)$ we mean a non-negative real function defined on $t>0$.

We say that $\phi(t)$ is quasi-increasing (or quasi-decreasing) if there is a positive constant $C$ such that if $t_{1}<t_{2}$ then $\phi\left(t_{1}\right) \leq C \phi\left(t_{2}\right)$ (or $\phi\left(t_{2}\right) \leq$ $\left.C \phi\left(t_{1}\right)\right)$.

The function $\phi$ is of lower type $\beta \in R$ if there is a constant $C>0$ such that

$$
\begin{equation*}
\phi(u v) \leq C u^{\beta} \phi(v) \text { for } u<1 \text { and all } v>0 . \tag{4}
\end{equation*}
$$

Analogously, $\phi(t)$ is of upper type $\alpha \in R$ if there is a constant $C>0$ such that

$$
\begin{equation*}
\phi(u v) \leq C u^{\alpha} \phi(v) \text { for } u \geq 1 \text { and all } v>0 . \tag{5}
\end{equation*}
$$

Clearly, the potential $t^{\alpha}$ is of lower and upper type $\alpha$. The functions $\max \left(t^{\alpha}, t^{\beta}\right)$ and $\min \left(t^{\alpha}, t^{\beta}\right)$, with $\alpha<\beta$, are both of lower type $\alpha$ and upper
type $\beta$. Also, $t^{\beta}\left(1+\log ^{+} t\right)$, with $\beta \geq 0$, is of lower type $\beta$ and of upper type $\beta+\epsilon$, for every $\epsilon>0$. Notice that (4) is equivalent to

$$
\phi(u v) \geq \frac{1}{C} u^{\beta} \phi(v), \text { for } u \geq 1 \text { and all } v>0
$$

and (5) is equivalent to

$$
\phi(u v) \geq \frac{1}{C} u^{\alpha} \phi(v), \text { for } u<1, v>0
$$

If $\phi(t)$ is of lower type $\beta$ and of upper type $\alpha$ then $\beta \leq \alpha$. Notice that if $\phi$ is of upper type $\alpha$ and $\gamma>\alpha$ then $\phi$ also is of upper type $\gamma$, thus, there is a right half line of upper types. Analogously if $\phi$ is of lower type $\beta$ and $\gamma<\beta$ then it also is of lower-type $\gamma$.

Two functions $\psi(t)$ and $\phi(t)$ are equivalent, and we denote it $\psi \simeq \phi$, if there are positive constants $C_{1}$ and $C_{2}$ such that $C_{1} \leq \phi / \psi \leq C_{2}$. Lower and upper types are invariant by equivalence of functions. That is, if $\phi$ is of lower (upper) type $\delta$ and $\psi \simeq \phi$ then $\psi$ is of lower (upper) type $\delta$.

If $\phi(t)$ is of lower type $\beta>0$ then $\phi(t) / t^{\gamma}$ is quasi-increasing for each $\gamma \leq \beta$. Nevertheless, $\tilde{\phi}_{\gamma}(t)=t^{\gamma} \sup _{s \leq t} \frac{\phi(s)}{s^{\gamma}}$ is a function equivalent to $\phi$ such that $\tilde{\phi}_{\gamma}(t) / t^{\gamma}$ is increasing. But, if $\phi$ also is of upper type $\alpha$ and $\gamma<\beta$ then there exists a function $\phi_{\gamma}$ equivalent to $\phi$ which is differentiable and such that $\phi_{\gamma}(t) / t^{\gamma}$ is increasing. More precisely, the function

$$
\begin{equation*}
\phi_{\gamma}(t)=t^{\gamma} \int_{0}^{t} \frac{\phi(u)}{u^{\gamma+1}} d u \tag{6}
\end{equation*}
$$

is such a function.
Analogously, if $\phi$ is of upper type $\alpha$ and $\delta \geq \alpha$ then

$$
\begin{equation*}
\bar{\phi}_{\delta}(t)=t^{\delta} \sup _{s \geq t} \frac{\phi(s)}{s^{\delta}} \tag{7}
\end{equation*}
$$

is a function equivalent to $\phi$ such that $\bar{\phi}_{\delta}(t) / t^{\delta}$ is decreasing. If, in addition, $\phi$ is of lower type $\beta$ and $\delta>\alpha$ then, $\ddot{\phi}_{\delta}(t)=t^{\delta} \int_{t}^{\infty} \frac{\phi(s)}{s^{\delta+1}} d s$ is a differentiable function, equivalent to $\phi$ such that $\ddot{\phi}_{\delta}(t) / t^{\delta}$ is decreasing.

Let denote $\mathcal{C}$ the class of all non-negative functions $\phi$ of positive lower type and upper type lower than 1. Notice that the existence of positive lower type implies the integrability in the origin of $\phi(t) / t$ and, on the other hand, upper type lower than 1 implies that $\phi(t) / t$ is decreasing. Let also denote $\mathcal{A}$ the class of functions $\phi(t)$ defined on $t>0$ such that

$$
\begin{equation*}
\phi(t)=\phi(1) e^{\int_{1}^{t} \frac{\eta(s)}{s} d s} \tag{8}
\end{equation*}
$$

where $\eta(t)$ is a measurable function defined on $t>0$ and $\beta \leq \eta(t) \leq \alpha$ for some $0<\beta \leq \alpha<1$

The following lemma, proved in [4], shows that there is an identification between the classes $\mathcal{A}$ and $\mathcal{C}$

Lemma 1. The class $\mathcal{A}$ is included in the class $\mathcal{C}$ and for every function in $\mathcal{C}$ there is an equivalent function in $\mathcal{A}$.
Moreover, if $\phi(t)=\phi(1) e^{\int_{1}^{t} \frac{\eta(s)}{s} d s}, s_{\phi}=\sup _{t>0} \eta(t)$ and $i_{\phi}=\inf _{t>0} \eta(t)$, then for $s<1$

$$
\begin{equation*}
\phi(t) s^{s_{\phi}} \leq \phi(s t) \leq \phi(t) s^{i_{\phi}}, \tag{9}
\end{equation*}
$$

and, for $s>1$,

$$
\begin{equation*}
\phi(t) s^{i_{\phi}} \leq \phi(s t) \leq \phi(t) s^{s_{\phi}} . \tag{10}
\end{equation*}
$$

Notice that then $\phi$ is of lower-type $i_{\phi}>0$ and upper-type $s_{\phi}<1$.

### 2.2 Spaces of homogeneous type and approximations to the identity

Given a set $X$ a real valued function $\delta(x, y)$ defined on $X \times X$ is a quasi-distance on $X$ if there exists a positive constant $A$ such that for all $x, y, z \in X$ it verifies:

$$
\begin{aligned}
& \delta(x, y) \geq 0 \text { and } \delta(x, y)=0 \text { if and only if } x=y \\
& \delta(x, y)=\delta(y, x) \\
& \delta(x, y) \leq A[\delta(x, z)+\delta(z, y)] .
\end{aligned}
$$

In a set $X$ endowed with a quasi-distance $\delta(x, y)$, the balls $B_{\delta}(x, r)=\{y$ : $\delta(x, y)<r\}$ form a basis of neighborhoods of $x$ for the topology induced by the uniform structure on $X$.
Let $\mu$ be a positive measure on a $\sigma$ - algebra of subsets of $X$ which contains the open set and the balls $B_{\delta}(x, r)$. The triple $X:=(X, \delta, \mu)$ is a space of homogeneous type if there exists a finite constant $A^{\prime}>0$ such that

$$
\mu\left(B_{\delta}(x, 2 r)\right) \leq A^{\prime} \mu\left(B_{\delta}(x, r)\right)
$$

for all $x \in X$ and $r>0$. Macías and Segovia in [18], showed that it is always possible to find a quasi-distance $d(x, y)$ equivalent to $\delta(x, y)$ and $0<\theta \leq 1$, such that

$$
\begin{equation*}
\left|d(x, y)-d\left(x^{\prime}, y\right)\right| \leq C r^{1-\theta} d\left(x, x^{\prime}\right)^{\theta} \tag{11}
\end{equation*}
$$

holds whenever $d(x, y)<r$ and $d\left(x^{\prime}, y\right)<r$. If $\delta$ satisfies (11) then we say that $X$ is of order $\theta$. Furthermore, $X$ is a normal space if $A_{1} r \leq$
$\mu\left(B_{\delta}(x, r)\right) \leq A_{2} r$ for every $x \in X$ and $r>0$ and some positive constants $A_{1}$ and $A_{2}$.

In this work $X:=(X, \delta, \mu)$ means a normal space of homogeneous type of order $\theta$ and $A$ denotes the constant of the triangular inequality associated to $\delta$.

Let now define an approximation to the identity as given in [2], [3](where it is given for non-compact supports) and [6]):
A family $\left\{S_{t}\right\}_{t>0}$ of operators is an approximation to the identity if there exist $\epsilon \leq \theta$ and $C, C_{1}$ and $C_{2}<\infty$ such that for all $t>0$ and all $x, x^{\prime}, y$ and $y^{\prime} \in X$, the kernel $s_{t}(x, y)$ of $S_{t}$, are functions from $X \times X$ into $\mathbb{R}$ satisfying:

$$
\begin{align*}
& s_{t}(x, y)=0 \text { for } \delta(x, y)>b_{1} t \text { and }\left|s_{t}(x, y)\right| \leq \frac{C_{1}}{t}  \tag{12}\\
& \frac{C_{2}}{t}<s_{t}(x, y) \text { if } \delta(x, y)<b_{2} t  \tag{13}\\
& \left|s_{t}(x, y)-s_{t}\left(x^{\prime}, y\right)\right|+\left|s_{t}(y, x)-s_{t}\left(y, x^{\prime}\right)\right| \leq C \frac{\delta\left(x, x^{\prime}\right)^{\epsilon}}{t^{1+\epsilon}}  \tag{14}\\
& \left|\left(s_{t}(x, y)-s_{t}\left(x^{\prime}, y\right)\right)-\left(s_{t}\left(x, y^{\prime}\right)-s_{t}\left(x^{\prime}, y^{\prime}\right)\right)\right| \leq C \frac{\delta\left(x, x^{\prime}\right)^{\epsilon} \delta\left(y, y^{\prime}\right)^{\epsilon}}{t^{1+2 \epsilon}},  \tag{15}\\
& \int s_{t}(x, y) d \mu(y)=\int s_{t}(x, y) d \mu(x)=1, \text { and }  \tag{16}\\
& { }_{t}(\boldsymbol{x}, y) \text { is continuously differentiable in } t . \tag{17}
\end{align*}
$$

Inequality (15) is not needed to our purposes but it follows from the construction made in [2] and, also, that the kernels may be chosen to be positive and symmetric, that is, for all $t>0, x$ and $y \in X$

$$
\begin{align*}
& s_{t}(x, y) \geq 0  \tag{18}\\
& s_{t}(x, y)=s_{t}(y, x) . \tag{19}
\end{align*}
$$

The family $\left\{S_{t}\right\}_{t>0}$ satisfies the condition $\lim _{t \rightarrow 0} S_{t}=I$ and $\lim _{t \rightarrow \infty} S_{t}=0$ in the strong operator topology on $L^{2}$.

In view of (12) to (19), the functions $q_{t}(x, y)=-t \frac{\partial}{\partial t} s_{t}(x, y)$ are symmetric and also satisfy (12), (14), (15) and

$$
\left.\int q_{t}(x, y) d \mu(y)=\int q_{t}(y, x) d \mu(y)=0 \text { for all } x \in X \text { and } t>0.20\right)
$$

Let

$$
\begin{equation*}
Q_{t} f=-t \frac{\partial}{\partial t} S_{t} \tag{21}
\end{equation*}
$$

be the operator defined by $Q_{t} f(x)=\int_{X} q_{t}(x, y) f(y) d \mu(y)$, for $f \in L_{\text {loc }}^{1}$ and $t>0$. The family $\left\{Q_{t}\right\}_{t>0}$ satisfies the condition $\int_{0}^{\infty} Q_{t} \frac{d t}{t}=I$ in the strong operator topology on $L^{2}$ in the sense that

$$
\lim _{\epsilon \rightarrow 0, R \rightarrow \infty}\left\|\int_{\epsilon}^{R} Q_{t} f \frac{d t}{t}-f\right\|_{2}=0
$$

The following Calderón type reproduction formula on spaces of homogeneous type plays a crucial roll in obtaining Littlewood-Paley characterizations of Besov and Triebel-Lizorkin spaces and proving continuity of the operators on them. A discrete version of it was proved in [8], and a continuous one, associated to a para-accretive function, in [3].
Theorem 2. Suppose that $\left\{S_{t}\right\}_{t>0}$ is an approximation to the identity and $\left\{Q_{t}\right\}_{t>0}$ are as in (21). Then there exists a family of operators $\left\{\tilde{Q}_{t}\right\}_{t>0}$ (or $\left.\left\{\hat{Q}_{t}\right\}_{t>0}\right)$ such that for all $f \in \mathcal{M}^{(\beta, \gamma)}$ with $0<\beta, \gamma<\epsilon$,

$$
\begin{equation*}
f=\int_{0}^{\infty} \tilde{Q}_{t} Q_{t} f \frac{d t}{t} \quad \text { or } f=\int_{0}^{\infty} Q_{t} \hat{Q}_{t} f \frac{d t}{t} \tag{22}
\end{equation*}
$$

where the integral converges in the norm of $L^{p}, 1<p<\infty$, and $\mathcal{M}^{\left(\beta^{\prime}, \gamma^{\prime}\right)}$, with $\beta^{\prime}<\beta$ and $\gamma^{\prime}<\gamma$. Moreover, $\tilde{Q}_{t}(x, y)$, the kernel of $\tilde{Q}_{t}$ satisfies the following estimates: for $\epsilon^{\prime}, 0<\epsilon^{\prime}<\epsilon$, there exists a constant $C$ such that

$$
\begin{align*}
& \left|\tilde{Q}_{t}(x, y)\right| \leq C \frac{t^{\epsilon^{\prime}}}{(t+\delta(x, y))^{1+\epsilon^{\prime}}} \\
& \left|\tilde{Q}_{t}(x, y)-\tilde{Q}_{t}\left(x^{\prime}, y\right)\right| \leq C\left(\frac{\delta\left(x, x^{\prime}\right)}{t+\delta(x, y)}\right)^{\epsilon^{\prime}} \frac{t^{\epsilon^{\prime}}}{(t+\delta(x, y))^{1+\epsilon^{\prime}}} \\
& \text { for } \delta\left(x, x^{\prime}\right) \leq \frac{1}{2 A} \delta(x, y)  \tag{23}\\
& \int \tilde{Q}_{t}(x, y) d \mu(y)=\int \tilde{Q}_{t}(x, y) d \mu(x)=0 \text { for all } t>0
\end{align*}
$$

### 2.3 Lipschitz functions, molecules, Besov and Triebel-Lizorkin spaces

Let now consider a positive and quasi-increasing function $\eta(t)$ defined on $t>0$ such that $\lim _{t \rightarrow 0} \eta(t)=0$. The Lipschitz space $\dot{\Lambda}^{\eta}$ is the class of all functions $f: X \rightarrow \mathbb{C}$ such that

$$
|f|_{\eta}=\sup _{x \neq y} \frac{|f(x)-f(y)|}{\eta(\delta(x, y))}<\infty
$$

The quantity $|f|_{\eta}$ defines a semi-norm on $\dot{\Lambda}^{\eta}$, since $|f|_{\eta}=0$ for all constants functions $f$. Given a ball $B$ in $X, \Lambda^{\eta}(B)$ denotes the set of functions $f \in \dot{\Lambda}^{\eta}$ with support in $B$. Since a function belonging to this space is bounded, the number

$$
\|f\|_{\eta}=\|f\|_{\infty}+|f|_{\eta},
$$

defines a norm that gives a Banach structure to $\Lambda^{\eta}(B)$. We say that a function $f$ belongs to $\Lambda_{0}^{\eta}$ iff $f \in \Lambda^{\eta}(B)$ for some ball $B$. The space $\Lambda_{0}^{\eta}$ is the inductive limit of the Banach spaces $\Lambda^{\eta}(B)$. Finally, $\left(\Lambda_{0}^{\eta}\right)^{\prime}$ will stand for the dual space of $\Lambda_{0}^{\eta}$.

Another suitable class of test functions of mean value 0 was defined in [8].

Let $0<\beta \leq 1, \gamma>0$ and $x_{0} \in X$ be fix. A function $f$ is a smooth molecule of type $(\beta, \gamma)$ of width $d$ centered in $x_{0}$, if there exists a constant $C>0$ such that

$$
\begin{align*}
& |f(x)| \leq C \frac{d^{\gamma}}{\left(d+\delta\left(x, x_{0}\right)\right)^{1+\gamma}}  \tag{24}\\
& \left|f(x)-f\left(x^{\prime}\right)\right| \leq C \delta\left(x, x^{\prime}\right)^{\beta}\left[\frac{d^{\gamma}}{\left(d+\delta\left(x, x_{0}\right)\right)^{1+\gamma}}+\frac{d^{\gamma}}{\left(d+\delta\left(x^{\prime}, x_{0}\right)\right)^{1+\gamma}}\right] \tag{25}
\end{align*}
$$

$$
\begin{equation*}
\int f(x) d \mu(x)=0 \tag{26}
\end{equation*}
$$

hold for every $x \in X$.
If the norm $\|f\|_{(\beta, \gamma)}$ is defined by the lowest of the constants appearing in (24) and (25), the set $\mathcal{M}^{\left(\beta, \gamma, x_{0}, d\right)}$ of all smooth molecules of type $(\beta, \gamma)$ of width $d$ centered in $x_{0}$ is a Banach space. By fixing $x_{0} \in X$ and $d=1$, that space will be named $\mathcal{M}^{(\beta, \gamma)}$, and its dual space, $\left(\mathcal{M}^{(\beta, \gamma)}\right)^{\prime}$. Along this work $<h, f>$ denotes the natural application of $h \in\left(\mathcal{M}^{(\beta, \gamma)}\right)^{\prime}$ to $f \in \mathcal{M}^{(\beta, \gamma)}$.

Local regularity of Besov and Triebel-Lizzorkin spaces we consider here is associated to a function $\psi$ having a representation as a quotient of quasiincreasing functions. More precisely, in the sequel, we consider $\psi=\psi_{1} / \psi_{2}$,
where $\psi_{1}(t)$ and $\psi_{2}(t)$ are quasi-increasing functions of upper types $s_{1}<\epsilon$ and $s_{2}<\epsilon$, respectively.

For $f \in\left(\mathcal{M}^{(\beta, \gamma)}\right)^{\prime}$, with $0<\beta, \gamma<\epsilon$, we define

$$
\begin{equation*}
\|f\|_{\dot{B}_{p}^{\psi, q}}=\left(\int_{0}^{\infty}\left(\frac{1}{\psi(t)}\left\|Q_{t} f\right\|_{p}\right)^{q} \frac{d t}{t}\right)^{\frac{1}{q}} \quad \text { if } 1 \leq p \leq \infty, 1 \leq q \leq \infty \tag{27}
\end{equation*}
$$

with the obvious change for the case $q=\infty$. By interchanging the order of the norms in $L^{p}$ and $l^{q}$ it is also defined the norm

$$
\begin{equation*}
\|f\|_{\dot{F}_{p}^{\psi, q}}=\left\|\left(\int_{0}^{\infty}\left(\frac{1}{\psi(t)}\left|Q_{t} f\right|\right)^{q} \frac{d t}{t}\right)^{\frac{1}{q}}\right\|_{L^{p}}, \text { if } 1<p, q<\infty . \tag{28}
\end{equation*}
$$

Notice that in any case, $\|f\|=0$ if, and only if, $Q_{t} f$ is the zero function for all $t>0$. But this is equivalent to having $S_{t} f$ constant. Finally this means that the distribution $f$ is a constant. Thus we work with modulo constants when considering (27) and 28 ; that is, $f \in\left(\mathcal{M}^{(\beta, \gamma)}\right)^{\prime} / \mathbb{C}$ in those equalities.

The Besov space of order $\psi, \dot{B}_{p}^{\psi, q}$, for $1 \leq p, q \leq \infty$, is the set of all $f \in\left(\mathcal{M}^{(\beta, \gamma)}\right)^{\prime}$, with $\beta>s_{1}$ and $\gamma>s_{2}$, such that

$$
\|f\|_{\dot{B}_{p}^{\psi, q}}<\infty \text { and }|\langle f, h\rangle| \leq C\|f\|_{\dot{B}_{p}^{\psi, q}}\|h\|_{(\beta, \gamma)}
$$

for all $h \in \mathcal{M}^{(\beta, \gamma)}$.
Analogously, the Triebel-Lizorkin space of order $\psi, \dot{F}_{p}^{\psi, q}$, with $1<p, q<\infty$, is the set of all $f \in\left(\mathcal{M}^{(\beta, \gamma)}\right)^{\prime}$, with $\beta>s_{1}$ and $\gamma>s_{2}$, such that

$$
\|f\|_{\dot{F}_{p}^{\psi, q}}<\infty, \text { and }|\langle f, h\rangle| \leq\|f\|_{\dot{F}_{p}^{\psi, q}}\|h\|_{(\beta, \gamma)},
$$

for all $h \in \mathcal{M}^{(\beta, \gamma)}$.
When $\psi(t)=t^{\alpha}$ we recover the spaces $\dot{B}_{p}^{\alpha, q}$ and $\dot{F}_{p}^{\alpha, q}$ with $-\epsilon<\alpha<\epsilon$, as defined in [3].
Applying Theorem 3 it follows that replacing the family $\left\{Q_{t}\right\}_{t>0}$ by a family $\left\{\tilde{Q}_{t}\right\}_{t>0}$ satisfying (12), (14) and (20), the resulting norms are equivalent to those defined in (27) and (28). Moreover, these norms are equivalent to those given in [10] or [11] in terms of the differences of a discrete approximation to the identity $\left\{S_{k}\right\}_{k \in Z}$.

By using the properties of the function $\psi$, and in the way it is proved in [8], follows that the classes $\dot{B}_{p}^{\psi, q}, 1 \leq p, q<\infty$ and $\dot{F}_{p}^{\psi, q}, 1<p, q<\infty$ are Banach spaces and their dual spaces are $\dot{B}_{p^{\prime}}^{1 / \psi, q^{\prime}}$ and $\dot{F}_{p^{\prime}}^{1 / \psi, q^{\prime}}$ respectively, with $1 / p+1 / p^{\prime}=1$ and $1 / q+1 / q^{\prime}=1$.
Moreover, the molecular space $\mathcal{M}^{(\beta, \gamma)}$ is embedded in $\dot{B}_{p}^{\psi, q}$ and $\dot{F}_{p}^{\psi, q}$ when $s_{1}<\beta$ and $s_{2}<\gamma$ and $\mathcal{M}^{\left(\epsilon^{\prime}, \epsilon^{\prime}\right)}$ is dense in $\dot{B}_{p}^{\psi, q}$ and $\dot{F}_{p}^{\psi, q}$ for all $\epsilon^{\prime}$, such that $\max \left(s_{1}, s_{2}\right)<\epsilon^{\prime}<\epsilon$.

### 2.4 Singular integral operators and $T 1$-theorem

A complex-valued function $K(x, y)$ defined on $\Omega=\{(x, y) \in X \times X: x \neq y\}$ is called a standard kernel if there exist $0<\epsilon \leq \theta$, and $C<\infty$ such that for all $x, y \in X$ with $x \neq y$,

$$
\begin{equation*}
|K(x, y)| \leq C \delta(x, y)^{-1}, \quad \text { for every } \quad x \neq y \tag{29}
\end{equation*}
$$

and

$$
\begin{equation*}
\left|K(x, y)-K\left(x^{\prime}, y\right)\right|+\left|K(y, x)-K\left(y, x^{\prime}\right)\right| \leq C \delta\left(x, x^{\prime}\right)^{\epsilon} / \delta(x, y)^{1+\epsilon} \tag{30}
\end{equation*}
$$

for $\delta(x, y)>2 A \delta\left(x, x^{\prime}\right)$.
A continuous linear operator $T: \Lambda_{0}^{\beta} \rightarrow\left(\Lambda_{0}^{\beta}\right)^{\prime}$ is a singular integral operator if there is a standard kernel $K$ such that

$$
<T f, g>=\iint K(x, y) f(y) g(x) d \mu(y) d \mu(x)
$$

for all $f, g \in \Lambda_{0}^{\beta}$ with supp $f \cap \operatorname{supp} g=\emptyset$. We then write $T \in C Z K(\epsilon)$.
A singular integral operator $T$ is a Calderón-Zygmund operator if it can be extended to a bounded operator on $L^{2}$. For such operators we write $T \in C Z 0$.

## 3 Integral and Derivative Operators of order $\phi$

To set the idea of the definitions given later in this section let first consider the kernels $|x-y|^{\alpha-n},-\infty<\alpha<n$ associated to the fractional integral and derivative operators $I_{\alpha}$ and $D_{\alpha}$ in the context of $R^{n}$. For $x \neq y$,

$$
\begin{align*}
|x-y|^{\alpha-n} & =(n-\alpha) \int_{|x-y|}^{\infty} t^{\alpha-n-1} d t \\
& =(n-\alpha) \int_{0}^{\infty} t^{\alpha-1} t^{-n} \chi_{B(0, t)}(|x-y|) d t \\
& =\omega_{n}(n-\alpha) \int_{0}^{\infty} t^{\alpha-1} s_{t}(x-y) d t \tag{31}
\end{align*}
$$

where $\omega_{n}$ is the measure of the unitary ball of $R^{n}$ and the family $s_{t}=$ $\omega_{n}^{-1} t^{-n} \chi_{B(0, t)}, t>0$ determines an approximation to the identity.

In the setting of spaces of homogeneous type, Gatto, Segovia and Vági, in [6], defined kernels $\delta_{\alpha}(x, y)^{\alpha-1}$ on this idea, by means of an approximation to the identity as defined in Section 2.2.

If we consider a function $\phi(t)$ instead of the potential $t^{\alpha},(\alpha>0)$, we aim to define the kernel associated to the integral operator of order $\phi$ resembling $\phi(\delta(x, y)) / \delta(x, y)$, as well as the kernel of the derivative operator should resemble $(\phi(\delta(x, y)) \delta(x, y))^{-1}$. Thus, in the spirit of the definitions given in [6], we defined in [10] the kernels associated to our operators.

Let remark that in [10] and [11] the definitions of the operators were given for $\phi \in \mathcal{C}$, that is a function of positive lower type and upper type less than 1 , and all the following results in fact are obtained for that class without any change. We restrict here to consider the class $\mathcal{A}$, which by Lemma 1 is identified with $\mathcal{C}$, since we are looking for the invertibility of the integral and derivative operators.

In the sequel, we consider $\phi \in \mathcal{A}$ as in (8). By the remark given at the end of section 2.1, $\phi$ satisfies (9) and (10) with $\alpha:=s_{\phi}<1$ as upper type and $\beta:=i_{\phi}>0$ as lower type, and $\eta(t)=\frac{t \phi^{\prime}(t)}{\phi(t)}$ is such that $\beta \leq \eta \leq \alpha$.

Let also consider a positive and symmetric approximation to the identity of order $\epsilon \leq \theta,\left\{S_{t}\right\}_{t>0}$, associated to the family of kernels $s_{t}(x, y), t>0$, as defined in section 2.2.

Set

$$
\begin{equation*}
K_{\phi}(x, y)=\int_{0}^{\infty} \frac{\phi(t) \eta(t)}{t} s_{t}(x, y) d t \quad \text { for } \quad x \neq y \tag{32}
\end{equation*}
$$

The application $K_{\phi}$ has the representation in terms of a quasi-metric we are seeking for. Indeed, since $\alpha<1$ then $\phi(t) / t$ is continuous, decreasing and invertible on $R^{+}$and, since the integral in (32) is positive then it is possible to define a non-negative application $\delta_{\phi}(x, y)$ as the unique solution of the equation

$$
\begin{aligned}
\frac{\phi\left(\delta_{\phi}(x, y)\right)}{\delta_{\phi}(x, y)} & =K_{\phi}(x, y) \text { for } x \neq y, \\
\text { and } \delta_{\phi}(x, y) & =0 \quad \text { for } x=y .
\end{aligned}
$$

Lemma 4. $\phi\left(\delta_{\phi}(x, y)\right) / \delta_{\phi}(x, y)$ is equivalent to $\phi(\delta(x, y)) / \delta(x, y)$ and, thus, $\delta_{\phi}(x, y)$ defines a quasi-metric equivalent to the natural quasi-metric $\delta$ of $X$.
Proof. From (12), the substitution $t=u \delta(x, y) / b_{1}$ and the left inequality in (9)

$$
\begin{align*}
\int_{0}^{\infty} \frac{\phi(t) \eta(t)}{t} s_{t}(x, y) d t & \leq c_{1} \alpha \int_{\delta(x, y) / b_{1}}^{\infty} \frac{\phi(t)}{t^{2}} d t \\
& \leq \frac{c_{1} b_{1} \alpha}{\delta(x, y)} \phi\left(\delta(x, y) / b_{1}\right) \int_{1}^{\infty} \frac{1}{u^{2-\alpha}} d u \\
& \leq C_{1} \frac{\phi(\delta(x, y))}{\delta(x, y)} \tag{33}
\end{align*}
$$

since $\alpha<1$ and $\phi\left(s / b_{1}\right) \leq \max \left(1,1 / b_{1}^{\alpha}\right) \phi(s)$ for all $s>0$.
On the other hand, by (13) and since the right inequality in (9) implies that $\phi$ is increasing,

$$
\begin{align*}
\int_{0}^{\infty} \frac{\phi(t) \eta(t)}{t} s_{t}(x, y) d t & \geq c_{2} \beta \int_{\delta(x, y) / b_{2}}^{\infty} \frac{\phi(t)}{t^{2}} d t \\
& \geq \frac{c_{2} b_{2} \beta}{\delta(x, y)} \phi\left(\delta(x, y) / b_{2}\right) \int_{1}^{\infty} \frac{1}{u^{2}} d u  \tag{34}\\
& \geq C_{2} \frac{\phi(\delta(x, y))}{\delta(x, y)} \tag{35}
\end{align*}
$$

since $\phi\left(s / b_{2}\right) \geq \min \left(1,1 / b_{2}^{\alpha}\right) \phi(s)$ for all $s>0$.
From (34) and (33),

$$
C_{2} \frac{\phi(\delta(x, y))}{\delta(x, y)} \leq \frac{\phi\left(\delta_{\phi}(x, y)\right)}{\delta_{\phi}(x, y)} \leq C_{1} \frac{\phi(\delta(x, y))}{\delta(x, y)}
$$

The equivalence between $\delta_{\phi}$ and $\delta$ now follows from the invertibility of $\phi(t) / t$ and the types of $\phi$.

Let now define the kernel associated to the derivative operator, (in fact, in this case only is required $\beta \geq 0$ and $\alpha$ finite):

$$
K_{1 / \phi}(x, y)=\int_{0}^{\infty} \frac{\eta(t)}{\phi(t) t} s_{t}(x, y) d t \text { for } x \neq y
$$

Reasoning as in the previous case, there is a quasi-distance $\delta_{1 / \phi}(x, y)$, equivalent to $\delta$, defined as the unique solution of

$$
\begin{aligned}
& \left(\phi\left(\delta_{1 / \phi}(x, y)\right) \delta_{1 / \phi}(x, y)\right)^{-1}=K_{1 / \phi}(x, y) \text { for } x \neq y \text { and } \\
& \delta_{1 / \phi}(x, y)=0 \text { for } x=y .
\end{aligned}
$$

Moreover, $K_{1 / \phi}(x, y)$ is equivalent to $(\phi(\delta(x, y)) \delta(x, y))^{-1}$.
Notice that the kernels $K_{\phi}(x, y)$ and $K_{1 / \phi}(x, y)$ are symmetric since $s_{t}(x, y)$ is.

To define the integral and derivative operators and study their action on Lipschitz, Besov and Triebel-Lizorkin spaces, one needs to know regularity conditions of their kernels. These properties, and also a cancellation one for $K_{\phi}$, are obtained from conditions (12) to (14), of $s_{t}$, and the types of $\phi$.

Lemma 5. Let $\phi$ be of lower-type $\beta$ and upper type $\alpha$.
If $\alpha<1$ then

$$
\left|K_{\phi}(x, y)-K_{\phi}\left(x^{\prime}, y\right)\right| \leq C \frac{\delta\left(x, x^{\prime}\right)^{\epsilon}}{\delta(x, y)^{1+\epsilon}} \phi(\delta(x, y)) \quad \text { if } \quad \delta(x, y) \geq 2 A \delta\left(x, x^{\prime}\right)
$$

and, if $\alpha<\epsilon$ and $\beta>0$ then

$$
\int_{X}\left[K_{\phi}(x, y)-K_{\phi}\left(x^{\prime}, y\right)\right] d \mu(y)=0, x \text { and } x^{\prime} \in X
$$

If $\beta \geq 0$ and $\alpha$ is finite then

$$
\left|K_{1 / \phi}(x, y)-K_{1 / \phi}\left(x^{\prime}, y\right)\right| \leq C \frac{\delta\left(x, x^{\prime}\right)^{\epsilon}}{\delta(x, y)^{1+\epsilon}} \frac{1}{\phi(\delta(x, y))} \quad \text { if } \quad \delta(x, y) \geq 2 A \delta\left(x, x^{\prime}\right)
$$

The Integral operator of functional order $\phi$ is defined as follows: For $f \in \dot{\Lambda}^{\xi} \cap L^{1}, \xi$ a quasi-increasing function of upper-type $s>0$,

$$
\begin{equation*}
I_{\phi} f(x)=\int_{X} K_{\phi}(x, y) f(y) d \mu(y) \tag{36}
\end{equation*}
$$

There is also an extension to $\dot{\Lambda}^{\xi}$ :
If $\alpha+s<\epsilon$ and $f \in \dot{\Lambda}^{\xi}$ then

$$
\begin{equation*}
\tilde{I}_{\phi} f(x):=\int_{X}\left(K_{\phi}(x, y)-K_{\phi}\left(x_{0}, y\right)\right) f(y) d \mu(y) \tag{37}
\end{equation*}
$$

for every $x \in X$ and an arbitrary fix $x_{0} \in X$.
If $f \in \dot{\Lambda}^{\xi} \cap L^{1}$, then $\tilde{I}_{\phi} f$ coincides with $I_{\phi} f$ as an element of $\dot{\Lambda}^{\xi \phi}$ since $\tilde{I}_{\phi} f(x)=I_{\phi} f(x)-I_{\phi} f\left(x_{0}\right)$.

The Derivative operator of functional order $\phi$ is defined in the following way:
For $f \in \dot{\Lambda}^{\xi} \cap L^{\infty}, \xi$ a function of positive lower-type $\lambda>\alpha$ and finite upper-type,

$$
\begin{equation*}
D_{\phi} f(x)=\int_{X} K_{1 / \phi}(x, y)(f(y)-f(x)) d \mu(y) \tag{38}
\end{equation*}
$$

Its extension to $\dot{\Lambda}^{\xi}$, $\xi$ of positive lower-type $\lambda>\alpha$ and finite upper-type, is given by

$$
\begin{equation*}
\tilde{D}_{\phi} f(x)=\int_{X}\left(K_{1 / \phi}(x, y)(f(y)-f(x))-K_{1 / \phi}\left(x_{0}, y\right)\left(f(y)-f\left(x_{0}\right)\right)\right) d \mu(y) \tag{39}
\end{equation*}
$$

for each $x \in X$ and an arbitrary, but fix, $x_{0} \in X$. If $f \in \dot{\Lambda}^{\xi} \cap L^{\infty}$, then $\tilde{D}_{\phi} f$ coincides with $D_{\phi} f$ as an element of $\dot{\Lambda}^{\xi / \phi}$, since $\tilde{D}_{\phi} f(x)=D_{\phi} f(x)-D_{\phi} f\left(x_{0}\right)$.

We can now state the continuity of the integral and derivative operator on Lipschitz spaces

Theorem 6. Let $\phi \in \mathcal{A}$, and denote $\alpha=s_{\phi}$ and $\beta=i_{\phi}$. Let also $\xi$ be $a$ quasi-increasing function of upper type $s, \alpha+s<\epsilon$.

If $f \in \dot{\Lambda}^{\xi} \cap L^{1}$ then there is a constant $C>0$, independent of $f$, such that

$$
\left|I_{\phi} f\right|_{\dot{\lambda}_{\xi \phi}} \leq C|f|_{\dot{\lambda} \xi} .
$$

If $f \in \dot{\Lambda}^{\xi}$ there is a constant $C>0$, independent of $f$, such that

$$
\left|\tilde{I}_{\phi} f\right|_{\dot{\Lambda} \xi \phi} \leq C|f|_{\dot{\Lambda} \xi} .
$$

Theorem 7. Let $\phi$ be like in the previous theorem and $\xi$ a quasi-increasing function of lower type $\lambda>\alpha$ and upper type $s<\epsilon+\beta$. If $f \in \dot{\Lambda}^{\xi} \cap L^{\infty}$ then

$$
\left\|D_{\phi} f\right\|_{\xi / \phi} \leq C\|f\|_{\xi} .
$$

If $f \in \dot{\Lambda}^{\xi}$,

$$
\left|\tilde{D}_{\phi} f\right|_{\xi / \phi} \leq C|f|_{\xi} .
$$

Since molecules are bounded and integrable Lipschitz functions, the integral and derivatives operators are well defined on them. Moreover, $I_{\phi}$ is a linear continuous operator from $M^{(\beta, \gamma)}$ to $\left(M^{\left(\beta^{\prime}, \gamma^{\prime}\right)}\right)^{\prime}$, for every $\beta, \gamma, \beta^{\prime}$ and $\gamma^{\prime}>0$ and $<I_{\phi} f, g>=<f, I_{\phi} g>$.
Analogously, if $s_{\phi}<\beta$ then $D_{\phi}$ is a linear continuous operator from $M^{(\beta, \gamma)}$ to $\left(M^{\left(\beta^{\prime}, \gamma^{\prime}\right)}\right)^{\prime}, \gamma, \gamma^{\prime}$ and $\beta^{\prime}>0$. Moreover, if also $s_{\phi}<\beta^{\prime}$ then $\left.<D_{\phi} f, g\right\rangle=<$ $D_{\phi} g, f>$.

We are then able to apply Theorems 2 and 3 to obtain decompositions of $I_{\phi}$ and $D_{\phi}$ on molecular spaces

$$
\begin{aligned}
& <I_{\phi} f, g>=\int_{0}^{\infty} \int_{0}^{\infty}<Q_{t} I_{\phi} Q_{s}\left(\hat{Q}_{s} f\right), \tilde{Q}_{t}^{*} g>\frac{d t}{t} \frac{d s}{s} \\
& <D_{\phi} f, g>=\int_{0}^{\infty} \int_{0}^{\infty}<Q_{t} D_{\phi} Q_{s}\left(\hat{Q}_{s} f\right), \tilde{Q}_{t}^{*} g>\frac{d t}{t} \frac{d s}{s}
\end{aligned}
$$

with Triebel-Lizorkin norms of $I_{\phi} f$ and $D_{\phi} f$ given by

$$
\begin{aligned}
\left\|I_{\phi} f\right\|_{\dot{F}_{p}^{\phi \psi, q}} & =\left\|\left(\int_{0}^{\infty}\left(\frac{1}{\phi(t) \psi(t)}\left|Q_{t} I_{\phi} f\right|\right)^{q} \frac{d t}{t}\right)^{1 / q}\right\|_{p} \\
& \leq\left\|\left(\int_{0}^{\infty}\left(\frac{1}{\phi(t) \psi(t)} \int_{0}^{\infty}\left|Q_{t} I_{\phi} Q_{s}\left(\hat{Q}_{s} f\right)\right| \frac{d s}{s}\right)^{q} \frac{d t}{t}\right)^{1 / q}\right\|_{p} \\
\left\|D_{\phi} f\right\|_{\dot{F}_{p}^{\psi / \phi, q}} & \leq\left\|\left(\int_{0}^{\infty}\left(\frac{\phi(t)}{\psi(t)} \int_{0}^{\infty}\left|Q_{t} D_{\phi} Q_{s}\left(\hat{Q}_{s} f\right)\right| \frac{d s}{s}\right)^{q} \frac{d t}{t}\right)^{1 / q}\right\|_{p}
\end{aligned}
$$

and analogous inequalities follow by interchanging the order of the norms $p$ and $q$ for the Besov norms of $I_{\phi}$ and $D_{\phi}$.

In this direction it was proved in ([10]) that
Theorem 8. Let $\phi \in \mathcal{A}$ and denote $\beta:=i_{\phi}>0$ and $\alpha:=s_{\phi}<\epsilon$. If $s_{1}+\alpha<\epsilon$ and $s_{2}+\alpha-\beta<\epsilon$, then there is a constant $C>0$ such that

$$
\begin{align*}
\left\|I_{\phi} f\right\|_{\dot{F}_{p}^{\phi \psi, q}} & \leq C\|f\|_{\dot{F}_{p}^{\psi, q},},  \tag{40}\\
\text { and }\left\|I_{\phi} f\right\|_{\dot{B}_{p}^{\phi \psi, q}} & \leq C\|f\|_{\dot{B}_{p}^{\psi, q}} . \tag{41}
\end{align*}
$$

If $s_{1}<\epsilon$ and $s_{2}+\alpha<\epsilon$ then there is a constant $C>0$ such that

$$
\begin{align*}
\left\|D_{\phi} f\right\|_{\dot{F}_{p}^{\psi / \phi, q}} & \leq C\|f\|_{\dot{F}_{p}^{\psi, q}},  \tag{42}\\
\text { and }\left\|D_{\phi} f\right\|_{\dot{B}_{p}^{\psi / \phi, q}} & \leq C\|f\|_{\dot{B}_{p}^{\psi, q}} . \tag{43}
\end{align*}
$$

where the range $1<p, q<\infty$ is considered for the scale of $F$ - spaces and $1 \leq p, q<\infty$ for the $B$-spaces and $\psi=\psi_{1} / \psi_{2}, \psi_{i}$ of upper-type $s_{i}, i=1,2$.

The key tools in the proof of the above inequalities are the following estimates on the family of operators $Q_{t} I_{\phi} Q_{s}$ and $Q_{t} D_{\phi} Q_{s}$, which are continuous versions of those obtained in [10].

$$
\begin{align*}
\left|Q_{t} I_{\phi} Q_{s} h(x)\right| & \leq C \phi(t \wedge s)\left(\frac{t}{s} \wedge \frac{s}{t}\right)^{\left(\epsilon-s_{\phi}\right)} M|h|(x)  \tag{44}\\
\left|Q_{t} D_{\phi} Q_{s} h(x)\right| & \leq C \frac{1}{\phi(t \wedge s)}\left(\frac{t}{s} \wedge \frac{s}{t}\right)^{\epsilon} M|h|(x) \tag{45}
\end{align*}
$$

where $M$ denotes the Hardy-Littlewood maximal operator and

$$
\begin{align*}
\left\|Q_{t} I_{\phi} Q_{s} h\right\|_{p} & \leq C \phi(t \wedge s)\left(\frac{t}{s} \wedge \frac{s}{t}\right)^{\left(\epsilon-s_{\phi}\right)}\|h\|_{p}  \tag{46}\\
\left\|Q_{t} D_{\phi} Q_{s} h\right\|_{p} & \leq C \frac{1}{\phi(t \wedge s)}\left(\frac{t}{s} \wedge \frac{s}{t}\right)^{\epsilon}\|h\|_{p} \tag{47}
\end{align*}
$$

for $1 \leq p<\infty$.
Let prove (40) and (41), the other two inequalities following with similar arguments.

By density arguments it is enough to prove those inequalities for $f \in$ $M^{(\epsilon, \epsilon)}$, where $\max \left(s_{1}, s_{2}\right)<\epsilon$.

We will also use the following inequalities. Since $\psi_{2}$ is quasi-increasing and $\psi_{1}$ is of upper-type $s_{1}$ then, for $v \geq 1$

$$
\begin{equation*}
\frac{1}{\psi(u)}=\frac{\psi_{2}(u)}{\psi_{1}(u)} \leq C v^{s_{1}} \frac{\psi_{2}(u v)}{\psi_{1}(u v)}=C \frac{v^{s_{1}}}{\psi(u v)}, \tag{48}
\end{equation*}
$$

and, since $\psi_{1}$ is quasi-increasing and $\psi_{2}$ is of upper-type $s_{2}$ then, for $v<1$

$$
\begin{equation*}
\frac{1}{\psi(u)} \leq C \frac{v^{-s_{2}}}{\psi(u v)}, \tag{49}
\end{equation*}
$$

Then, by the substitution $t=u$ and $s=u v$,

$$
\begin{aligned}
\left\|I_{\phi} f\right\|_{\dot{F}_{p}^{\phi u, q}} \leq & \|\left(\int_{0}^{\infty}\left(\frac{1}{\phi(u) \psi(u)} \int_{1}^{\infty}\left|Q_{u} I_{\phi} Q_{u v}\left(\hat{Q}_{u v} f\right)\right| \frac{d v}{v}\right)^{q} \frac{d u}{u}\right)^{1 / q} \\
& +\left(\int_{0}^{\infty}\left(\frac{1}{\phi(u) \psi(u)} \int_{0}^{1}\left|Q_{u} I_{\phi} Q_{u v}\left(\hat{Q}_{u v} f\right)\right| \frac{d v}{v}\right)^{q} \frac{d u}{u}\right)^{1 / q} \|_{p} \\
= & \left\|S_{1}+S_{2}\right\|_{p}
\end{aligned}
$$

Applying (44) and (48) we get

$$
S_{1}(x) \leq C\left(\int_{0}^{\infty}\left(\int_{1}^{\infty} \frac{v^{-\left(\epsilon-s_{\phi}-s_{1}\right)}}{\psi(u v)} M\left|\hat{Q}_{u v} f\right|(x) \frac{d v}{v}\right)^{q} \frac{d u}{u}\right)^{1 / q}
$$

On the other hand, by (44), (49) and the right side of (9)

$$
S_{2}(x) \leq C\left(\int_{0}^{\infty}\left(\int_{0}^{1} \frac{v^{\left(\epsilon-s_{\phi}+i_{\phi}-s_{2}\right)}}{\psi(u v)} M\left|\hat{Q}_{u v} f\right|(x) \frac{d v}{v}\right)^{q} \frac{d u}{u}\right)^{1 / q}
$$

By Minkowski's inequality and hypothesis $s_{\phi}+s_{1}<\epsilon$ and $s_{\phi}-i_{\phi}+s_{2}<\epsilon$,

$$
S_{1}(x)+S_{2}(x) \leq C\left(\int_{0}^{\infty}\left(\frac{M\left|\hat{Q}_{u} f\right|(x)}{\psi(u)}\right)^{q} \frac{d u}{u}\right)^{1 / q}
$$

for every $x \in X$. Since $1<p, q<\infty$, we are able to apply the FeffermanStein vector valued maximal inequality to get that

$$
\left\|S_{1}+S_{2}\right\|_{p} \leq C\left\|\left(\int_{0}^{\infty}\left(\frac{\left|\hat{Q}_{u} f\right|}{\psi(u)}\right)^{q} \frac{d u}{u}\right)^{1 / q}\right\|_{p} \leq C\|f\|_{\dot{F}_{p}^{u, q}}
$$

since the family of operators $\left\{\hat{Q}_{t}\right\}_{t>0}$ determines equivalent norms on Besov and Triebel-Lizorkin spaces.

On the other hand,

$$
\begin{aligned}
\left\|I_{\phi} f\right\|_{\dot{B}_{p}^{\phi \psi, q}} \leq & \left(\int_{0}^{\infty}\left(\frac{1}{\phi(u) \psi(u)} \int_{1}^{\infty}\left\|Q_{u} I_{\phi} Q_{u v}\left(\hat{Q}_{u v} f\right)\right\|_{p} \frac{d v}{v}\right)^{q} \frac{d u}{u}\right)^{1 / q} \\
& +\left(\int_{0}^{\infty}\left(\frac{1}{\phi(u) \psi(u)} \int_{0}^{1}\left\|Q_{u} I_{\phi} Q_{u v}\left(\hat{Q}_{u v} f\right)\right\|_{p} \frac{d v}{v}\right)^{q} \frac{d u}{u}\right)^{1 / q} \\
= & S_{1}+S_{2}
\end{aligned}
$$

By (46) and (48),

$$
S_{1} \leq C\left(\int_{0}^{\infty}\left(\int_{1}^{\infty} v^{-\left(\epsilon-s_{\phi}-s_{1}\right)} \frac{\left\|\left(\hat{Q}_{u v} f\right)\right\|_{p}}{\psi(u v)} \frac{d v}{v}\right)^{q} \frac{d u}{u}\right)^{1 / q}
$$

By (46), (49) and the right side of (9),

$$
S_{2} \leq C\left(\int_{0}^{\infty}\left(\int_{0}^{1} v^{\left(\epsilon-s_{\phi}+i_{\phi}-s_{2}\right)} \frac{\left\|\left(\hat{Q}_{u v} f\right)\right\|_{p}}{\psi(u v)} \frac{d v}{v}\right)^{q} \frac{d u}{u}\right)^{1 / q}
$$

From Minkowski's inequality and the hypothesis on $s_{1}$ and $s_{2}$,

$$
S_{1}+S_{2} \leq C\|f\|_{\dot{B}_{p}^{\psi, q} \cdot \diamond}
$$

As we mentioned at the beginning of this work, in the context of $R^{n}$ the integral and derivative operators $I_{\alpha}$ and $D_{\alpha}$ are inverse one of the other on Besov and Triebel-Lizorkin spaces. Even though this is not true in the context of spaces of homogeneous type, Gatto, Segovia and Vgi showed in
[6] that $T_{\alpha}=D_{\alpha} \circ I_{\alpha}$ is a Calderón-Zygmund operator and also that it is invertible in $L^{2}$-and the same result also applies to $\dot{F}_{p}^{\alpha, q}$ and $\dot{B}_{p}^{\alpha, q}$ - for small values of $\alpha$.

Focusing on the composition operator $T_{\phi}=D_{\phi} \circ I_{\phi}$, it was proved in [11]:

Theorem 9. If $\max \left(s_{1}, s_{2}\right)+s_{\phi}<\epsilon$ then $T_{\phi}=D_{\phi} \circ I_{\phi}$ is a CalderónZygmund operator bounded on $\dot{F}_{p}^{\psi, q}$ and $\dot{B}_{p}^{\psi, q}$, whose associated kernel is

$$
K(x, y)=\int K_{1 / \phi}(x, z)\left(K_{\phi}(z, y)-K_{\phi}(x, y)\right) d \mu(z)
$$

Since $I_{\phi}$ and $D_{\phi}$ are self-adjoint then the adjoint operator $T_{\phi}^{*}$ coincides with $S_{\phi}=I_{\phi} \circ D_{\phi}$. In this way and from the above theorem follows that $S_{\phi}$ also is a $C Z O$ whose kernel is $\tilde{K}(x, y)=K(y, x)$.

In a work to appear soon it is proved that for small values of $s_{\phi}$ the operators $T_{\phi}$ and $S_{\phi}$ are invertible on Besov and Triebel-Lizorkin spaces, providing the key for the invertibility of $I_{\phi}$ between spaces of local regularity $\psi$ and those of the same scale and local regularity $\psi \phi$ and, reciprocally, the invertibility of $D_{\phi}$ between spaces of regularity $\psi$ and those of regularity $\psi / \phi$, provided that $s_{\phi}$ is small enough. We use these facts to obtain an identification between the space $L^{p, \phi}=\left\{f \in L^{p}: D_{\phi} f \in L^{p}\right\}$ and the inhomogeneous Triebel-Lizorkin space $F_{p}^{\phi, 2}$, with $1<p<\infty$ and $s_{\phi}$ small. Moreover, for the values of $s_{\phi}$ such that $T_{\phi}$ is invertible, we also prove that its inverse $T_{\phi}^{-1}$ is a Calderón Zygmund operator. The proof is the same for $S_{\phi}$. It can therefore be said that $T_{\phi}$ and $S_{\phi}$ 'almost are' the identity operator.

## References

[1] R.R.Coifman, G.Weiss Analyse harmonique noncommutative sur certains espaces homogénes, Lectures Notes in Math., 242, Berlin, Springer-Verlag, 1971.
[2] G. David, J.L.Journé, S. Semmes, Opérateurs de Calderón-Zygmund, fonction para-accretive et interpolation, Rev. Mat. Iberoamericana 1, 1985, 1-56.
[3] D.Deng, Y.-S.Han Calderón type reproducing formula on spaces of homogeneous type, Science in China (Series A), 38, n.1, 17-28, 1995.
[4] B. Iaffei, Espacios Lipschitz generalizados y operadores invariantes por traslaciones, Thesis. Universidad Nacional del Litoral, Argentina, 1996.
[5] M. Frazier, B. Jawerth, G. Weiss, Littlewood-Paley Theory and the Study of Function Spaces, Conference Board of the Mathematical Sciences, CBMS, Regional Conference Series in Mathematics, Number 79, 1991.
[6] A.E. Gatto, C. Segovia, S.Vági, On fractional differentiation and integration on spaces of homogencous type, Rev. Mat. Iberoamericana, 12, n.2, 111-145, 1996.
[7] A.E. Gatto, S.Vági, On Sobolev Spaces of Fractional Order and $\epsilon-$ Families of Operators on Spaces of Homogeneous Type, Studia Math 133 (1), 1999, 19-27.
[8] Y.-S. Han, E.T.Sawyer, Littlewood-Paley theory on spaces of homogeneous type and the classical function spaces, Memoirs of the Amer. Math. Soc., 110, N.530, 1994.
[9] S.I.Hartzstein, B.E.Viviani, T1 theorems on generalized Besov and Triebel-Lizorkin spaces over spaces of homogeneous type, Revista de la Unión Matemática Argentina, 42, No. 1, 2000, 51-73.
[10] S.I.Hartzstein, B.E.Viviani, Integral and Derivative Operators of Functional Order on Generalized Besov and Triebel-Lizorkin Spaces on Spaces of Homogeneous Type, Comment. Math. Univ. Carolinae 43, 4 (2002), 723-754.
[11] S.I.Hartzstein, B.E.Viviani, On the Composition of the Integral and Derivative Operators of functional order, Comment.Math.Univ.Carolinae 44, 1 (2003), 99-120.
[12] H. Triebel, Theory of function spaces, Monographs in Math., 78, Birkhauser, Verlag, Basel, 1983.
[13] J. Peetre, Sur les espaces de Besov, C.R. Acad. Sci. Paris, Ser A-B 264 (1967), 281-283.
[14] J. Peetre, On spaces of Triebel-Lizorkin type, Ark. Mat. 13 (1975), 123-130.
[15] J. Peetre, New thoughts on Besov spaces, Duke Univ. Math. Series, Durham, NC, 1976.
[16] M.M. Rao, Z.D. Ren, Theory of Orlicz Spaces Marcel Dekker 1991.
[17] M.A. Krasnosel'skii, Ya.B. Ruticki, Convex Functions and Orlicz Spaces (translation), Groningen 1961. MR 23:A4016.
[18] R.A. Macías, C. Segovia, Lipschitz functions on spaces of homogeneous type, Adv. in Math. 33, 1979, 257-270.

## Author's address:

Silvia I. Harzstein<br>Departamento Matemática<br>Fac. de Ingeniería Química<br>Univ. Nac. del Litoral<br>Santiago del Estero 2829, 3000 Santa Fe, Argentina.<br>E-mail: silviah@math.unl.edu.ar

# Universality in Orlicz spaces 

Francisco L. Hernández


#### Abstract

We describe some properties concerning the symmetric and lattice structure of the Orlicz function spaces and its arrangement invariant subspaces over probabilistic spaces. We compare with the behavior for Orlicz sequence spaces and Orlicz function spaces over $\mathbb{R}^{+}$. We also study the existence of universal Orlicz funcion spaces with previously fixed Boyd indices.


## 1 Introduction

The aim of these notes is to survey several universal and structure properties of Orlicz spaces and rearrangement invariant (r.i.) Banach spaces. Our goal is to study isomorphic embeddings of $L^{p}$-spaces $(1<p<\infty)$ into separable Orlicz function spaces and r.i. function spaces.

The study of the isomorphic structure of separable rearrangement invariant function spaces has been developed in the Memoirs of Johnson, Maurey, Schechtman and Tzafriri [13] and Kalton [15] (see also Lindenstrauss and Tzafriri [20]). It is well-known that isomorphic embeddings of $L^{p}$-spaces into separable r.i. function spaces on $[0,1]$ are abundant in the case $1 \leq p<2$ (by using probabilistic techniques). On the other side, in the opposite case $p>2$ there is a strong shortage of separable r.i. function spaces on $[0,1]$ containing isomorphic copies of $L^{p}$-spaces. Thus it holds that the existence of an isomorphic embedding of $L^{p}$ (for $p>2$ ) into a separable r.i. function space $E[0,1]$ implies in fact the existence of a sublattice of the r.i. function space $E[0,1]$ which is lattice-isomorphic to $L^{p}$ (see [7]).

The existence (or non-existence) of universal and complementably universal spaces inside predetermined classes of Banach spaces has been studied intensely in many contexts. It is well-know the universality of the space $C[0,1]$ for the class of all separable Banach spaces (i.e. every separable Banach space is isomorphic to a subspace of $C[0,1]$ ), but it is not complementably universal. Let us recall also that there is no a reflexive Banach

[^4]space universal for the class of all separable reflexive Banach spaces and that there does not exist a separable super-reflexive Banach space universal for the class of all $\ell^{p}$-spaces for $1<p<\infty$ (see [1], [24]).

In the setting of Orlicz spaces, the existence of universal sequence spaces with prefixed estimates was proved by Lindenstrauss and Tzafriri in [L-T1] finding Orlicz sequence spaces $\ell^{F_{\alpha, \beta}}$, with arbitrary prefixed indices $1 \leq$ $\alpha<\beta<\infty$, in which every Orlicz sequence space $\ell^{G}$ with indices between $\alpha$ and $\beta$ is isomorphic to a (complemented) subspace of the space $\ell^{F_{\alpha, \beta}}$. Universal Orlicz function spaces on the unbounded interval $(0, \infty)$ were studied by C. Ruiz and the author in [11] showing that the spaces $L^{\alpha}+L^{\beta}$ are lattice-universal for the class of all Orlicz function spaces $L^{G}(0, \infty)$ with Boyd indices strictly between $\alpha$ and $\beta$, i.e. every space $L^{G}(0, \infty)$ is latticeisomorphic to a sublattice of the space $L^{\alpha}+L^{\beta}$.

Recently the existence of universal Orlicz function spaces $L^{F}[0,1]$ on the $[0,1]$-interval has been proved in [10], jointly with B. Rodriguez-Salinas. The construction requires some combinatorial facts as well as to consider the uncountable discrete case. We describe here these results explaining the connection of this topic with the study of discrete Orlicz spaces $\ell^{F}(I)$ with uncountable symmetric basis, more precisely with isomorphic embeddings of $\ell^{p}(\Gamma)$-spaces into Orlicz spaces $\ell^{F}(I)$ for uncountable sets $\Gamma \subset I$ and $1<p<\infty$.

## 2 Orlicz function spaces on [0,1]

The notation and terminology follow the classical monographs [L-T 77, 79]. Let us begin recalling that an Orlicz space $L^{F}[0,1]$ generated by a Young function $F$ is the space of all measurable functions on $[0,1]$ such that

$$
I_{F}(s f)=\int_{0}^{1} F(s|f|) d \lambda<\infty
$$

for some $s>0$, endowed with the Luxemburg norm $\|f\|=\inf \{s>0$ : $\left.I_{F}(f / s) \leq 1\right\}$. The space $L^{F}[0,1]$ is separable if and only if the function $F$ satisfies the growth $\Delta_{2}$-condition at $\infty$ (i.e. $\left.\limsup _{x \rightarrow \infty} F(2 x) / F(x)<\infty\right)$.

The study of the structure of separable Orlicz function spaces $L^{F}[0,1]$ on the $[0,1]$-interval was initiated by Lindestrauss and Tzafriri in [18] introducing the useful $E_{F, 1}^{\infty}, C_{F, 1}^{\infty}$ sets and using the Kadec-Pelczynski disjointification method. Let us recall that

$$
E_{F, 1}^{\infty}=\overline{\left\{\frac{F(t x)}{F(t)}: t>1\right\}}
$$

and $C_{F, 1}^{\infty}=\overline{\operatorname{conv}}\left(E_{F, 1}^{\infty}\right)$, which are compact sets in $C[0,1]$. The characterization of $\ell^{p}$-subspaces inside an space $L^{F}[0,1]$ is the following:

$$
L^{F}[0,1] \underset{\sim}{\supset} \ell^{p} \Longleftrightarrow p \in\left[\alpha_{F}^{\infty}, \beta_{F}^{\infty}\right] \cup\left(\beta_{F}^{\infty}, 2\right) \cup\{2\}
$$

where $\alpha_{F}^{\infty}$ and $\beta_{F}^{\infty}$ denote the lower and upper Matuszewska-Orlicz indices of the function $F$ at $\infty$ which estimate the grade of $p-$ convexity and $q$-concavity. In this result the case $p=2$ is clear, by using the Rademacher functions and Khintchine inequality. And the case of $p$ belonging to the interval $\left(\beta_{F}^{\infty}, 2\right)$ is obtained via $p-s t a b l e$ random variables.

The behavior of complemented $\ell^{p}$-subspaces in Orlicz function spaces is more involved. Let us denote by $P_{L^{F}}$ the set of $p^{\prime}$ 's such that $L^{F}[0,1]$ has a complemented $\ell^{p}$-subspace. The geometry of the sets $P_{L^{F}}$ can be varied. Thus the set $P_{L^{F}}$ can be reduced to be just the singleton $\{2\}$ or more general, it can be any closed subset of an interval $[\alpha, \beta]$ union $\{2\}$. It remain open to know whether the set $P_{L^{F}}$ is always a closed set.

We pass to discuss isomorphic embeddings of the function spaces $L^{p}[0,1]$ into separable Orlicz function spaces $L^{F}[0,1]$ and separable r.i. function spaces $E[0,1]$ i.e. when $L^{F}[0,1] \underset{\sim}{\supset} L^{p}$. There is no ambiguity in this notation since $L^{p}$-spaces on $[0,1]$ and on $[0, \infty)$ are lattice-isomorphic. The structure of r.i. function spaces on $[0,1]$ is quite more rigid than those in the sequence case.

Theorem 1. ([13]) Let $E[0,1]$ be an r.i. function space which does not contain isomorphic copies of $\ell_{n}^{\infty}$ uniformly and $F[0,1]$ be a separable r.i. function space $\left(\neq L_{2}\right)$ with non trivial indices. If $E[0,1] \underset{\sim}{\sim} F[0,1]$ then either $E[0,1] \supset F[0,1]$ or the Haar basis of $F[0,1]$ is equivalent to a disjoint sequence in $E[0,1]$.

In particular for Orlicz spaces $L^{F}[0,1]$, due to the impossibility of embedding isomorphically Orlicz function spaces into sequence spaces, we have that $L^{F}[0,1] \underset{\sim}{\supset} L^{G}[0,1]\left(\neq L^{2}\right) \quad \Longrightarrow \quad L^{F}[0,1] \supset L^{G}[0,1]$. It follows that reflexive Orlicz function spaces on $[0,1]$ have a unique representation as Orlicz function spaces: $L^{F}[0,1] \approx L^{G}[0,1] \Longrightarrow L^{F}[0,1]=L^{G}[0,1]$. On the other hand, an Orlicz function space $L^{F}[0,1]$ cannot contain complemented copies of other Orlicz function spaces: if $L^{F}[0,1]{\underset{\sim}{\sim}}_{c} L^{G}[0,1]\left(\neq L^{2}\right)$ then $L^{F}[0,1]=L^{G}[0,1]$.

A general result by Kalton (obtained also in [14] and [13] under stronger conditions) claims the following:

Theorem 2. ([15]) A separable r.i. function space $E[0,1]$ contains an isomorphic copy of $L^{1}$ if and only if $E[0,1]=L^{1}[0,1]$, up to an equivalent renorming.

This is obtained in two steps: first, a separable r.i. function space $E[0,1]$ containing an isomorphic copy of $L^{1}$ must also contain a lattice-isomorphic copy of $L^{1}$; from this it is deduced next that $E[0,1]=L^{1}[0,1]$ up to an equivalent renorming, i.e.

$$
E[0,1] \underset{\sim}{\supset} L^{1} \Longrightarrow E[0,1] \underset{\sim}{\supset} L_{\ell}^{1} \Longrightarrow E[0,1]=L^{1}[0,1]
$$

A useful criteria for the lattice-embedding of function spaces into r.i. function spaces $E(I)$ over an interval $I$ was given in [13]. Let us denote by $\sum_{F, 1}^{\infty}$ the set of Orlicz functions $G$ of the form

$$
G(x) \cong \int_{0}^{\infty} \frac{F(s x)}{F(s)} d \mu(s)
$$

for $x>1$, where $\mu$ is a probability measure on $(0, \infty)$ such that

$$
\int_{0}^{\infty} \frac{1}{F(s)} d \mu(s) \leq 1
$$

Theorem 3. ([13]) Given a separable Orlicz function space $L^{F}[0,1]$, if $G \in$ $\sum_{F, 1}^{\infty}$ then $L^{F}[0,1] \underset{\sim}{\sim} L^{G}[0,1]$. Furthermore, if $L^{F}[0,1]$ is $p$-convex for some $p>2$ and $E[0,1]\left(\neq L^{2}\right)$ is an r.i. function space which embeds isomorphically into $L^{F}[0,1]$, then $E[0,1]=L^{G}[0,1]$, up to an equivalent renorming, for some Orlicz function $G \in \sum_{F, 1}^{\infty}$.

## 3 Universal Orlicz function spaces on the [ $0, \infty$ )interval

The structure of separable Orlicz function spaces $L^{F}[0, \infty)$ on the interval $[0, \infty)$ has several peculiar and interesting properties which have been studied in [21] and in [8]. For example, the characterization of $\ell^{p}$-subspaces in terms of the Matuszewska-Orlicz indices of the function $F$ at 0 and at $\infty$ is now:

$$
L^{F}[0, \infty) \underset{\sim}{\supset} \ell^{p} \Longleftrightarrow p \in\left[\alpha_{F}, \beta_{F}\right] \cup\left[\alpha_{F}^{\infty}, \beta_{F}^{\infty}\right] \cup\left(\beta_{F}^{\infty}, 2\right) \cup\{2\} \cup\left(\beta_{F}^{\infty}, \alpha_{F}\right)
$$

Universal Orlicz function spaces $L^{F}[0, \infty)$ with prefixed indices were given in [11] showing that the classical spaces $L^{\alpha}+L^{\beta}$, regarded as the Orlicz function spaces $L^{x^{\alpha} \wedge x^{\beta}}[0, \infty)$, are universal in the following sense:

Theorem 4. Given $1 \leq \alpha<\beta<\infty$, the space $L^{\alpha}+L^{\beta}$ is lattice universal for the class of all Orlicz function spaces $L^{G}[0, \infty)$ with indices strictly between $\alpha$ and $\beta$, i.e.

$$
L^{\alpha}+L^{\beta} \underset{\sim}{\supset} L^{G}[0, \infty)
$$

In particular $L^{\alpha}+L^{\beta}$ contain an isomorphic copy of $L^{p}$ for every $\alpha<p<$ $\beta$. The proof uses some "interpolation" arguments connecting the behavior of a function near 0 and near $\infty$ in order to represent every Orlicz function $G$ in an integral form with respect to the function $x^{\alpha} \wedge x^{\beta}$ (hence this method does not work in the $[0,1]$-case)

The embedding behavior into $L^{\alpha}+L^{\beta}$ is varied in the extreme cases of Orlicz spaces $L^{G}[0, \infty)$ with one of their indices equal to $\alpha$ or $\beta$. For example if $\alpha<2<\beta$ or if $2<\alpha<\beta$, the space $L^{\beta}$ is not isomorphic to any subspace of $L^{\alpha}+L^{\beta}([6],[7])$

Let us recall also here the first result of this nature: the existence of universal Orlicz sequence spaces $\ell^{F}$ with prefixed indices, proved by Lindenstrauss and Tzafriri.

Theorem 5. ([17]) There exist Orlicz sequence spaces $\ell^{F}$ with prefixed indices $1 \leq \alpha<\beta<\infty$, such that every Orlicz sequence space $\ell^{G}$ with indices between $\alpha$ and $\beta$ is isomorphic to a complemented subspace of the space $\ell^{F}$

There is uniqueness of these universal Orlicz sequence spaces $\ell^{F}$ with prefixed indices. This follows from the complementation fact by using the Pelczynski decomposition method. These universal spaces $\ell^{F}$ provide examples of Banach spaces with a symmetric basis which are isomorphic to their dual spaces (different from $\ell^{2}$ ).

## 4 Universal Orlicz function spaces $L^{F}[0,1]$

The study of isomorphic embeddings of $L^{p}$-spaces into separable r.i. function spaces on $[0,1]$ leads to distinguish essentially two different cases: the 2 -concave case and the opposite.

In the 2-concave case there is a big amount of separable r.i.function spaces containing isomorphically scales of $L^{p}$-spaces for $p \leq 2$. This is a well-known fact and requires some probabilistic tools (Poisson process, or $p$-stable variables and ultrapowers). Thus ([13] Section 8, [20] p.212):

Let $E[0,1]$ be an r.i.function space. If the function $x^{-1 / p} \in E[0,1]$ for some $1<p<2$, then $E[0,1] \underset{\sim}{\supset} L^{p} \quad$ (isometrically)

In particular for $1 \leq q \leq p \leq 2$, we have that the spaces $L^{q}[0,1] \underset{\sim}{\sim}$ $L^{p}$ (isometrically), a classical result of Bretagnolle, Dacunha-Castelle and Krivine [3]. Other classical function spaces with symmetric structure can be isomorphically embedded into the spaces $L^{q}[0,1]$ for $1 \leq q<2$. Thus for Orlicz function spaces we have: $L^{1}[0,1] \underset{\sim}{\supset} L^{F}[0,1]$ holds if and only if the Orlicz space $L^{F}[0,1]$ is 2-concave (Bretagnolle and Dacunha-Castelle [2], Schütt [23]). See also the recent survey by Dilworth [5].

The non 2-concave case on $[0,1]$. In contrast with the above case, there is a strong shortage of separable r.i. function spaces on $[0,1]$ containing isomorphically scales of $L^{p}$-spaces, for $p>2$. Recall for example that for the spaces $L^{q}[0,1]$ (with $q>2$ ), the Banach lattices which can be isomorphically embedded into $L^{q}[0,1]$ are essentially $L^{q}(\mu)$-spaces for some suitable measures $\mu(c f$. [20] p.202). In particular, the r.i. function spaces on $[0,1]$ which are isomorphic to a subspace of $L^{q}[0,1]$, for $q>2$, are just $L^{q}[0,1]$ or $L^{2}[0,1]$ ([13] p.41).

Given any $2<p<\infty$, no example was known of a separable r.i. function space $E[0,1]\left(\neq L^{p}[0,1]\right)$ such that $E[0,1] \underset{\sim}{\supset} L^{p}$. The impossibility of finding such examples inside the class of Lorentz function spaces was showed by Carothers ([4]). This shortage of r.i. function spaces with this property is related to the following fact proved by Kalton and the author ([7] p.827):

Theorem 6. Let $E[0,1]$ be a r.i.function space with some concavity and $p$ convex for some $p>2$. If a r.i. function space $F[0,1]\left(\neq L^{2}\right)$ is isomorphic to a subspace of $E[0,1]$ then $F[0,1]$ is lattice-isomorphic to a sublattice of $E[0,1]$.

In particular, the existence of an isomorphic embedding of an $L^{p}$-space for $p>2$ into a separable r.i. function space $E[0,1]$ implies that there exists also a lattice - isomorphic embedding of $L^{p}$ into $E[0,1]$ (i.e. for $2<p<$ $\left.\infty, E[0,1] \underset{\sim}{\supset} L^{p} \Longrightarrow E[0,1] \underset{\sim}{\sim}{\underset{\ell}{e}}^{p}\right)$

The existence of separable Orlicz function spaces $L^{F}[0,1]$ containing lattice-isomorphically scales of $L^{p}$ spaces for different values of $p$ has been proved in [H-R 98]. A more general result on universality given in [10] is the following:

Theorem 7. Given $1<\alpha<\beta<\infty$, there exists an Orlicz function space $L^{F_{\alpha, \beta}}[0,1]$, with indices $\alpha_{F_{\alpha, \beta}}^{\infty}=\alpha$ and $\beta_{F_{\alpha, \beta}}^{\infty}=\beta$, such that every $\alpha$-convex $\beta$-concave Orlicz function space $L^{G}[0,1]$ is lattice-isomorphic to a sublattice of $L^{F_{\alpha, \beta}}[0,1]$.

Thus the spaces $L^{F_{\alpha, \beta}}[0,1]$ verify that $L^{F_{\alpha, \beta}}[0,1] \underset{\sim_{\ell}}{\supset} L^{G}[0,1]$, in particular for every $\alpha \leq p \leq \beta$ we have $L^{F_{\alpha, \beta}}[0,1] \underset{\sim}{\sim_{\ell}} L^{p}$.

Let us denote by $P_{F}$ the set of scalars $p$ such that $L^{p}$ embeds lattice isomorphically into $L^{F}[0,1]$, i.e. $P_{F}:=\left\{p>1: L^{F}[0,1] \underset{\sim}{\supset} L^{p}\right\}$. We consider a third parameter (different from the Matuszewska-Orlicz indices): the "inclusion" index $\gamma_{F}^{\infty}$ associated to $F$. Define

$$
\gamma_{F}^{\infty}:=\limsup _{x \rightarrow \infty} \frac{\log F(x)}{\log x}
$$

It turns out that $\gamma_{F}^{\infty}=\inf \left\{p>1: L^{p}[0,1] \subset L^{F}[0,1]\right\}$
It is easy to check that $\alpha_{F}^{\infty} \leq \gamma_{F}^{\infty} \leq \beta_{F}^{\infty}$. And these inequalities can be strict. Clearly if $p$ is a scalar such that $L^{F}[0,1] \underset{\sim_{\ell}}{\supset} L^{p}$, then $\gamma_{F}^{\infty} \leq p \leq \beta_{F}^{\infty}$. Hence $P_{F} \subset\left[\gamma_{F}^{\infty}, \beta_{F}^{\infty}\right]$

Many natural Orlicz functions $F$ satisfy that the set $P_{F}$ is just the empty set (for example submultiplicative functions at $\infty$ ). On the other hand notice that the universal Orlicz function spaces $L^{F_{\alpha, \beta}}[0,1]$, given in Theorem 4.1, have inclusion index $\gamma_{F_{\alpha, \beta}}^{\infty}=\alpha_{F_{\alpha, \beta}}^{\infty}$ and that in this case the sets $P_{F}$ reach their biggest possible size filling all of the interval $[\alpha, \beta]$, i.e. $P_{F}=[\alpha, \beta]$.

The "size" of the sets $P_{F}$ can be arbitrarily small comparing with the size of the interval $[\alpha, \beta]$ (see [9]):
Theorem 8. Let $1<\alpha<\gamma \leq \beta<\infty$. There exists an Orlicz function space $L^{F}[0,1]$ with indices $\alpha_{F}^{\infty}=\alpha, \gamma_{F}^{\infty}=\gamma$ and $\beta_{F}^{\infty}=\beta$ such that $L^{p}$ is lattice-isomorphic to a sublattice of $L^{F}[0,1]$ for every $p \in\left[\gamma_{F}^{\infty}, \beta_{F}^{\infty}\right]$.

Thus for these spaces $L^{F}[0,1]$ we have $P_{F}=[\gamma, \beta] \subset[\alpha, \beta]$. Let us also mention that the sets $P_{F}$ are not always closed: Given $1<\alpha \leq \gamma<$ $\beta<\infty$. There exists a $\beta$-concave Orlicz function space $L^{F}[0,1]$ with indices $\alpha_{F}^{\infty}=\alpha, \gamma_{F}^{\infty}=\gamma$ and $\beta_{F}^{\infty}=\beta$ such that $L^{F}[0,1] \underset{\sim_{\ell}}{\supset_{\ell}} L^{p}$ if and only if $p \in\left[\gamma_{F}^{\infty}, \beta_{F}^{\infty}\right)$

It is a open question whether or not the sets $P_{F}$ are always convex.
The method used in the proof of above Theorems involves some combinatorial facts and properties of Banach spaces with uncountable symmetric basis. We pass to discribe this.

## 5 Uncountable symmetric basis and universality

The structure of Banach spaces with an uncountable symmetric basis has a behavior quite different to the countable case. Recall that a family of
vectors $\left(e_{i}\right)_{i \in I}$ in a Banach space $E$ is a symmetric basis if it is an unconditional basis and for every pair $\left(i_{k}\right)$ and $\left(i_{j}\right)$ of sequences of different elements indices in $I$ we have that $\left(e_{i_{k}}\right)$ and $\left(e_{i_{j}}\right)$ are equivalent basic sequences.

Using renorming arguments, Troyanski [25] proved that: if a Banach space $E$ with an uncountable symmetric basis $\left(e_{i}\right)_{i \in I}$ contains an isomorphic copy of $\ell^{1}(\Gamma)$ for some uncountable $\Gamma \subset I$, then $E=\ell^{1}(I)$. A similar result holds with the space $c_{0}(\Gamma)$ : if a Banach space $E$ with an uncountable symmetric basis verifies $E \underset{\sim}{\supset} c_{0}(\Gamma)$ for some uncountable $\Gamma \subset I$ then $E=$ $c_{0}(I)$. This was also proved by Troyanski in [25].

A natural question is the possible extensions of Troyanski's result on $\ell^{1}(\Gamma)$-embeddings to the case $1<p<\infty$, i.e. whether there exist Banach spaces $\left(\neq \ell^{p}(I)\right)$ with an uncountable symmetric basis containing an isomorphic copy of $\ell^{p}(\Gamma)$ for uncountable $\Gamma$. The answer is yes, and the first examples were certain non-reflexive Orlicz spaces with symmetric basis given by Troyanski and the author in [12].

Fixed a discrete Orlicz space $\ell^{F}(I)$, we consider the set $\sum_{F, 1}$ of all Orlicz functions

$$
G(x)=\int_{0}^{1} \frac{F(s x)}{F(s)} d \mu(s) \quad(\text { for } 0<x<1)
$$

where $\mu$ is a probability measure on $(0,1]$. The following criteria is useful ([22], [12]):

Theorem 9. Given an Orlicz space $\ell^{F}(I)$ with the function $F$ satisfying the $\Delta_{2}$-condition at 0 . Then $\ell^{F}(I)$ contains an isomorphic copy of $\ell^{G}(\Gamma)$ for $\Gamma \subset$ Iuncountable sets if and only if $G \in \sum_{F, 1}$.

We indicate now the method used in the proofs of the existence of universal spaces and Orlicz function spaces $L^{G}[0,1]$ containing a lattice-isomorphic copy of $L^{p}$ with prefixed indices $\alpha_{G}^{\infty}=\alpha$ and $\beta_{G}^{\infty}=\beta$ and $\alpha<p<\beta$.

Let us first point out that this result can be quite easily deduced after solving a related problem for the uncountable discrete case, i.e. the existence of discrete Orlicz spaces $\ell^{F}(I)$ containing an isomorphic copy of $\ell^{p}(\Gamma)$ for uncountable $\Gamma \subset I$. Indeed, by transfer arguments, we consider some $r>\beta$ and define then an Orlicz function $G$ near $\infty$ by

$$
G(x):=x^{r} F(1 / x)
$$

where $F$ is a certain Orlicz function defined near 0 such that $\ell^{F}(I) \underset{\sim}{\supset} \ell^{p}(\Gamma)$ for uncountable $\Gamma \subset I$. Using Theorem 2.3, the criteria for lattice embedding $L^{p}$ spaces into Orlicz function spaces $L^{G}[0,1]$ given in terms of the set $\sum_{G, 1}^{\infty}$, can be applied to get the $L^{p}$-embedding.

We now focus on the construction of discrete Orlicz spaces $\ell^{F}(I)$ such that $\ell^{F}(I) \supset \ell^{p}(\Gamma)$ for uncountable $\Gamma \subset I$ and prefixed indices. A crucial point in doing this is the existence of series of positive terms with the following "shift uniformly bounded" property:

Lemma 10. There exist sequences $\left(\alpha_{n}\right)_{n=0}^{\infty}$ and $\left(\varepsilon_{n}\right)_{n=0}^{\infty}$ of positive numbers with $\sum_{n=0}^{\infty} \alpha_{n}=\infty$ and constants $A>0$ and $B>0$ such that

$$
A \leq \sum_{n=0}^{\infty} \alpha_{n} \varepsilon_{n+k} \leq B
$$

for every natural $k=0,1,2, \ldots$.
The existence of these sequences $\left(\alpha_{n}\right)$ and $\left(\varepsilon_{n}\right)$ is proved using the following combinatorial fact: given an arbitrary sequence $\left(h_{i}\right)_{i=0}^{\infty}$ of positive integers with $h_{0}=1$, there exists a set of couples of positive integers $\left\{\left(m_{j}, k_{i}\right)\right\}$ with $m_{i}>k_{i}$ such that for each $n$ :
(i) there exist precisely $h_{n}$ couples ( $m_{j}, k_{i}$ ) such that $m_{j}-k_{i}=n$.
(ii) there exist at most $(n+2)^{2}$ couples $\left(m_{j}, k_{i}\right)$ such that $k_{i}-m_{j}=n$.

Let us indicate now other steps of the proof of $\ell^{F}(I) \underset{\sim}{\supset} \ell^{p}(\Gamma)$. We can assume $\alpha=1<p<\beta=p+\varepsilon$ The other cases can be deduced from this using $q$-concavification and $r$-convexification reductions and properties of the $\sum_{F, 1}$ sets. We consider the function

$$
f=\sum_{n=0}^{\infty} \varepsilon_{n} \chi_{\left(2^{-(n+1)}, 2^{-n}\right]}
$$

where the sequence $\left(\varepsilon_{n}\right)$ is given by the Lemma, and define the Orlicz function $F$ at 0 by

$$
F(x)=\int_{0}^{x}(x-t) t^{p-2} f(t) d t
$$

It turns out that the function $f$ satisfies the following key property:

$$
A \leq \sum_{n=0}^{\infty} \alpha_{n} f\left(\frac{x}{2^{n}}\right) \leq B
$$

for $0<x \leq 1$. From this, the following inequalities are obtained by integration

$$
A \frac{x^{p}}{p(p-1)} \leq \sum_{n=0}^{\infty} \alpha_{n} 2^{p n} F\left(\frac{x}{2^{n}}\right) \leq B \frac{x^{p}}{p(p-1)}
$$

for $0 \leq x \leq 1$.
Thus, if we consider the discrete measure $\mu$ on $(0,1]$ defined by $\mu\left(2^{-n}\right):=$ $\alpha_{n} 2^{p n} F\left(2^{-n}\right)$, we deduce that the function $G$, defined by

$$
G(x)=\int_{0}^{1} \frac{F(x t)}{F(t)} d \mu \quad(0 \leq x \leq 1)
$$

satisfies that $x^{p} \cong G \in \sum_{F, 1}$. Hence, using Theorem 5.2, we deduce that the Orlicz space $\ell^{F}(I)$ verifies $\ell^{F}(I) \underset{\sim}{\supset} \ell^{p}(\Gamma)$ for uncountable $\Gamma \subset I$.

Finally, using properties of the sequence $\left(\varepsilon_{n}\right)$ constructed in the Lemma, the associated indices of the Orlicz function $F$ can be computed to obtain $\alpha_{F}=1$ and $\beta_{F}=p+\varepsilon$.

A more general result is the existence of universal Orlicz discrete spaces $\ell^{F}(I)$ with prefixed indices given also in [10]:

Theorem 11. Let $1<\alpha<\beta<\infty$. There exists an Orlicz space $\ell^{F_{\alpha, \beta}}(I)$, with indices $\alpha_{F_{\alpha, \beta}}=\alpha$ and $\beta_{F_{\alpha, \beta}}=\beta$, such that $\ell^{F_{\alpha, \beta}}(I)$ contains an isomorphic copy of any $\alpha$-convex $\beta$-concave Orlicz space $\ell^{G}(\Gamma)$ with $\Gamma \subset I$ arbitrary sets.

This theorem provides in particular new examples of universal Orlicz sequence spaces $\ell^{F_{\alpha, \beta}}$ with prefixed indices $\alpha$ and $\beta$ different from the given by Lindenstrauss and Tzafriri in [17].

## Acknowledgement

I would like to thank the organizers of the Seminario de Analisis Funcional de la Universidad de Sevilla for the invitation and kind hospitality.

## References

[1] J.Bourgain, On separable Banach spaces universal for all separable reflexive spaces, Proc. Amer. Math. Soc. 79 (1980), 241-246.
[2] J.J. Bretagnolle, D. Dacunha-Castelle, Applications de l'etude de certain formes lineaires aleatoires an plongement d'espaces de Banach dans des espaces L ${ }^{p}$. Ann. Scient. Ec. Norm. Sup. 2 (1969), 437-480.
[3] J.J. Bretagnolle, D. Dacunha-Castelle, J.L. Krivine, Lois stables et espaces $L_{p}$. Ann. Ins. H. Poicare, Probab. Statist. 2 (1966), 231-259.
[4] N.L. Carothers, Rearrangement invariant subspaces of Lorentz function spaces II. Rocky Mountain J. Math. 17 (1987), 607-616.
[5] S.J. Dilworth, Special Banach lattices and their applicatios. Handbook of the Geometry of Banach spaces, Vol 1 (2001), 497-532. Edited by W.B. Johnson and J. Lindenstrauss. Elseiver
[6] A. Garcia del Amo, F.L. Hernández, On embedding of function spaces into $L^{p}+L^{q}$. Contemporary Math. 144 (1993), 107-113.
[7] F.L. Hernández, N.J. Kalton, Subspaces of rearrangement invariant spaces. Canadian J. of Math. 48 (1996), 794-833.
[8] F.L. Hernández, B. Rodríguez-Salinas, Remarks on the Orlicz function spaces $L^{\varphi}[0, \infty)$. Math. Nachr. 156 (1992), 225-232.
[9] F.L. Hernández, B. Rodríguez-Salinas, Lattice-embedding scales of $L^{p_{-}}$ spaces into Orlicz spaces. Israel J. Math. 104 (1998), 191-220.
[10] F.L. Hernández, B. Rodríguez-Salinas, Lattice-universal Orlicz spaces on probability spaces. Israel J. Math. 133 (2003), 9-28.
[11] F.L. Hernández, C. Ruiz, Universal classes of Orlicz function spaces. Pacific J. Math. 155 (1992), 87-98.
[12] F.L. Hernández, S. Troyanski, On the representation of uncountable symmetric basic sets and its applications. Studia Math. 107 (1993), 287-304.
[13] W.B. Johnson, B. Maurey, G. Schechtmann, L. Tzafriri, Symmetric structures in Banach spaces. Memoires Amer. Math. Soc. 217 (1979), 1-298.
[14] N.J. Kalton, Embedding $L^{1}$ in a Banach lattice. Israel J. Math. 32 (1979), 209-220.
[15] N.J. Kalton, Lattice structures in Banach spaces. Memoires Amer. Math. Soc. 493 (1993), 1-92.
[16] J. Lindenstrauss, L. Tzafriri, On Orlicz sequence spaces . Israel J. Math. 10 (1971), 379-390.
[17] J. Lindenstrauss, L. Tzafriri, On Orlicz sequence spaces II. Israel J. Math. 11 (1972), 355-379.
[18] J. Lindenstrauss, L. Tzafriri, On Orlicz sequence spaces III. Israel J. Math. 14 (1973), 368-389.
[19] J. Lindenstrauss, L. Tzafriri, Classical Banach spaces I. Sequence spaces. Springer-Verlag, Berlin, 1977.
[20] J. Lindenstrauss, L. Tzafriri, Classical Banach spaces II. Function spaces. Springer-Verlag, Berlin 1979.
[21] N.J. Nielsen, On the Orlicz function spaces $L^{M}[0, \infty)$. Israel J. Math. 20 (1975), 237-259.
[22] B. Rodríguez-Salinas, Subspaces with symmetric basis in the Orlicz spaces $\ell^{F}(I)$. Comment. Math. Prac. Math. 36 (1996), 201-222.
[23] C. Schütt, On the embedding of 2 -concave Orlicz spaces into $L^{1}$. Studia Math. 113 (1995), 73-80.
[24] W. Szlenk, The non-existence of a separable reflexive Banach space universal for all separable reflexive Banach spaces. Studia Math. 30 (1968), 53-61.
[25] S. Troyanski, On non-separable Banach spaces with symmetric basis. Studia Math. 53 (1975), 253-263.

## Author's address:

Francisco J. Hernández

Departamento de Analisis Matematico
Facultad de Matematicas, Universidad Complutense de Madrid 28040 Madrid (Spain)
E-mail: pacoh@mat.ucm.es

# Projection constants, isometric imbeddings and spherical designs 

Hermann König


#### Abstract

We derive upper estimates for projection constants of finite-dimensional normed spaces and show that the bounds are attained for spaces with unit balls generated by certain spherical designs. The extremal spaces, however, are non-unique, in general. We also discuss applications of spherical design techniques to the problem of isometric imbeddings of enclidean spaces into $l_{p}$-spaces if $p \in 2 \mathbb{N}$.


## 1 Introduction and main results on projection constants

Projections of minimal norm onto subspaces of a given Banach space are useful in approximation theory and in functional analysis when one considers continuous linear extensions of operators given on subspaces only. Given a (closed) subspace $X$ of a Banach space $Z$, the relative projection constant of $X$ in $Z$ is
$\lambda(X, Z):=\inf \left\{\|P\| \mid P: Z \rightarrow X \subseteq Z\right.$ a linear projection onto $\left.X, P^{2}=P\right\}$, the (absolute) projection constant of $X$ is

$$
\lambda(X):=\sup \{\lambda(X, Z) \mid Z \text { is a Banach space containing } \mathrm{X} \text { as a subspace }\} .
$$

The scalar field will be always $\mathbb{K} \in\{\mathbb{R}, \mathbb{C}\}$. Any separable Banach space can be imbedded isometrically into $l_{\infty}$; for any such imbedding one has $\lambda(X)=\lambda\left(X, l_{\infty}\right): l_{\infty}$ is the "worst" superspace with maximal relative projection constant. For finite-dimensional spaces $X, \operatorname{dim} X=: n$, one has by Kadets-Snobar [8], $\lambda(X) \leq \sqrt{n}$. In fact, a stronger estimate is known. Let $\langle\cdot, \cdot\rangle$ denote the standard scalar product on $\mathbb{K}^{n}$ and $\|\cdot\|_{2}=\sqrt{<\cdot, \cdot\rangle}$ be the euclidean norm. Vectors $x_{1}, \ldots, x_{N} \in \mathbb{K}^{n}$ are called equiangular provided that $\left\|x_{i}\right\|_{2}=1,\left|<x_{i}, x_{j}>\right|=\alpha<1$ for all $1 \leq i \neq j \leq N$ holds. Let $N(n):=n(n+1) / 2$ if $\mathbb{K}=\mathbb{R}$ and $N(n):=n^{2}$ if $\mathbb{K}=\mathbb{C}$. It
is easy to see that no more than $N(n)$ equiangular vectors can exist in $\mathbb{K}^{n}$ since the orthogonal projections onto the lines $\mathbb{K} x_{i}$ turn out to be linearly independent as operators over $\mathbb{R}$, cf. Lemmens-Seidel [14]. Let us define functions

$$
\begin{aligned}
g_{\mathbb{R}}: \mathbb{N} \rightarrow \mathbb{R}, g_{\mathbb{R}}(n) & =(2+(n-1) \sqrt{n+2}) /(n+1) \\
g_{\mathbb{C}}: \mathbb{N} \rightarrow \mathbb{R}, g_{\mathbb{C}}(n) & =(1+(n-1) \sqrt{n+1}) / n,
\end{aligned}
$$

and denote $g=g_{\mathbb{K}}$ for $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$. It is easy to see that $g(n)<\sqrt{n}$ for all $n \in \mathbb{N}$ and $g_{\mathbb{R}}(n)=\sqrt{n}-\frac{1}{\sqrt{n}}+0\left(\frac{1}{n}\right)$ as well as $g_{\mathbb{C}}(n)=\sqrt{n}-\frac{1}{2 \sqrt{n}}+0\left(\frac{1}{n}\right)$ for large $n \in \mathbb{N}$. The following strengthening of the Kadets-Snobar bound is valid, cf. [10], [11]:

Theorem 1. (a) The projection constant of any n-dimensional normed space $X_{n}$ satisfies

$$
\begin{equation*}
\lambda\left(X_{n}\right) \leq g(n)<\sqrt{n} \tag{1}
\end{equation*}
$$

(b) Given $\mathbb{K}$ and $n \in \mathbb{N}$, there exist $n$-dimensional spaces $X_{n}$ for which the bound (1) is attained if and only if there exist $N(n)$ equiangular vectors $x_{1}, \ldots, x_{N(n)} \in \mathbb{K}^{n}$. In this case, extremal spaces $X_{n}$ can be realized as subspaces of $l_{\infty}^{N(n)}$ or $l_{1}^{N(n)}$ by defining the norm of $X_{n}=\left(\mathbb{K}^{n},\| \|\right)$ as

$$
\|x\|_{\infty}=\sup _{1 \leq j \leq N(n)}\left|<x, x_{j}>\right|
$$

or

$$
\|x\|_{1}=\sum_{j=1}^{N(n)}\left|<x, x_{j}>\right| .
$$

Unless $\mathbb{K}=\mathbb{R}$ and $n=2$, both spaces are non-isometric and both have maximal projection constant. The orthogonal projection is the minimal projection.

Let us remark that $N(n)$ equiangular vectors are known to exist for $\mathbb{K}=\mathbb{R}: n=2,3,7,23$.
$\mathbb{K}=\mathbb{C}: n=2,3,4,8$, cf. [14], [18].
In the real case, for $n>3$, a necessary condition for the existence of $N(n)$ equiangular vectors is $n=(2 m+1)^{2}-2, m \in \mathbb{N}$, cf. [14]. For $m=$ $3, n=47$, they do not exist, however (E. Bannai, personal communication). It is unlikely that $N(n)$ equiangular vectors exist for other $n$ than $2,3,7,23$ if $\mathbb{K}=\mathbb{R}$. On the other hand, it is conjectured that $N(n)=n^{2}$ equiangular vectors always exist in $\mathbb{C}^{n}$. This was checked numerically up to $n \leq 45$ within error tolerance of $10^{-8}$.

For $\mathbb{K}=\mathbb{R}, n=2, N(n)=3$ the extremal 2-dimensional subspace of $l_{\infty}^{3}$ or $l_{1}^{3}$ is the one with sum of coordinates being zero, in both cases the unit ball is a regular hexagon. For $\mathbb{K}=\mathbb{R}, n=3, N(n)=6$, the equiangular vectors in $\mathbb{R}^{3}$ are the six diagonals of the regular icosahedron.

The norm $\|\cdot\|_{\infty}$ thus has the regular dodecahedron as its unit ball, it is dual to the icosahedron. The norm $\|\cdot\|_{1}$ yields a different unit ball which may be described as $D \cap \phi I$ where $D$ is a regular dodecahedron and $I$ the icosahedron having the midpoints of the faces of $D$ as its vertices and $\phi=(1+\sqrt{5}) / 2$ is the "golden ratio", $D \cap \phi I$ has 12 regular pentagons and 20 regular triangles as its faces.

Even though the existence of $N(n)=n^{2}$ equiangular vectors in $\mathbb{C}^{n}$ has not been proved, there exist $n^{2}-n+1$ such vectors if $n$ is an odd prime power. As a consequence, the bound in (1) is almost attained:

Proposition 2. Let $n=p^{m}+1$ be a prime power plus 1 and $N=n^{2}-n+1$. Then there exist complex $n$-dimensional subspaces $X_{n}$ of $\mathbb{C}^{N}$, which when considered as subspaces of $l_{\infty}^{N}$ or $l_{1}^{N}$, satisfy

$$
\begin{equation*}
\lambda\left(X_{n}\right) \geq g(n)-\frac{1}{2 \sqrt{n}} . \tag{2}
\end{equation*}
$$

The $N(n)$ vectors $x_{j}$ of Theorem 1 form a spherical 4-design on $S^{n-1} \subseteq$ $\mathbb{K}^{n}$; this means that

$$
\begin{equation*}
\int_{S^{n-1}} p(y) d \sigma(y)=\frac{1}{N} \sum_{j=1}^{N} p\left(x_{j}\right) \tag{3}
\end{equation*}
$$

holds for $N=N(n)$ and $p$ being an even polynomial of degree 4. If one adds the vectors $\left(-x_{j}\right)$ and replaces $N$ by $2 N$, one has equality (3) for all polynomials of degree 5 .

Since $\frac{g(n)}{\sqrt{n}} \rightarrow 1$, the spaces $X_{n}$ of Proposition 2 satisfy $\lim _{n \rightarrow \infty} \frac{\lambda\left(X_{n}\right)}{\sqrt{n}}=1$. An improvement of this estimate can be given for spaces with 1 -symmetric spaces: It was shown in [9] that there is $c<1$ such that for any $n$-dimensional space $X_{n}$ with an 1-symmetric basis one has

$$
\lambda\left(X_{n}\right) \leq c \sqrt{n} .
$$

For example, if $X_{n}=l_{p}^{n}, 1 \leq p \leq 2$, one knows that $\lim _{n \rightarrow \infty} \frac{\lambda\left(l_{p}^{n}\right)}{\sqrt{n}}=\sqrt{\frac{2}{\pi}}$ $(<1)$ if $\mathbb{K}=\mathbb{R}$ and $\frac{\sqrt{\pi}}{2}(<1)$ if $\mathbb{K}=\mathbb{C}$.

## 2 Trace-duality

As with many optimization questions, it is useful to consider a dual formulation of the problem.

Lemma 3. Let $n<N$ and $X_{n} \subseteq l_{\infty}^{N}$ be an n-dimensional subspace. Than there exists a map $u: l_{\infty}^{N} \rightarrow l_{\infty}^{N}$ with $u\left(X_{n}\right) \subseteq X_{n}$ such that

$$
\lambda\left(X_{n}\right)=\operatorname{tr}\left(u: X_{n} \rightarrow X_{n}\right) \text { and } \sum_{j=1}^{N}\left\|u e_{j}\right\|_{\infty}=1
$$

Here $e_{j}$ denote the standard unit vectors in $l_{\infty}^{N}$.
Proof. By compactness in finite dimensions, there exists a projection $P_{0}$ : $l_{\infty}^{N} \rightarrow X_{n} \subseteq l_{\infty}^{N}$ onto $X_{n}$ of minimal norm. Thus $\left\|P_{0}\right\|=\lambda\left(X_{n}\right)=: \lambda$. Consider the space $\mathcal{L}\left(l_{\infty}^{N}, l_{\infty}^{N}\right)$ of linear operators on $l_{\infty}^{N}$, equipped with the operator norm. The sets

$$
\begin{aligned}
A=\left\{S \in \mathcal{L}\left(l_{\infty}^{N}, l_{\infty}^{N}\right) \quad \mid\right. & \|S\|<\lambda\} \\
B=\left\{P \in \mathcal{L}\left(l_{\infty}^{N}, l_{\infty}^{N}\right) \quad \mid\right. & P=P_{0}+\sum_{i=1}^{m} x_{i}^{*}(\cdot) x_{i} \text { for some } m \in \mathbb{N} \\
& \left.x_{1}, \ldots, x_{m} \in X_{n}, x_{1}^{*}, \ldots, x_{m}^{*} \in X_{n}^{\perp} \subseteq\left(l_{\infty}^{N}\right)^{*}\right\}
\end{aligned}
$$

are convex and disjoint since $B$ consists of projections onto $X_{n}$ and $\|P\| \geq \lambda$ for any projection $P$. Since $A$ is open, by Hahn-Banach there exists a functional $\phi \in \mathcal{L}\left(l_{\infty}^{N}, l_{\infty}^{N}\right)^{*}$ of norm $\|\phi\|=1$ such that $\phi\left(P_{0}\right) \in \mathbb{R}$ and for all $S \in A$ and $P \in B$ we have

$$
\operatorname{Re} \phi(S)<\lambda \leq \operatorname{Re} \phi(P)
$$

In particular, $\phi\left(P_{0}\right)=\lambda$. By trace-duality, $\phi$ is represented by $u \in$ $\mathcal{L}\left(l_{\infty}^{N}, l_{\infty}^{N}\right)$ as $\phi=\operatorname{tr}\left(u\right.$.). Since the operator norm of $\omega \in \mathcal{L}\left(l_{\infty}^{N}, l_{\infty}^{N}\right)$ is just

$$
\begin{gathered}
\|\omega\|=\sup _{1 \leq j \leq N}\left\|\omega e_{j}\right\|_{1}, \text { the dual norm is given by } \\
\|\phi\|=\|\operatorname{tr}(u .)\|=\sum_{j=1}^{N}\left\|u e_{j}\right\|_{\infty}=1
\end{gathered}
$$

For any $x \in X_{n}, x^{*} \in X_{n}^{\perp}$ we know for $P=P_{0}+x^{*}()$.

$$
\begin{aligned}
\lambda & \leq \operatorname{Re} \phi(P)=\operatorname{Re}\left[\phi\left(P_{0}\right)+\phi\left(^{*}(.) x\right)\right]=\lambda+\operatorname{Retr}\left(u \circ\left(x^{*}(.) x\right)\right) \\
& =\lambda+\operatorname{Re} x^{*}(u(x))
\end{aligned}
$$

Hence $\operatorname{Re} x^{*}(u(x)) \geq 0$ for all $x^{*} \in X_{n}^{\perp}, x \in X_{n}$ which yields $x^{*}(u x)=0$, i.e.[?] $u\left(X_{n}\right) \subseteq X_{n}$. Further

$$
\lambda\left(X_{n}\right)=\lambda=\phi\left(P_{0}\right)=\operatorname{tr}\left(u P_{0}\right)=\operatorname{tr}\left(u: X_{n} \rightarrow X_{n}\right)
$$

and as seen before $\sum_{j=1}^{N}\left\|u e_{j}\right\|_{\infty}=1$.
Proposition 4. Let $n, N \in \mathbb{N}$ and $n<N$. Then

$$
\begin{align*}
& \sup \left\{\lambda\left(X_{n}, l_{\infty}^{N}\right) \mid X_{n} \text { is an } n \text {-dimensional subspace of } l_{\infty}^{N}\right\} \\
& =n \sup \left\{\sum_{j, k=1}^{N} \mu_{j} \mu_{k}\left|<x_{j}, x_{k}>\right|\right\} \tag{4}
\end{align*}
$$

where the second supremum is taken over all discrete probability measures [?] $\mu=\left(\mu_{j}\right)_{j=1}^{N} \in\left(\mathbb{R}_{+}\right)^{N},\|\mu\|_{1}=1$ and over all sets of vectors $x_{j} \in S^{n-1}$ such that

$$
I d_{n}=n \sum_{j=1}^{N} \mu_{j}<\cdot, x_{j}>x_{j} \text { on } \mathbb{K}^{n}
$$

Both supremes are, in fact, maxima. Given extremal elements $\left(x_{j}, \mu_{j}\right)$ for the right side of (4) - where we may assume that all $\mu_{j} \neq 0$ - an $n$ dimensional space $X_{n}$ with maximal projection constant may be defined by its norm $\|x\|=\sup _{1 \leq j \leq N}\left|<x, x_{j}>\right|$ as a subspace of $l_{\infty}^{N}$ or $\|x\|=$ $\sum_{j=1}^{N} \mu_{j}\left|<x, x_{j}>\right|$ as a subspace of $l_{1}^{N}(\mu)$.

Proof. " $\leq$ ": Assume that $X_{n}$ has maximal projection constant among $n$ dimensional subspaces of $l_{\infty}^{N}$, i.e. $\lambda\left(X_{n}, l_{\infty}^{N}\right)=\lambda\left(X_{n}\right)$ attains the left supremum. Choose $u \in \mathcal{L}\left(l_{\infty}^{N}, l_{\infty}^{N}\right)$ as in Lemma 3 with $u\left(X_{n}\right) \subseteq X_{n}$ and $\sum_{j=1}^{N}\left\|u e_{j}\right\|_{\infty}=$ $1, \lambda\left(X_{n}\right)=\operatorname{tr}\left(u: X_{n} \rightarrow X_{n}\right)$. Let $\mu_{j}:=\left\|u e_{j}\right\|_{\infty}$. Hence $\mu=\left(\mu_{j}\right)_{j=1}^{N} \in$ $\left(\mathbb{R}_{+}\right)^{N}$ is a discrete probability measure. Consider $X_{n} \subseteq l_{\infty}^{N}=\mathbb{K}^{N}$ as an
algebraic subspace of $l_{2}^{N}(\mu)=\mathbb{K}^{N}$ and let $f_{1}, \ldots, f_{n}$ be an orthonormal basis of $X_{n}$ under the norm of $l_{2}^{N}(\mu),\|\xi\|_{2}=\left(\sum_{j=1}^{N}\left|\xi_{j}\right|^{2} \mu_{j}\right)^{\frac{1}{2}}$. Then

$$
\lambda\left(X_{n}\right)=\operatorname{tr}\left(u: X_{n} \rightarrow X_{n}\right)=\sum_{i=1}^{n}<u f_{i}, f_{i}>_{l_{2}^{N}(\mu)}
$$

$$
=\sum_{i=1}^{n} \sum_{j=1}^{N} u f_{i}(j) \overline{f_{i}(j)} \mu_{j} \leq \sum_{j=1}^{N} \mu_{j}\left|\sum_{i=1}^{n} u f_{i}(j) \overline{f_{i}(j)}\right|
$$

$$
\leq \sum_{j=1}^{N} \mu_{j}\left\|\sum_{i=1}^{n} \overline{f_{i}(j)} u f_{i}\right\|_{\infty}=\sum_{j=1}^{N} \mu_{j}\left\|u\left(\sum_{i=1}^{n} \overline{f_{i}(j)} f_{i}\right)\right\|_{\infty}
$$

But for $\xi \in X_{n},\|u(\xi)\|_{\infty}=\left\|u\left(\sum_{k=1}^{N} \xi_{k} e_{k}\right)\right\|_{\infty} \leq \sum_{k=1}^{N}\left|\xi_{k}\right|\left\|u e_{k}\right\|_{\infty}=\sum_{k=1}^{N}\left|\xi_{k}\right| \mu_{k}$. Thus

$$
\lambda\left(X_{n}\right) \leq \sum_{j=1}^{N} \sum_{k=1}^{N} \mu_{j} \mu_{k}\left|\sum_{i=1}^{n} \overline{f_{i}(j)} f_{i}(k)\right|,
$$

taking $\xi=\sum_{i=1}^{n} \overline{f_{i}(j)} f_{i} \in X_{n}$. Let $x_{j}=\frac{1}{\sqrt{n}}\left(f_{i}(j)\right)_{i=1}^{n} \in \mathbb{K}^{n}$. Then

$$
\lambda\left(X_{n}\right) \leq n \sum_{j, k=1}^{N} \mu_{j} \mu_{k}\left|<x_{j}, x_{k}>\right|
$$

where $\sum_{j=1}^{N} \mu_{j}<\cdot, x_{j}>x_{j}$ is a multiple of the Identity on $\mathbb{K}^{n}$ since $f_{1}, \ldots, f_{n}$ was a $\mu$-orthogonal basis. Since

$$
\begin{aligned}
\operatorname{tr}\left(\sum_{j=1}^{N} \mu_{j}<\cdot, x_{j}>x_{j}\right) & =\sum_{j=1}^{N} \mu_{j}\left\langle x_{j}, x_{j}\right\rangle=\sum_{i=1}^{n} \sum_{j=1}^{N}\left|x_{j}(i)\right|^{2} \mu_{j} \\
& =\frac{1}{n} \sum_{i=1}^{n}\left\|f_{i}\right\|_{l_{2}^{N}(\mu)}^{2}=1
\end{aligned}
$$

the multiple is $\frac{1}{n}$, i.e.

$$
I d_{n}=n \sum_{j=1}^{N} \mu_{j}<\cdot, x_{j}>x_{j}
$$

and $\sum_{j=1}^{N} \mu_{j}\left\|x_{j}\right\|_{2}^{2}=1$. Hence

$$
\begin{equation*}
\lambda\left(X_{n}\right) \leq n \sup \left\{\sum_{j, k=1}^{N} \mu_{j} \mu_{k}\left|<x_{j}, x_{k}>| | I d_{n}=n \sum_{j=1}^{N} \mu_{j}<\cdot, x_{j}>x_{j}\right\}\right. \tag{5}
\end{equation*}
$$

It is shown in [10] by a rather lenghty argument using Lagrange multipliers that the right hand supremum in (5) is, in fact, attained for vectors $x_{j}$ which have constant $l_{2}$-norm $\mu$-a.e. We will not give this argument here. Assuming thus $\mu_{j} \neq 0$ for all $j=1, \cdots N$ (otherwise $X_{n}$ is a subspace of some $l_{\infty}^{M}$ with $\left.M<N\right)$, we get in the extremal case $\left\|x_{1}\right\|_{2}=\cdots=\left\|x_{N}\right\|$. From $\sum_{j=1}^{N} \mu_{j}\left\|x_{j}\right\|_{2}^{2}=1$ we infer that $\left\|x_{j}\right\|_{2}=1$, i.e. $x_{j} \in S^{n-1}$. This proves the inequality " $\leq$ ".
$" \geq$ ": We only give a sketch of the argument, for details see [10]. Assume that $\left(x_{j}, \mu_{j}\right)_{j=1}^{N}$ attains the right hand supremum in (5), call this

$$
\Lambda=n \sum_{j, k=1}^{N} \mu_{j} \mu_{k}\left|<x_{j}, x_{k}>\right|
$$

The Lagrange multiplier approach yields that the map

$$
u:=\left(\operatorname{sign}\left(<x_{j}, x_{k}>\right) \mu_{k}\right)_{j, k=1}^{N}: l_{\infty}^{N} \rightarrow l_{\infty}^{N}
$$

is a multiple of the identity on $X_{n}:=\operatorname{Span}\left[f_{1}, \ldots, f_{n}\right]$ where $f_{i}(k):=\sqrt{n} x_{k i}$, $x_{k}=\left(x_{k i}\right)_{i=1}^{n} \in \mathbb{K}^{n}$. Hence for all $j=1, \ldots, N ; i=1 \ldots n$

$$
\sum_{k=1}^{N} \operatorname{sign}\left(<x_{j}, x_{k}>\right) \mu_{k} x_{k i}=\gamma x_{j i}
$$

Multiplying with $\bar{x}_{j i}$ and summing over $i$ yields together with $x_{j} \in S^{n-1}$ that

$$
\sum_{k=1}^{N}\left|<x_{j}, x_{k}>\right| \mu_{k}=\gamma
$$

independent of $j \in\{1, \ldots, N\}$. Thus the multiple $\gamma$ is equal to $\gamma=\Lambda / n$. If
$[?] P: l_{\infty}^{N} \rightarrow X_{n} \subseteq l_{\infty}^{N}$ is any projection onto $X_{n}$,

$$
\begin{aligned}
\Lambda & =\operatorname{tr}\left(u: X_{n} \rightarrow X_{n}\right)=\operatorname{tr}\left(u P: l_{\infty}^{N} \rightarrow l_{\infty}^{N}\right) \leq\|P\| \sum_{k=1}^{N}\left\|u e_{k}\right\|_{\infty} \\
& =\|P\| \sum_{k=1}^{N} \mu_{k}=\|P\| .
\end{aligned}
$$

Hence $X_{n}$ is an $n$-dimensional subspace of $l_{\infty}^{N}$ with projection constant $\geq \Lambda$. As a subspace of $l_{\infty}^{N}$, it inherits the norm

$$
\begin{equation*}
\frac{1}{\sqrt{n}}\left\|\sum_{i=1}^{n} \alpha_{i} \overline{f_{i}}\right\|_{\infty}=\sup _{1 \leq k \leq N}\left|\sum_{i=1}^{n} \alpha_{i} \overline{x_{k i}}\right|=\sup _{1 \leq k \leq N}\left|<x_{k}, \alpha>\right| \tag{6}
\end{equation*}
$$

where $\alpha=\left(\alpha_{i}\right)_{i=1}^{n} \in \mathbb{K}^{n}$ is the natural coordinate vector. By some duality argument using $u^{*}$ it can be seen that also $\operatorname{Span}\left[f_{1}, \ldots, f_{n}\right]$ as a subspace of $l_{1}^{N}(\mu)$, i.e. with norm $\sum_{k=1}^{N}\left|<x_{k}, \alpha\right\rangle \mid \mu_{k}$, has projection constant $\Lambda$. This proves the reverse inequality and shows how the vectors $x_{k}$ are related to the construction of extremal spaces $X_{n}$. The norm given by (6) on $X_{n}$ shows that the dual unit ball of $X_{n}$ (as a subspace of $l_{\infty}^{N}$ ) is just the absolutely convex hull of the vectors $x_{1}, \ldots, x_{N} \in \mathbb{K}^{n}$.

## 3 Estimating the projection constant

We now turn to the proof of Theorem 1 using the duality result of Proposition 4. For this we use a simple but useful Lemma due to Sidelnikov [4].

Lemma 5. Let $\mu=\left(\mu_{j}\right)_{j=1}^{N} \in\left(\mathbb{R}_{+}\right)^{N}$ be a discrete probability measure and consider $N$ points $x_{1}, \ldots, x_{N} \in S^{n-1}$ on the sphere in $\mathbb{K}^{n}$. Let $\sigma$ denote the normalized surface measure on $S^{n-1}$. Then for every even integer $2 k \in 2 \mathbb{N}$

$$
\sum_{j, l=1}^{N}\left|<x_{j}, x_{l}>\left.\right|^{2 k} \mu_{j} \mu_{l} \geq \int_{S^{n-1}} \int_{S^{n-1}}\right|<x, y>\left.\right|^{2 k} d \sigma(x) d \sigma(y)
$$

In the complex case, express the integrand in real variables and integrate over $S^{2 n-1}(\mathbb{R})$. Thus the discrete $(2 k)$-th moment is always bigger than the continuous moment with respect to the surface measure. For other powers this is not true, in general.

Proof. (for $\mathbb{K}=\mathbb{R}$ ). For $x \in \mathbb{R}^{n}$, denote by $x^{\otimes 2 k}=x \otimes \cdots \otimes x \in \mathbb{R}^{n^{2 k}}$ the ( $2 k$ )-fold tensor product. Then for $x, y \in \mathbb{R}^{n}$

$$
<x^{\otimes 2 k}, y^{\otimes 2 k}>_{\mathbb{R}^{n^{2 k}}}=<x, y>_{\mathbb{R}^{n}}^{2 k}=\left|<x, y>_{\mathbb{R}^{n}}\right|^{2 k}
$$

Given $\left(x_{j}, \mu_{j}\right)$ define the tensor

$$
\zeta:=\sum_{j=1}^{N} \mu_{j} x_{j}^{\otimes 2 k}-\int_{S^{n-1}} x^{\otimes 2 k} d \sigma(x) \in \mathbb{R}^{n^{2 k}} .
$$

The integral is a vector integral which can be defined coordinate-wise. Clearly (3) $0 \leq<\zeta, \zeta>_{\mathbb{R}^{n^{2 k}}}=\sum_{j, l=1}^{N} \mu_{j} \mu_{l}\left|<x_{j}, x_{l}>\right|^{2 k}$
$-2 \sum_{j=1}^{N} \mu_{j} \int_{S^{n-1}}\left|<x_{j}, x>\left.\right|^{2 k} d \sigma(x)+\int_{S^{n-1}} \int_{S^{n-1}}\right|<x, y>\left.\right|^{2 k} d \sigma(x) d \sigma(y)$.
Since $\sigma$ is rotation invariant and $x_{j} \in S^{n-1}$, the first integral does not depend on $j \in\{1, \ldots, N\}$. We may just take $x_{j}=e$ to be a standard unit vector, similarly the inner integral in the last term does not depend on $y$. Thus with $\sum_{j=1}^{N} \mu_{j}=1$

$$
0 \leq<\zeta, \zeta>=\sum_{j, l=1}^{N} \mu_{j} \mu_{l}\left|<x_{j}, x_{l}>\left.\right|^{2 k}-\int_{S^{n-1}}\right|<e, x>\left.\right|^{2 k} d \sigma(x)
$$

which proves Lemma 5.
Clearly $\int_{S^{n-1}}|\langle e, x\rangle|^{2 k} d \sigma(x)=: c_{n k}$ depends only on $n$ and $k$ and can be calculated easily using polar coordinates. In particular, one has

$$
\begin{equation*}
c_{n 1}=\frac{1}{n}, c_{n 2}(\mathbb{R})=\frac{3}{n(n+2)}, c_{n 2}(\mathbb{C})=\frac{2}{n(n+1)} \tag{7}
\end{equation*}
$$

Equality in Lemma 5 is equivalent to $\zeta=0$. Expressing $\zeta$ in coordinates, this means that for any monomial $p(x)=x_{1}^{\alpha_{1}} \cdots x_{n}^{\alpha_{n}}$ of total degree $\alpha_{1}+$ $\cdots+\alpha_{n}=2 k$ we have

$$
\begin{equation*}
\sum_{j=1}^{N} \mu_{j} p\left(x_{j}\right)=\int_{S^{n-1}} p(x) d \sigma(x) \tag{8}
\end{equation*}
$$

Since these monomials form a basis of all polynomials $p$ which are homogeneous of degree ( $2 k$ ) in $n$ variables, these are also integrated exactly by the cubature rule (8). Any even polynomial of degree $\leq 2 k$ is a sum of polynomials homogeneous of degree ( $2 l$ ) where $0 \leq l \leq k$. Since these may be multiplied by $\langle x, x\rangle^{k-l}$ on the sphere $S^{n-1}$, (8) integrates all even polynomials of degree $\leq 2 k$ : the vectors $\left(x_{j}\right)$ and the measure $\left(\mu_{j}\right)$ constitute a spherical design of degree $2 k$ in $n$ variables.

Proof of Theorem 1. i) Let $X_{n}$ be an $n$-dimensional normed space. By approximation, we may assume that $X_{n} \subseteq l_{\infty}^{N}$ for some finite $N \in \mathbb{N}$. To estimate $\lambda\left(X_{n}\right)$ from above, using Proposition 4, we have to bound any expression of the form

$$
n \sum_{j, k=1}^{N} \mu_{j} \mu_{k}\left|<x_{j}, x_{k}>\right|
$$

by $g(n)$ where $x_{j} \in S^{n-1}$ and $\mu_{j}$ are such that

$$
\begin{equation*}
I d_{n}=n \sum_{j=1}^{N} \mu_{j}<\cdot, x_{j}>x_{j} \tag{9}
\end{equation*}
$$

Let $\alpha=1 / \sqrt{n+2}$ if $\mathbb{K}=\mathbb{R}$ and $\alpha=1 / \sqrt{n+1}$ if $\mathbb{K}=\mathbb{C}$. For $t \in[-1,1]$,

$$
\begin{gather*}
(|t|-\alpha)^{2}=\left(\left(t^{2}-\alpha^{2}\right) /(|t|+\alpha)\right)^{2} \geq\left(t^{2}-\alpha^{2}\right)^{2} /(1+\alpha)^{2} \\
|t| \leq \gamma_{0}+\gamma_{2} t^{2}-\gamma_{4} t^{4}, t \in[-1,1] \tag{10}
\end{gather*}
$$

where

$$
\gamma_{0}=\frac{\alpha}{2}-\frac{\alpha^{3}}{2(1+\alpha)^{2}}, \gamma_{2}=\frac{1}{2 \alpha}+\frac{\alpha}{(1+\alpha)^{2}}, \gamma_{4}=\frac{1}{2 \alpha(1+\alpha)^{2}}
$$

are $\geq 0$. Equality in (10) holds if and only if $|t|=\alpha$ or 1 . Therefore

$$
\begin{aligned}
& n \sum_{j, k=1}^{N} \mu_{j} \mu_{k} \mid<x_{j}, x_{k}>1 \\
& \quad \leq n \gamma_{0}+n \gamma_{2} \sum_{j, k=1}^{N} \mu_{j} \mu_{k}\left|<x_{j}, x_{k}>\left.\right|^{2}-n \gamma_{4} \sum_{j, k=1}^{N} \mu_{j} \mu_{k}\right|<x_{j}, x_{k}>\left.\right|^{4} .
\end{aligned}
$$

By (7), $\sum_{j, k=1}^{N} \mu_{j} \mu_{k}\left|<x_{j}, x_{k}>\right|^{2}=1 / n$. Using Lemma 5 , we find that

$$
n \sum_{j, k=1}^{N} \mu_{j} \mu_{k}\left|<x_{j}, x_{k}>\right| \leq n \gamma_{0}+\gamma_{2}-n \gamma_{4} c_{n 2}(\mathbb{K})=g(n)
$$

with $c_{n 2}(\mathbb{K})$ as in (7). The last inequality is by calculation, $g$ is the function given in Theorem 1. Actually, the value of $\alpha$ chosen is such that it minimizes the expression $n \gamma_{0}+\gamma_{2}-n \gamma_{4} c_{n 2}(\mathbb{K})$; it depends on $\mathbb{K}$ since $c_{n 2}(\mathbb{K})$ depends on $\mathbb{K} \in\{\mathbb{R}, \mathbb{C}\}$. Thus $\lambda\left(X_{n}\right) \leq g(n)$ as claimed.
ii) As for the case of equality, clearly needed is for all $j, k=1 \ldots N$ that $[?]\left|<x_{j}, x_{k}>\right| \in\{\alpha, 1\}$, i.e. for $j \neq k$ that $\left|<x_{j}, x_{k}>\right|=\alpha$ : the lines spanned by the vectors $x_{j}$ should be equiangular. There are at most $N(n)$ equiangular vectors in $\mathbb{K}^{n}$ with $N(n)=n(n+1) / 2$ if $\mathbb{K}=\mathbb{R}$ and $N(n)=n^{2}$ if $\mathbb{K}=\mathbb{C}$ since the orthogonal projections onto the lines spanned by the $x_{j}$ 's, $P_{j}:=<\cdot, x_{j}>x_{j}$, are linearly independent as operators:

We know that

$$
<P_{j}, P_{k}>:=\operatorname{tr}\left(P_{j} P_{k}\right)=\left|<x_{j}, x_{k}>\right|^{2}=\alpha^{2}
$$

and hence the Gram matrix

$$
\left.G=\left(<P_{j}, P_{k}\right\rangle\right)_{j, k=1}^{N}=\left(\begin{array}{cc}
1 & \alpha^{2} \\
\ddots & \\
\alpha^{2} & 1
\end{array}\right)
$$

has determinant $\neq 0$ since $\alpha<1$. Being hermitean, there are only at most $N(n)$ linearly independent $P_{j}$ 's. In the case of equality, therefore by using the Cauchy-Schwartz inequality

$$
\begin{aligned}
g(n) & =n\left(\sum_{j, k=1}^{N} \mu_{j} \mu_{k} \alpha+\sum_{j=1}^{N} \mu_{j}^{2}(1-\alpha)\right) \\
& \geq n \alpha+n / N\left(\sum_{j=1}^{N} \mu_{j}\right)^{2}(1-\alpha)=n \alpha+n / N(1-\alpha) \\
& \geq n \alpha+n / N(n)(1-\alpha)=g(n) .
\end{aligned}
$$

The last equality again is by direct calculation. Since this is a chain of equalities, $N=N(n)$ must be maximal and all $\mu_{j}$ 's equal to $\mu_{j}=\frac{1}{N}$. An extremal space $X_{n}$ can hence be realized by considering the norm

$$
\|x\|=\sup _{1 \leq j \leq N}\left|<x, x_{j}>\right|
$$

or, dually, by

$$
\|x\|=\sum_{j=1}^{N}\left|<x, x_{j}>\right| .
$$

The matrix $P=n / N\left(<x_{j}, x_{k}>\right)_{j, k=1}^{N}=1$ gives the orthogonal projection onto $X_{n}$ with $\|P\|=\frac{n}{N}(1+(N-1) \alpha)=g(n)$.

Since there are only a few cases of known sets of equiangular vectors in $\mathbb{K}^{n}$ of maximal possible size $N(n)$, it is useful to study examples which almost give the maximal number and yield spaces with almost maximal projection constant. We can do this in certain complex dimensions. For this, we need the following number theoretic fact, cf. [6]
Lemma 6. Let $n=p^{m}+1$ be a prime power plus 1 and $N=n^{2}-n+1$. Then there exist $d_{1}, \ldots, d_{n} \in\{0, \ldots, N-1\}$ such that all differences $\left(d_{l}-d_{m}\right)$ modulo $N(l \neq m)$ are all different and yield all $n(n-1)=N-1$ integers between 1 and $N-1$.

Using this, we now construct complex spaces to give the
Proof of Proposition 2. Take $d_{1}, \ldots, d_{n}$ as in Lemma 6 and let with $N=$ $n^{2}-n+1$

$$
x_{k}:=n^{-1 / 2}\left(\exp \left(\frac{2 \pi i}{N} d_{l} k\right)\right)_{l=1}^{n} \in S^{n-1}(\mathbb{C})
$$

for $k=1, \ldots, N$. The vectors $x_{k}$ are equiangular since for $j \neq k$

$$
\begin{aligned}
n^{2}\left|<x_{j}, x_{k}>\right|^{2} & =\left|\sum_{l=1}^{n} \exp \left(\frac{2 \pi i}{N} d_{l}(j-k)\right)\right|^{2}, \quad \theta:=j-k \neq 0 \\
& =\sum_{l, m=1}^{n} \exp \left(\frac{2 \pi i}{N}\left(d_{l}-d_{m}\right) \theta\right) \\
& =n+\sum_{l \neq m} \exp \left(\frac{2 \pi i}{N}\left(d_{l}-d_{m}\right) \theta\right) \\
& =n+\sum_{k=1}^{N-1} \exp \left(\frac{2 \pi i}{N} k \theta\right)=n-1
\end{aligned}
$$

hence $\left|<x_{j}, x_{k}>\right|=\sqrt{n-1} / n$ for all $1 \leq j \neq k \leq N=n^{2}-n+1$. Defining $f_{l}=\left(x_{k l}\right)_{k=1}^{N} \in \mathbb{C}^{N}$, we have for $l \neq m$

$$
<f_{l}, f_{m}>=\sum_{k=1}^{N} \frac{1}{n} \exp \left(\frac{2 \pi i}{N}\left(d_{l}-d_{m}\right) k\right)=\frac{N}{n} \delta_{l m}
$$

and thus

$$
I d_{n}=\frac{n}{N} \sum_{k=1}^{N}<\cdot, x_{k}>x_{k} \text { on } \mathbb{C}^{n}
$$

Let $X_{n}=\operatorname{Span}\left(f_{1}, \ldots, f_{n}\right) \subseteq l_{\infty}^{N}$. Considered as a subspace of $l_{2}^{N},\left(\sqrt{\frac{n}{N}} f_{l}\right)_{l=1}^{n}$ is an orthonormal basis of $X_{n}$ and $P:=\frac{n}{N}\left(\left\langle x_{j}, x_{k}\right\rangle\right)_{j, k=1}^{N}$ gives the orthogonal projection onto $X_{n}$. Clearly

$$
\left.\lambda\left(X_{n}\right) \leq\|P\|=\frac{n}{N} \sup _{1 \leq j \leq N} \sum_{k=1}^{N}\left|<x_{j}, x_{k}\right\rangle \right\rvert\,=\frac{n}{N}\left(1+(N-1) \frac{\sqrt{n-1}}{n}\right)=: h(n)
$$

To show that $\lambda\left(X_{n}\right)=h(n)$, we consider

$$
u:=\left(<x_{j}, x_{k}>\right)_{j, k=1}^{N}-\left(1-\frac{\sqrt{n-1}}{n}\right) I d_{N}: l_{\infty}^{N} \rightarrow l_{\infty}^{N}
$$

Then for all $j, k=1, \ldots, N,\left|u_{j k}\right|=\sqrt{n-1} / n$ and

$$
\left.u\right|_{X_{n}}=\left(\frac{N}{n}-1+\frac{\sqrt{n-1}}{n}\right) I d_{X_{n}}
$$

Hence

$$
\operatorname{tr}\left(u: X_{n} \rightarrow X_{n}\right)=N-n+\sqrt{n-1}
$$

and

$$
\sum_{j=1}^{N}\left\|u e_{j}\right\|_{\infty}=\sum_{j=1}^{N} \sup _{1 \leq k \leq N}\left|u_{j k}\right|=N \frac{\sqrt{n-1}}{n} .
$$

This implies that for any projection $Q: l_{\infty}^{N} \rightarrow X_{n} \subseteq l_{\infty}^{N}$ onto $X_{n}$

$$
\begin{aligned}
N-n+\sqrt{n-1} & =\operatorname{tr}\left(u: X_{n} \rightarrow X_{n}\right)=\operatorname{tr}(u Q) \\
& \leq\|Q\| \sum_{j=1}^{N}\left\|u e_{j}\right\|_{\infty}=\|Q\| N \frac{\sqrt{n-1}}{n}
\end{aligned}
$$

therefore

$$
\lambda\left(X_{n}\right) \geq \frac{N-n+\sqrt{n-1}}{N \sqrt{n-1}} n=h(n)
$$

the last equality again verified by direct calculation. It is easily checked that [?] $h(n) \leq g(n) \leq h(n)+\frac{1}{2 \sqrt{n}}$ so that (2) holds. Since $u$ is symmetric, a similar construction works in $l_{1}^{N}$ instead of $l_{\infty}^{N}$ by duality.

## 4 Isometric imbeddings of euclidean spaces into $l_{p}$-spaces

We now turn to the problem of imbedding euclidean space isometrically into $l_{p}^{N}$-spaces where similar spherical designs techniques are useful: again equiangular vectors and the Sidelnikov lemma are needed.

It has been known for a long time - see e.g. Lyubich [13] that even the 2-dimensional Hilbert space $l_{2}^{2}$ does not imbed isometrically into $l_{p}^{N}$ for some finite $N \in \mathbb{N}$ unless $p$ is an even integer. In fact the same is true even for imbedding $l_{2}^{2}$ into the infinite-dimensional $l_{p}$, see Delbaen, Jarchow and Pelczynski [2]. However, for even $p=2 k \in 2 \mathbb{N}$, isometric imbeddings of $l_{2}^{n}$ into $l_{p}^{N}$ exist provided $N=N(n, k)$ is sufficiently large. The easiest example of imbedding $l_{2}^{2}$ into $l_{4}^{3}$ is a consequence of the equality

$$
x^{4}+\left(-\frac{x}{2}+\frac{\sqrt{3}}{2} y\right)^{4}+\left(-\frac{x}{2}-\frac{\sqrt{3}}{2} y\right)^{4}=\frac{9}{8}\left(x^{2}+y^{2}\right)^{2},(x, y) \in \mathbb{R}^{2}
$$

which means that with $f_{1}=(1,-1 / 2,-1 / 2), f_{2}=(0, \sqrt{3 / 2},-\sqrt{3 / 2})$

$$
\left\|x f_{1}+y f_{2}\right\|_{4}=\sqrt[4]{9 / 8}\|(x, y)\|_{2}
$$

holds and $X_{2}=\operatorname{Span}\left[f_{1}, f_{2}\right]=\left\{z \in \mathbb{R}^{3} \mid \sum_{i=1}^{3} z_{i}=0\right\} \subseteq l_{4}^{3}$ is hilbertian. In general, the following holds:

Proposition 7. Let $n, k \in \mathbb{N}_{\geq 2}$. Then there exists $N \in \mathbb{N}$ such that $l_{2}^{n} \hookrightarrow l_{2 k}^{N}$ imbeds isometrically. In fact, $N$ can be chosen such that

$$
L(n, k) \leq N \leq U(n, k)
$$

where

$$
\left\{\begin{array}{lll}
L(n, k)=\binom{n+k-1}{k}, & U(n, k)=\binom{n+2 k-1}{2 k} & \mathbb{K}=\mathbb{R} \\
L(n, k)=\binom{n+\left[\frac{k+1}{2}\right]-1}{\left[\frac{k+1}{2}\right]}\binom{n+\left[\frac{k}{2}\right]-1}{\left[\frac{k}{2}\right]}, & U(n, k)=\binom{n+k-1}{k}^{2} & \mathbb{K}=\mathbb{C}
\end{array}\right.
$$

Clearly, $L(n, k) \sim n^{k}, U(n, k) \sim n^{2 k}$ as $n \rightarrow \infty$.
Proof. We give the proof for the upper bound yielding the possibility of isometric imbedding, for $\mathbb{K}=\mathbb{R}$, taken from [16]. If $\sigma$ denotes again the
normalized Haar measure on $S^{n-1}$, the rotation invariance of $\sigma$ implies for $x \in \mathbb{R}^{n}$

$$
\int_{S^{n-1}}|<x, y>|^{2 k} d \sigma(y)=c_{n k}\|x\|_{2}^{2 k} .
$$

Approximating the integral by Riemann sums, we see that $c_{n k}\|\cdot\|^{2 k}$ is in the closed convex hull of the polynomials $P=\left\{\langle\cdot, y\rangle^{2 k} \mid y \in S^{n-1}\right\}$ in the positive cone of the ( $2 k$ )-homogeneous polynomials of degree $2 k$ in $n$ variables $P_{2 k, n}^{\mathrm{hom}}$. Being a finite-dimensional space, the convex hull is closed and hence Carathéodory's theorem yields that $N$ may be chosen smaller than $N \leq U(n, k)=\operatorname{dim} P_{2 k, n}^{\text {hom }}=\binom{n+2 k-1}{2 k}$.

The connection between isometric imbeddings and spherical designs is given by the following result found e.g. in [7] or partly in [4], [15] or [17].

Proposition 8. Let $n, k, N \in \mathbb{N}_{\geq 2}$. Then the following are equivalent:
(1) There exists an isometric imbedding $l_{2}^{n} \hookrightarrow l_{2 k}^{N}$.
(2) There exist points $x_{1}, \ldots, x_{N} \in S^{n-1}$ and $\left(\mu_{j}\right)_{j=1}^{N} \subset \mathbb{R}_{+}, \sum_{j=1}^{N} \mu_{j}=1$ such that

$$
\sum_{j, l=1}^{N} \mu_{j} \mu_{l}\left|<x_{j}, x_{l}>\left.\right|^{2 k}=\int_{S^{n-1}} \int_{S^{n-1}}\right|<x, y>\left.\right|^{2 k} d \sigma(x) d \sigma(y)=: c_{n k}
$$

(3) There exist points $x_{1}, \ldots, x_{N} \in S^{n-1}$ and $\left(\mu_{j}\right)_{j=1}^{N} \subset \mathbb{R}_{+}, \sum_{j=1}^{N} \mu_{j}=1$ such that for all even polynomials of degree $2 k$ in $n$ variables $p \in P_{2 k, n}^{\text {even }}$

$$
\begin{equation*}
\sum_{j=1}^{N} \mu_{j} p\left(x_{j}\right)=\int_{S^{n-1}} p(x) d \sigma(x) \tag{11}
\end{equation*}
$$

Thus ( $x_{j}, \mu_{j}$ ) yield a cubature formula with $N$ nodes for even polynomials of degree $2 k$ which constitutes a spherical design of degree $2 k$.

Proof. (1) $\Rightarrow(2)$. Any isometric imbedding $i: l_{2}^{n} \rightarrow l_{2 k}^{N}$ has the form

$$
x \in l_{2}^{n} \mapsto\left(<x, z_{j}>\right)_{j=1}^{N} \in l_{2 k}^{N},\|x\|_{2}=\left(\sum_{j=1}^{N}\left|<x, z_{j}>\right|^{2 k}\right)^{1 / 2 k}
$$

Let $x_{j}:=z_{j} /\left\|z_{j}\right\|_{2}, \mu_{j}:=\left\|z_{j}\right\|_{2}^{2 k} / \sum_{l=1}^{N}\left\|z_{l}\right\|_{2}^{2 k}$ and $d=\sum_{l=1}^{N}\left\|z_{l}\right\|_{2}^{2 k}$. By rotation invariance of $\sigma$, for $x \in S^{n-1}$

$$
\begin{aligned}
& \sum_{j=1}^{N} \mu_{j}\left|<x, x_{j}>\right|^{2 k}=d^{-1}\|x\|_{2}^{2 k}=d^{-1} \\
& \quad=d^{-1} \int_{S^{n-1}}\|y\|_{2}^{2 k} d \sigma(y)=\sum_{j=1}^{N} \mu_{j} \int_{S^{n-1}}\left|<y, x_{j}>\right|^{2 k} d \sigma(y)=c_{n k}
\end{aligned}
$$

Take $x=x_{m}$ and sum over $m=1, \ldots, N$.
$(2) \Rightarrow(3)$. This was proved in the remark following the proof of Lemma 5 since (2) means that $\zeta=0$.
$(3) \Rightarrow(1)$. Apply (3) to $p=<\cdot, x>^{2 k}$ for fixed $x \in \mathbb{K}^{n}$.
The lower bound in Proposition 7 is now an immediate consequence of (3): If $N$ could be chosen $<\binom{n+k-1}{k}=\operatorname{dim} P_{k, n}^{h o m}$, for any set of vectors $\left(x_{j}\right)_{j=1}^{N} \subseteq S^{n-1}$ there would be $p \in P_{k, n}^{\mathrm{hom}}$ with $p\left(x_{j}\right)=0$ for all $j$. Then $p^{2} \in P_{2 k, n}^{\mathrm{hom}}$ and for any $\left(\mu_{j}\right)$

$$
\sum_{j=1}^{N} \mu_{j} p\left(x_{j}\right)^{2}=0 \neq \int_{S^{n-1}} p(y)^{2} d \sigma(y)
$$

in contradiction with (3).
Clearly, one is interested in cubature formulas with a minimal number $N$ of nodes or equivalently in imbeddings into $l_{2 k}^{N}$-spaces of smallest dimension $N$. Thus let

$$
N(n, k):=\min \left\{N \mid \exists \text { an isometric imbedding } l_{2}^{n} \rightarrow l_{2 k}^{N}\right\}
$$

and we know that $L(n, k) \leq N(n, k) \leq U(n, k)$. Looking at (3) in Proposition 8 , dimension reasons might indicate that the upper bound may be the right order. However, Dvoretzky's theorem as in [3] might give hope that the lower bound could be sharp. Designs $\left(x_{j}, \mu_{j}\right)$ attaining this lower bound $N(n, k)=L(n, k)$ are called tight. There exists a charaterization of tight designs by Bannai [1] and Hoggar [5] which we cite as.

Proposition 9. Let $n, k \in \mathbb{N}_{\geq 2}$ and assume that $N(n, k)=L(n, k)=: N$. Then there exist points $\left(x_{j}\right)_{j=1}^{N} \subseteq S^{n-1}$ such that (11) holds with $\mu_{j}=1 / N$ and for any $1 \leq j \neq l \leq N$ the scalar products $\left|<x_{j}, x_{l}>\right|$ are zeros of $a$
fixed polynomial $C_{n k}$ of degree $k$ in one variable. For $k=2$ this is equivalent to the existence of $n(n+1) / 2(\mathbb{R})$ or $n^{2}(\mathbb{C})$ equiangular vectors in $\mathbb{K}^{n}$ which was discussed before.

The polynomials $C_{n k}$ are given in terms of Jacobi polynomials for $\mathbb{K}=\mathbb{R}$ as $C_{n k}=P_{k}^{((n-1) / 2,(n-1) / 2)}$. They are even/odd depending on whether $k$ is even/odd. For instance, $C_{n 2}(t)=(n+2) t^{2}-1$ so that one needs $\left|\left\langle x_{j}, x_{l}\right\rangle\right|$ to be equal to $\left|\left\langle x_{j}, x_{l}\right\rangle\right|=1 / \sqrt{n+2}$ : the vectors are equiangular and $N=L(n, 2)=n(n+1) / 2$. Unfortunately, not too many examples where $N(n, k)=L(n, k)$ holds, are known. In fact, they do not exist for $k>5, n>2$. So we have to be satisfied with examples where $N(n, k)$ does not deviate too much from the lower bound. In the complex case, the almost extremal number of equiangular vectors constructed in the proof of Proposition 2 provides an example for $k=2, \mathbb{K}=\mathbb{C}$. Here $N=n^{2}+1$ while $L(n, 2)=n^{2}$ :

Proposition 10. Let $n=p^{t}+1$ be a prime power plus one. Then there exists a complex isometric imbedding $l_{2}^{n} \rightarrow l_{4}^{n^{2}+1}$.

There are also $\mathrm{O}\left(n^{2}\right)$-examples in the real case, cf. [7].
Proof. Let $d_{1}, \ldots, d_{n}$ be as in Lemma 6 and consider again the $M=n^{2}-n+1$ vectors

$$
x_{j}:=n^{-1 / 2}\left(\exp \left(\frac{2 \pi i}{N} d_{m} j\right)\right)_{m=1}^{n} \in S^{n-1}(\mathbb{C}) \subseteq \mathbb{C}^{n}
$$

As shown before, they are equiangular, $\left|\left\langle x_{j}, x_{l}\right\rangle\right|=\sqrt{n-1} / n$ for all $1 \leq j \neq l \leq M$. We would like to check condition (2) of Proposition 8 to get isometric imbeddings into $l_{4}^{N}$; then $c_{n 2}(\mathbb{C})=\frac{2}{n(n+1)}$. Taking $\mu_{j}=1 / M$ and the above vectors gives a slightly larger value on the left side of (2) than $c_{n 2}(\mathbb{C})$. However, adding the standard unit vectors $e_{1}, \ldots, e_{n}$ as the vectors $x_{M+1}, \ldots, x_{N}$ with $N:=n^{2}+1$ and letting $\mu_{j}:=\frac{n}{n+1} \frac{1}{M}$ if $1 \leq j \leq M$ as well as $\mu_{j}:=\frac{n}{n+1} \frac{1}{n^{2}}$ if $M<j \leq N$, one checks by direct calculation that (2) of Proposition 8 is satisfied for this slightly larger set of vectors. Interestingly enough, the $\mu_{j}$ 's attain 2 different values, although both are of very similar size.

For imbeddings into $l_{2 k}^{N}$ with $k \geq 3$ the minimal order of $N(n, k)$ for $n \rightarrow \infty$ is unknown. A recent result of Kuperberg [12] improves the O $\left(n^{2 k}\right)$-bound to $\mathrm{O}\left(n^{2 k-1}\right)$; it might possibly be strengthened to $\mathrm{O}\left(n^{2 k-2}\right)$ which would also fit with Proposition 10. However, the constants involved in the O-term are bigger than 1 in Kuperberg's construction using BCH-codes.

## Acknowledgement

These are notes of a series of lectures given in the Seminario de Análisis Matemático de la Universidad de Sevilla in March 2004. I express my gratitude to Prof. Rafael Villa-Caro and the people at the Department of Mathematics for the invitation and the hospitality I enjoyed there.

## References

[1] E. Bannai, On extremal finite sets in the sphere and other matric spaces, Lond. Math. Soc. Lect. Notes Ser., 131 (1986), 13-38.
[2] F. Delbaen, H. Jarchow, A. Pelczynski, Subspaces of $L_{p}$ isometric to subspaces of $l_{p}$, Positivity, 2 (1998), 339-367.
[3] T. Figiel, J. Lindenstrauss, V.D. Milman, The dimension of almost spherical sections of convex bodies, Acta Math. , 139 (1977), 53-94.
[4] J.M. Goethals, J.J. Seidel, Cubature formulas, polytopes, and spherical designs, in: The Geometric Vein, ed. C. Davis et alt., 203-218, Springer (1981).
[5] S.G. Hoggar, t-designs in projective spaces, Europ. Journ. Combin., 3 (1982) 233-254.
[6] H. Halberstam, K.F. Roth, Sequences, Springer (1983).
[7] H. König, Isometric imbeddings of enclidean spaces into finite dimensional $l_{p}$-spaces, Banach Center Public. , 34 (1995), 79-87.
[8] I.M. Kadets, M.G. Snobar, Certain functionals on the Minkowski compactum, Math. Notes, 10 (1971), 694-696.
[9] H. König, C. Schtt, N. Tomczak-Jaegermann, Projection constants of symmetric spaces and variants of Khintchine's inequality, Journ. Reine Angew. Math., 511 (1999), 1-42.
[10] H. König, N. Tomczak-Jaegermann, Norms of minimal projections, Journ. Funct. Anal., 119 (1994), 253-280.
[11] H. König, N. Tomczak-Jaegermann, Spaces with maximal projection constants, Studia Math., 159 (2003), 357-372.
[12] G. Kuperberg, Numerical cubature using error-correcting codes, Preprint Feb. 2004, arXiv: math. NA/0402047v1 4 Feb 2004
[13] Yu. Lyubich, On the boundary spectrum of a contraction in Minkowsky spaces, Siberian Math. Journ., 11 (1970), 271-279.
[14] P.W.H. Lemmens, J.J. Seidel, Equiangular lines, Journ. Algebra, 24 (1973), 494-512.
[15] Yu. Lyubich, L. Vaserstein, Isometric imbeddings between classical Banach spaces, cubature formulas, and spherical designs, Geom. Dedic. , 47 (1993), 327-362.
[16] V.D. Milman, A few observations on the connections between local theory and some other fields, Lect. Notes in Math., 1317 (1988), 283289.
[17] B. Reznick, Sums of even powers of real linear forms, Mem. AMS, 96 (1992), no. 463.
[18] J.M. Renes, R. Blume-Kohout, A. J. Scott, C. M. Caves, Symmetric informationally complete quantum measurements, Preprint Oct 2003, arXiv: quant-ph/0310075 v1 13 Oct 2003.

## Author's address:

H. König<br>Mathematisches Seminar<br>Universitt Kiel 24098 Kiel, Germany<br>E-mail: hkoenig@math.uni-kiel.de

# Weighted norm inequalities for singular integral operators 

J.M. Martell *


#### Abstract

We study weighted norm inequalities for singular integral operators with different smoothness conditions assumed on the kernels. The weakest one is the so-called classical Hörmander condition, which is an $L^{1}$ regularity, and the strongest is given by a Hölder or Lipschitz smoothness. Between them we have some kind of $L^{r}$-regularity, $1<$ $r \leq \infty$. We will present some results that are known for singular integrals with these kernels. We will be focused on studying Coifman's inequality: $$
\int_{\mathbb{R}^{n}}|T f(x)|^{p} w(x) d x \leq C \int_{\mathbb{R}^{n}} M f(x)^{p} w(x) d x
$$ for any $0<p<\infty$ and $w \in A_{\infty}$, where $T$ is a singular integral operator with kernel satisfying a Hölder regularity condition and $M$ is the Hardy-Littlewood maximal function. We will see that such an inequality is no longer true when the hypotheses on the kernel are relaxed. This is the case for kernels satisfying the Hörmander condition. For the intermediate regularity conditions some positive and negative results of this kind are shown. In these cases the operator on the right hand side is changed in such a way that it can measure the singularity of $T$. Some of the results we will present are in a collaboration paper with Carlos Pérez and Rodrigo Trujillo-González.


## 1 Introduction.

Some of the most significant and studied operators in Harmonic Analysis are the Hardy-Littlewood maximal function, the Hilbert transform and the Riesz transforms. The first one is defined as the supremum of the averages of

[^5]the function over all the cubes $Q \subset \mathbb{R}^{n}$ (with sides parallel to the coordinate axes in the sequel), that is,
$$
M f(x)=\sup _{Q \ni x} \frac{1}{|Q|} \int_{Q}|f(y)| d y .
$$

The Hilbert transform is defined in $\mathbb{R}$ and the Riesz transforms are the analogs in $\mathbb{R}^{n}, n \geq 2$, and they are given in the following way

$$
H f(x)=p \cdot v \cdot \int_{\mathbb{R}} \frac{f(y)}{x-y} d y, \quad \quad R_{j} f(x)=p \cdot v \cdot \int_{\mathbb{R}^{n}} \frac{x_{j}-y_{j}}{|x-y|^{n+1}} f(y) d y
$$

These integrals have to be defined in such a way they make sense. Note that the kernels $1 / x$ in $\mathbb{R}$ or $x_{j} /|x|^{n+1}$ in $\mathbb{R}^{n}, n \geq 2$, are singular and they are not locally integrable at the origin and this is the reason why the integrals are understood in a principal value sense. The Hardy-Littlewood maximal function is very related to Hilbert or Riesz transforms since it controls them as we will see later. Studying maximal operators turns out to be easier and this control might be crucial to understand the singular integral operators.

A generalization of the Hilbert or Riesz transforms is given by the following convolution type operators

$$
T f(x)=p \cdot v \cdot \int_{\mathbb{R}^{n}} K(x-y) f(y) d y
$$

with kernel $K$ having bounded Fourier transform $\widehat{K} \in L^{\infty}\left(\mathbb{R}^{n}\right)$. Thus, $T$ is a linear and bounded operator on $L^{2}\left(\mathbb{R}^{n}\right)$. Further generalizations can be considered with two-variable kernels that do not give a convolution type operator and some of them play an important role in Analysis. Nevertheless, we are going to concentrate in the simplest case on which the operators are of convolution type, the reader is referred, for instance, to [5] for the general case.

Coming back to the singular integral operators defined above, so far we only know that they are continuous in $L^{2}\left(\mathbb{R}^{n}\right)$. To get better properties on $T$ some conditions can be imposed about the size or the smoothness of $K$. The size condition of the kernel that generalizes the case of the Hilbert or Riesz transforms is $|K(x)| \leq A|x|^{-n}$. Note that this decay has a problem of integrability both at the origin and at infinity. For the operators that we want to consider this condition will not be assumed, we will be focused on different smoothness conditions on $K$ and the results that can be achieved by assuming them. The regularity conditions will be scaled in the Lebesgue
spaces and we will use the notation $H_{r}, 1 \leq r \leq \infty$. The weakest one is the so-called Hörmander condition

$$
\begin{equation*}
\sup _{y \in \mathbb{R}^{n}} \int_{|x|>c|y|}|K(x-y)-K(x)| d x<\infty \tag{1}
\end{equation*}
$$

which is understood as an $L^{1}$-regularity. A singular integral operator with kernel satisfying $\left(H_{1}\right)$ is of weak type $(1,1)$ and bounded on $L^{p}\left(\mathbb{R}^{n}\right), 1<$ $p<\infty$. This a classical result obtained by Calderón and Zygmund in the 50 's, see [2]. The main tool for this proof is the Calderón-Zygmund decomposition of the function into a good and a bad part. This decomposition is performed by means of the Hardy-Littlewood maximal operator, fact that reflects the connection between this maximal function and the singular integral operators.

If $\left(H_{1}\right)$ is the weakest regularity assumption, the strongest one will be of Hölder or Lipschitz type, namely,

$$
|K(x-y)-K(x)| \leq C \frac{|y|^{\alpha}}{|x|^{\alpha+n}}, \quad \text { whenever }|x|>c|y|, \quad\left(H_{\infty}^{*}\right)
$$

for some $c>1$ and $0<\alpha \leq 1$. The reason why we have used ( $H_{\infty}^{*}$ ) rather than $\left(H_{\infty}\right)$ will be clear later - we keep this latter notation for an $L^{\infty}$ condition that is weaker-. Note that this condition implies $\left(H_{1}\right)$ and also that the kernels of the Hilbert or Riesz transforms satisfy $\left(H_{\infty}^{*}\right)$ with $c=2$ and $\alpha=1$. Indeed, they verify an estimate that is better: $|\nabla K(x)| \leq$ $A|x|^{-(n+1)}$. We will see after a while that $\left(H_{\infty}^{*}\right)$ is key when weighted norm inequalities are studied.

Between $\left(H_{1}\right)$ and ( $H_{\infty}^{*}$ ) the following variant of the Hörmander condition can be considered: let $1 \leq r \leq \infty$, we say that the kernel $K$ verifies the $L^{r}$-Hörmander condition, if there are $c, C_{r}>0$ such that for any $y \in \mathbb{R}^{n}$ and $R>c|y|$

$$
\begin{equation*}
\sum_{m=1}^{\infty}\left(2^{m} R\right)^{\frac{n}{r^{\prime}}}\left(\int_{2^{m} R<|x| \leq 2^{m+1} R}|K(x-y)-K(x)|^{r} d x\right)^{\frac{1}{r}} \leq C_{r} \tag{r}
\end{equation*}
$$

in the case $r<\infty$, and

$$
\sum_{m=1}^{\infty}\left(2^{m} R\right)^{n} \sup _{2^{m} R<|x| \leq 2^{m+1} R}|K(x-y)-K(x)| \leq C_{\infty}
$$

when $r=\infty$. We will use the notation $\left(H_{r}\right)$ for the previous conditions and $H_{r}$ for the classes of kernels satisfying them, the same is applied to $\left(H_{\infty}^{*}\right)$.

This definition is implicit in the work of D. Kurtz and R. Wheeden [8], where it is shown that the classical Dini condition for $K$ implies that $K \in H_{r}$ (see [8, p. 359]). Later on, these classes $H_{r}$ were considered in [12] and [13]. In fact, in this last paper the $L^{r}$-Hörmander condition plays an essential role when studying rough singular integral operators. Namely, for such an operator $T$, one can write $T=\sum T_{j}$ where the kernel of $T_{j}$ satisfies the $L^{r}$-Hörmander condition with constant growing linearly in $j$.

Our aim is twofold. Firstly, we will review the weighted norm estimates that are known for the singular integral operators with the kernels in the previous classes. We will study how sharp they are. Secondly, we present some lack of weighted norm inequalities when the kernels are less regular. In particular, for $K \in H_{1}$ we are going to provide some counterexamples on which the expected weighted norm inequalities do not hold. To prove these negative results we will use some extrapolation results taken from [4].

The source of this presentation is the paper [10] written in collaboration with C. Pérez and R. Trujillo-González to whom the author wants to express his gratitude.

## 2 Weighted norm inequalities and Coifman's type estimates

In what follows a weight $w$ is a non-negative locally integrable function. As usual $L^{p}(w)$ will denote the $L^{p}$ space with the underlying measure $w(x) d x$.

Muckenhoupt in [11] found some classes of weights when he characterized the boundedness of the Hardy-Littlewood maximal function in weighted Lebesgue spaces. The classes $A_{p}, 1 \leq p<\infty$, are defined as

$$
\begin{gather*}
\left(\frac{1}{|Q|} \int_{Q} w(x) d x\right)\left(\frac{1}{|Q|} \int_{Q} w(x)^{1-p^{\prime}} d x\right)^{p-1} \leq C<\infty, \text { for } p>1,  \tag{p}\\
\frac{1}{|Q|} \int_{Q} w(x) d x \leq C w(x), \quad \text { for a.e. } x \in Q
\end{gather*}
$$

The class $A_{1}$ can be equivalently defined as $M w(x) \leq C w(x)$ a.e. We also remind that $A_{\infty}=\bigcup_{p \geq 1} A_{p}$. The result proved in [11] establishes that $M$ maps $L^{1}(w)$ into $L^{1, \infty}(w)$ if and only if $w \in A_{1}$ and that $w$ is bounded on $L^{p}(w), 1<p<\infty$, if and only if $w \in A_{p}$.

On the other hand, Coifman's inequality, see [3], states a precise control of Calderón-Zygmund operators $T$ with kernel $K \in H_{\infty}^{*}$ in terms of $M$ :

$$
\begin{equation*}
\int_{\mathbb{R}^{n}}|T f(x)|^{p} w(x) d x \leq C \int_{\mathbb{R}^{n}} M f(x)^{p} w(x) d x \tag{C}
\end{equation*}
$$

for any $0<p<\infty$ and $w \in A_{\infty}$. Thus, we can get, for instance, that $T$ is bounded on $L^{p}(w)$ for $w \in A_{p}, p>1$. Similar estimates hold replacing the $L^{p}(w)$ norms in both sides by the weak norms in $L^{p, \infty}(w)$ which, for $p=1$, yields that $T: L^{1}(w) \longrightarrow L^{1, \infty}(w)$ for $w \in A_{1}$. Coifman proved $(C)$ by establishing a good $-\lambda$ inequality relating $T$ and $M$. There is another approach using the sharp maximal function (see [1] for details of this technique). Recall that

$$
M^{\#} f(x)=\sup _{Q \ni x} \frac{1}{|Q|} \int_{Q}\left|f(x)-f_{Q}\right| d x,
$$

where $f_{Q}$ stands for the average of $f$ over $Q$, and that

$$
M_{\delta}^{\#} f(x)=M^{\#}\left(|f|^{\delta}\right)(x)^{1 / \delta}
$$

Then, for $T$ with $K \in H_{\infty}^{*}$ we have the pointwise estimate $M_{\delta}^{\#}(T f)(x) \leq$ $C_{\delta} M f(x), 0<\delta<1$. This fact plus Fefferman-Stein inequality for $M$ and $M^{\#}$ (proved as well by means of a good- $\lambda$ inequality) also yield Coifman's inequality $(C)$. There is still another approach with no use at all of the good- $\lambda$ technique, this way combines ideas from [9] and [4], we will give more details later.

When $K$ is less regular, say $K \in H_{r}$ for $1<r \leq \infty$, some substitutes of $(C)$ arise. Now the operator is worse and it is expectable to get a bigger maximal function on the right hand side. Let us set $M_{q} f(x)=M\left(|f|^{q}\right)(x)^{1 / q}$ and note that $M f(x) \leq M_{q} f(x)$ for $1 \leq q<\infty$. In [12], [13] we can find the pointwise estimate $M^{\#}(T f)(x) \leq c_{r} M_{r^{\prime}} f(x)$ whenever $K \in H_{r}, 1<r<\infty$. Then, the following Coifman's type inequality holds

$$
\begin{equation*}
\int_{\mathbb{R}^{n}}|T f(x)|^{p} w(x) d x \leq C \int_{\mathbb{R}^{n}} M_{r^{\prime}} f(x)^{p} w(x) d x \tag{1}
\end{equation*}
$$

for any $0<p<\infty$ and $w \in A_{\infty}$. As a direct consequence, we have that $T$ is bounded on $L^{p}(w)$, if $w \in A_{p / r^{\prime}}$ for $r^{\prime}<p<\infty$, or if $w^{1-p^{\prime}} \in A_{p^{\prime} / r^{\prime}}$ for $1<p<r$, or if $w^{r^{\prime}} \in A_{p}$ for $1<p<\infty$. The case $p=r^{\prime}$ follows by interpolation with change of measure and by the reverse Hölder property (see [12] for more details).

When $T$ is a singular integral operator with kernel in the class $H_{\infty}$, then we get $(C)$, or what is the same, (1) with $M$ in place of $M_{r^{\prime}}$. As a consequence, $T$ is bounded on $L^{p}(w)$ for $w \in A_{p}, 1<p<\infty$. In this case the proof of $(C)$ is also obtained by proving the pointwise estimate $M_{\delta}^{\#}(T f)(x) \leq C_{\delta} M f(x), 0<\delta<1$. For more examples of this kind the reader is referred to [1]. We remark that this gives an improvement of (C) since, as we noted, $H_{\infty}^{*} \subsetneq H_{\infty}$. An explicit example can be easily adapted from the proof of Theorem 7 by taking $K=\chi_{B_{1}(0)} \in H_{\infty}$ but it is not in $H_{\infty}^{*}$.

These positive results drive us to the following questions:

- Is it possible to get similar estimates for $r=1$, in other words, what kind of weighted estimates can be proved when the kernel is in $H_{1}$ ?
- For $1<r<\infty$, can we replace $M_{r^{\prime}}$ in (1) by the pointwise smaller operator $M_{t}$ with $1 \leq t<r^{\prime}$ ?
- Is the operator $T$ bounded on $L^{p}(w)$ for every $1<p<\infty$ and for every $w \in A_{p}$ or, even more, for $w \in A_{1}$ ?

We are going to show that the answer to each of the above questions is negative.

## 3 Extrapolation for $A_{\infty}$ weights

One of the main ingredients to negatively answer the latter questions will be some extrapolation results taken from [4]. We will see that to disprove $(C)$ or (1), or their weak type-weak type analogs, it suffices to show that they fail for just one exponent $p_{0}$.

In what follows $G$ and $S$ are two operators defined on some class of smooth functions $\mathcal{S}$ such that $G f \geq 0, S f \geq 0$ for $f \in \mathcal{S}$. When we write an estimate like

$$
\begin{equation*}
\|G f\|_{L^{p}(w)} \leq C\|S f\|_{L^{p}(w)}, \tag{2}
\end{equation*}
$$

we always understand that it holds for any $f \in \mathcal{S}$ such that the left hand side is finite and that $C$ depends only upon the $A_{\infty}$ constant of $w$ and $p$. We are not assuming any linearity or sublinearity on the operators, the only thing we need is that they are reasonably defined: $G f$ and $S f$ are measurable functions for any $f \in \mathcal{S}$. Indeed, one can formulate the result in terms of pairs of functions since the operators play no role. This is the approach
used in [4] and its generality is extensively used there to deal with several implications, among them we remark those vector-valued that arise almost automatically.

Theorem 1 ([4]). Let $G, S$ be as above. Consider the following estimates:
(a) $\|G f\|_{L^{p_{0}}(w)} \leq C\|S f\|_{L^{p_{0}}(w)}$, for some $0<p_{0}<\infty$ and all $w \in A_{\infty}$.
(b) $\|G f\|_{L^{p}(w)} \leq C\|S f\|_{L^{p}(w)}$, for all $0<p<\infty$ and all $w \in A_{\infty}$.
(c) $\|G f\|_{L^{p}(w)} \leq C\|S f\|_{L^{p}(w)}$, for all $0<p<p_{0}$, for some $p_{0}$, and all $w \in A_{1}$.
(d) $\|G f\|_{L^{p_{0}, \infty}(w)} \leq C\|S f\|_{L^{p_{0}, \infty}(w)}$, for some $0<p_{0}<\infty$ and all $w \in$ $A_{\infty}$.
(e) $\|G f\|_{L^{p, \infty}(w)} \leq C\|S f\|_{L^{p, \infty}(w)}$, for some $0<p<\infty$ and all $w \in A_{\infty}$. Then,

$$
(a) \Longleftrightarrow(b) \Longleftrightarrow(c) \Longrightarrow(e) \quad \text { and } \quad(d) \Longleftrightarrow(e)
$$

The reader is referred to the original source [4] for a complete account of this technique and also for a great deal of examples that can be used to exploit the latter result.

## 4 Negative results

Now we have the ingredients needed to answer the questions posed above.
Theorem 2 ([10]). Let $1 \leq r<\infty$. There exists a singular integral operator $T$ with kernel in $H_{r}$ for which the following estimates do not hold:
(i) $\int_{\mathbb{R}^{n}}|T f(x)|^{p} w(x) d x \leq C \int_{\mathbb{R}^{n}} M_{t} f(x)^{p} w(x) d x$, for $0<p<\infty, w \in$ $A_{\infty}$ and $1 \leq t<r^{\prime}$.
(ii) $\|T f\|_{L^{p, \infty}(w)} \leq C\left\|M_{t} f\right\|_{L^{p, \infty}(w)}$, for $0<p<\infty, w \in A_{\infty}, 1 \leq t<r^{\prime}$.
(iii) $\int_{\mathbb{R}^{n}}|T f(x)|^{p} w(x) d x \leq C \int_{\mathbb{R}^{n}} M_{t} f(x)^{p} M w(x) d x$, for $0<p \leq 1$, $w$ an arbitrary weight (that is, a non-negative locally integrable function) and $1 \leq t<r^{\prime}$.
(iv) $\int_{\mathbb{R}^{n}}|T f(x)|^{p} w(x) d x \leq C \int_{\mathbb{R}^{n}}|f(x)|^{p} w(x) d x$, where, either $1<p<$ $r^{\prime}, w \in A_{1} ;$ or $1<p<\infty, w \in A_{p}$.

Remark 3. Note that for kernels satisfying just the classical Hörmander condition $\left(H_{1}\right)$, none of the maximal operators $M_{t}$ can be written in the right hand side of $(i),(i i)$ or $(i i i)$. Observe that no weighted estimate as (iv) holds even for the best class of weights $A_{1}$. In short, no weighted norm estimate is satisfied in general for operators with kernels satisfying the classical Hörmander condition $\left(H_{1}\right)$. Some other results in this direction are given in [6].

Remark 4. As we have just mentioned $\left(H_{1}\right)$ is not sufficient for showing weighted norm inequalities for $T$. However, it has recently obtained that $\left(H_{1}\right)$ yields the boundedness of the supremum of the truncated integrals, see [7].

Remark 5. The estimates in (i) say that both (1) and the pointwise estimate $M^{\#}(T f)(x) \leq c_{r} M_{r^{\prime}} f(x)$ are sharp. Note also, that in (iv) the range of exponents $1<p<r^{\prime}$ and $w \in A_{1}$ is optimal, since for $r^{\prime} \leq p<\infty$ and $w \in A_{1} \subset A_{p / r^{\prime}}, T$ is bounded on $L^{p}(w)$ as mentioned before.

Remark 6. The importance of (iii) is given by the following argument. A. Lerner has recently obtained the following estimate

$$
\int_{\mathbb{R}^{n}}|T f(x)| w(x) d x \leq C \int_{\mathbb{R}^{n}} M f(x) M w(x) d x
$$

for a singular integral operator $T$ with kernel satisfying $\left(H_{\infty}^{*}\right)$ and for any arbitrary weight $w$. His proof is not based on the good- $\lambda$ technique but uses the so called local sharp maximal function of F. John. Pushing Lerner techniques one can get the same estimate with exponents $0<p \leq 1$. Taking in particular $w \in A_{1}$ which means $M w(x) \leq C w(x)$ we get

$$
\int_{\mathbb{R}^{n}}|T f(x)|^{p} w(x) d x \leq C \int_{\mathbb{R}^{n}} M f(x)^{p} w(x) d x
$$

for any $0<p \leq 1$ and for any $w \in A_{1}$. Applying Theorem 1 to the latter estimate, which corresponds to $(c)$, we eventually get Coifman's inequality. We would like to emphasize that this combination of [9] and [4] has not used the good- $\lambda$ technique and provides a new proof of $(C)$.

The proof of Theorem 2 will be a consequence of the extrapolation technique in [4], Theorem 1 above, plus the following negative result for power weights.

Theorem 7 ([10]). Let $1 \leq r<\infty, 1 \leq p<r^{\prime},-n<\alpha<-n p / r^{\prime}$ and $w_{\alpha}(x)=|x|^{\alpha}$. There exists a singular integral operator $T$ with kernel in $H_{r}$ for which the following estimate does not hold:

$$
\begin{equation*}
\|T f\|_{L^{p, \infty}\left(w_{\alpha}\right)} \leq C\|f\|_{L^{p}\left(w_{\alpha}\right)} . \tag{3}
\end{equation*}
$$

This negative result should be compared with the following positive result: let $r, p$ be as in the theorem and let $-n p / r^{\prime}<\alpha \leq 0$, then the following estimate holds

$$
\begin{equation*}
\|T f\|_{L^{p}(w)} \leq C\|f\|_{L^{p}(w)} \tag{4}
\end{equation*}
$$

where $w(x)=|x|^{\alpha}$. This arises essentially from the results by Watson [13] using interpolation with change of measures.

Next, we are going to sketch the proof of Theorem 2 and the counterexample for Theorem 7 will be given afterwards.

Proof of Theorem 2. The estimate in (iii) with $w \in A_{1}$, that is, with $M w(x) \leq$ $C w(x)$ a.e., implies (i), since, in Theorem 1, (a) and (c) are equivalent. On the other hand, by Theorem $1,(i)$ yields (ii). So, if we show that (ii) leads to a contradiction then $(i)$ and (iii) have to failed. Furthermore, by the extrapolation result Theorem 1, it suffices to get some fixed exponent $p_{0}$ for which the weak type-weak type (ii) does not hold. Fix $1 \leq t<r^{\prime}$ and $w \in A_{1} \subset A_{\infty}$. Then we take any $p_{0}$ such that $t<p_{0}<r^{\prime}$. Assume that (ii) holds, then

$$
\|T f\|_{L^{p_{0}, \infty}(w)} \leq C\left\|M_{t} f\right\|_{L^{p_{0}, \infty}(w)} \leq C\left\|M\left(|f|^{t}\right)\right\|_{L^{\frac{p_{0}}{t}}(w)}^{\frac{1}{t}} \leq C\|f\|_{L^{p_{0}}(w)},
$$

where in the latter estimate we have used that $p_{0} / t>1$ and that $w \in A_{1}$, so that $M$ is bounded on $L^{\frac{p_{0}}{t}}(w)$. Note that this estimate says that $T$ is bounded from $L^{p_{0}}(w)$ to $L^{p_{0}, \infty}(w)$ for any $w \in A_{1}$ where $1<p_{0}<r^{\prime}$. In particular, this estimate holds for the $A_{1}$-weight $w(x)=|x|^{\alpha}$ with $-n<$ $\alpha<-n p_{0} / r^{\prime}$, contradicting Theorem 7 .

It remains to show that (iv) does not hold. When $1<p<r^{\prime}$ and $w \in A_{1}$, Theorem 7 is contradicted since the weights $w_{\alpha}$ are in $A_{1}$. In the other case, $1<p<\infty$ and $w \in A_{p}$. If the estimate holds for some $p_{0}$ and any $w \in A_{p_{0}}$ then, by the Rubio de Francia extrapolation theorem (see [5, p. 141]), the estimate will be valid for all $1<p<\infty$ and $w \in A_{p}$ which will contradict again Theorem 7.

Proof of Theorem 7. We briefly present the counterexample leaving the details to the reader, (see [10]). Let $\beta>0$ and consider the kernel $K(x)=$
$k(|x|)$ where

$$
k(t)=t^{-\frac{n}{r}}\left(\log \frac{e}{t}\right)^{-\frac{1+\beta}{r}} \chi_{(0,1)}(t)
$$

Note that $K \in L^{r}\left(\mathbb{R}^{n}\right)$. Take $0 \neq \eta \in \mathbb{R}^{n}$ far enough from the origin, for instance $|\eta|=4$. We define the kernel $\widetilde{K}(x)=K(x-\eta)$ and the operator $T$ as

$$
T f(x)=\widetilde{K} * f(x)=\int_{\mathbb{R}^{n}} K(x-\eta-y) f(y) d y
$$

Observe that $\widetilde{K} \in L^{r}\left(\mathbb{R}^{n}\right) \cap L^{1}\left(\mathbb{R}^{n}\right)$ and hence the operator $T$ is bounded on $L^{q}\left(\mathbb{R}^{n}\right)$ for every $1 \leq q \leq \infty$. Just by using that $K \in L^{r}\left(\mathbb{R}^{n}\right)$ and that it is supported in the unit ball, we can show that $\widetilde{K} \in H_{r}$ (see [10]). Note that when $r=1$, since $\widetilde{K} \in L^{1}\left(\mathbb{R}^{n}\right)$, we automatically have $\widetilde{K} \in H_{1}$. Assume that $T$ maps $L^{p}\left(w_{\alpha}\right)$ into $L^{p, \infty}\left(w_{\alpha}\right)$ and take
$0<\varepsilon<-\alpha-\frac{n}{r^{\prime}} p \quad$ and $\quad f(x)=|x+\eta|^{\frac{-n+\varepsilon}{p}} \chi_{B_{1}(-\eta)}(x) \in L^{p}\left(\mathbb{R}^{n}\right)$.
If $x \in B_{1}(-\eta)$ then $3<|x|<5$ and therefore

$$
\begin{aligned}
\sup _{\lambda>0} \lambda w\left\{x \in \mathbb{R}^{n}:|T f(x)|>\lambda\right\}^{\frac{1}{p}} & \leq C\left(\int_{\mathbb{R}^{n}}|f(x)|^{p}|x|^{\alpha} d x\right)^{\frac{1}{p}} \\
& \leq C 3^{\frac{\alpha}{p}}\left(\int_{\mathbb{R}^{n}}|f(x)|^{p} d x\right)^{\frac{1}{p}}<+\infty
\end{aligned}
$$

The contradiction arises here because one can show that the left hand side of this estimate is infinity.

Acknowledgements. The author would like to thank Prof. C. Pérez for the opportunity of giving the talk contained in this paper, for his hospitality and for many interesting discussions.

## References

[1] J. Álvarez, C. Pérez, Estimates with $A_{\infty}$ weights for various singular integral operators, Boll. Un. Mat. Ital. A (7) 8 (1994), no. 1, 123-133.
[2] A.P. Calderón, A. Zygmund, On the existence of certain singular integrals, Acta Math. 88 (1952), 85-139.
[3] R. Coifman, Distribution function inequalities for singular integrals, Proc. Acad. Sci. U.S.A. 69 (1972), 2838-2839.
[4] D. Cruz-Uribe, J.M. Martell, C. Pérez, Extrapolation results for $A_{\infty}$ weights and applications, to appear in Journal of Functional Analysis.
[5] J. Duoandikoetxea, Fourier Analysis, American Math. Soc., Grad. Stud. Math. 29, Providence, RI, 2000.
[6] S. Hofmann, Singular integrals with power weights, Proc. Amer. Math. Soc. 110 (1990), no. 2, 343-353.
[7] L. Grafakos, Estimates for maximal singular integrals, to appear in Colloq. Math.
[8] D.S. Kurtz, R.L. Wheeden, Results on weighted norm inequalities for multipliers, Trans. Amer. Math. Soc. 255 (1979), 343-362.
[9] A.K. Lerner, Weighted norm inequalities for the local sharp maximal function, to appear in J. Fourier Anal. Appl.
[10] J.M. Martell, C. Pérez and R. Trujillo-González, Lack of weighted estimates for some singular integral operators, to appear in Trans. Amer. Math. Soc.
[11] B. Muckenhoupt, Weighted norm inequalities for the Hardy maximal function, Trans. Amer. Math. Soc. 165 (1972), 207-226.
[12] J.L. Rubio de Francia, F.J. Ruiz, J.L. Torrea, Calderón-Zygmund theory for operator-valued kernels, Adv. Math. 62 (1986), 7-48.
[13] D.K. Watson, Weighted estimates for singular integrals via Fourier transform estimates, Duke Math. J. 60 (1990), no. 2, 389-399.

## Author's address:

José María Martell<br>Departamento de Matemáticas, C-XV<br>Universidad Autónoma de Madrid<br>28049 Madrid, Spain<br>E-mail: chema.martell@uam.es

# Weighted norm inequalities and extrapolation 

José María Martell *


#### Abstract

We presente a very general extrapolation principle for weights in the classes of Muckenhoupt which provides a method to obtain weighted norm inequalities in Lebesgue and more general function spaces, and also weighted modular inequalities. Vector-valued estimates are derived almost automatically. We will exploit this technique paying special attention to operators that are controlled in weighted Lebesgue spaces by the Hardy-Littlewood maximal function or, more in general, by its iterations. This is the case for regular Calderón-Zygmund operators and their commutators with bounded mean oscillation functions. We will show that these operators behave as the corresponding maximal operator that controls them. Some of the results we will present are in collaboration papers with David Cruz-Uribe and Carlos Pérez, also with Guillermo Curbera, José García-Cuerva and Carlos Pérez.


## 1 Introduction.

We start by introducing some of the needed background. Consider the Hardy-Littlewood maximal function in $\mathbb{R}^{n}$ defined as

$$
M f(x)=\sup _{Q \ni x} \frac{1}{|Q|} \int_{Q}|f(y)| d y
$$

where the cubes $Q \subset \mathbb{R}^{n}$ are always considered with their sides parallel to the coordinate axes. This operator is bounded on $L^{p}$ for every $1<p \leq \infty$ and it maps $L^{1}$ into $L^{1, \infty}$. One can change the underlying measure in the Lebesgue spaces by introducing a weight $w$, which is a non-negative

[^6]$$
\left(\frac{1}{|Q|} \int_{Q} w(x) d x\right)\left(\frac{1}{|Q|} \int_{Q} w(x)^{1-p^{\prime}} d x\right)^{p-1} \leq C
$$
when $1<p<\infty$, and, for $p=1$,
$$
\frac{1}{|Q|} \int_{Q} w(x) d x \leq C w(x), \quad \text { for a.e. } x \in Q
$$

This latter condition can be rewritten in terms of the Hardy-Littlewood maximal function: $w \in A_{1}$ if and only if $M w(x) \leq C w(x)$ for a.e. $x \in \mathbb{R}^{n}$. The class $A_{\infty}$ is defined as $A_{\infty}=\bigcup_{p \geq 1} A_{p}$.

Muckenhoupt in [13] proved that the weighted norm inequalities of the Hardy-Littlewood maximal function are characterized by the classes $A_{p}$, namely, $M$ maps $L^{1}(w)$ into $L^{1, \infty}(w)$ if and only if $w \in A_{1}$ and $M$ is bounded on $L^{p}(w), 1<p<\infty$, if and only if $w \in A_{p}$.

Let $T$ be an operator which is defined on some class of nice functions $\mathcal{D}_{T}$. Let us point out that nothing else is assumed on $T$, in particular, $T$ does not have to be linear or quasilinear. We assume that there exists $0<p_{0}<\infty$ such that $M$ controls $T$ on $L^{p_{0}}(w)$ for all $w \in A_{\infty}$, that is, for all $w \in A_{\infty}$

$$
\begin{equation*}
\int_{\mathbb{R}^{n}}|T f(x)|^{p_{0}} w(x) d x \leq C \int_{\mathbb{R}^{n}} M f(x)^{p_{0}} w(x) d x, \quad f \in \mathcal{D}_{T} \tag{1}
\end{equation*}
$$

whenever the left-hand side is finite. The aim of this paper is to show that from this assumption one can prove that $T$ satisfies weighted norm inequalities on Lebesgue spaces and function spaces, and weighted modular inequalities as $M$ does. Besides, all these estimates admit vector-valued extensions. In other words we are able to show that most of the weighted estimates that $M$ satisfies can be proved for $T$. We also see that similar results are obtained when the operator $T$ is controlled by a given iteration of the Hardy-Littlewood maximal function. We will apply the results obtained to Calderón-Zygmund operators with standard kernel which are controlled by $M$ (see Coifman's estimate (11)). We will also consider the commutators of these operators with bounded mean oscillation functions. In this case, the appropriate operators to be written in the right-hand side are the iterations of the Hardy-Littlewood maximal functions.

To work with this kind of estimates we collect the extrapolation results obtained in [4] and [5]. Before that, we introduce some notation: as mentioned, there is no assumption on the operator $T$ and in (1) one can replace $M$ by any other given operator. In fact, the operators do not need to appear explicitly and one can work with pairs of functions. In what follows, $\mathcal{F}$ is a family of ordered pairs of non-negative measurable functions $(f, g)$. If we say that for some $p_{0}, 0<p_{0}<\infty$, and $w \in A_{\infty}$,

$$
\begin{equation*}
\int_{\mathbb{R}^{n}} f(x)^{p_{0}} w(x) d x \leq C \int_{\mathbb{R}^{n}} g(x)^{p_{0}} w(x) d x, \quad(f, g) \in \mathcal{F} \tag{2}
\end{equation*}
$$

we always mean that (2) holds for any $(f, g) \in \mathcal{F}$ such that the left hand side is finite, and that the constant $C$ depends only upon $p$ and the $A_{\infty}$ constant of $w$. We will make similar abbreviated statements involving other function norms or quasi-norms, or even modular type estimates; they will be always interpreted in the same way. Note that using this notation, (1) is (2) with $\mathcal{F}$ consisting of the pairs $(|T f|, M f)$ for $f \in \mathcal{D}_{T}$.

In [4] it is shown that starting from (2) one can extrapolate and the same estimate holds for the full range of exponents $0<p<\infty$ and for all $w \in A_{\infty}$. In that paper it is also proved that the spaces $L^{p}(w)$ can be replaced by the Lorentz spaces $L^{p, q}(w)$ for all $0<p<\infty$ and $0<q \leq \infty$. This was generalized in [5] obtaining that (2) implies estimates on very general rearrangement invariant quasi-Banach function spaces (RIQBFS in the sequel) and also very general weighted modular inequalities. Furthermore, the fact that one can work with general families $\mathcal{F}$ allows one to prove, in an almost automatic way, that all these estimates extend to sequencevalued functions. The next result collects all these extrapolation results. The needed background is collected in Section 2.

Theorem 1 ([4], [5]). Let $\mathcal{F}$ be a family of ordered pairs of non-negative, measurable functions $(f, g)$. Assume that there exists $0<p_{0}<\infty$ such that

$$
\begin{equation*}
\int_{\mathbb{R}^{n}} f(x)^{p_{0}} w(x) d x \leq C \int_{\mathbb{R}^{n}} g(x)^{p_{0}} w(x) d x, \quad(f, g) \in \mathcal{F} \tag{3}
\end{equation*}
$$

for all $w \in A_{\infty}$ and whenever the left-hand side is finite. Then, for all $(f, g) \in \mathcal{F}$ and all $\left\{\left(f_{j}, g_{j}\right)\right\}_{j} \subset \mathcal{F}$ we have the following estimates:
(a) Lebesgue spaces, [4]: For all $0<p, q<\infty$ and $w \in A_{\infty}$,

$$
\|f\|_{L^{p}(w)} \leq C\|g\|_{L^{p}(w)}, \quad\left\|\left(\sum_{j}\left(f_{j}\right)^{q}\right)^{\frac{1}{q}}\right\|_{L^{p}(w)} \leq C\left\|\left(\sum_{j}\left(g_{j}\right)^{q}\right)^{\frac{1}{q}}\right\|_{L^{p}(w)}
$$

(b) Rearrangement invariant quasi-Banach function spaces, [5]: Let $\mathbb{X}$ be a RIQBFS such that $\mathbb{X}$ is $p$-convex for some $0<p \leq 1-$ equivalently $\mathbb{X}^{r}$ is Banach for some $r \geq 1$ - and with upper Boyd index $q_{\mathbb{X}}<\infty$. Then for all $0<q<\infty$ and $w \in A_{\infty}$ we have

$$
\|f\|_{\mathbb{X}(w)} \leq C\|g\|_{\mathbb{X}(w)}, \quad\left\|\left(\sum_{j}\left(f_{j}\right)^{q}\right)^{\frac{1}{q}}\right\|_{\mathbb{X}(w)} \leq C\left\|\left(\sum_{j}\left(g_{j}\right)^{q}\right)^{\frac{1}{q}}\right\|_{\mathbb{X}(w)}
$$

(c) Modular inequalities, [5]: Let $\phi \in \Phi$ with $\phi \in \Delta_{2}$ and suppose that there exist some exponents $0<r, s<\infty$ such that $\phi\left(t^{r}\right)^{s}$ is quasiconvex. Then for all $0<q<\infty$ and all $w \in A_{\infty}$,

$$
\begin{aligned}
\int_{\mathbb{R}^{n}} \phi(f(x)) w(x) d x & \leq C \int_{\mathbb{R}^{n}} \phi(g(x)) w(x) d x \\
\int_{\mathbb{R}^{n}} \phi\left(\left(\sum_{j} f_{j}(x)^{q}\right)^{\frac{1}{q}}\right) w(x) d x & \leq C \int_{\mathbb{R}^{n}} \phi\left(\left(\sum_{j} g_{j}(x)^{q}\right)^{\frac{1}{q}}\right) w(x) d x
\end{aligned}
$$

Furthermore, for $\mathbb{X}$ as before one can also get that $\phi(f)$ is controlled by $\phi(g)$ on $\mathbb{X}(w)$. In particular, taking $\mathbb{X}=L^{1, \infty}$ we have the following weak-type modular inequalities

$$
\begin{aligned}
\sup _{\lambda} \phi(\lambda) w\{x: f(x)>\lambda\} & \leq C \sup _{\lambda} \phi(\lambda) w\{x: g(x)>\lambda\}, \\
\sup _{\lambda} \phi(\lambda) w\left\{x:\left(\sum_{j} f_{j}(x)^{q}\right)^{\frac{1}{q}}>\lambda\right\} & \leq C \int_{\mathbb{R}^{n}} \phi\left(\left(\sum_{j} g_{j}(x)^{q}\right)^{\frac{1}{q}}\right) w(x) d x,
\end{aligned}
$$

for all $w \in A_{\infty}$.
We will use this result starting with (1) which will allow us to obtain inequalities for $T$ using those that are known for $M$. The advantage of this method is that once (1) is known, no property of $T$ is used and everything reduces to prove estimates for $M$.

The plan of the paper is as follows. The next section is devoted to introduce the needed background. In Section 3 we study those operators that satisfy (1): we will present a collection of weighted estimates for the Hardy-Littlewood maximal function to show that $T$ behaves in the same way. Finally, in Section 4 we consider operators with a higher degree of singularity in the sense that the operator appearing in the right hand side of (1) is an iteration of the Hardy-Littlewood maximal function. We will establish weighted estimates for $T$ as a consequence of the extrapolation results. We will pay special attention to those estimates near $L^{1}$.

## 2 Preliminaries

In this section we present the needed background.

### 2.1 Basics on Function Spaces

We collect several basic facts about rearrangement invariant quasi-Banach function spaces (RIQBFS). We start with the Banach case. For a complete account the reader is referred to [1]. Let $(\Omega, \Sigma, \mu)$ be a $\sigma$-finite non-atomic measure space. We write $\mathcal{M}$ for the set of measurable functions and $\mathcal{M}^{+}$ for the non-negative ones. Given a Banach function norm $\rho$ we the Banach function space $\mathbb{X}=\mathbb{X}(\rho)$ as

$$
\mathbb{X}=\left\{f \in \mathcal{M}:\|f\|_{\mathbb{X}}=\rho(|f|)<\infty\right\}
$$

The associate space of $\mathbb{X}$ is the space $\mathbb{X}^{\prime}$ defined by the Banach function norm $\rho^{\prime}$ :

$$
\rho^{\prime}(f)=\sup \left\{\int_{\Omega} f g d \mu: g \in \mathcal{M}^{+}, \rho(g) \leq 1\right\} .
$$

Note that, by definition, it follows that for all $f \in \mathbb{X}, g \in \mathbb{X}^{\prime}$ the following generalized Hölder's inequality holds:

$$
\int_{\Omega}|f g| d \mu \leq\|f\|_{\mathbb{X}}\|g\|_{\mathbb{X}^{\prime}}
$$

The distribution function $\mu_{f}$ of a measurable function $f$ is

$$
\mu_{f}(\lambda)=\mu\{x \in \Omega:|f(x)|>\lambda\}, \quad \lambda \geq 0 .
$$

A Banach function space $\mathbb{X}$ is rearrangement invariant if $\rho(f)=\rho(g)$ for every pair of functions $f, g$ which are equimeasurable, that is, $\mu_{f}=\mu_{g}$. In this case, we say that the Banach function space $\mathbb{X}=\mathbb{X}(\rho)$ is rearrangement invariant. It follows that $\mathbb{X}^{\prime}$ is also rearrangement invariant. The decreasing rearrangement of $f$ is the function $f^{*}$ defined on $[0, \infty)$ by

$$
f^{*}(t)=\inf \left\{\lambda \geq 0: \mu_{f}(\lambda) \leq t\right\}, \quad t \geq 0
$$

The main property of $f^{*}$ is that it is equimeasurable with $f$, that is,

$$
\mu\{x \in \Omega:|f(x)|>\lambda\}=\left|\left\{t \in \mathbb{R}^{+}: f^{*}(t)>\lambda\right\}\right| .
$$

This allows one to obtain a representation of $\mathbb{X}$ on the measure space $\left(\mathbb{R}^{+}, d t\right)$. That is, there exists a RIBFS $\overline{\mathbb{X}}$ over $\left(\mathbb{R}^{+}, d t\right)$ such that $f \in \mathbb{X}$ if and only if
$f^{*} \in \overline{\mathbb{X}}$, and in this case $\|f\|_{\mathbb{X}}=\left\|f^{*}\right\|_{\overline{\mathbb{X}}}$ (Luxemburg's representation theorem, see $[1$, p. 62]). Furthermore, the associate space $\mathbb{X}$ ' of $\mathbb{X}$ is represented in the same way by the associate space $\overline{\mathbb{X}^{\prime}}$ of $\overline{\mathbb{X}}$, and so $\|f\|_{\mathbb{X}^{\prime}}=\left\|f^{*}\right\|_{\overline{\mathbb{X}}^{\prime}}$.

From now on let $\mathbb{X}$ be rearrangement invariant Banach function spaces (RIBFS) in $\left(\mathbb{R}^{n}, d x\right)$ and let $\overline{\mathbb{X}}$ be its corresponding RIBFS in $\left(\mathbb{R}^{+}, t\right)$.

Next, we define the Boyd indices of $\mathbb{X}$, which are closely related to some interpolation properties, see [1, Ch. 3] for more details. First we introduce the dilation operator

$$
D_{t} f(s)=f(s / t), \quad 0<t<\infty, \quad f \in \overline{\mathbb{X}},
$$

and its norm $h_{\mathbb{X}}(t)=\left\|D_{t}\right\|_{\mathcal{B}(\overline{\mathbb{X}})}$ where $\mathcal{B}(\overline{\mathbb{X}})$ denotes the space of bounded linear operators on $\overline{\mathbb{X}}$. Then, the lower and upper Boyd indices are defined respectively by

$$
\begin{aligned}
& p_{\mathbb{X}}=\lim _{t \rightarrow \infty} \frac{\log t}{\log h_{\mathbb{X}}(t)}=\sup _{1<t<\infty} \frac{\log t}{\log h_{\mathbb{X}}(t)}, \\
& q_{\mathbb{X}}=\lim _{t \rightarrow 0^{+}} \frac{\log t}{\log h_{\mathbb{X}}(t)}=\inf _{0<t<1} \frac{\log t}{\log h_{\mathbb{X}}(t)}
\end{aligned}
$$

We have that $1 \leq p_{\mathbb{X}} \leq q_{\mathbb{X}} \leq \infty$. The relationship between the Boyd indices of $\mathbb{X}$ and $\mathbb{X}^{\prime}$ is the following: $p_{\mathbb{X}^{\prime}}=\left(q_{\mathbb{X}}\right)^{\prime}$ and $q_{\mathbb{X}^{\prime}}=\left(p_{\mathbb{X}}\right)^{\prime}$, where, as usual, $p$ and $p^{\prime}$ are conjugate exponents.

Take $w$ an $A_{\infty}$-weight on $\mathbb{R}^{n}$. We use the standard notation $w(E)=$ $\int_{E} w(x) d x$. The distribution function and the decreasing rearrangement with respect to $w$ are given by

$$
w_{f}(\lambda)=w\left\{x \in \mathbb{R}^{n}:|f(x)|>\lambda\right\} ; \quad f_{w}^{*}(t)=\inf \left\{\lambda \geq 0: w_{f}(\lambda) \leq t\right\} .
$$

We define the weighted version of the space $\mathbb{X}$ :

$$
\mathbb{X}(w)=\left\{f \in \mathcal{M}:\left\|f_{w}^{*}\right\|_{\overline{\mathbb{X}}}<\infty\right\}
$$

and the norm associated to it $\|f\|_{\mathbb{X}(w)}=\left\|f_{w}^{*}\right\|_{\overline{\mathbb{X}}}$. By construction $\mathbb{X}(w)$ is a Banach function space built over $\mathcal{M}\left(\mathbb{R}^{n}, w(x) d x\right)$. By doing the same procedure with the associate spaces we can see that the associate space $\mathbb{X}(w)^{\prime}$ coincides with the weighted space $\mathbb{X}^{\prime}(w)$.

Given a Banach function space $\mathbb{X}$, for each $0<r<\infty$, as in [7], we define

$$
\mathbb{X}^{r}=\left\{f \in \mathcal{M}:|f|^{r} \in \mathbb{X}\right\}=\left\{f \in \mathcal{M}:\|f\|_{\mathbb{X}^{r}}=\left\||f|^{r}\right\|_{\mathbb{X}}^{\frac{1}{\mathbb{X}}}\right\}
$$

Note that this notation is natural for the Lebesgue spaces since $L^{r}$ coincides with $\left(L^{1}\right)^{r}$. If $\mathbb{X}$ is a RIBFS and $r \geq 1$ then, $\mathbb{X}^{r}$ still is a RIBFS but, in general, for $0<r<1$, the space $\mathbb{X}^{r}$ is not necessarily Banach. Note that in the same way we can also define powers of weighted spaces and we have $(\mathbb{X}(w))^{r}=\mathbb{X}^{r}(w)$.

In this paper we work with spaces $\mathbb{X}$ so that $\mathbb{X}=\mathbb{Y}^{s}$ for some RIBFS $\mathbb{Y}$ and some $0<s<\infty$. The space $\mathbb{X}$ is in particular a rearrangement quasiBanach space (RIQBFS in the sequel), see [6] or [12] for more details. Let us observe that another equivalent approach consists in introducing first the quasi-Banach case and then one restricts the attention to those RIQBFS for which a large power is a Banach space. This latter property turns out to be equivalent to the fact that the RIQBFS $\mathbb{X}$ is $p$-convex for some $0<p \leq 1$, that is, there exists $C$ such that for all $N \geq 1$ and $f_{1}, \cdots, f_{N} \in \mathbb{X}$, all

$$
\left\|\left(\sum_{j=1}^{N}\left|f_{j}\right|^{p}\right)^{\frac{1}{p}}\right\|_{\mathbb{X}} \leq C\left(\sum_{j=1}^{N}\left\|f_{j}\right\|_{\mathbb{X}}^{p}\right)^{\frac{1}{p}} .
$$

In this case, after renorming if necessary, one has that $\mathbb{X}^{\frac{1}{p}}$ is a RIBFS.
Regarding the statement of Theorem 1 we have to make several remarks.
Remark 2. Note that in (b) of Theorem 1 we have restricted ourselves to the case of $\mathbb{X} p$-convex with $q_{\mathbb{X}}<\infty$. As we have just mentioned, this means that $\mathbb{X}^{r}$ is a Banach space (with $r=1 / p$ ). Thus, by Lorentz-Shimogaki's theorem (see [11], [16] and [1, p. 54]) $q_{\mathbb{X}}<\infty$ is equivalent to the boundedness of the Hardy-Littlewood maximal function on $\left(\mathbb{X}^{r}\right)^{\prime}$.

Remark 3. Theorem 1 part (b) can be equivalently formulated in terms of RIBFS rather than quasi-Banach spaces. The conclusion would be as follows:

Then, for all RIBFS $\mathbb{X}$ such that $q_{\mathbb{X}}<\infty$-or equivalently, that the Hardy-Littlewood maximal function is bounded on $\mathbb{X}^{\prime}$-, all $p$ such that $0<p<\infty$, and all $w \in A_{\infty}$, we have

$$
\|f\|_{\mathbb{X}^{p}(w)} \leq C\|g\|_{\mathbb{X}^{p}(w)}, \quad(f, g) \in \mathcal{F},
$$

and the corresponding vector-valued inequalities also hold.
The equivalence is based on the fact that if $\mathbb{Y}=\mathbb{X}^{r}$ then $q_{\mathbb{Y}}=r \cdot q_{\mathbb{X}}$.
Remark 4. The formulation given in (b) of Theorem 1 and the equivalent one presented in the previous remark reflect that there are two different
points of view: suppose that one wants to get estimates in $L^{\frac{1}{2}}$. The first formulation consists in looking at the RIQBFS $\mathbb{X}=L^{\frac{1}{2}}$ which has the property that $\mathbb{X}^{2}=L^{1}$ is a Banach space. This convexity allows us to apply Theorem 1 to $\mathbb{X}$. Alternatively one can start from $\mathbb{X}=L^{1}$ which is a RIBFS and by the second formulation get estimates in $\mathbb{X}^{p}$ for all $0<p<\infty$, and in particular in $\mathbb{X}^{\frac{1}{2}}=L^{\frac{1}{2}}$.

Some examples of RIQBFS are Lebesgue spaces, classical Lorentz spaces, Lorentz $\Lambda$-spaces, Orlicz spaces, Marcinkiewicz spaces, etc, see [5] for more details. In some of these examples, the Boyd indices can be computed very easily, for instance if $\mathbb{X}$ is $L^{p}, L^{p, q}, L^{p}(\log L)^{\alpha}$ or $L^{p, q}(\log L)^{\alpha}$ (where $0<p<\infty, 0<q \leq \infty, \alpha \in \mathbb{R})$ then $p_{\mathbb{X}}=q_{\mathbb{X}}=p$. In this cases, it is easy to compute the powers of $\mathbb{X}$ and one obtains

$$
\left(L^{p, q}\right)^{r}=L^{p r, q r}, \quad\left(L^{p, q}(\log L)^{\alpha}\right)^{r}=L^{p r, q r}(\log L)^{\alpha},
$$

note the same applies to $L^{p}=L^{p, p}$ and $L^{p}(\log L)^{\alpha}=L^{p, p}(\log L)^{\alpha}$.

### 2.2 Basics on modular inequalities

We introduce some notation, the terminology used is taken from [9] and [15]. Let $\Phi$ be the set of functions $\phi:[0, \infty) \longrightarrow[0, \infty)$ which are nonnegative, increasing and such that $\phi\left(0^{+}\right)=0$ and $\phi(\infty)=\infty$. If $\phi \in \Phi$ is convex we say that $\phi$ is a Young function. An $N$-function (from nice Young function) $\phi$ is a Young function such that

$$
\lim _{t \rightarrow 0^{+}} \frac{\phi(t)}{t}=0 \quad \text { and } \quad \lim _{t \rightarrow \infty} \frac{\phi(t)}{t}=\infty
$$

The function $\phi \in \Phi$ is said to be quasi-convex if there exists a convex function $\widetilde{\phi}$ and $a \geq 1$ such that

$$
\begin{equation*}
\widetilde{\phi}(t) \leq \phi(t) \leq a \widetilde{\phi}(a t), \quad t \geq 0 . \tag{4}
\end{equation*}
$$

We say that $\phi \in \Phi$ satisfies the $\Delta_{2}$ condition, we will write $\phi \in \Delta_{2}$, if $\phi$ is doubling, that is, if

$$
\phi(2 t) \leq C \phi(t), \quad t \geq 0 .
$$

Given $\phi \in \Phi$ we define the complementary function $\bar{\phi}$ by

$$
\bar{\phi}(s)=\sup _{t>0}\{s t-\phi(t)\}, \quad s \geq 0 .
$$

By definition we have Young's inequality

$$
\begin{equation*}
s t \leq \phi(s)+\bar{\phi}(t), \quad s, t \geq 0 . \tag{5}
\end{equation*}
$$

When $\phi$ is an $N$-function, then $\bar{\phi}$ is an $N$-function too, and we have the following

$$
\begin{equation*}
t \leq \phi^{-1}(t) \bar{\phi}^{-1}(t) \leq 2 t, \quad t \geq 0 \tag{6}
\end{equation*}
$$

The lower and upper dilation indices of $\phi \in \Phi$ are defined respectively by
$i_{\phi}=\lim _{t \rightarrow 0^{+}} \frac{\log h_{\phi}(t)}{\log t}=\sup _{0<t<1} \frac{\log h_{\phi}(t)}{\log t}, I_{\phi}=\lim _{t \rightarrow \infty} \frac{\log h_{\phi}(t)}{\log t}=\inf _{1<t<\infty} \frac{\log h_{\phi}(t)}{\log t}$,
where

$$
h_{\phi}(t)=\sup _{s>0} \frac{\phi(s t)}{\phi(s)}, \quad t>0
$$

see [10] and [9]. Observe that $0 \leq i_{\phi} \leq I_{\phi} \leq \infty$. It is easy to see that if $\phi$ is quasi-convex, then $i_{\phi} \geq 1$. If $\phi$ is an $N$-function, then we have that the indices for $\phi$ and $\bar{\phi}$ satisfy the following: $i_{\bar{\phi}}=\left(I_{\phi}\right)^{\prime}$ and $I_{\bar{\phi}}=\left(i_{\phi}\right)^{\prime}$.

These indices give, among other things, information about the growth of $\phi$ in terms of power functions. Indeed, if $0<i_{\phi} \leq I_{\phi}<\infty$, given $\varepsilon$ small enough, we have for all $t \geq 0$

$$
\begin{aligned}
& \phi(\lambda t) \leq C_{\varepsilon} \lambda^{I_{\phi}+\varepsilon} \phi(t), \quad \text { for } \quad \lambda \geq 1, \\
& \phi(\lambda t) \leq C_{\varepsilon} \lambda^{i_{\phi}-\varepsilon} \phi(t), \quad \text { for } \quad \lambda \leq 1 .
\end{aligned}
$$

It is clear then, that $\phi \in \Delta_{2}$ if and only if $I_{\phi}<\infty$.
Remark 5. We would like to stress the analogy between the hypotheses of Theorem 1 parts (b) and (c). The facts that $\mathbb{X}^{r}$ is Banach for some $r \geq 1$ and $\phi\left(t^{r}\right)^{s}$ is quasi-convex for some $0<r, s<\infty$ play the same role. Indeed in the proofs these properties are used to ensure the existence of a dual space and a complementary function which allow one to perform a duality argument in both cases. On the other hand, in (b) one assumes that $q_{\mathrm{X}}<\infty$ and in (c) it is supposed that $\phi \in \Delta_{2}$ which, as mentioned, means $I_{\phi}<\infty$. So, in both cases, we are assuming the finiteness of the upper indices. In the proofs, these conditions are used to assure that the Hardy-Littlewood maximal function is bounded on the dual of $\mathbb{X}^{r}$ and also it satisfies a modular inequality with respect to the complementary function of $\phi\left(t^{r}\right)^{s}$.

Remark 6. As in Remark 3, one can reformulate part (c) in Theorem 1 in the following way: one can start with an $N$-function $\phi$ such that $I_{\phi}<\infty$, or
equivalently, $M$ satisfies a modular inequality with respect to $\bar{\phi}$, and then get weighted modular inequalities with respect to the functions $\phi\left(t^{r}\right)^{s}$ for all $0<r, s<\infty$.

Some examples to whom these results can be applied are $\phi(t)=t^{p}$, $\phi(t)=t^{p}\left(1+\log ^{+} t\right)^{\alpha}, \phi(t)=t^{p}\left(1+\log ^{+} t\right)^{\alpha}\left(1+\log ^{+} \log ^{+} t\right)^{\beta}$ with $0<p<$ $\infty$ and $\alpha, \beta \in \mathbb{R}$. In all these cases one can see that $i_{\phi}=I_{\phi}=p$ and also that $\phi\left(t^{r}\right)$ is quasi-convex for $r$ large enough.

## 3 Operators controlled by the Hardy-Littlewood maximal function

We are going to apply Theorem 1 to (1) in order to get all those inequalities for the pairs $(|T f|, M f)$. Then next goal consists in proving weighted norm inequalities for $T$ as a consequence of the ones known for $M$.

We already know that $M$ maps $L^{p}(w)$ into $L^{p}(w)$ for all $w \in A_{p}, 1<p<$ $\infty$, and $L^{1, \infty}(w)$ into $L^{1}(w)$ for all $w \in A_{1}$. Regarding the vector-valued inequalities it is also known that $M$ satisfies the corresponding $\ell^{q}$-valued weighted estimates for $1<q<\infty$. In order to show that $M$ satisfies vectorvalued estimates on RIQBFS or of modular type we will use the following inequality, see [5]: if $1<q<\infty$, we have for all $0<p<\infty$, and all $w \in A_{\infty}$

$$
\begin{equation*}
\left\|\left(\sum_{j}\left(M f_{j}\right)^{q}\right)^{\frac{1}{q}}\right\|_{L^{p}(w)} \leq C\left\|M\left(\left(\sum_{j}\left|f_{j}\right|^{q}\right)^{\frac{1}{q}}\right)\right\|_{L^{p}(w)} \tag{7}
\end{equation*}
$$

This allows us to use Theorem 1 with the pairs given by this estimate and therefore the vector-valued inequalities for $M$ follows from its scalar estimates. Next, we collect the weighted vector-valued inequalities obtained for $M$ by this method:

Theorem 7. Let $\mathbb{X}$ be a RIQBFS which is $p$-convex for some $p>0$ and let $\phi \in \Phi$ be a quasi-convex function.
(i) If $1<p_{\mathbb{X}} \leq \infty$, for all $w \in A_{p_{\mathbb{X}}}$ we have

$$
\begin{equation*}
\|M f\|_{\mathbb{X}(w)} \leq C\|f\|_{\mathbb{X}(w)} \tag{8}
\end{equation*}
$$

(ii) If $1<p_{\mathbb{X}} \leq q_{\mathbb{X}}<\infty$ we have that for all $1<q<\infty$ and for all $w \in A_{p_{\mathbb{X}}}$, $M$ satisfies the following weighted vector-valued inequality

$$
\begin{equation*}
\left\|\left(\sum_{j}\left(M f_{j}\right)^{q}\right)^{\frac{1}{q}}\right\|_{\mathbb{X}(w)} \leq C\left\|\left(\sum_{j}\left|f_{j}\right|^{q}\right)^{\frac{1}{q}}\right\|_{\mathbb{X}(w)} \tag{9}
\end{equation*}
$$

(iii) For all $w \in A_{i_{\phi}}$,

$$
\begin{aligned}
\int_{\mathbb{R}^{n}} \phi(M f(x)) w(x) d x & \leq C \int_{\mathbb{R}^{n}} \phi(C|f(x)|) w(x) d x, \quad \text { if } 1<i_{\phi} \leq \infty \\
\sup _{\lambda} \phi(\lambda) w\{x: M f(x)>\lambda\} & \leq C \int_{\mathbb{R}^{n}} \phi(C|f(x)|) w(x) d x, \quad \text { if } i_{\phi}=1
\end{aligned}
$$

(iv) If $\phi \in \Delta_{2}$ (or, what is the same, $\left.I_{\phi}<\infty\right)$, for all $1<q<\infty$, $M$ satisfies the following vector-valued weighted modular inequalities: for all $w \in A_{i_{\phi}}$,
$\int_{\mathbb{R}^{n}} \phi\left(\left(\sum_{j} M f_{j}(x)^{q}\right)^{\frac{1}{q}}\right) w(x) d x \leq C \int_{\mathbb{R}^{n}} \phi\left(\left(\sum_{j}\left|f_{j}(x)\right|^{q}\right)^{\frac{1}{q}}\right) w(x) d x$
if $1<i_{\phi}<\infty$, and if $i_{\phi}=1$ we have the weak-type modular inequality
$\sup _{\lambda} \phi(\lambda) w\left\{x:\left(\sum_{j} M f_{j}(x)^{q}\right)^{\frac{1}{q}}>\lambda\right\} \leq C \int_{\mathbb{R}^{n}} \phi\left(\left(\sum_{j}\left|f_{j}(x)\right|^{q}\right)^{\frac{1}{q}}\right) w(x) d x$.

Remark 8. This result can be seen as an extension of the classical Theorem of Lorentz-Shimogaki (see [11], [16] and [1, p. 54]) which states that the Hardy-Littlewood maximal function is bounded on a RIBFS $\mathbb{X}$ if and only if $p_{\mathbb{X}}>1$. Note that Theorem 7 contains weighted, vector-valued and modular extensions of this result.

Remark 9. As in Remark 5 one can see the analogy between the hypotheses of parts $(i)$, (ii) and respectively (iii) and (iv). Note, for instance, that we have obtained weighted vector-valued inequalities for $M$ on $\mathbb{X}$ provided $1<p_{\mathbb{X}} \leq q_{\mathbb{X}}<\infty$ and $w \in A_{p_{\mathbb{X}}}$. Analogously, $M$ satisfies strong weighted modular inequalities with respecto to $\phi$ whenever $1<i_{\phi} \leq I_{\phi}<\infty$. Note that this same comment applies to Corollaries 10 and 13 below.

The proof of Theorem 7 can be found in [5]. The first part is obtained directly, while (ii) follows by $(i)$ and by extrapolation applying (b) in Theorem 1 to (7). Part ( $i i i$ ) can be proved directly using the convexity properties of $\phi$. This inequality was first consider in [8] under slightly hypotheses, see also [9]. Part (iv) can be shown as before from (iii) and by applying (c) in Theorem 1 to (7). Similar results are proved by different methods in [9].

The following result shows that if $T$ satisfies (1), then $T$ behaves as the Hardy-Littlewood maximal function in terms of the weighted estimates.

Corollary 10. Let $T$ be an operator defined in some class of nice functions $\mathcal{D}_{T}$. Assume that there is $0<p_{0}<\infty$ such that

$$
\begin{equation*}
\int_{\mathbb{R}^{n}}|T f(x)|^{p_{0}} w(x) d x \leq C \int_{\mathbb{R}^{n}} M f(x)^{p_{0}} w(x) d x, \quad f \in \mathcal{D}_{T} \tag{10}
\end{equation*}
$$

for all $w \in A_{\infty}$ and whenever the left-hand side is finite. Then the pairs $(|T f|, M f)$, for $f \in \mathcal{D}_{T}$, satisfy all the estimates contained in Theorem 1. Hence, for all $1<p, q<\infty$ and all $w \in A_{p}$

$$
\|T f\|_{L^{p}(w)} \leq C\|f\|_{L^{p}(w)}, \quad\left\|\left(\sum_{j}\left|T f_{j}\right|^{q}\right)^{\frac{1}{q}}\right\|_{L^{p}(w)} \leq C\left\|\left(\sum_{j}\left|f_{j}\right|^{q}\right)^{\frac{1}{q}}\right\|_{L^{p}(w)}
$$

If $w \in A_{1}$ and $1<q<\infty$ we have

$$
\|T f\|_{L^{1, \infty}(w)} \leq C\|f\|_{L^{1}(w)},\left\|\left(\sum_{j}\left|T f_{j}\right|^{q}\right)^{\frac{1}{q}}\right\|_{L^{1, \infty}(w)} \leq C\left\|\left(\sum_{j}\left|f_{j}\right|^{q}\right)^{\frac{1}{q}}\right\|_{L^{1}(w)}
$$

Furthermore, let $\mathbb{X}$ be a RIQBFS such that $\mathbb{X}$ is $p$-convex for some $0<p<1$ and such that $1<p_{\mathbb{X}} \leq q_{\mathbb{X}}<\infty$. Then, $T$ satisfies (8) and (9). On the other hand, let $\phi \in \Phi$ be a quasi-convex function such that $\phi \in \Delta_{2}$, (or, what is the same, $\left.I_{\phi}<\infty\right)$. Then, $T$ satisfies the weighted modular inequalities contained in (iii) and (iv) of Theorem 7.

Remark 11. This result extends the classical Theorem of Boyd (see [2] and [1, p. 154]) on which it is obtained that the Hilbert transform is bounded on a RIBFS $\mathbb{X}$ if and only if $1<p_{\mathbb{X}} \leq q_{\mathbb{X}}<\infty$. As we see below, Coifman's inequality (11) implies that the Hilbert transform satisfies (10) and so all the weighted estimates in Corollary 10 hold. Furthermore, any CalderónZygmund operator can be controlled by the Hardy-Littlewood maximal function (see (11) below) and therefore we obtain this family of weighted estimates for this class of operators. Thus, we are extending Boyd's result in the way that the class of operators is wider, we get weighted estimates, modular inequalities and also all of them admit vector-valued versions.

Remark 12. In addition to the previous remark, notice that Corollary 10 can be applied to operators which are not necessarily linear or quasilinear, this means that the general interpolation results can not be used. Thus, it is not clear how to get estimates on RIQBFS following the classical ways (see [1]). The idea behind this latter comment is that estimates for $T$ are proved through $M$, for which classical interpolation results can be employed. On the other hand, it should be pointed out that it is not clear how to interpolate between estimates like (10) -even if the operator $T$ is linear- since $M$ appears in the right-hand side.

Corollary 10 follows directly from Theorem 1 applied to (10) and then by using the weighted estimates for the Hardy-Littlewood maximal function contained in Theorem 7. For the estimates in $L^{1, \infty}$ one can apply (b) in Theorem 1 with $\mathbb{X}=L^{1, \infty}$ and then employ the well known weak type vector-valued inequalities for $M$. Another possible way consists in taking $\phi(\lambda)=\lambda$ for which $i_{\phi}=1$ and then one can use (iii) and (iv) in Theorem 7 with $T$ in place of $M$.

The main example of operators satisfying (10) is given by CalderónZygmund operators $T$ which are bounded linear operators on $L^{2}$ such that

$$
T f(x)=\int_{\mathbb{R}^{n}} K(x, y) f(y) d y, \quad \text { for a.e. } x \notin \operatorname{supp} f
$$

where the kernel $K$ satisfies the standard estimates

$$
|K(x, y)| \leq \frac{A}{|x-y|^{n}}
$$

and

$$
\left|K(x, y)-K\left(x, y^{\prime}\right)\right|+\left|K(y, x)-K\left(y^{\prime}, x\right)\right| \leq A \frac{\left|y-y^{\prime}\right|^{\tau}}{|x-y|^{n+\tau}}, \quad|x-y|>2\left|y-y^{\prime}\right|
$$

for some $A, \tau>0$. These operators satisfy Coifman's inequality, see [3]:

$$
\begin{equation*}
\int_{\mathbb{R}^{n}}|T f(x)|^{p} w(x) d x \leq C \int_{\mathbb{R}^{n}} M f(x)^{p} w(x) d x \tag{11}
\end{equation*}
$$

for all $0<p<\infty$ and all $w \in A_{\infty}$ and all $f \in C_{0}^{\infty}$ such that the left hand-side is finite. This means that we can apply Corollary 10 obtaining all the weighted estimates contained there.

## 4 Operators controlled by iterations of the HardyLittlewood maximal function

In this section we consider operators that are controlled by iterations of the Hardy-Littlewood maximal function. Suppose that we have as before some operator $T$ defined in $\mathcal{D}_{T}$ such that there exists $0<p_{0}<\infty$ and for all $w \in A_{\infty}$,

$$
\begin{equation*}
\int_{\mathbb{R}^{n}}|T f(x)|^{p_{0}} w(x) d x \leq C \int_{\mathbb{R}^{n}} M^{m+1} f(x)^{p_{0}} w(x) d x, \quad f \in \mathcal{D}_{T} \tag{12}
\end{equation*}
$$

whenever the left-hand side is finite and where $M^{m+1}$ is the Hardy-Littlewood maximal operator iterated $m+1$-times with $m \geq 1$ (note that the case $m=0$ was considered in the previous section). As done before, this implies that $T$ is controlled by $M^{m+1}$ in all the senses of Theorem 1 . Note that in terms of weighted estimates, $M^{m+1}$ and $M$ behave in the same way provided the space is not "close" to $L^{1}$, that is, $M^{m+1}$ satisfies all the estimates in Theorem 7 but the weak-type modular estimates in (iii) and (iv). This implies some of the inequalities in Corollary 10 but one has to be careful at the end-point $p=1$. Let us first state the result that one can get as a direct consequence of the extrapolation technique and we will study later the issues with the end-point estimates.

Corollary 13. Let $T$ be an operator defined in some class of nice functions $\mathcal{D}_{T}$. Assume that there are an integer $m \geq 1$ and $0<p_{0}<\infty$ such that for all $w \in A_{\infty}$

$$
\begin{equation*}
\int_{\mathbb{R}^{n}}|T f(x)|^{p_{0}} w(x) d x \leq C \int_{\mathbb{R}^{n}} M^{m+1} f(x)^{p_{0}} w(x) d x, \quad f \in \mathcal{D}_{T} \tag{13}
\end{equation*}
$$

whenever the left-hand side is finite. Then the pairs $\left(|T f|, M^{m+1} f\right)$, for $f \in \mathcal{D}_{T}$, satisfy all the estimates contained in Theorem 1. Hence, for all $1<p, q<\infty$ and all $w \in A_{p}$

$$
\|T f\|_{L^{p}(w)} \leq C\|f\|_{L^{p}(w)}, \quad\left\|\left(\sum_{j}\left|T f_{j}\right|^{q}\right)^{\frac{1}{q}}\right\|_{L^{p}(w)} \leq C\left\|\left(\sum_{j}\left|f_{j}\right|^{q}\right)^{\frac{1}{q}}\right\|_{L^{p}(w)}
$$

Furthermore, let $\mathbb{X}$ be a RIQBFS such that $\mathbb{X}$ is $p$-convex for some $0<p<1$ and such that $1<p_{\mathbb{X}} \leq q_{\mathbb{X}}<\infty$. Then, $T$ satisfies (8) and (9). On the other hand, let $\phi \in \Phi$ be a quasi-convex function such that $\phi \in \Delta_{2}$, (or, what is the same, $\left.I_{\phi}<\infty\right)$. Then, if $1<i_{\phi}<\infty, T$ satisfies the first estimate in (iii) and the first estimate in (iv) of Theorem 7.

To prove this result we first observe that $M^{m+1}$ satisfy all these estimates since $M$ does. Then, using Theorem 1 as in the previous section the proof is completed.

We now study the behavior of $M^{m+1}$ near $L^{1}$ to eventually show that $T$ satisfies the same estimates. In terms of RIQBFS the natural end-point estimate for the Hardy-Littlewood maximal inequality is the boundedness of $M$ from $L^{1}$ to $L^{1, \infty}$ which turns out to be also a weak-type modular inequality. To find the natural spaces and modular inequalities for $M^{m+1}$, we first consider the function

$$
\varphi_{m}(t)=\frac{t}{\left(1+\log ^{+} t\right)^{m}}, \quad t>0
$$

which is increasing, quasi-concave (that is, $\varphi_{m}(t) / t$ is decreasing) and satisfies that $\varphi_{m}\left(0^{+}\right)=0$. We can define the Marcinkiewicz type space $\widetilde{\mathbb{M}}_{\varphi_{m}}$ by the quasi-norm

$$
\|f\|_{\tilde{\mathbb{M}}_{\varphi_{m}}}=\sup _{t} \varphi_{m}(t) f^{*}(t) .
$$

Thus $\mathbb{X}=\widetilde{\mathbb{M}}_{\varphi_{m}}$ is a RIQBFS such that $\mathbb{X}^{r}$ is a Banach space for any $r>1$ and $p_{\mathbb{X}}=q_{\mathbb{X}}=1$, see [5]. We note that this allows us to use Theorem 1 with $\mathbb{X}$. This space plays the role of $L^{1, \infty}$ as we see below.

To deal with the modular inequalities we introduce the function

$$
\psi_{m}(t)=t\left(1+\log ^{+} t\right)^{m}, \quad t>0
$$

Note that $\psi$ is an increasing convex function with $\psi\left(0^{+}\right)=0$ and $\psi \in \Delta_{2}$.
For $M^{k+1}$ we have the following end-point estimates:
Proposition 14. Let $m \geq 1$. Then,

$$
M^{m+1}: L(\log L)^{m} \longrightarrow \tilde{\mathbb{M}}_{\varphi_{m}}
$$

and

$$
\left|\left\{x \in \mathbb{R}^{n}: M^{m+1} f(x)>\lambda\right\}\right| \leq C \int_{\mathbb{R}^{n}} \psi_{m}\left(\frac{|f(x)|}{\lambda}\right) d x
$$

Furthermore, for any $w \in A_{1}$ we have the weighted estimates

$$
M^{m+1}: L(\log L)^{m}(w) \longrightarrow \widetilde{\mathbb{M}}_{\varphi_{m}}(w)
$$

and

$$
w\left\{x \in \mathbb{R}^{n}: M^{m+1} f(x)>\lambda\right\} \leq C \int_{\mathbb{R}^{n}} \psi_{m}\left(\frac{|f(x)|}{\lambda}\right) w(x) d x .
$$

These estimates are the analogs in terms of RIQBFS and modular inequalities of the weak type $(1,1)$ of $M$. As before, we can show that the operator $T$ satisfies the same estimates.

Corollary 15. Let $T$ be an operator as in Corollary 13 satisfying (13). Then, for all $w \in A_{1}$

$$
T: L(\log L)^{m}(w) \longrightarrow \widetilde{\mathbb{M}}_{\varphi_{m}}(w)
$$

and

$$
w\left\{x \in \mathbb{R}^{n}:|T f(x)|>\lambda\right\} \leq C \int_{\mathbb{R}^{n}} \psi_{m}\left(\frac{|f(x)|}{\lambda}\right) w(x) d x
$$

To prove the first estimate we only need to apply Theorem 1 , part (b), with the pairs $\left(|T f|, M^{m+1} f\right)$ for $f \in \mathcal{D}_{T}$ and $\mathbb{X}=\widetilde{\mathbb{M}}_{\varphi_{m}}$, and then Proposition 14. Note that as mentioned $\mathbb{X}$ is a RIQBFS with the property that $\mathbb{X}^{r}$ is Banach for every $r>1$ and also $p_{\mathbb{X}}=q_{\mathbb{X}}=1$. Observe that the class of weights $A_{1}$ is natural since $p_{\mathbb{X}}=1$.

The modular inequality is not so automatic. Define the function $\phi_{m}(t)=$ $\frac{1}{\psi_{m}(1 / t)}$ and observe that $\phi_{m} \in \Phi$ is such that $\phi_{m} \in \Delta_{2}$ (indeed, $i_{\phi_{m}}=I_{\phi_{m}}=$ 1) and $\phi\left(t^{r}\right)$ is quasi-convex for some large $r$. Then, we can apply ( $c$ ) in Theorem 1 with $\phi_{m}$ and Proposition 14 to obtain

$$
\begin{align*}
w\left\{x \in \mathbb{R}^{n}:|T f(x)|>1\right\} & \leq \sup _{t} \phi_{m}(t) w\left\{x \in \mathbb{R}^{n}:|T f(x)|>t\right\} \\
& \leq C \sup _{t} \phi_{m}(t) w\left\{x \in \mathbb{R}^{n}: M^{m+1} f(x)>t\right\} \\
& \leq C \sup _{t} \phi_{m}(t) \int_{\mathbb{R}^{n}} \psi_{m}\left(\frac{|f(x)|}{t}\right) w(x) d x \\
& \leq C \sup _{t} \phi_{m}(t) \psi_{m}\left(\frac{1}{t}\right) \int_{\mathbb{R}^{n}} \psi_{m}(|f(x)|) w(x) d x \\
& \leq C \int_{\mathbb{R}^{n}} \psi_{m}(|f(x)|) w(x) d x \tag{14}
\end{align*}
$$

where we have used that $\psi_{m}$ is submultiplicative. If the operator $T$ is linear, (14) implies the desired estimate by homogeneity. Otherwise, we observe that we have proved this estimate starting from (13) which for any $\lambda>0$ implies

$$
\int_{\mathbb{R}^{n}}\left(\frac{|T f(x)|}{\lambda}\right)^{p_{0}} w(x) d x \leq C \int_{\mathbb{R}^{n}}\left(\frac{M^{m+1} f(x)}{\lambda}\right)^{p_{0}} w(x) d x, \quad f \in \mathcal{D}_{T}
$$

with $C$ independent of $\lambda$. This induces a new family of pairs of functions given by $\left(|T f| / \lambda, M^{m+1} f / \lambda\right)$ to whom we can apply (14) to conclude as desired

$$
w\left\{x \in \mathbb{R}^{n}:|T f(x)|>\lambda\right\} \leq C \int_{\mathbb{R}^{n}} \psi_{m}\left(\frac{|f(x)|}{\lambda}\right) w(x) d x
$$

where $C$ does not depend on $\lambda>0$.
The main example of operators satisfying (13) is given by the commutators of Calderón-Zygmund operators with bounded mean oscillation functions. Let $T$ be a Calderón-Zygmund operator with standard kernel as before. Let $b$ be a function of bounded mean oscillation, that is,

$$
\sup _{Q} \frac{1}{|Q|} \int_{Q}\left|b(x)-b_{Q}\right| d x<\infty
$$

where $b_{Q}$ stands for the average of $b$ on $Q$. Then we define the first order commutator

$$
C_{b}^{1} f(x)=[b, T] f(x)=b(x) T f(x)-T(b f)(x)
$$

and for $m \geq 2$, the $m$-order commutator $C_{b}^{m} f(x)=\left[b, C_{b}^{m-1}\right] f(x)$. In this way we have

$$
C_{b}^{m} f(x)=\int_{\mathbb{R}^{n}}(b(x)-b(y))^{m} K(x, y) f(y), \quad \text { for a.e. } x \notin \operatorname{supp} f .
$$

Note that this definition makes sense for $m \geq 0$ and the commutator of order 0 is nothing but $T$. The maximal operator that controls the commutator $C_{b}^{m}$ is $M^{m+1}$ which is the Hardy-Littlewood maximal function iterated $m+1$ times, namely, in [14] it is shown that

$$
\begin{equation*}
\int_{\mathbb{R}^{n}}\left|C_{b}^{m} f(x)\right|^{p} w(x) d x \leq C \int_{\mathbb{R}^{n}} M^{m+1} f(x)^{p} w(x) d x \tag{15}
\end{equation*}
$$

for every $0<p<\infty$ and $w \in A_{\infty}$ and all $f \in C_{0}^{\infty}$ such that the left handside is finite. Thus, Corollaries 13 and 15 can be applied and we obtain all those weighted estimates for $C_{b}^{m}$.

## Acknowledgement

This paper is based on a talk given in the Departamento de Análisis Matemático de la Universidad de Málaga in January 2004. The author would like to thank Prof. F. Martín-Reyes and Prof. M. Lorente for the invitation and their hospitality.

This paper has intended to present some of the most interesting results contained in [4] and [5]. The author would like to express his deep gratitude to the coauthors D. Cruz-Uribe, G.P. Curbera, J. García-Cuerva and C. Pérez, for many interesting discussions and all what was learnt from them.

## References

[1] C. Bennett and R. Sharpley, Interpolation of Operators, Academic Press, New York, 1988.
[2] D.W. Boyd, The Hilbert transform on rearrangement-invariant spaces, Canad. J. Math. 19 (1967), 599-616.
[3] R.R. Coifman, Distribution function inequalities for singular integrals, Proc. Acad. Sci. U.S.A. 69 (1972), 2838-2839.
[4] D. Cruz-Uribe, J.M. Martell and C. Pérez, Extrapolation results for $A_{\infty}$ weights and applications, to appear in J. Funct. Anal.
[5] G.P. Curbera, J. García-Cuerva, J.M. Martell and C. Pérez, Extrapolation with Weights, Rearrangement Invariant Function Spaces, Modular inequalities and applications to Singular Integrals, Preprint 2004.
[6] L. Grafakos and N.J. Kalton, Some remarks on multilinear maps and interpolation, Math. Ann. 319 (2001), 151-180.
[7] W.B. Johnson and G. Schechtman, Sums of independent random variables in rearrangement invariant function spaces, Ann. Probab. 17 (1989), 789-808.
[8] R.A. Kerman and R.A. Torchinsky, Integral inequalities with weights for the Hardy maximal function, Studia Math. 71 (1982), 277-284.
[9] V. Kokilashvili and M. Krbec, Weighted Inequalities in Lorentz and Orlicz Spaces, World Scientific, 1991.
[10] S.G. Krein, Ju.I. Petunin and E.M. Semenov, Interpolation of Linear Operators, Amer. Math. Soc., Rhode Island, 1982.
[11] G.G. Lorentz, Majorants in spaces of integrable functions, Amer. J. Math. 77 (1955), 484-492.
[12] S.J. Montgomery-Smith, The Hardy operator and Boyd indices, Interaction between functional analysis, harmonic analysis, and probability (Columbia, MO, 1994), 359-364, Lecture Notes in Pure and Appl. Math. 175, Dekker, New York, 1996.
[13] B. Muckenhoupt, Weighted norm inequalities for the Hardy maximal function, Trans. Amer. Math. Soc. 165 (1972), 207-226.
[14] C. Pérez, Sharp estimates for commutators of singular integrals via iterations of the Hardy-Littlewood maximal function, J. Fourier Anal. Appl. 3 (1997), 743-756.
[15] M.M. Rao and Z.D. Ren, Theory of Orlicz spaces, Monographs and Textbooks in Pure and Applied Mathematics 146, Marcel Dekker, Inc., New York, 1991.
[16] T. Shimogaki, Hardy-Littlewood majorants in function spaces, J. Math. Soc. Japan 17 (1965), 365-373.

## Author's address:

José María Martell<br>Departamento de Matemáticas, C-XV<br>Universidad Autónoma de Madrid<br>28049 Madrid, Spain<br>E-mail: chema.martell@uam.es

# Littlewood-Paley-Stein theory for semigroups and its applications to the characterization of Banach spaces 

Teresa Martinez


#### Abstract

We study a generalization of the theory of Littlewood-Paley for semigroups acting on $L^{p}$-spaces of functions with values in uniformly convex or uniformly smooth Banach spaces. We characterize, in the vector-valued context, the validity of the boundedness inequalities for the generalized Littlewood-Paley $g$-function, defined for the subordinated Poisson semigroup of a symmetric diffusion semigroup, in terms of the type and cotype properties of the underlying Banach space. We see that in the case of the classical Poisson semigroups (in the torus and in the Euclidean setting), this theory is more satisfactory and easiest to handle, due to the application of the theory of vector-valued Calderón-Zygmund singular integrals.


## 1 Introduction

In these notes we give an overview of the results contained in [12] and [19], and specially of the main techniques involved in their proofs. In fact, we are not only interested in the study of the geometrical properties of Banach spaces, but also in the connection and use of the different tools that will appear: spectral theory, semigroup theory, probability theory, real variable.

Let us recall some very well known facts. Denote by $\mathbb{T}=[-\pi, \pi]$ the torus, and let $f$ be a function in $L^{1}(\mathbb{T})$. For notation simplicity, let $f$ also be its harmonic extension to the whole disc:

$$
f\left(r e^{i \theta}\right)=P_{r} * f(\theta)
$$

where

$$
P_{r}(\theta)=\frac{1}{2 \pi} \frac{1-r^{2}}{1+r^{2}-2 r \cos \theta}
$$

2000 Mathematics Subject Classification. 46B20; 42B25,42A61
Key words and phrases. Littlewood-Paley theory, semigroups, uniformly convex or smooth Banach spaces, vector-valued Calderón-Zygmund operators
is the Poisson kernel for the disc. The classical Littlewood-Paley $g$-function is defined, for $f \in L^{p}(\mathbb{T}), 1 \leq p \leq \infty$ as

$$
G f(\theta)=\left(\int_{0}^{1}\left\|(1-r) \nabla P_{r} * f(\theta)\right\|^{2} \frac{d r}{1-r}\right)^{1 / 2}
$$

where

$$
\begin{equation*}
\left\|\nabla P_{r} * f(\theta)\right\|=\left\|\left(\frac{\partial P_{r}}{\partial r}, \frac{1}{r} \frac{\partial P_{r}}{\partial \theta}\right)\right\|_{\ell^{2}}=\left(\left|\frac{\partial P_{r}}{\partial r} * f(\theta)\right|^{2}+\left|\frac{1}{r} \frac{\partial P_{r}}{\partial \theta} * f(\theta)\right|^{2}\right)^{1 / 2} \tag{1}
\end{equation*}
$$

It is a classical fact that for every $p \in(1, \infty)$, there exist constants $c$ and $C$ depending only on $p$ such that

$$
\begin{equation*}
c\|f\|_{L^{p}(\mathbb{T})} \leq|\hat{f}(0)|+\|G f\|_{L^{p}(\mathbb{T})} \leq C\|f\|_{L^{p}(\mathbb{T})} \tag{2}
\end{equation*}
$$

If instead of considering scalar-valued functions, we deal with functions taking values in a Banach space $\mathcal{B}$, the definition of the $g$-function given above is still valid, just replacing absolute values by norms in $\mathcal{B}$ in (1). In this case, it is also very well known (see [8] and [15]), that the equivalence (2) holds if and only if the Banach space is isomorphic to a Hilbert space. However, one of the inequalities can still hold in non Hilbertian spaces. Thus, when talking about vector-valued functions, we will be interested in just one of the inequalities in (2). The results mentioned in this notes are a consequence of the tight relationship between vector-valued Harmonic Analysis and the Geometry of Banach spaces. The validity of an inequality for vector-valued functions often gives a new characterization of a known property of the spaces, or introduces a new class of them.

## 2 Lusin type and cotype properties, their connection with Probability: martingale type and cotype properties

Let us define, for $q \in[1, \infty)$ and $f \in L_{\mathcal{B}}^{1}(\mathbb{T})$, the classical space of Bochner integrable functions on $\mathbb{T}$, the generalized "Littlewood-Paley $g$-function" as

$$
\begin{equation*}
G_{q} f(z)=\left(\int_{0}^{1}(1-r)^{q}\left\|\nabla P_{r} * f(z)\right\|_{\mathcal{B}}^{q} \frac{d r}{1-r}\right)^{1 / q} \tag{3}
\end{equation*}
$$

Then $\mathcal{B}$ is said to be of Lusin cotype $q$ (resp. Lusin type $q$ ) if there exist $p \in(1, \infty)$ and a positive constant $C$ such that

$$
\left\|G_{q} f\right\|_{L^{p}(\mathbb{T})} \leq C\|f\|_{L_{\mathcal{B}}^{p}(\mathbb{T})} \quad\left(\text { resp. }\|f\|_{L_{\mathcal{B}}^{p}(\mathbb{T})} \leq C\left(\|\hat{f}(0)\|_{\mathcal{B}}+\left\|G_{q} f\right\|_{L^{p}(\mathbb{T})}\right)\right)
$$

It is not difficult to see that if $\mathcal{B}$ is of Lusin cotype $q$ (resp. Lusin type $q$ ), then $2 \leq q \leq \infty$ (resp. $1 \leq q \leq 2$ ). In these notes, we will be mainly interested in the first inequality. In [12] and [19] the case of the Lusin type property is also treated.

It is proved in [19] that the definition of cotype property above is independent of $p$, that is, if one of the inequalities above holds for one $p \in(1, \infty)$, then so does it for every $p \in(1, \infty)$ (with a different constant depending on $p)$. Also, it is shown to be equivalent to the boundedness of $G_{q}$ from $L_{\mathcal{B}}^{1}(\mathbb{T})$ into $L^{1, \infty}(\mathbb{T})$, and from $H_{\mathcal{B}}^{1}(\mathbb{T})$ into $L^{1}(\mathbb{T})$. The main result of [19] states in particular that a Banach space $\mathcal{B}$ is of Lusin cotype $q$ if and only if $\mathcal{B}$ is of martingale cotype $q$ (as defined bellow).

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\left\{\mathcal{F}_{n}\right\}_{n \geq 1}$ be a non-decreasing sequence of sub- $\sigma$-fields of $\mathcal{F}$ such that $\mathcal{F}=\sigma\left(\cup \mathcal{F}_{n}\right)$ (such a sequence will be called a stochastic basis). Given a Banach space $\mathcal{B}$, a sequence $f=$ $\left\{f_{n}\right\}_{n \geq 1}$ of $\mathcal{B}$-valued random variables is a $\mathcal{B}$-valued martingale relative to $\left\{\mathcal{F}_{n}\right\}$ if each $f_{n}$ is an integrable $\mathcal{F}_{n}$-measurable function and $E_{n}\left(f_{n+1}\right)=$ $E\left(f_{n+1} \mid \mathcal{F}_{n}\right)=f_{n} . E_{n}$ will denote the operator defined as the conditional expectation to the sub- $\sigma$-field $\mathcal{F}_{n}$. For every martingale $f=\left\{f_{n}\right\}_{n \geq 1}$ we shall denote $d_{k} f$ the "increments" of the martingale $f: d_{k} f=f_{k}-f_{k-1}$, $k \geq 1, f_{0}=0$, in such a way that $f_{n}=\sum_{k=1}^{n} d_{k} f, d_{k} f$ is $\mathcal{F}_{k}$-measurable, integrable and $E_{k}\left(d_{k+1} f\right)=0, k \geq 1$. For a complete account on $\mathcal{B}$-valued martingales see [4].

A Banach space $\mathcal{B}$ is said to be of martingale cotype $q, 2 \leq q<\infty$, or in short, $M$-cotype $q$, if there exist a constant $C$ such that for any $\mathcal{B}$-valued martingale $f=\left\{f_{n}\right\}$

$$
\begin{equation*}
\left\|S_{q} f\right\|_{L^{q}}=\left\|\left(\sum_{n=1}^{\infty}\left\|d_{n} f\right\|_{\mathcal{B}}^{q}\right)^{1 / q}\right\|_{L^{q}} \leq C \sup _{n \geq 1}\left\|f_{n}\right\|_{L_{\mathcal{B}}^{q}}^{q} \tag{4}
\end{equation*}
$$

Every Banach space is of $M$-cotype $q=\infty$. The definition (and the analog for martingale type property) is due to G. Pisier [13]. Non-trivial $M$-cotype $q<\infty$ is a geometrical property of the space, implying superreflexivity, and it happens (this is Pisier's renorming theorem) if and only if the space admits an equivalent uniformly convex norm with modulus of convexity of power type $q$, see [13], [14]. Pisier, see [13], proved that a space is of $M$-cotype $q$ if and only if for every (or, equivalently, for some) $p, 1 \leq p<\infty,\left\|S_{q} f\right\|_{L^{p}} \leq C\left\|f^{*}\right\|_{L^{p}}$, where $f^{*}(\omega)=\sup _{n \geq 1}\left\|f_{n}(\omega)\right\|_{\mathcal{B}}$ stands for Doob's maximal function of $f$. It is also possible to characterize $M$ cotype $q$ in terms of the boundedness of $S_{q}$ from $L_{\mathcal{B}}^{1}$ into $L^{1, \infty}$ (being these spaces the corresponding ones for martingales), $H_{\mathcal{B}}^{1}$ into $L^{1}$ and also from
$B M O$-type spaces (see also [10] and the references therein for the details).
The proof in [19] of the equivalence between martingale and Lusin cotype properties uses, on one hand, the connection between the Poisson kernel in the disc and the law of a certain random variable associated to a Brownian motion. More precisely, to prove that martingale cotype property implies Lusin cotype property, it is used that $P_{\sqrt{r}}(\theta-\phi)$ gives the law in the set $\sqrt{r} \mathbb{T}$ of $B_{\tau_{\sqrt{r}}} \mid B_{\tau_{r}}=r e^{i \theta}$, where $\left\{B_{t}\right\}_{t \geq 0}$ is a standard Brownian motion in the disc, and $\tau_{r}=\inf \left\{t>0:\left|B_{t}\right|=r\right\}$. The proof of the converse is very technical and involves a careful analysis of $G_{q} f$ for $f$ with a highly lacunar Fourier series.

An important feature of the proof of the equivalence between Lusin and martingale cotype properties is that the class of spaces remains unchanged if we substitute $G_{q}$ by the "partial" generalized $g$-functions

$$
\begin{align*}
G_{q}^{1} f(\theta) & =\left(\int_{0}^{1}\left\|(1-r) \frac{\partial P_{r}}{\partial r} * f(\theta)\right\|_{\mathcal{B}}^{q} \frac{d r}{1-r}\right)^{1 / q}  \tag{5}\\
G_{q}^{2} f(\theta) & =\left(\int_{0}^{1}\left\|\frac{1-r}{r} \frac{\partial P_{r}}{\partial \theta} * f(\theta)\right\|_{\mathcal{B}}^{q} \frac{d r}{1-r}\right)^{1 / q} \tag{6}
\end{align*}
$$

## 3 Vector-valued singular integrals

To prove the other characterizations of Lusin cotype property, namely, that the definition given above is independent of $p$, and that it is equivalent to the boundedness of $G_{q}$ from $L_{\mathcal{B}}^{1}(\mathbb{T})$ into $L^{1, \infty}(\mathbb{T})$, and from $H_{\mathcal{B}}^{1}(\mathbb{T})$ into $L^{1}(\mathbb{T})$, vector-valued Calderón-Zygmund singular integrals are used in [19]. The key point is that $G_{q}$ can be seen as the norm of a vector-valued CalderónZygmund operator. Let us first of all recall the definition of such an object (see, for example [7]).

Definition 1. Given $\mathcal{B}_{\infty}, \mathcal{B}_{\in}$ a pair of Banach spaces, let $T$ be a linear operator defined in $L_{c, \mathcal{B}_{\infty}}^{\infty}$ and taking values in the space of $\mathcal{B}_{\epsilon}$-valued and strongly measurable functions on $\mathbb{T}$ satisfying
(a) $T$ extends to a bounded operator either from $L_{\mathcal{B}_{\infty}}^{q}(\mathbb{T})$ into $L_{\mathcal{B}_{\epsilon}}^{q}(\mathbb{T})$ for some $1<q<\infty$, or from $L_{\mathcal{B}_{\infty}}^{q}(\mathbb{T})$ into $L_{\mathcal{B}_{\in}}^{q, \infty}(\mathbb{T})$,
(b) there exists a $\mathcal{L}\left(\mathcal{B}_{\infty}, \mathcal{B}_{\in}\right)$-valued measurable function $K$, defined in the complement of the diagonal in $\mathbb{T} \times \mathbb{T}$, such that for every function $f \in L_{\mathcal{B}_{\infty}}^{\infty}$,

$$
T f(\theta)=\int_{\mathbb{T}} K(\theta, \phi) f(\phi) d \phi
$$

for all $\theta$ outside the support of $f$,
(c) the function $K$ satisfies the estimates:

$$
\begin{aligned}
\|K(\theta, \phi)\| & \leq C|\theta-\phi|^{-1} \\
\left\|\partial_{\theta} K(\theta, \phi)\right\|+\left\|\partial_{\phi} K(\theta, \phi)\right\| & \leq C|\theta-\phi|^{-2}
\end{aligned}
$$

for all $(\theta, \phi), \theta \neq \phi$.
In the case of our $G_{q}$ function, we have

$$
T: L_{\mathcal{B}_{\infty}}^{\infty} \longrightarrow L_{\mathcal{B}_{\epsilon}}^{0}, \quad \mathcal{B}_{\infty}=\mathcal{B}, \quad \mathcal{B}_{\in}=\mathcal{L}_{\ell_{\mathcal{B}}}^{\mathrm{L}}\left([1, \infty], \frac{\lceil\nabla}{(\infty-\nabla)}\right)
$$

and thus

$$
\begin{equation*}
G_{q} f(\theta)=\|T f(\theta)\|_{L_{\ell_{\mathcal{B}}^{q}}^{q}\left([0,1], \frac{d r}{(1-r)}\right)} \tag{7}
\end{equation*}
$$

where

$$
T f(\theta)=(1-r)\left(\frac{\partial P_{r}}{\partial r}, \frac{1}{r} \frac{\partial P_{r}}{\partial \theta}\right) * f(\theta)
$$

We have a similar expression for $G_{q}^{1}$ and $G_{q}^{2}$. One can eventually prove that the kernels of the corresponding operators $T$ satisfy the conditions in Definition 1. For such a kind of operators, a lot is known about their boundedness properties and the next one is a cyclic theorem that is intended to collect the folklore about the boundedness properties of vector-valued Calderón-Zygmund operators. In the theorem bellow, the space $H_{\mathcal{B}}^{1}$ is defined in the atomic sense. Namely, we say that a function $a \in L_{\mathcal{B}}^{\infty}(\mathbb{T})$ is an atom if there exists an interval $I$ containing the support of $a$, and such that $\|a\|_{L_{\mathcal{B}}^{\infty}(\mathbb{T})} \leq|I|$, and $\int_{I} a(\theta) d \theta=0$. We also consider atoms without cancellation, which are simply $\mathcal{B}$-valued functions $a$ such that $\|a(x)\|_{\mathcal{B}} \leq 1$ (see [2]). Then, we say that a function $f$ is in $H_{\mathcal{B}}^{1}$, if it admits a decomposition $f=\sum_{i} \lambda_{i} a_{i}$, where $a_{i}$ are $\mathcal{B}$-valued atoms and $\sum_{i}\left|\lambda_{i}\right|<\infty$. We define $\|f\|_{H_{\mathcal{B}}^{1}}=\inf \left\{\sum_{i}\left|\lambda_{i}\right|\right\}$, where the infimum runs over all those such decompositions. Let us also recall that, given a Banach space $\mathcal{B}$ the space $\mathrm{BMO}_{\mathcal{B}}(\mathbb{T})$ is the space of $\mathcal{B}$-valued functions $f$ defined on the torus such that $\|f\|_{\mathrm{BMO}_{\mathcal{B}(\mathbb{T})}}=\sup _{I} \frac{1}{|I|} \int_{I}\left\|f(\theta)-f_{I}\right\|_{\mathcal{B}} d \theta$, where $f_{I}=\frac{1}{|I|} \int_{I} f(\theta) d \theta$ and the supremum is taken over the intervals $I \subset \mathbb{T}$.

Theorem 2. Let $T$ be a Calderón-Zygmund operator with an associated kernel $K$. Assume that there exists $\Lambda \in \mathcal{L}\left(\mathcal{B}_{\infty}, \mathcal{B}_{\epsilon}\right)$ such that for all $c \in \mathcal{B}_{\infty}$ we have $T(c)(x)=\Lambda(c)$ for almost every $x \in \mathbb{T}$. Let $S$ be defined as $S(f)=$ $\|T(f)\|_{\mathcal{B}_{\epsilon}}$. Then, the following statements are equivalent:
i) The operator $T$ maps $L_{\mathcal{B}_{\infty}}^{\infty}$ into $\mathrm{BMO}_{\mathcal{B}_{\epsilon}}$.
ii) The operator $S$ maps $L_{\mathcal{B}_{\infty}}^{\infty}$ into BMO.
iii) The operator $T$ maps $H_{\mathcal{B}_{\infty}}^{1}$ into $L_{\mathcal{B}_{\epsilon}}^{1}$. Equivalently $S$ maps $H_{\mathcal{B}_{\infty}}^{1}$ into $L^{1}$.
iv) The operator $T$ maps $L_{\mathcal{B}_{\infty}}^{p}$ into $L_{\mathcal{B}_{\epsilon}}^{p}$, for any (or equivalently, for some) $p, 1<p<\infty$. Equivalently $S$ maps $L_{\mathcal{B}_{\infty}}^{p}$ into $L^{p}$ (for any, or for some $p \in(1, \infty))$.
v) The operator $T$ maps $\mathrm{BMO}_{\mathcal{B}_{\infty}}$ into $\mathrm{BMO}_{\mathcal{B}_{\epsilon}}$.
vi) The operator $S$ maps $\mathrm{BMO}_{\mathcal{B}_{\infty}}$ into BMO .
vii) The operator $T$ maps $L_{\mathcal{B}_{\infty}}^{1}$ into $L_{\mathcal{B}_{\epsilon}}^{1, \infty}$. Equivalently, $S$ maps $L_{\mathcal{B}_{\infty}}^{1}$ into $L^{1, \infty}$.

This theorem is valid in a finite measure space. The case of infinite measure (such as $\mathbb{R}^{n}$ endowed with Lebesgue's measure) needs some modifications that we will comment on later. The condition $T(c)(x)=\Lambda(c)$ is crucial for the proof. If the operator does not verify it, it is very easy to see that the theorem is false. Consider, for instance $T f(x)=g(x) f(x)$ with $g \in L^{\infty}(\mathbb{T})$. This operator trivially sends $L_{\mathcal{B}}^{p}(\mathbb{T})$ into $L_{\mathcal{B}}^{p}(\mathbb{T})$ and it is clearly a Calderón-Zygmund operator with kernel $K(x, y)=0$. But it is known (it is a result due to Stegenga, see [16]), that to be bounded in BMO, the necessary and sufficient condition is that

$$
\frac{1}{|I|} \int_{I}\left|g(x)-g_{I}\right| d x \leq \frac{1}{\log \left(|I|^{-1}\right)} \quad \text { for every } \quad I
$$

which in particular requires that $g \in \mathrm{VMO}$.
The proof of Theorem 2 is completely standard (see, for example [7] and [9]), and the point where the condition $T(c)(x)=\Lambda(c)$ is used is to obtain that the boundedness in $L^{p}$ of the operator implies the boundedness in BMO (see also [12] for the details).

Since clearly $P_{r} * c=1$ for every constant function $c$, we have $G_{q}^{1} c=$ $G_{q}^{2} c=G_{q} c=0$, and we are in the hypothesis of the theorem. Thus, we obtain the following theorem, that extends the results in [19].

Theorem 3. Given a Banach space, $\mathcal{B}$, and $q \geq 2$, the following statements are equivalent:
i) $\mathcal{B}$ is of Lusin cotype $q$,
ii) $G_{q}$ maps $L_{\mathcal{B}}^{\infty}(\mathbb{T})$ into $\mathrm{BMO}(\mathbb{T})$ boundedly,
iii) $G_{q}$ maps $\mathrm{BMO}_{\mathcal{B}}(\mathbb{T})$ into $\mathrm{BMO}(\mathbb{T})$ boundedly,

The equivalence holds also with the same statements with $G_{q}^{1}$ or $G_{q}^{2}$ in place of $G_{q}$.

## 4 Another point of view: semigroups

In this section, we introduce another approach to the problem, by using semigroup theory. For instance, let us observe that with the change of variables $r=e^{-t}$ in the Poisson kernel of the disc, $\mathcal{P}_{t} f(\theta)=P_{e^{t}} * f(\theta)$ is (except, maybe, for a constant), is the subordinated Poisson semigroup corresponding to the Laplacian in $\mathbb{T}\left(\partial^{2} / \partial \theta^{2}\right)$. And the operator $G_{q}$ is almost the generalized $g$-function associated to the Laplacian, as will be defined bellow.

The semigroup associated to the Laplacian in $\mathbb{T}$ is an example of a symmetric diffusion semigroup. Although it is a well known concept, let us recall here its definition (see [17], [3], [6] and [20] for a general account on semigroup theory). These semigroups are collections of linear operators $\left\{\mathcal{I}_{t}\right\}_{t \geq 0}$ defined on $L^{p}(\Omega, d \mu)$ for every $p, 1 \leq p \leq \infty$, where $(\Omega, d \mu)$ is any positive measure space, satisfying the properties of a semigroup

$$
\begin{equation*}
\mathcal{T}_{0}=\mathrm{Id}, \quad \mathcal{T}_{t} \mathcal{T}_{s}=\mathcal{I}_{t+s}, \quad \lim _{t \rightarrow 0} \mathcal{T}_{t} f=f \text { in } L^{2} \text { for every } f \in L^{2}, \tag{8}
\end{equation*}
$$

together with the specific conditions of being

- contractions in all $L^{p}:\left\|\mathcal{T}_{t} f\right\|_{p} \leq\|f\|_{p}, p \in[1, \infty]$
- selfadjoint in $L^{2}, \mathcal{T}_{t}^{*}=\mathcal{T}_{t}$,
- Markovian, $\mathcal{T}_{t} 1=1$, and
- positive $\mathcal{T}_{t} f \geq 0$ if $f \geq 0$.

The infinitesimal generator of any semigroup $\mathcal{T}_{t}$ acting on some space of functions $\mathcal{X}$, is the operator $A$, defined as

$$
\lim _{t \rightarrow 0} \frac{\mathcal{T}_{t} f-f}{t}=A f
$$

for $f$ in a suitable dense class of functions (for example, the infinitely differentiable functions with compact support) in $\mathcal{X}$. If the operator $A$ is bounded in $\mathcal{X}$, the associated semigroup is given by the formal series $\mathcal{T}_{t}=e^{t A}=\sum_{n=0}^{\infty} \frac{t^{n} A^{n}}{n!}$, which in fact converges in norm. In most cases, we are interested in unbounded operators $A$ (such as the Laplacian in $L_{\mathbb{T}}^{2}$ ). The general theory on semigroups states that under certain conditions, for example if $A$ is closed, densely defined and $\left\|(\lambda-A)^{-1}\right\| \leq 1 / \lambda$ for every $\lambda>0$, we still have that $A$ is the infinitesimal generator of a semigroup. Formally, we will denote this semigroup also as $\mathcal{T}_{t}=e^{t A}$. Formally, then it holds that $\partial_{t} \mathcal{T}_{t} f=A \mathcal{T}_{t} f, \mathcal{T}_{0} f=f$ for $f$ satisfying certain conditions (this calculation can be made rigurous). For this reason, $\left\{\mathcal{I}_{t}\right\}$ is called the heat semigroup of $A$.

Let us recall some typical examples of symmetric diffusion semigroups:
Example 4. The most classical example of a symmetric diffusion semigroup is the one generated by the Laplacian in $\mathbb{R}^{n}$, $n \geq 1$, with the Lebesgue's measure. It is well known that the corresponding heat semigroup is given by the heat kernel,

$$
T_{t} f(x)=e^{t \Delta} f(x)=\frac{1}{(4 \pi t)^{n / 2}} \int_{\mathbb{R}^{n}} e^{\frac{|x-y|^{2}}{4 t}} f(y) d y
$$

Observe that the kernel of $T_{t}$ is a Gaussian density. It is very well known the connection between semigroups and Markov processes, and in fact what the expression of $T_{t}$ says is that, under certain conditions on $f$, the solution $u(t, x)$ to the equation $\partial_{t} u(t, x)=\Delta u(t, x), u(0, x)=f$ is $u(t, x)=T_{t} f(x)=$ $E^{x}\left(f\left(B_{2 t}\right)\right)$, where $\left\{B_{t}\right\}$ is a Brownian motion and the expectation $E^{x}$ is taken with respect to the law of the Brownian motion started at $x$. We will not comment further on this very interesting connection between Probability and PDE's and we refer to [5] for a detailed treatment of this topic.

Example 5. Another operator generating a symmetric diffusion semigroup (see [17]) is the Ornstein-Uhlenbeck operator, $A=\frac{1}{2} \Delta-x \cdot \nabla$, in $\left(\mathbb{R}^{n}, d \gamma(x)\right)$, $n \geq 1$, where $d \gamma(x)=\pi^{-n / 2} e^{-|x|^{2}} d x$ is the Gaussian measure. The action of this semigroup is most commonly expressed as

$$
e^{-t A} f(x)=M_{r} f(x)=\frac{1}{\pi^{n / 2}} \int_{\mathbb{R}^{n}} K_{r}(x, y) f(y) d y
$$

where $r=e^{-t}$ and

$$
K_{r}(x, y)=\frac{1}{\left(1-r^{2}\right)^{n / 2}} \exp \left(-\frac{|y-r x|^{2}}{1-r^{2}}\right), \quad 0<r<1
$$

is called the Mehler kernel.

Example 6. An example of an operator that generates a semigroup, but not a symmetric diffusion one (because it is not Markovian) is the Harmonic Oscillator, $A=\Delta-|x|^{2}$, in $\left(\mathbb{R}^{n}, d x\right)$.

From the heat semigroup associated to an operator $A$, we can define a number of semigroups, which are called subordinated semigroups. We will be interested in the Poisson subordinated semigroup, which is defined, by using spectral techniques, as

$$
\begin{equation*}
\mathcal{P}_{t} f=\frac{1}{\sqrt{\pi}} \int_{0}^{\infty} \frac{e^{-u}}{\sqrt{u}} \mathcal{T}_{t^{2} / 4 u} f d u=\frac{t}{2 \sqrt{\pi}} \int_{0}^{\infty} \frac{e^{-t^{2} / 4 u}}{u^{3 / 2}} \mathcal{T}_{u} f d u \tag{9}
\end{equation*}
$$

It is easy to see that if $\left\{\mathcal{I}_{t}\right\}$ is a symmetric diffusion semigroup, so is $\left\{\mathcal{P}_{t}\right\}$, and with some more effort (see [3]), it can be seen that if $\mathcal{T}_{t}=e^{t A}$, then $\mathcal{P}_{t}=e^{-t \sqrt{-A}}$. "Heuristically", this formula can be understood by using a well known formula for the Gamma function (see the book by Folland, Introduction to Partial Differential Equations for a proof of it)

$$
\mathcal{P}_{t}=\frac{1}{\sqrt{\pi}} \int_{0}^{\infty} \frac{e^{-u}}{\sqrt{u}} e^{\frac{t^{2}}{4 u} A} d u=\frac{1}{\sqrt{\pi}} \int_{0}^{\infty} \frac{e^{-u}}{\sqrt{u}} e^{-\beta^{2} / 4 u} d u=e^{-\beta},
$$

with $\beta^{2}=t^{2}(-A)$, which is "positive". Also, formally we can differentiate twice in the formula for $\mathcal{P}_{t}$ (this calculation can be also made rigourously for $f$ satisfying certain properties), and see that it satisfies $\partial_{t}^{2} \mathcal{P}_{t} f+A \mathcal{P}_{t} f=0$, $\mathcal{P}_{0} f=f$, that is, the Laplace equation for $A$.

Example 7. In particular, in the situation of Example 4, when $A$ is the Laplace operator on $\mathbb{R}^{n}$ with the Lebesgue's measure, $\mathcal{P}_{t} f(x)$ should be the harmonic extension to $(0, \infty) \times \mathbb{R}^{n}$ of $f$. In fact, it can be proved that

$$
P_{t} f(x)=e^{-t \sqrt{-\Delta}} f(x)=C_{n} \int_{\mathbb{R}^{n}} \frac{t}{\left(t^{2}+|x-y|^{2}\right)^{(n+1) / 2}} f(y) d y
$$

where the kernel

$$
P_{t}(x-y)=\frac{C_{n} t}{\left(t^{2}+|x-y|^{2}\right)^{(n+1) / 2}}
$$

is the Poisson kernel for the upper half space. As in the semigroup of Example 4, this kernel is a density, and the process associated to these densities is the Cauchy process, which is called the subordinated process to Brownian motion.

Let us recall that we are interested in defining a generalized $g$-function for functions with values in Banach spaces. Since all the operators $\mathcal{T}_{t}$ and $\mathcal{P}_{t}$ are positive bounded operators in $L^{p}(\Omega, d \mu)$, they have a straightforward extension to $L_{\mathcal{B}}^{p}(\Omega, d \mu)$ for every Banach space $\mathcal{B}$, with the same norm. Namely, let $f=\sum_{k=1}^{K} v_{k} \varphi_{k}$ be a function in the tensor product $\mathcal{B} \otimes \mathcal{L} \vee(\otimes, \beta,\lceil\mu)$, we define the vector valued extension of the operators as $\mathcal{T}_{t} f=\sum_{k=1}^{K} v_{k} \mathcal{T}_{t} \varphi_{k}$. This extension verifies, for $f$ in the tensor product, that $\left\|\mathcal{T}_{t} f\right\|_{L_{\mathcal{B}}^{p}} \leq\|f\|_{L_{\mathcal{B}}^{p}}$. The boundedness for all functions in $L_{\mathcal{B}}^{p}$ follows from the density of the tensor product in $L_{\mathcal{B}}^{p}$.

Given a symmetric diffusion semigroup, we define the generalized $g$ function associated to it as

$$
\begin{equation*}
\mathfrak{G}_{q}(f)(x)=\left(\int_{0}^{\infty}\left\|t \frac{\partial \mathcal{P}_{t} f}{\partial t}\right\|_{\mathcal{B}}^{q} \frac{d t}{t}\right)^{1 / q} \tag{10}
\end{equation*}
$$

And in the particular case that the symmetric diffusion semigroup is the one of Examples 4 and 7, we define

$$
\mathcal{G}_{q} f(x)=\left(\int_{0}^{\infty}\left\|t \nabla P_{t} * f(x)\right\|_{\ell_{\mathcal{B}}^{2}}^{q} \frac{d t}{t}\right)^{1 / q}
$$

where

$$
\begin{equation*}
\left\|t \nabla P_{t} * f(x)\right\|_{\ell_{\mathcal{B}}^{2}}=t\left(\left|\frac{\partial P_{t}}{\partial t} * f(x)\right|^{2}+\sum_{i=1}^{n}\left|\frac{\partial P_{t}}{\partial x_{i}} * f(x)\right|^{2}\right)^{1 / 2} \tag{11}
\end{equation*}
$$

and the "partial" generalized $g$-functions

$$
\begin{aligned}
\mathcal{G}_{q}^{1} f(x) & =\left(\int_{0}^{\infty}\left\|t \frac{\partial P_{t}}{\partial t} * f(x)\right\|_{\mathcal{B}}^{q} \frac{d t}{t}\right)^{1 / q} \\
\mathcal{G}_{q}^{2} f(x) & =\left(\int_{0}^{\infty}\left\|t \nabla_{x} P_{t} * f(x)\right\|_{\ell_{\mathcal{B}}^{2}}^{q} \frac{d t}{t}\right)^{1 / q}
\end{aligned}
$$

where

$$
\left\|t \nabla_{x} P_{t} * f(x)\right\|_{\ell_{\mathcal{B}}^{2}}=t\left(\sum_{i=1}^{n}\left|\frac{\partial P_{t}}{\partial x_{i}} * f(x)\right|^{2}\right)^{1 / 2}
$$

In the case of the Laplacian in the torus, the definition we have given in (5) is the classical one, although it does not fully coincide with the corresponding one in (10). This last one, after the change of variables $r=e^{-t}$ gives

$$
\tilde{G}_{q}^{1} f(\theta)=\left(\int_{0}^{1}\left(r \log \frac{1}{r}\right)^{q-1}\left\|\frac{\partial P_{r}}{\partial r} * f(x)\right\|_{\mathcal{B}}^{q} d r\right)^{1 / q}
$$

and it can be seen that the boundedness in $L^{p}$ of $\tilde{G}_{q}$ and $G_{q}$ are equivalent: for $r$ far away from 1 , the operators are trivially bounded in $L^{p}$, and for $r$ close to $1, r \log \frac{1}{r} \sim 1-r$ (see [12]).

The following theorem characterizes Lusin cotype property in terms of the boundedness in $L^{p}$ of $\mathfrak{G}_{q}$.

Theorem 8. Given a Banach space, $\mathcal{B}$, and $2 \leq q<\infty$, the following sentences are equivalent:
i) $\mathcal{B}$ is of Lusin cotype $q$.
ii) For every symmetric diffusion semigroup $\left\{\mathcal{T}_{t}\right\}_{t \geq 0}$ with subordinated semigroup $\left\{\mathcal{P}_{t}\right\}_{t \geq 0},\left\|\mathfrak{G}_{q} f\right\|_{L^{p}(\Omega, d \mu)} \leq C\|f\|_{L_{\mathcal{B}}^{p}(\Omega, d \mu)}$, for every (or, equivalently, for some) $p \in(1, \infty)$.

The main ideas of the proof of this theorem are contained in [17]. Proving that ii) implies i) is easy, since it is enough to consider the generalized $g$ function associated to the semigroup $\mathcal{T}_{t}=e^{t \Delta}$ in the torus, and make the change of variables $r=e^{-t}$, to get the boundedness of $G_{q}^{1}$. Let us sketch the proof of the converse implication, which is quite interesting by its own right. It involves three steps (the details can be found in [12]). In the first one, and by several changes of variables, we reduce the problem of the boundedness of $\mathfrak{G}_{q}$ to the boundedness of another " $g$-function" involving only means of the operators $\mathcal{T}_{t}$ :

$$
\mathfrak{G}_{q} f(\omega) \leq C\left(\int_{0}^{\infty}\left\|t \frac{\partial M_{t} f}{\partial t}\right\|_{\mathcal{B}}^{q} \frac{d t}{t}\right)^{1 / q}
$$

where

$$
M_{t} f(\omega)=\frac{1}{t} \int_{0}^{t} \mathcal{T}_{s} f(\omega) d s
$$

The second step takes advantage of this function. By first observing that for $t$ away from 0 and $\infty$, everything inside is real analytic, we can discretize the integrals and derivatives appearing in the former expression: for fixed $\varepsilon$ and $m$

$$
\int_{\varepsilon}^{m \varepsilon}\left\|t \frac{\partial M_{t} f}{\partial t}\right\|_{\mathcal{B}}^{q} \frac{d t}{t} \sim \sum_{n=1}^{m}(n \varepsilon)^{q-1}\left\|\left.\frac{\partial M_{t} f}{\partial t}\right|_{t=n \varepsilon}\right\|_{\mathcal{B}}^{q} \varepsilon
$$

and

$$
\left.\frac{\partial M_{t} f}{\partial t}\right|_{t=n \varepsilon} \sim \frac{M_{(n+1) \varepsilon} f-M_{n \varepsilon} f}{\varepsilon} \sim \frac{1}{(n+1) \varepsilon^{2}} \sum_{j=0}^{n} \mathcal{T}_{j \varepsilon} f \varepsilon-\frac{1}{n \varepsilon^{2}} \sum_{j=0}^{n-1} \mathcal{T}_{j \varepsilon} f \varepsilon
$$

where the $\sim$ sign means that the $L^{p}$ norms of the terms on both sides are equivalent. Thus, we get that

$$
\left\|\mathfrak{G}_{q} f\right\|_{L^{p}} \leq\left\|\left(\sum_{n=1}^{m}\left\|\frac{1}{n+1} \sum_{j=0}^{n} \mathcal{T}_{j \varepsilon} f-\frac{1}{n} \sum_{j=0}^{n-1} \mathcal{T}_{j \varepsilon} f\right\|_{\mathcal{B}}^{q}\right)^{1 / q}\right\|_{L^{p}}
$$

and our aim is to obtain that the right-hand side of the former inequality is controlled by $\|f\|_{L_{\mathcal{B}}^{p}}$ independently of $m$ and $\varepsilon$. For this, the third step identifies, in some sense, the right-hand of the former expression with the generalized square function of martingales (defined in (4)) of some martingale, and then apply that Lusin cotype and martingale cotype are equivalent properties of the space. We shall need the following result due to Rota, see Chapter V of [17]. Let $Q$ be a linear operator on $L^{p}(\Omega, \beta, d \mu)$ satisfying the axioms
i) $\|Q f\|_{L^{p}} \leq\|f\|_{L^{p}}$ for every $p, 1 \leq p \leq \infty$,
ii) $Q=Q^{*}$ in $L^{2}$,
iii) $Q f \geq 0$ for every $f \geq 0$,
iv) $Q 1=1$.

Theorem 9. For any $Q$ as above, there exist a measure space $(M, \mathcal{F}, d m)$, a collection of $\sigma$-fields $\cdots \subset \mathcal{F}_{n+1} \subset \mathcal{F}_{n} \subset \cdots \subset \mathcal{F}_{1} \subset \mathcal{F}_{0} \subset \mathcal{F}$, and another $\sigma$-algebra $\hat{\mathcal{F}} \subset \mathcal{F}$ such that
a) there exists an isomorphism $i:(\Omega, \beta, d \mu) \rightarrow(M, \hat{\mathcal{F}}, d m)$ (which induces an isomorphism between $L^{p}$ spaces, also denoted by $i, i(f)(m)=f\left(i^{-1} m\right)$ ),
b) for every $f \in L^{p}(M, \hat{\mathcal{F}}, d m)$, we have

$$
Q^{2 n}\left(i^{-1} f\right)(x)=\hat{E}\left(E_{n}(f)\right)(i x), \quad x \in \Omega
$$

where $\hat{E}(f)=E(f \mid \hat{\mathcal{F}})$ and $E_{n}(f)=E\left(f \mid \mathcal{F}_{n}\right)$.
This theorem holds in the scalar valued case. For the vector valued case, the validity of the second statement is a consequence of the extension of positive contractive operators (as in the argument of page 159). Since we have symmetric diffusion semigroups, every operator $\mathcal{T}_{\varepsilon / 2}$ with $\varepsilon>0$ verifies the hypothesis in Theorem 9, and therefore $\mathcal{T}_{j \varepsilon}=\left(\mathcal{T}_{\varepsilon / 2}\right)^{2 j}=E\left(\cdot \mid \mathcal{F}_{j}\right)$. Calling $\sigma_{n}=\frac{E_{0}+\cdots+E_{n}}{n+1}$, we have that

$$
\begin{aligned}
\left\|\mathfrak{G}_{q} f(\omega)\right\|_{L^{p}} & \leq C\left\|\left(\sum_{n=1}^{m}\left\|\frac{1}{n+1} \sum_{j=0}^{n} \mathcal{T}_{j \varepsilon} f-\frac{1}{n} \sum_{j=0}^{n-1} \mathcal{T}_{j \varepsilon} f\right\|_{\mathcal{B}}^{q}\right)^{1 / q}\right\|_{L^{p}} \\
& \leq\left\|\left(\sum_{n=0}^{\infty} n^{q-1}\left\|\left(\sigma_{n}-\sigma_{n-1}\right) f\right\|_{\mathcal{B}}^{q}\right)^{1 / q}\right\|_{L^{p}}
\end{aligned}
$$

The last expression is not yet $S_{q} f$ for some martingale $f$, but after some more tedious calculations, we can find the desired $S_{q} f$ and obtain that the last term is bounded by $C_{p, q}\|f\|_{L_{\mathcal{B}}^{p}}$.

## 5 Characterizations in $\mathbb{R}^{n}$

So far we have seen that the Lusin cotype property of a Banach space $\mathcal{B}$ can be characterized in terms of the boundedness of the generalized $g$ function of all the symmetric diffusion semigroups, and also by using just one of them, the generalized $g$-function associated to the Laplacian operator in the torus. The natural question now is wether we can find some more semigroups with the same property, and the answer is yes: the natural candidate, the generalized $g$-function associated to the Laplacian in $\mathbb{R}^{n}, \mathcal{G}_{q}$, also characterizes the Lusin cotype property, as it is stated in the following result.

Theorem 10. Given a Banach space, $\mathcal{B}$, and $q \geq 2$, the following sentences are equivalent:
i) $\mathcal{B}$ is of Lusin cotype $q$.
ii) For every, or equivalently, for some $n>1,\left\|\mathcal{G}_{q} f\right\|_{L^{p}\left(\mathbb{R}^{n}\right)} \leq C\|f\|_{L_{\mathcal{B}}^{p}\left(\mathbb{R}^{n}\right)}$ for every (or, equivalently, for some) $p \in(1, \infty)$.
iii) $\left\|\mathcal{G}_{q} f\right\|_{L^{p}(\mathbb{R})} \leq C\|f\|_{L_{\mathcal{B}}^{p}(\mathbb{R})}$ for every (or, equivalently, for some) $p \in$ $(1, \infty)$.

The equivalence of the same statements holds for $\mathcal{G}_{q}^{1}$ or $\mathcal{G}_{q}^{2}$ in place of $\mathcal{G}_{q}$.
Remark 11. It is also true that the Lusin cotype property can be characterized in terms of the Ornstein-Uhlenbeck semigroup (see Example 5), in the same way as we have seen in the former theorem for the classical heat semigroup in $\mathbb{R}^{n}$. Wether non-Markovian semigroups, as the one in Example 6, also characterize Lusin cotype property or not, is an open problem (see [12]).

Let us give the main ideas of the proof of Theorem 10 in the case of $\mathcal{G}_{q}^{1}$ (for the other operators the proof is similar). That the first statement implies the second one is just taking the semigroup associated to the Laplacian in $\mathbb{R}^{n}$ in Theorem 8. For the rest of the implications, the key point is again
considering the generalized $g$-function $\mathcal{G}_{q}^{1}$ in $\mathbb{R}^{n}, n \geq 1$, as the norm of certain vector-valued Calderón-Zygmund operator. Namely,

$$
\begin{equation*}
\mathcal{G}_{q}^{1} f(x)=\|T f(x)\|_{L_{\mathcal{B}}^{q}\left((0, \infty), \frac{d t}{t}\right)} \tag{12}
\end{equation*}
$$

where $T$ is the operator sending $\mathcal{B}$-valued functions defined on $\left(\mathbb{R}^{n}, d x\right)$ into $L_{\mathcal{B}}^{q}\left((0, \infty), \frac{d t}{t}\right)$-valued functions on $\left(\mathbb{R}^{n}, d x\right)$, given by

$$
T f(x)=t \frac{\partial P_{t}}{\partial t} * f(x)
$$

The definition of a Calderón-Zygmund operator on $\mathbb{R}^{n}$ is similar to the one given for the case of the torus, see Definition 1: given $\mathcal{B}_{\infty}, \mathcal{B}_{\in}$ a pair of Banach spaces, let $T$ be a linear operator defined in $L_{c, \mathcal{B}_{\infty}}^{\infty}$ and taking values in the space of $\mathcal{B}_{\epsilon}$-valued and strongly measurable functions on $\mathbb{R}^{n}$ such that $T$ extends to a $(q, q)$ strong or weak type operator for some $1<q<\infty$, and such that there exists a $\mathcal{L}\left(\mathcal{B}_{\infty}, \mathcal{B}_{\in}\right)$-valued measurable function $K$, defined in the complement of the diagonal in $\mathbb{R}^{n} \times \mathbb{R}^{n}$, such that for every function $f \in L_{\mathcal{B}_{\infty}}^{\infty}, T f(y)=\int_{\mathbb{R}^{n}} K(x, y) f(y) d y$, for all $x$ outside the support of $f$, satisfying the estimates

$$
\begin{aligned}
\|K(x, y)\| & \leq C|x-y|^{-n}, \\
\left\|\nabla_{x} K(x, y)\right\|+\left\|\nabla_{y} K(x, y)\right\| & \leq C|x-y|^{-n-1}, \quad \text { for all } \quad x \neq y .
\end{aligned}
$$

For these operators, Theorem 2 still holds, although in this new infinite measure setting, it requires some modifications. First, let us recall the $\mathrm{BMO}_{\mathcal{B}}$ and $H_{\mathcal{B}}^{1}$ spaces on $\mathbb{R}^{n}$. Let $\mathcal{B}$ be a Banach space. $\mathrm{BMO}_{\mathcal{B}}\left(\mathbb{R}^{n}\right)$ is the space of $\mathcal{B}$-valued functions $f$ defined on $\mathbb{R}^{n}$ such that
where $f_{Q}=\frac{1}{|Q|} \int_{Q} f(x) d x$ and the supremum is taken over the cubes $Q \subset \mathbb{R}^{n}$ with sides parallel to the axis. The space $H_{\mathcal{B}}^{1}$ is defined in the atomic sense. Namely, we say that a function $a \in L_{\mathcal{B}}^{\infty}\left(\mathbb{R}^{n}\right)$ is a $\mathcal{B}$-atom if there exists a cube $Q \subset \mathbb{R}^{n}$ containing the support of $a$, and such that $\|a\|_{L_{\mathcal{B}}^{\infty}\left(\mathbb{R}^{n}\right)} \leq|Q|^{-1}$, and $\int_{Q} a(x) d x=0$. Then, we say that a function $f$ is in $H_{\mathcal{B}}^{1}\left(\mathbb{R}^{n}\right)$ if it admits a decomposition $f=\sum_{i} \lambda_{i} a_{i}$, where $a_{i}$ are $\mathcal{B}$-valued atoms and $\sum_{i}\left|\lambda_{i}\right|<\infty$. We define $\|f\|_{H_{\mathcal{B}}^{1}}=\inf \left\{\sum_{i}\left|\lambda_{i}\right|\right\}$, where the infimum runs over all those such decompositions (see [2]).

For operators defined in $\mathbb{R}^{n}$, Theorem 2 still holds when the origin spaces $L^{\infty}$ and BMO are considered only with functions of compact support. Bennet, De Vore and Sharpley (see [1]) proved that for a function in BMO,
the most classical singular integral operator, the (non centered) maximal function is, either infinite almost everywhere, or a function in BMO. This dichotomy holds for many singular integral operators. For the ones we are handling, it is very easy to see that $\mathcal{G}_{q}^{2}\left(\chi_{(0, \infty)}\right)(x)=\infty$ a.e. $x \in \mathbb{R}$. It is tedious but not difficult to check that $T$ in (12) is a vector-valued CalderónZygmund operator in $\mathbb{R}^{n}$. Also, it is clear that for any constant $c, \mathcal{G}_{q}^{1} c(x)=0$. Thus, to prove that statement ii) in Theorem 10 implies statement iii), it is enough to see that the boundedness from $L_{c}^{\infty}\left(\mathbb{R}^{n}\right)$ into $\operatorname{BMO}\left(\mathbb{R}^{n}\right)$ of $\mathcal{G}_{q}^{1}$ for $n>1$ implies the same boundedness property of the corresponding operator for $n=1$.

To this end, consider $\tilde{x}=\left(x_{2}, \ldots, x_{n}\right) \in \mathbb{R}^{n-1}$, and $h \in L_{c, \mathcal{B}}^{\infty}(\mathbb{R})$, and define $f(x)=h\left(x_{1}\right) \chi_{[0,1]^{n-1}}(\tilde{x})$, where $x=\left(x_{1}, x_{2}, \ldots, x_{n}\right) \in \mathbb{R}^{n}$. The symmetric diffusion semigroup generated by the Laplacian on $\mathbb{R}^{n}$ is given by convolution with the Gaussian density. Then we have

$$
\begin{aligned}
\mathcal{T}_{t} f(x) & =\int_{\mathbb{R}^{n}} \frac{1}{(4 \pi t)^{n / 2}} e^{-\frac{|x-y|^{2}}{4 t}} f(y) d y \\
& =C_{0} \int_{\mathbb{R}} \frac{1}{(4 \pi t)^{1 / 2}} e^{-\frac{\left|x_{1}-y_{1}\right|^{2}}{4 t}} h\left(y_{1}\right) d y_{1}=C_{0} \mathcal{T}_{t}^{1} h\left(x_{1}\right),
\end{aligned}
$$

where $\mathcal{T}_{t}^{1}$ is the heat kernel in $\mathbb{R}$. If we denote by $\mathcal{P}_{t}^{1}$ the Poisson semigroup subordinated to $\mathcal{T}_{t}^{1}$ on $\mathbb{R}$ and by $P_{t}^{1}$ the Poisson kernel on $\mathbb{R}$, the formula of the subordinated semigroup (9) implies that $P_{t} * f(x)=\mathcal{P}_{t} f(x)=$ $C_{0} \mathcal{P}_{t}^{1} h\left(x_{1}\right)=C_{0} P_{t}^{1} * h\left(x_{1}\right)$, and therefore $\mathcal{G}_{q}^{1} f(x)=C_{0} \mathcal{G}_{q}^{1} h\left(x_{1}\right)$. Now, for every interval $I \subset \mathbb{R}$ consider $Q=I^{n}$ the cube in $\mathbb{R}^{n}$ whose sides are the interval $I$. Then,

$$
\frac{1}{|Q|} \int_{Q} \mathcal{G}_{q}^{1} f(x) d x=\frac{1}{|I|^{n}} \int_{I^{n}} C_{0} \mathcal{G}_{q}^{1} h\left(x_{1}\right) d x_{1} \ldots d x_{n}=\frac{C_{0}}{|I|} \int_{I} \mathcal{G}_{q}^{1} h\left(x_{1}\right) d x_{1}
$$

Therefore, and also by using similar arguments,

$$
\frac{1}{|Q|} \int_{Q}\left|\mathcal{G}_{q}^{1} f(x)-\left(\mathcal{G}_{q}^{1} f\right)_{Q}\right| d x=\frac{C_{0}}{|I|} \int_{I}\left|\mathcal{G}_{q}^{1} h\left(x_{1}\right)-\left(\mathcal{G}_{q}^{1} h\right)_{I}\right| d x_{1} .
$$

Hence,

$$
\left\|\mathcal{G}_{q}^{1} h\right\|_{\mathrm{BMO}_{(\mathbb{R})}}=\frac{1}{C_{0}}\left\|\mathcal{G}_{q}^{1} f\right\|_{\mathrm{BMO}_{\left(\mathbb{R}^{n}\right)}} \leq C\|f\|_{L_{\mathfrak{B}}^{\infty}\left(\mathbb{R}^{n}\right)}=C\|h\|_{L_{\mathcal{B}}^{\infty}(\mathbb{R})}
$$

Finally, to prove that statement iii) in Theorem 10 implies that the space is of Lusin cotype $q$, we need to compare the generalized $g$-functions in $\mathbb{R}^{n}$
and in the torus. The first step consists in observing that only the part of the kernel for $x$ and $t$ near 0 in the case of $\mathcal{G}_{q}^{1}$ and the part of the kernel with $\theta$ near 0 and $r$ near 1 for $G_{q}^{1}$, play a rôle in the boundedness of these generalized $g$-functions. In fact, one can prove

$$
\begin{aligned}
& \left\|\mathcal{G}_{q}^{1} f\right\|_{L^{p}(\mathbb{R})} \leq\|f\|_{L_{\mathcal{B}}^{p}(\mathbb{R})} \Longleftrightarrow \\
& \left\|\left[t \frac{\partial P_{t}}{\partial t} \chi_{(0, \varepsilon)}(t) \chi_{(0, \varepsilon)}(|x|)\right] * f\right\|_{L_{\left.L_{\mathcal{B}}^{p}(0, \infty), \frac{d t}{t}\right)}(\mathbb{R})} \leq C_{\varepsilon}\|f\|_{L_{\mathcal{B}}^{p}(\mathbb{R})} \\
& \left\|G_{q}^{1} f\right\|_{L^{p}(\mathbb{T})} \leq\|f\|_{L_{\mathcal{B}}^{p}(\mathbb{T})} \Longleftrightarrow \\
& \left\|\left[(1-r) \frac{\partial P_{r}}{\partial r} \chi_{(1-\delta, 1)}(r) \chi_{(0, \delta)}(\theta)\right] * f\right\|_{L_{L_{\mathcal{B}}^{p}\left((0,1), \frac{d r}{1}\right)}(\mathbb{T})} \leq C_{\delta}\|f\|_{L_{\mathcal{B}}^{p}(\mathbb{T})}
\end{aligned}
$$

With these restrictions, the change of variables $r=e^{-t}, \theta=x$ and the following equivalences: for $0<t<\varepsilon, 0<\theta<2 \delta: 1-e^{-t} \sim t, 1-e^{-2 t} \sim 2 t$, $\sin ^{2}(\theta / 2) \sim \theta^{2} / 4$, it is not difficult to see that

$$
\left[t \frac{\partial P_{t}}{\partial t} \chi_{(0, \varepsilon)}(t) \chi_{(0, \varepsilon)}(|x|)\right] * f(x) \sim\left[(1-r) \frac{\partial P_{r}}{\partial r} \chi_{(0,1-\delta)}(r) \chi_{(0, \delta)}(\theta)\right] * f(\theta)
$$

where the symbol $\sim$ means that the difference is bounded in $L^{p}$ (see [12] for the details of the proof).

Once we have Theorem 10, a new application of the suitable version of Theorem 2, gives us the following corollary.

Corollary 12. Given a Banach space, $\mathcal{B}$, and $q \geq 2$, the following sentences are equivalent:
i) $\mathcal{B}$ is of Lusin cotype $q$,
ii) $\mathcal{G}_{q}$ maps $L_{c, \mathcal{B}}^{\infty}\left(\mathbb{R}^{n}\right)$ into $\operatorname{BMO}\left(\mathbb{R}^{n}\right)$ boundedly,
iii) $\mathcal{G}_{q}$ maps $\mathrm{BMO}_{c, \mathcal{B}}\left(\mathbb{R}^{n}\right)$ into $\mathrm{BMO}\left(\mathbb{R}^{n}\right)$ boundedly,
iv) $\mathcal{G}_{q}$ maps $H_{\mathcal{B}}^{1}\left(\mathbb{R}^{n}\right)$ into $L^{1}\left(\mathbb{R}^{n}\right)$ boundedly,
v) $\mathcal{G}_{q}$ maps $L_{\mathcal{B}}^{1}\left(\mathbb{R}^{n}\right)$ into $L^{1, \infty}\left(\mathbb{R}^{n}\right)$ boundedly,
and they are also equivalent to the same statements ii$)-\mathrm{v}$ ) with $\mathcal{G}_{q}^{1}$ or $\mathcal{G}_{q}^{2}$ in place of $\mathcal{G}_{q}$, for every (or, equivalently, for some) $n \geq 1$.

## 6 Characterizations in terms of almost sure finiteness of the operators

The characterizations of Lusin cotype property we have given so far have been in terms of the boundedness between several spaces of the generalized $g$-functions. We can give a different kind of characterizations, in terms of the finiteness of the operators, as it is stated in the following theorem.

Theorem 13. Given a Banach space $\mathcal{B}$, the following statements are equivalent:
i) $\mathcal{B}$ is of Lusin cotype $q$.
ii) For any $f \in L_{\mathcal{B}}^{1}(\mathbb{T}), G_{q}^{1} f(z)<\infty$ for almost every $z \in \mathbb{T}$.
iii) For any $f \in L_{\mathcal{B}}^{1}(\mathbb{R}), \mathcal{G}_{q}^{1} f(x)<\infty$ for almost every $x \in \mathbb{R}^{n}$ for some (or, equivalently, for any) $n \geq 1$.
The equivalence holds also when in statement ii) we replace $G_{q}^{1}$ by $G_{q}^{2}$ or by $G_{q}$, and also if in statement iii) we replace $\mathcal{G}_{q}^{1}$ by $\mathcal{G}_{q}^{2}$ or by $\mathcal{G}_{q}$.

By the results in [19], the statement $i$ ) is equivalent to the weak type $(1,1)$ of either $G_{q}^{1}, G_{q}^{2}$ or $G_{q}$. By Corollary 12 statement i) is equivalent to the weak $(1,1)$ boundedness of either $\mathcal{G}_{q}, \mathcal{G}_{q}^{1}$ or $\mathcal{G}_{q}^{2}$. Then, statements ii) and iii) clearly follow from i). Thus, the only non trivial part is ii) implies i) and also that iii) implies i). Observe that to proving them is equivalent to get them for the $L^{q}$-vector valued operators $T$ whose norm in $L^{q}$ are the $g$-functions (as in (7) and (12)). Let us first prove ii) implies i). To this end, observe that

$$
\begin{equation*}
G_{q}^{1}(f)(z)=\|T f(z)\|_{L_{\ell_{\mathcal{B}}^{2}}^{q}\left((0,1), \frac{d r}{1-r}\right)}=\sup _{\varepsilon>0}\left\|T^{\varepsilon} f(z)\right\|_{\left.L_{\ell_{\mathcal{B}}^{2}}^{q}(0,1), \frac{d r}{1-r}\right)} \tag{13}
\end{equation*}
$$

where $T^{\varepsilon}$ is the operator that sends $\mathcal{B}$-valued functions into $L_{\mathcal{B} \times \mathcal{B}}^{q}\left((0,1), \frac{d r}{1-r}\right)$ valued functions given by

$$
T^{\varepsilon} f(z)=\left[(1-r) \chi_{(\varepsilon, 1-\varepsilon)}(r) \frac{\partial P_{r}}{\partial r}\right] * f(\theta)
$$

Since for $r \in(\varepsilon, 1-\varepsilon), 1+r^{2}-2 r \cos (\theta-t) \geq 1+r^{2}-2 r \geq \varepsilon^{2}$, we have that $\left|\partial P_{r} / \partial r\right| \leq C_{\varepsilon}$, and therefore

$$
\left\|\chi_{(\varepsilon, 1-\varepsilon)}(r) \frac{\partial P_{r}}{\partial r} * f(\theta)\right\|_{\mathcal{B}} \leq C_{\varepsilon} \int_{\mathbb{T}}\|f(t)\|_{\mathcal{B}} d t .
$$

In particular, the norms in $L_{L_{\mathcal{B}}^{q}\left((0,1), \frac{d r}{1-r}\right)}$ of the operators $T^{\varepsilon}$ are continuous-in-measure sublinear operators. By (13), $G_{q}^{1}$ is the supremum of these norms and we are assuming that it is finite almost everywhere. By Banach's continuity principle (Proposition VI.1.4, [7]), $G_{q}$ is a continuous-in-measure operator. Next, we apply Stein's theorem. A proof of the scalar version can be found in Section VI. 2 of [7], and by using the ideas there, one can prove the following vector-valued version.

Lemma 14. Let $G$ be a locally compact group with Haar measure $\mu, \mathcal{B}$ be a Banach space of Rademacher type $p_{0}$ and let

$$
T: L_{\mathcal{B}}^{p}(G) \longrightarrow L^{0}(G)
$$

be a sublinear, continuous in measure and invariant under left translations operator. Then, for every compact $K$ subset of $G$, there exists a constant $C_{K}$ such that

$$
\mu(\{x \in K:|T f(x)|>\lambda\}) \leq C_{K}\left(\frac{\|f\|_{L_{\mathcal{B}}^{p}}}{\lambda}\right)^{q}
$$

with $q=\inf \left\{p, p_{0}\right\}$. In particular, if the group $G$ is compact, $T$ is of weak type $(p, q)$.

Let us recall that every Banach space is of Rademacher type 1. Then, $G_{q}^{1}$ is of weak type $(1,1)$, because it is clearly sublinear and it is given by a convolution, which is invariant under translations.

The proof for the case of iii) implies i) is analogous, by applying the suitable version of Stein's results, concretely a corollary of Stein's theorem that can be found in the scalar valued version in Section VI. 2 of [7] (and whose extension to the vector valued case is made in the same way as for Lemma 14).

We can also characterize Lusin cotype property in terms of the finiteness almost everywhere of the generalized square function of martingales.

Theorem 15. Given a Banach space $\mathcal{B}$, the following sentences are equivalent:
i) $\mathcal{B}$ has martingale cotype $q, 2 \leq q<\infty$,
ii) If $f$ is a martingale bounded in $L_{\mathcal{B}}^{1}$, then $S_{q} f<\infty$ almost everywhere.

For the proof of this theorem, we will use martingale transform operators. Let $\mathcal{B}_{\infty}$ and $\mathcal{B}_{\in}$ be two Banach spaces, $(\Omega, \mathcal{F}, P)$ be a probability space, and $\left\{\mathcal{F}_{n}\right\}_{n \geq 1}$ be a stochastic basis. A multiplying sequence $v=\left\{v_{n}\right\}_{n \geq 1}$ is a sequence of random variables with values in the space of linear continuous applications between the Banach spaces, $v_{n}: \Omega \longrightarrow \mathcal{L}\left(\mathcal{B}_{\infty}, \mathcal{B}_{\epsilon}\right)$, such that each $v_{n}$ is $\mathcal{F}_{n-1}$-measurable, and it is uniformly bounded, $\sup _{n \geq 1}\left\|v_{n}\right\|_{L_{\mathcal{C}\left(\mathcal{B}_{\infty}, \mathcal{B}_{\epsilon}\right)}^{\infty}}<$ $\infty$. Given such a multiplying sequence, define $T$ the martingale transform operator given by $v$, as the one assigning to each martingale $f$ the transformed martingale $T f$, defined as $(T f)_{n}=\sum_{k=1}^{n} v_{k} d_{k} f$. It is proved in [11] that a for a martingale transform operator several boundedness properties are equivalent. In particular, weak type $(1,1)$ holds for $T$ if and only if it is of strong type $(p, p)$.

$$
\begin{equation*}
\lambda P\left\{(T f)^{*}>\lambda\right\} \leq C\|f\|_{L_{\mathcal{B}_{\infty}}^{1}} \Longleftrightarrow\left\|(T f)^{*}\right\|_{L^{p}} \leq C_{p}\|f\|_{L_{\mathcal{B}_{\infty}}^{p}} \tag{14}
\end{equation*}
$$

It is also proved in [11] that, if $T$ is a translation invariant martingale transform operator such that each term of its multiplying sequence $\left\{v_{k}\right\}_{k \geq 1} \subset$ $\mathcal{L}\left(\mathcal{B}_{\infty}, \mathcal{B}_{\epsilon}\right)$ is a constant operator from $\mathcal{B}_{\infty}$ into $\mathcal{B}_{\epsilon}, v_{k}(\omega)=v_{k} \in \mathcal{L}\left(\mathcal{B}_{\infty}, \mathcal{B}_{\epsilon}\right)$, such that

$$
\begin{equation*}
f^{*} \in L^{1} \Longrightarrow T f \text { converges a.e., } \tag{15}
\end{equation*}
$$

it also verifies the inequalities in (14). $T$ is translation invariant if for any $k_{0} \in \mathbb{N}$, the sequence $\left\{v_{k}^{k_{0}}\right\}_{k \geq 1}, v_{k}^{k_{0}}=v_{k_{0}+k}$, defines a martingale transform operator $T_{k_{0}}$ such that for any martingale $f$ bounded in $L_{\mathcal{B}_{\infty}}^{1}$,

$$
\left\|(T f)_{n}\right\|_{\mathcal{B}_{\epsilon}}=\left\|\sum_{k=1}^{n} v_{k} d_{k} f\right\|_{\mathcal{B}_{\epsilon}}=\left\|\sum_{k=1}^{n} v_{k_{0}+k} d_{k} f\right\|_{\mathcal{B}_{\epsilon}}=\left\|\left(T_{k_{0}} f\right)_{n}\right\|_{\mathcal{B}_{\epsilon}} .
$$

Now, let $Q_{q}$ be the martingale transform operator mapping $\mathcal{B}$-valued martingales into $\ell_{\mathcal{B}}^{q}$-valued martingales defined by the multipliying sequence $\left\{v_{k}\right\}_{k \geq 1}$, such that each $v_{k}$ is constant and for any $b \in \mathcal{B}, v_{k}(b)=\left(0,{ }_{\cdots}-\cdots\right)$ $, 0, b, 0, \ldots)$ is a vector in $\ell_{\mathcal{B}}^{q}$. Given a $\mathcal{B}$-valued martingale, we have

$$
\begin{gathered}
\left(Q_{q} f\right)_{n}=\sum_{k=1}^{n} v_{k} d_{k} f=\left(d_{1} f, d_{2} f, \ldots, d_{n} f, 0, \ldots\right) \in \ell_{\mathcal{B}}^{q} \\
\left\|\left(Q_{q} f\right)_{n}\right\|_{\ell_{\mathcal{B}}^{q}}=\left(\sum_{k=1}^{n}\left\|d_{k} f\right\|_{\mathcal{B}}^{q}\right)^{1 / q}, \quad\left(Q_{q} f\right)^{*}=\left\|\left(Q_{q} f\right)_{n}\right\|_{\ell_{\mathcal{B}}^{q}}=\sup _{n} S_{q} f .
\end{gathered}
$$

We can now proceed with the proof of Theorem 15. That i) implies ii) is obvious, as a consequence of the weak type $(1,1)$ of $S_{q}$. To prove the
converse implication, we use that (15) implies the inequalities (14), applied to $T f=Q_{q} f$, which is translation invariant and for $f^{*} \in L^{1}$,

$$
\left\|\left(Q_{q} f\right)_{n}-\left(Q_{q} f\right)_{m}\right\|_{\ell_{\mathcal{B}}^{q}}=\left(\sum_{k=m+1}^{n}\left\|d_{k} f\right\|_{\mathcal{B}}^{q}\right)^{1 / q} \longrightarrow 0 \quad \text { a.e. as } n, m \rightarrow \infty
$$

since it is the tail of a convergent series (by ii)).

## 7 Characterizations by the Lusin area function

The classical Lusin area function in the torus is defined as

$$
A f(z)=\left(\int_{\Gamma(z)}(1-r)^{2}\left\|\nabla P_{r} * f(\xi)\right\|^{2} \frac{r d \xi d r}{(1-r)^{2}}\right)^{1 / 2}
$$

where $\Gamma(z)$ is the convex domain with vertex $z=e^{i \theta}$ and width 1 , given by

$$
\Gamma(z)=\left\{r \xi: 0 \leq r \leq 1, \xi=e^{i t},|\theta-t| \leq 1-r\right\}
$$

and the integrand is defined as in (1). It is proved in [19] that Lusin cotype property is also equivalent to the same inequalities already seen for $G_{q}$, but with the generalized Lusin area function $A_{q}$ instead of $G_{q}$. The generalized Lusin area function of $f \in L_{\mathcal{B}}^{p}(\mathbb{T}), p \in(1, \infty)$ is defined for $z \in \mathbb{T}$ as

$$
A_{q} f(z)=\left(\int_{\Gamma(z)}(1-r)^{q}\left\|\nabla P_{r} * f(\xi)\right\|_{\mathcal{B}}^{q} \frac{r d \xi d r}{(1-r)^{2}}\right)^{1 / q}
$$

where $\Gamma(z)$ is as above. As in the classical case, it holds (see [19] and the references therein) that for every function $f$ and every $\theta \in \mathbb{T}$,

$$
\begin{equation*}
G_{q} f(\theta) \leq C_{q} A_{q} f(\theta) \tag{16}
\end{equation*}
$$

and also a simple calculation gives

$$
\begin{equation*}
\left\|A_{q} f\right\|_{L^{q}(\mathbb{T})} \leq C_{q}\left\|G_{q} f\right\|_{L^{q}(\mathbb{T})} \tag{17}
\end{equation*}
$$

where $C_{q}$ is a constant just depending on $q$. In $\mathbb{R}^{n}$, the generalized Lusin area function is defined as

$$
\mathcal{A}_{q}(f)(x)=\left(\iint_{\Gamma(x)} t^{q}\left\|\nabla P_{t} * f(y)\right\|_{\ell_{\mathfrak{B}}^{2}}^{q} \frac{d y d t}{t^{n+1}}\right)^{1 / q}
$$

where $\Gamma(x)$ is the cone with vertex $x$ and width 1 ,

$$
\Gamma(x)=\left\{(t, y) \in \mathbb{R}_{+}^{n+1}: t \geq 0,|x-y| \leq t\right\}
$$

and the integrand is defined as in (11). As in the classical case, it also holds (see [18]) that for every function $f$ and every $x \in \mathbb{R}^{n}$,

$$
\begin{equation*}
\mathcal{G}_{q} f(x) \leq C_{q} \mathcal{A}_{q} f(x) \tag{18}
\end{equation*}
$$

and also a simple calculation gives

$$
\begin{equation*}
\left\|\mathcal{A}_{q} f\right\|_{L^{q}\left(\mathbb{R}^{n}\right)} \leq C_{q}\left\|\mathcal{G}_{q} f\right\|_{L^{q}\left(\mathbb{R}^{n}\right)}, \tag{19}
\end{equation*}
$$

where $C_{q}$ is a constant just depending on $q$. The following result states the characterizations of Lusin cotype property in terms of Lusin area function.

Theorem 16. Given a Banach space $\mathcal{B}$ and $q \in[2, \infty)$, the following sentences are equivalent, when they hold for some or, equivalently, for any $n \geq 1$.
i) $\mathcal{B}$ is of Lusin cotype $q$.
ii) $\mathcal{A}_{q}$ maps $L_{c, \mathcal{B}}^{\infty}\left(\mathbb{R}^{n}\right)$ into $\operatorname{BMO}\left(\mathbb{R}^{n}\right)$.
iii) $\mathcal{A}_{q}$ maps $H_{\mathcal{B}}^{1}\left(\mathbb{R}^{n}\right)$ into $L^{1}\left(\mathbb{R}^{n}\right)$.
iv) $\mathcal{A}_{q}$ maps $L_{\mathcal{B}}^{p}\left(\mathbb{R}^{n}\right)$ into $L_{\mathcal{B}}^{p}\left(\mathbb{R}^{n}\right)$ for any (or, equivalently for some) $p \in$ $(1, \infty)$.
v) $\mathcal{A}_{q}$ maps $\mathrm{BMO}_{c, \mathcal{B}}\left(\mathbb{R}^{n}\right)$ into $\mathrm{BMO}\left(\mathbb{R}^{n}\right)$.
vi) $\mathcal{A}_{q}$ maps $L_{\mathcal{B}}^{1}\left(\mathbb{R}^{n}\right)$ into $L^{1, \infty}\left(\mathbb{R}^{n}\right)$.
vii) For every $f \in L_{\mathcal{B}}^{1}\left(\mathbb{R}^{n}\right), \mathcal{A}_{q} f(x)<\infty$ for almost every $x \in \mathbb{R}^{n}$.

These statements are also equivalent to the same ones with the Lusin area function in the torus instead of $\mathcal{A}_{q}$.

To prove this theorem, observe that the Lusin area function can be seen as the convolution of the $\mathcal{B}$-valued function $f$ with the $L_{\ell_{\mathcal{B}}^{2}}^{q}\left([0, \infty) \times B, \frac{d z d t}{t}\right)$ function $k_{t}(x)$, where $B$ denotes the unit ball in $\mathbb{R}^{n}$, given by

$$
k_{t}(x)=\left(\left(\psi_{z}\right)^{t}(x),\left(\psi_{z}^{1}\right)^{t}(x), \ldots,\left(\psi_{z}^{n}\right)^{t}(x)\right)
$$

where

$$
\psi(x)=C \frac{|x|^{2}-n}{\left(|x|^{2}+1\right)^{\frac{n+3}{2}}}, \quad \psi^{j}(x)=C \frac{x_{j}}{\left(|x|^{2}+1\right)^{\frac{n+3}{2}}} .
$$

and for any $x, z \in \mathbb{R}^{n}, f_{z}(x)=f(x-z)$, and for $\eta>0, f^{\eta}(x)=\frac{1}{\eta^{n}} f\left(\frac{x}{t}\right)$. This function satisfies the assumptions of a regular Calderón-Zygmund kernel, so the equivalence of statements $\mathbf{i i})-\mathrm{vi}$ ) is a consequence of Theorem 2. The equivalence of this statements with statement i) is due to Theorem 10 and the inequalities (18) and (19). Finally, statement vi) clearly implies vii), and by (18) and Theorem 13 we get that vii) implies i).

## Acknowledgement

This paper is based on a talk given in the Departamento de Análisis Matemático de la Universidad de Sevilla in March 2004. The author would like to thank Prof. Carlos Pérez for the invitation and their hospitality.

As it was mentioned in the introduction, this paper has intended to present some of the most interesting results contained in [12] and [19]. The author would like to express his deep gratitude to the coauthors José Luis Torrea and Quanhua Xu , for many interesting discussions and all what was learnt from them.

## References

[1] C. Bennet, R.A. DeVore, R. Sharpley, Weak- $L^{\infty}$ and BMO, Ann. of Math. 113 (1981), 601-611.
[2] O. Blasco, J. García-Cuerva, Hardy classes of Banach-space-valued distributions, Math. Nachr. 132 (1987), 57-65.
[3] E.B. Davies, One-parameter diffusion semigroups, L.M.S. Monographs, 15, Academic Press, London, 1980.
[4] J. Diestel, J.J. Uhl, Vector Measures. Math. Surveys, 15 (1977), Amer. Math. Soc.
[5] R. Durrett, Brownian motion and martingales in analysis, Wadsworth Mathematics Series, Wadsworth International Group, Belmont, CA, 1984.
[6] J.A. Goldstein, Semigroups of Linear Operators and Applications, Oxford Mathematical Monographs, Oxford Science Publications, Oxford, 1985.
[7] J. García-Cuerva, J.L. Rubio de Francia, Weighted norm inequalities and related topics, North-Holland Mathematics Studies, 116, Elsevier Science Publishers, 1985.
[8] S. Kwapien, Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients, Studia Math. 44 (1972), 583-595.
[9] J.L. Journé, Calderón-Zygmund operators, Pseudo-differential operators and the Cauchy integral of Calderón, Lecture Notes in Math. 994, Springer-Verlag, Berlin, 1983.
[10] T. Martínez, Uniform convexity in terms of martingale $H^{1}$ and $B M O$ spaces, Houston Journal of Mathematics, 27, (2001) 461-447.
[11] T. Martinez, J.L. Torrea, Operator-valued martingale transforms Tohoku Mathematical Journal, 52, (2000) 449-474.
[12] T. Martinez, J.L. Torrea, Q. Xu, Vector-valued Littlewood-Paley-Stein theory for semigroups, preprint.
[13] G. Pisier, Martingales with values in uniformly convex spaces, Israel J. Math., 20, (1975), 326-350.
[14] G. Pisier, Probabilistic methods in the geometry of Banach spaces, Lecture Notes in Mathematics, vol. 1206, Springer-Verlag, Berlin. 1986.
[15] J.L. Rubio de Francia, Martingale and integral transforms of Banach space valued functions, in Conference on Probability and Banach spaces, Zaragoza, 1985, Lecture Notes in Math. 1221, 195-222, SpringerVerlag, Berlin, 1986.
[16] D.A. Stegenga, Bounded Toeplitz operators on $H^{1}$ and applications of the duality between $H^{1}$ and the functions of bounded mean oscillations, Amer. J. Math. 98 (1976), 573-589.
[17] E.M Stein, Topics in Harmonic Analysis Related to the Littlewood-Paley Theory, Princeton University Press and the University of Tokio Press, Princeton, New Jersey, 1970.
[18] E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton University Press, Princeton, New Jersey, 1970.
[19] Q. Xu, Littlewood-Paley theory for functions with values in uniformly convex spaces, J. reine angew. Math. 504 (1998), 195-226.
[20] K. Yosida, Functional Analysis, Springer-Verlag Berlin Heidelberg New York, 1968.

## Author's address:

Teresa Martinez<br>Departamento de Matemáticas<br>Faculdad de Ciencias<br>Universidad Autónoma de Madrid<br>28049 Madrid, Spain<br>E-mail: teresa.martinez@uam.es

# Fixed point theory the Picard operators technique 

Adrian Petruşel


#### Abstract

The purpose of this synthesis is to present several results for singlevalued and multivalued operators using the abstract technique of (weakly) Picard operators.


## 1 Notations and basic notions

Throughout this paper, the standard notations and terminologies in nonlinear analysis are used. For the convenience of the reader we recall some of them.

Let $X$ be a nonempty set. Then:

$$
\mathcal{P}(X)=\{Y \mid Y \text { is a subset of } X\}, P(X)=\{Y \in \mathcal{P}(X) \mid Y \text { is nonempty }\}
$$

Let $f: X \rightarrow X$ be an operator. Then $f^{0}:=1_{X}, f^{1}:=f, \ldots, f^{n+1}=$ $f \circ f^{n}, n \in \mathbb{N}$ denote the iterate operators of $f$.

Also, by $F_{f}:=\{x \in X \mid x=f(x)\}$ we will denote the fixed point set of the operator $f$, while the set of all nonempty invariant subsets of $f$ will be denoted by $I(f)$, i. e. $I(f):=\{Y \in P(X) \mid f(Y) \subset Y\}$.

Let $(X, d)$ be a metric space. Then $\delta(Y):=\sup \{d(a, b) \mid a, b \in Y\}$ and

$$
\begin{gathered}
P_{b}(X):=\{Y \in P(X) \mid \delta(Y)<+\infty\}, P_{c l}(X):=\{Y \in P(X) \mid Y \text { is closed }\}, \\
P_{c p}(X):=\{Y \in P(X) \mid Y \text { is compact }\}, P_{b, c l}(X):=P_{b}(X) \cap P_{c l}(X) .
\end{gathered}
$$

Let $(X, d),\left(Z, d^{\prime}\right)$ be metric spaces and $T: X \rightarrow P(Z)$ be a multivalued operator. Then, the symbol GrafT $:=\{(x, z) \in X \times Z \mid z \in T(x)\}$ denotes the graph of $T$. A multivalued operator $T$ is called closed if GrafT is a closed set. The multivalued operator $T$ is called upper semicontinuous

[^7](briefly u.s.c.) on $X$ if and only if $T^{+}(V):=\{x \in X \mid T(x) \subset V\}$ is open, for each open set $V \subset Z$ and it is said to be lower semicontinuous (briefly l.s.c.) on $X$ if and only if $T^{-}(W):=\{x \in X \mid T(x) \cap W \neq \emptyset\}$ is open, for each open set $W \subset Z$. If $T$ is u.s.c. and l.s.c. on $X$ then it is called continuous on $X$.

The operator $\hat{T}: P(X) \rightarrow P(Z)$, defined by:

$$
\hat{T}(Y):=\bigcup_{x \in Y} T(x), \text { for } Y \in P(X)
$$

is called the fractal operator generated by $T$. If $T$ is a continuous multivalued operator with compact values, then $\hat{T}$ is continuous too.

If $T: X \rightarrow P(X)$ then the set of all nonempty invariant subsets of $T$ will be denoted by $I(T):=\{Y \in P(X) \mid T(Y) \subset Y\}$.

A sequence of successive approximations of $T$ starting from $x \in X$ is a sequence $\left(x_{n}\right)_{n \in \mathbb{N}}$ of elements of $X$ with $x_{0}=x, x_{n+1} \in T\left(x_{n}\right)$, for $n \in \mathbb{N}$.

Throughout the paper $F_{T}:=\{x \in X \mid x \in T(x)\}$ denotes the fixed point set $T$, while $(S F)_{T}:=\{x \in X \mid\{x\}=T(x)\}$ is the strict fixed point set of $T$.

Let $X$ be a nonempty set. Denote $s(X):=\left\{\left(x_{n}\right)_{n \in N} \mid x_{n} \in X, n \in N\right\}$.
Let $c(X) \subset s(X)$ a subset of $s(X)$ and $\operatorname{Lim}: c(X) \rightarrow X$ an operator. By definition the triple ( $X, c(X), \operatorname{Lim})$ is called an L-space (Fréchet [12]) if the following conditions are satisfied:
(i) If $x_{n}=x, \forall n \in N$, then $\left(x_{n}\right)_{n \in N} \in c(X)$ and $\operatorname{Lim}\left(x_{n}\right)_{n \in N}=x$.
(ii) If $\left(x_{n}\right)_{n \in N} \in c(X)$ and $\operatorname{Lim}\left(x_{n}\right)_{n \in N}=x$, then for all subsequences, $\left(x_{n_{i}}\right)_{i \in N}$, of $\left(x_{n}\right)_{n \in N}$ we have that $\left(x_{n_{i}}\right)_{i \in N} \in c(X)$ and $\operatorname{Lim}\left(x_{n_{i}}\right)_{i \in N}=x$.

By definition an element of $c(X)$ is convergent sequence and the limit of this sequence is $x:=\operatorname{Lim}\left(x_{n}\right)_{n \in N}$ and we write $x_{n} \rightarrow x$ as $n \rightarrow \infty$.

In what follow we denote an L-space by $(X, \rightarrow)$.
The following (generalized) functionals are important for the main sections of the paper.

## The gap functional

$$
\begin{aligned}
& \text { (1) } D: \mathcal{P}(X) \times \mathcal{P}(X) \rightarrow \mathbb{R}_{+} \cup\{+\infty\} \\
& D(A, B)= \begin{cases}\inf \{d(a, b) \mid a \in A, b \in B\}, & A \neq \emptyset \neq B \\
0, & A=\emptyset=B \\
+\infty, & \text { otherwise }\end{cases}
\end{aligned}
$$

## $\delta$ generalized functional

$$
\text { (2) } \delta: \mathcal{P}(X) \times \mathcal{P}(X) \rightarrow \mathbb{R}_{+} \cup\{+\infty\}
$$

$$
\delta(A, B)= \begin{cases}\sup \{d(a, b) \mid a \in A, b \in B\}, & A \neq \emptyset \neq B \\ 0, & \text { otherwise }\end{cases}
$$

## The excess generalized functional

$$
\begin{aligned}
& \text { (3) } \rho: \mathcal{P}(X) \times \mathcal{P}(X) \rightarrow \mathbb{R}_{+} \cup\{+\infty\} \\
& \rho(A, B)= \begin{cases}\sup \{D(a, B) \mid a \in A\}, & A \neq \emptyset \neq B \\
0, & A=\emptyset \\
+\infty, & B=\emptyset \neq A\end{cases}
\end{aligned}
$$

## Pompeiu-Hausdorff generalized functional

$$
\begin{aligned}
& \text { (4) } H: \mathcal{P}(X) \times \mathcal{P}(X) \rightarrow \mathbb{R}_{+} \cup\{+\infty\} \\
& H(A, B)= \begin{cases}\max \{\rho(A, B), \rho(B, A)\}, & A \neq \emptyset \neq B \\
0, & A=\emptyset=B \\
+\infty, & \text { othewise }\end{cases}
\end{aligned}
$$

For more details and basic results concerning the above notions see for example [5], [6], [8], [15], [16], [17], [20], [43], etc.

## 2 Picard operators and weakly Picard operators

Definition 1. (I.A. Rus [37]) Let $(X, \rightarrow)$ be an L-space. An operator $f$ : $X \rightarrow X$ is, by definition, a Picard operator if:
(i) $F_{f}=\left\{x^{*}\right\}$;
(ii) $f^{n}(x) \rightarrow x^{*}$ as $n \rightarrow \infty$, for all $x \in X$.

Example 2. Let $(X, d)$ be a complete metric space and $f: X \rightarrow X$ an acontraction, i. e. $a \in] 0,1[$ and $d(f(x), f(y)) \leq a \cdot d(x, y)$, for each $x, y \in X$. Then the operator $f$ is Picard. (Banach-Caccioppoli)

Example 3. Let $(X, d)$ be a compact metric space and $f: X \rightarrow X$ satisfying $d(f(x), f(y))<d(x, y)$, for all $x, y \in X$ with $x \neq y$. Then the operator $f$ is Picard. (Nemytzki-Edelstein)

Example 4. Let $(X, d)$ be a complete generalized metric space $(d(x, y) \in$ $\left.\mathbb{R}_{+}^{m}\right)$ and $A \in M_{m m}\left(\mathbb{R}_{+}\right)$, such that, $A^{n} \rightarrow 0$ as $n \rightarrow \infty$. If $f: X \rightarrow X$ is an $A$-contraction, i. e., $d(f(x), f(y)) \leq A d(x, y)$, for all $x, y \in X$, then it is Picard operator. (Perov)

Example 5. Let $(X, d)$ be a complete metric space and $f: X \rightarrow X$ be a Meir-Keeler type operator, i. e. for each $\eta>0$ there exists $\delta>0$ such that $x, y \in X, \eta \leq d(x, y)<\eta+\delta$ we have $d(f(x), f(y))<\eta$. Then $f$ is a Picard operator. (Meir-Keeler)

Another important concept is:
Definition 6. Let $(X, \rightarrow)$ be an L-space. By definition, $f: X \rightarrow X$ is called a weakly Picard operator if the sequence $\left(f^{n}(x)\right)_{n \in N}$ converges for all $x \in X$ and the limit (which may depend on $x$ ) is a fixed point of $f$.

Example 7. Let $(X, d)$ be a complete metric space and $f: X \rightarrow X$ such that $f$ is closed and there is a $\in] 0,1\left[\right.$ with the property $d\left(f(x), f^{2}(x)\right) \leq$ $a \cdot d(x, f(x))$, for each $x \in X$. Then $f$ is a weakly Picard operator.

Example 8. Let $(X, d)$ be a complete metric space, $f: X \rightarrow X$ an operator and $\varphi: X \rightarrow \mathbb{R}_{+}$a function. We suppose that:
(i) the operator $f$ satisfies the Caristi condition with respect to $\varphi$, i.e.,

$$
d(x, f(x)) \leq \varphi(x)-\varphi(f(x)), \text { for all } x \in X
$$

(ii) the operator $f$ is closed.

Then $f$ is a weakly Picard operator.
In I. A. Rus [41] (see Theorem 4.2) the following characterization theorem for the class of weakly Picard operators was proved:

Theorem 9. Let $(X, \rightarrow)$ be an L-space and $f: X \rightarrow X$ an operator. The operator $f$ is weakly Picard operator if and only if there exists a partition of $X, X=\bigcup_{\lambda \in \Lambda} X_{\lambda}$, such that
(a) $X_{\lambda} \in I(f), \forall \lambda \in \Lambda$;
(b) the restriction of $f$ to $X_{\lambda},\left.f\right|_{X_{\lambda}}: X_{\lambda} \rightarrow X_{\lambda}$, is a Picard operator, for all $\lambda \in \Lambda$.

In [41] the basic theory of Picard and weakly Picard operators is presented.

## 3 Multivalued weakly Picard operators

In a similar way with the singlevalued case, the following notion was introduced.

Definition 10. (Rus-Petruşel-Sîntămărian [45]) Let $(X, \rightarrow)$ be an L-space. Then $T: X \rightarrow P(X)$ is a multivalued weakly Picard operator (briefly MWP operator) if for each $x \in X$ and each $y \in T(x)$ there exists a sequence $\left(x_{n}\right)_{n \in \mathbb{N}}$ in $X$ such that:
i) $x_{0}=x, x_{1}=y$
ii) $x_{n+1} \in T\left(x_{n}\right)$, for all $n \in \mathbb{N}$
iii) the sequence $\left(x_{n}\right)_{n \in \mathbb{N}}$ is convergent and its limit is a fixed point of $T$.

Example 11. Let $(X, d)$ be a complete metric space and $T: X \rightarrow P_{c l}(X) a$ Reich type multivalued operator, $i$. e. there exist $\alpha, \beta, \gamma \in \mathbb{R}_{+}$with $\alpha+\beta+$ $\gamma<1$ such that
$H(T(x), T(y)) \leq \alpha d(x, y)+\beta D(x, T(x))+\gamma D(y, T(y))$, for all $x, y \in X$.
Then $T$ is a MWP operator. (Reich [35])
Let us remark that if $\beta=\gamma=0$, then $T$ is said to be a multivalued $\alpha$-contraction.

Example 12. Let $(X, d)$ be a complete metric space and $T_{1}, T_{2}: X \rightarrow$ $P_{c l}(X)$ for which there exists $\left.\alpha \in\right] 0, \frac{1}{2}[$ such that

$$
H\left(T_{1}(x), T_{2}(y)\right) \leq a\left[D\left(x, T_{1}(x)\right)+D\left(y, T_{2}(y)\right)\right]
$$

for each $x, y \in X$.
Then $T_{1}$ and $T_{2}$ are MWP operators. (Sintămărian [47])
Example 13. Let $(X, d)$ be a complete metric space and $T: X \rightarrow P_{c l}(X)$. Suppose that there exist $\alpha, \beta \in \mathbb{R}_{+}$, with $\alpha+2 \beta<1$ such that: $\rho(T(x), T(y)) \leq$ $\alpha d(x, y)+\beta(D(x, T(x))+D(y, T(y)))$ for all $x, y \in X$.

Then $T$ is a MWP operator. (Wang [51])
Example 14. (Petruşel [31]) Let ( $X, d$ ) be a generalized complete metric space, (i. e. $\left.d(x, y) \in \mathbb{R}_{+}^{m}\right)$ and $T: X \rightarrow P_{c l}(X)$ be a multivalued $A$ contraction, i.e. there exists $A \in \mathcal{M}_{m m}(\mathbb{R})$ such that $A^{n} \rightarrow 0, n \rightarrow \infty$ and for each $x, y \in X$ and each $u \in T(x)$ there exists $v \in T(y)$ such that $d(u, v) \leq \operatorname{Ad}(x, y)$.

Then $T$ is a MWP operator.
Example 15. Let $X$ be a complete gauge space and consider $T: X \rightarrow$ $P_{c l}(X)$. Suppose there exist constants $a_{1}=\left\{a_{1, \alpha}\right\}_{\alpha \in \Lambda} \in\left[0,1\left[^{\Lambda}, a_{2}=\right.\right.$ $\left\{a_{2, \alpha}\right\}_{\alpha \in \Lambda} \in\left[0,1\left[^{\Lambda}, a_{3}=\left\{a_{3, \alpha}\right\}_{\alpha \in \Lambda} \in\left[0,1\left[^{\Lambda}, a_{4}=\left\{a_{4, \alpha}\right\}_{\alpha \in \Lambda} \in\left[0,1\left[^{\Lambda}\right.\right.\right.\right.\right.\right.$, $a_{5}=\left\{a_{5, \alpha}\right\}_{\alpha \in \Lambda} \in\left[0,1\left[{ }^{\Lambda}\right.\right.$ such that for every $\alpha \in \Lambda$ and every $x, y \in X$ the following condition is satisfied:

$$
\begin{array}{r}
H_{\alpha}(T(x), T(y)) \leq a_{1, \alpha} D_{\alpha}(x, T(x))+a_{2, \alpha} D_{\alpha}(y, T(y))+a_{3, \alpha} D_{\alpha}(y, T(x))+ \\
a_{4, \alpha} D_{\alpha}(x, T(y))+a_{5, \alpha} d(x, y), \text { where for every } \alpha \in \Lambda, a_{1, \alpha}+a_{2, \alpha}+a_{3, \alpha}+a_{4, \alpha}+
\end{array}
$$

$a_{5, \alpha}<1, a_{1, \alpha}+a_{4, \alpha}+a_{5, \alpha}>0, a_{2, \alpha}+a_{3, \alpha}+a_{5, \alpha}>0$ and with either $a_{1, \alpha}=$ $a_{2, \alpha}$ or $a_{3, \alpha}=a_{4, \alpha}$.

Then $T$ is a MWP operator. (Agarwal-O'Regan [1])
In [31] other examples of MWP operators are presented.
Let $(X, d)$ be a metric space and $T_{1}, T_{2}: X \rightarrow P(X)$ be two multi-valued operators such that the fixed points sets $F_{T_{1}}$ and $F_{T_{2}}$ are nonempty and there exists $\eta>0$ with the property $H\left(T_{1}(x), T_{2}(x)\right) \leq \eta$, for all $x \in X$. The data dependence problem for the fixed point set of a multivalued operator is to estimate $H\left(F_{T_{1}}, F_{T_{2}}\right)$.

Several partial answers to this problem are given in Lim [24], Wang [51], Rus [41], Rus-Mureşan [42] and Rus-Petruşel-Sîntămărian [44].

In what follows we shall study the data dependence problem for a special class of multivalued weakly Picard operators.

Let us recall the following important notion:
Definition 16. Let $(X, d)$ be a metric space and $T: X \rightarrow P(X)$ be an MWP operator. Then we define the multivalued operator $T^{\infty}: \operatorname{Graf}(T) \rightarrow P\left(F_{T}\right)$ by the formula $T^{\infty}(x, y)=\left\{z \in F_{T} \mid\right.$ there exists a sequence of successive approximations of $T$ starting from $(x, y)$ that converges to $z\}$.

An important concept is given by the following definition:
Definition 17. Let $(X, d)$ be a metric space and $T: X \rightarrow P(X)$ an MWP operator. Then $T$ is a c-multivalued weakly Picard operator (briefly cMWP operator) if and only if there exists a selection $t^{\infty}$ of $T^{\infty}$ such that $d\left(x, t^{\infty}(x, y)\right) \leq c d(x, y)$, for all $(x, y) \in \operatorname{Graf}(T)$.

Further on we shall present some examples of $c$-MWP operators.
Example 18. Let $(X, d)$ be a complete metric space and $T: X \rightarrow P_{c l}(X)$ be a multivalued a-contraction $(0<a<1)$. Then $T$ is a $c-M W P$ operator, where $c=(1-a)^{-1}$.

Example 19. Let $(X, d)$ be a complete metric space and $T: X \rightarrow P_{c l}(X)$ be a multivalued Reich type operator. Then $T$ is a $c-M W P$ operator, where $c=(1-\gamma)[1-(\alpha+\beta+\gamma)]^{-1}$.

An important abstract result is the following:
Theorem 20. Let $(X, d)$ be a metric space and $T_{1}, T_{2}: X \rightarrow P(X)$ be two multivalued operators. We suppose that:
i) $T_{i}$ is a $c_{i}-M W P$ operator, for $i \in\{1,2\}$
ii) there exists $\eta>0$ such that $H\left(T_{1}(x), T_{2}(x)\right) \leq \eta$, for all $x \in X$. Then

$$
H\left(F_{T_{1}}, F_{T_{2}}\right) \leq \eta \max \left\{c_{1}, c_{2}\right\}
$$

Proof. Let $t_{i}: X \rightarrow X$ be a selection of $T_{i}, i \in\{1,2\}$. We remark that

$$
H\left(F_{T_{1}}, F_{T_{2}}\right) \leq \max \left\{\sup _{x \in F_{T_{2}}} d\left(x, t_{1}^{\infty}\left(x, t_{1}(x)\right)\right), \sup _{x \in F_{T_{1}}} d\left(x, t_{2}^{\infty}\left(x, t_{2}(x)\right)\right)\right\} .
$$

Let $q>1$. Then by we can choose $t_{i}$, for $i \in\{1,2\}$, such that

$$
d\left(x, t_{1}^{\infty}\left(x, t_{1}(x)\right)\right) \leq c_{1} q H\left(T_{2}(x), T_{1}(x)\right), \text { for all } x \in F_{T_{2}},
$$

and

$$
d\left(x, t_{2}^{\infty}\left(x, t_{2}(x)\right)\right) \leq c_{2} q H\left(T_{1}(x), T_{2}(x)\right), \text { for all } x \in F_{T_{1}} .
$$

Thus, we have

$$
H\left(F_{T_{1}}, F_{T_{2}}\right) \leq q \eta \max \left\{c_{1}, c_{2}\right\} .
$$

Letting $q \searrow 1$, the proof is complete.
Let us consider some consequences of this result.
Theorem 21. Let $(X, d)$ be a complete metric space and $T: X \rightarrow P_{c l}(X)$ be a multivalued operator satisfying the following two assumptions:
a) there exist $\alpha, \beta, \gamma, \mu \in \mathbb{R}_{+}$with $\alpha+\beta+\gamma+2 \mu<1$ such that:

$$
H(T x, T y) \leq \alpha d(x, y)+\beta d(x, T x)+\gamma D(y, T y)+\mu D(x, T y)
$$

for each $x \in X$ and each $y \in T(x)$
b)
i) $T$ is closed, or
ii) there exists a continuous function $\psi: \mathbb{R}_{+}^{5} \rightarrow \mathbb{R}_{+}$such that:

1) $H(T x, T y) \leq \psi(d(x, y), D(x, T x), D(y, T y), D(x, T y), D(y, T x))$, for each $x, y \in X$
2) $\psi(0,0, r, r, 0)<r$ if $r>0$
3) If $u_{1}, u_{2}, v_{1}, v_{2} \in \mathbb{R}$ with $u_{1} \leq u_{2}$ and $v_{1} \leq v_{2}$ then $\psi\left(u, u_{1}, v, w, v_{1}\right) \leq$ $\psi\left(u, u_{2}, v, w, v_{2}\right)$ for all $u, v, w \in \mathbb{R}_{+}$.

Then $T$ is a $c-M W P$ operator. (with $c=\frac{1-\gamma-\mu}{1-\alpha-\beta-\gamma-2 \mu}$.)
Remark 22. Theorem 21 generalizes several known results in the literature, such as Nadler [29], Covitz-Nadler [11], Reich [35], I. A. Rus [38], etc.

## 4 Multivalued Picard operators

The notion of multivalued Picard operator is now presented.
Definition 23. (Petruşel-Rus [34]) Let ( $X, d$ ) be a complete metric space and $T: X \rightarrow P(X)$. By definition, $T$ is called a multivalued Picard operator (MP operator) if:
(i) $(S F)_{T}=F_{T}=\left\{x^{*}\right\}$
(ii) $T^{n}(x) \rightarrow\left\{x^{*}\right\}$, as $n \rightarrow \infty$, for each $x \in X$.

Example 24. Let $T: X \rightarrow P_{c l}(X)$ be a multivalued $\delta$-Reich type operator, i.e. there exist $\alpha, \beta, \gamma \in \mathbb{R}_{+}$with $\alpha+\beta+\gamma<1$ such that: $\delta(T(x), T(y)) \leq$ $\alpha d(x, y)+\beta D(x, T(x))+\gamma D(y, T(y))$, for all $x, y \in X$. Then $T$ is a $M P$ operator. (Reich [35])
Definition 25. (Tarafdar-Yuan [50], see [54]) Let $X$ be a topological space and $T: X \rightarrow P_{c l}(X)$. Then $T$ is said to be a topological contraction if:
a) $T$ is u.s.c.
b) for each $Y \in P_{c l}(X)$ with $T(Y)=Y$ it follows that $Y$ is a singleton, i. e. $Y=\left\{x^{*}\right\}$.

Then we have:
Theorem 26. If $(X, d)$ is a compact metric space and $T: X \rightarrow P_{c l}(X)$ is a l.s.c. topological contraction, then $T$ is a MP operator.

Proof. $T$ being continuous it follows that $\hat{T}$ is a continuous operator from $P_{c p}(X)$ to itself. Moreover $\hat{T}$ is a (singlevalued) topological contraction (since $\hat{T}(Y)=Y$ implies $T(Y)=Y$ and so $Y=\left\{x^{*}\right\}$ ) and so taking into account of Corollary 9.3.16. from Yuan [54], page 559, we obtain that there exists an unique fixed point $A^{*} \in P_{c l}(X)$ of $\hat{T}$ and $T^{n}(Y)$ converges to $A^{*}$, for each $Y \in P_{c l}(X)$.

On the other side, from Tarafdar-Yuan theorem [50], (see Yuan [54], page 559, Theorem 9.3.14.) applied to $T$, we have that there exists an unique $x^{*} \in X$ such that $T\left(x^{*}\right)=\left\{x^{*}\right\}=\bigcap_{n \geq 0} T^{n}(X)$. Let $x \in F_{T}$ be arbitrary. Then $x \in T(x) \subset T^{2}(x) \subset \cdots \subset T^{n}(x) \subset \cdots$. So, $x \in T^{n}(x) \subset T^{n}(X)$, for each $n \in \mathbb{N}$. Hence $x=x^{*}$ and so $F_{T}=(S F)_{T}=A^{*}=\left\{x^{*}\right\}$.

## 5 Iterated function systems

If $f_{i}, i \in\{1, \ldots, m\}$ are continuous operators of $X$ into itself, then a nonempty compact set $Y$ in $X$ is said to be self-similar if it satisfies the
condition $Y=\bigcup_{i=1}^{m} f_{i}(Y)$. Obviously, we may regard the above relation as a fixed point problem for the operator $T_{f}:\left(P_{c p}(X), H\right) \rightarrow\left(P_{c p}(X), H\right)$ defined by $T_{f}(Y)=\bigcup_{i=1}^{m} f_{i}(Y)$. Then $f=\left\{f_{1}, f_{2}, \ldots, f_{m}\right\}$ is said to be an iterated function system (briefly IFS).

The Hausdorff dimension of a self-similar set $Y$ is not, in general, an integer. For this reason, $Y$ is a fractal and $P_{c p}(X)$ is called the space of fractals. The mathematical study of self-similar sets in connection with the mathematics of fractals was initiated by Mandelbrot and then developed by Barnsley, Hutchinson and Hata.

Definition 27. Let $f_{i}: X \rightarrow X, i \in\{1, \ldots, m\}$ be a finite family of continuous operators. Let us define $T_{f}:\left(P_{c p}(X), H\right) \rightarrow\left(P_{c p}(X), H\right)$ by $T_{f}(Y)=\bigcup_{i=1}^{m} f_{i}(Y)$. Then, $T_{f}$ is the Barnsley-Hutchinson operator generated by the iterated function system $f=\left(f_{1}, f_{2}, \ldots, f_{m}\right)$.

Theorem 28. (Barnsley-Hutchinson [7], [18]) Let ( $X, d$ ) be a complete metric space and $f_{i}: X \rightarrow X$ be $a_{i}$-contractions, for $i \in\{1,2, \ldots, m\}$. Then the Barnsley-Hutchinson operator $T_{f}$ generated by the iterated function system $f=\left\{f_{1}, f_{2}, \ldots, f_{m}\right\}$ is a $\max _{1 \leq i \leq m}\left(a_{i}\right)$-contraction on the complete metric space $\left(P_{c p}(X), H\right)$ and has a unique fixed point (i. e. a self-similar set for $f$ ) $A^{*} \in P_{c p}(X)$. Moreover, for each compact subset $A_{0}$ of $X$ the sequence of successive approximations $\left(T_{f}^{n}\left(A_{0}\right)\right)_{n \in N}$ converges to $A^{*}$.
Remark 29. By definition, the set $A^{*}$ is called the attractor of the system $f=\left(f_{1}, f_{2}, \ldots, f_{m}\right)$. Hence, previous theorem is an existence result of an attractor.

Remark 30. The similarity dimension $d_{s}\left(A^{*}\right)$ of a self-similar set $A^{*}$ corresponding to an iterated functions system $f=\left\{f_{1}, f_{2}, \ldots, f_{m}\right\}$, where $f_{i}$ is an $a_{i}$-contraction, for each $i \in\{1,2, \ldots, m\}$, is defined as the unique positive root of the equation $\sum_{i=1}^{m} a_{i}^{d}=1$. It is known that $d_{s}\left(A^{*}\right) \geq d_{H}\left(A^{*}\right)$, for each self-similar set $A^{*}$, where $d_{H}\left(A^{*}\right)$ denotes the Hausdorff dimension of $A^{*}$. The answer when the previous inequality becomes an equality was given by Hutchinson.

Definition 31. Let $X$ be a Banach space. The operator $f: X \rightarrow X$ is said to be an $a$-similar contraction if $a \in] 0,1[$ and $\|f(x)-f(y)\|=a \cdot\|x-y\|$, for each $x, y \in X$.

Let $X=\mathbb{R}^{n}$. Then the IFS $f=\left(f_{1}, f_{2}, \ldots, f_{m}\right)$ satisfies the open set condition if there exists a nonempty bounded open set $U$ of $\mathbb{R}^{n}$ such that $f_{i}(U) \subset U$, for each $i \in\{1,2, \ldots, m\}$ and $f_{i}(U) \cap f_{j}(U)=\emptyset$, for $i \neq j$.

Theorem 32. For a self-similar set $A^{*} \in P_{c p}\left(\mathbb{R}^{n}\right)$ defined by a family of similar contractions which satisfies the open set condition, we have that $d_{s}\left(A^{*}\right)=d_{H}\left(A^{*}\right)$.

Example 33. Consider the mappings $f_{1}, f_{2}: \mathbb{R} \rightarrow \mathbb{R}$ defined by:

$$
f_{1}(x)=\frac{x}{3}, \quad f_{2}(x)=1-\frac{x}{3}
$$

then it is easily to see that $f_{1}, f_{2}$ are $\frac{1}{3}$-contraction mappings, having an unique fixed point $A^{*}$. The set $A^{*}$ is self-similar. Moreover, it is a fractal, namely the well-known Cantor set, $i$. e.

$$
A^{*}=\left\{x \in[0,1] \left\lvert\, x=\sum_{i=1}^{\infty} \frac{a_{n}}{3^{n}}\right., a_{n} \in\{0,2\}\right\} .
$$

The similarity dimension $d$ is the positive solution of the equation: $\left(\frac{1}{3}\right)^{d}+$ $\left(\frac{1}{3}\right)^{d}=1$ and we obtain $d=\operatorname{dim}_{S}\left(A^{*}\right)=\frac{\log 2}{\log 3}$.

Example 34. Let $f_{1}, f_{2}: \mathbb{C} \rightarrow \mathbb{C}$ be mappings

$$
f_{1}(z)=\omega z, \quad f_{2}(z)=\bar{\omega}(\bar{z}-1)+1
$$

where $\omega=\frac{1}{2}+i \frac{\sqrt{3}}{6}$. Then $f_{1}, f_{2}$ are $\frac{1}{\sqrt{3}}$-contraction mappings having as unique self-similar set the set $A^{*}$, which is the Koch curve. This set is a fractal too and its similarity dimension is $\operatorname{dim}_{S}\left(A^{*}\right)=\frac{\log 4}{\log 3}$.

For other examples and related results see the nice book of Yamaguti, Hata and Kigami [53].

It is well known (see Example 5) that a Meir-Keeler type operator on a complete metric space is a Picard operator.

Theorem 35. (Petruşel [33]) Let ( $X, d$ ) be a complete metric space and $f_{i}: X \rightarrow X$, for $i \in\{1,2, \ldots, m\}$ are Meir-Keeler type operators. Then the fractal operator $T_{f}:\left(P_{c p}(X), H\right) \rightarrow\left(P_{c p}(X), H\right)$ is a Meir-Keller type operator and hence $F_{T_{f}}=\left\{A^{*}\right\}$ and $\left(T_{f}^{n}(A)\right)_{n \in \mathbb{N}}$ converges to $A^{*}$, for each $A \in P_{c p}(X)$.

Proof. We shall prove that for each $\eta>0$ there is $\delta>0$ such that the following implication holds

$$
\eta \leq H(A, B)<\eta+\delta \text { we have } H(T(A), T(B))<\eta
$$

Let us consider $A, B \in P_{c p}(X)$ such that $\eta \leq H(A, B)<\eta+\delta$.
If $u \in T(A)$ then there exists $j \in\{1, \ldots, m\}$ and $x \in A$ such that $u=f_{j}(x)$.

For $x \in A$ we can choose $y \in B$ such that $d(x, y) \leq H(A, B)<\eta+\delta$. We have the following alternative:

If $d(x, y) \geq \eta$ then $\eta \leq d(x, y)<\eta+\delta$ implies $d\left(f_{j}(x), f_{j}(y)\right)<\eta$. Hence $D(u, T(B)) \leq d\left(u, f_{j}(y)\right)<\eta$.

On the other hand, if $d(x, y)<\eta$ then from the definition of the MeirKeeler type operator we have $d\left(f_{j}(x), f_{j}(y)\right)<d(x, y)<\eta$ and again the conclusion $D(u, T(B))<\eta$.

Because $T(A)$ is compact we have that $\rho(T(A), T(B))<\eta$.
Interchanging the roles of $T(A)$ and $T(B)$ we obtain $\rho(T(B), T(A))<\eta$ and hence $H(T(A), T(B))<\eta$, showing the fact that $T$ is a Meir-Keelertype operator. From Meir-Keeler fixed point result (Example 3) we obtain that there exists an unique $A^{*} \in P_{c p}(X)$ such that $T\left(A^{*}\right)=A^{*}$.

Remark 36. It is an open problem to compute the Hausdorff dimension of the self-similar set in the above theorem.

## 6 Iterated multifunction systems

Let $(X, d)$ be a complete metric space and $F_{1}, \ldots, F_{m}: X \rightarrow P(X)$ be multivalued operators.

The system $F=\left(F_{1}, \ldots, F_{m}\right)$ is called an iterated multifunction system (briefly IMS).

If $F=\left(F_{1}, \ldots, F_{m}\right)$ is an IMS such that each $F_{i}: X \rightarrow P_{c p}(X)$ is u. s. c., for $i \in\{1, \ldots, m\}$ then $\tilde{F}: X \rightarrow P_{c p}(X)$ given by the formula $\tilde{F}(x)=\bigcup_{i=1}^{m} F_{i}(x), x \in X$ is called the Barnsley-Hutchinson multifunction generated by the IMS $F$.

In the same setting, an operator $T_{F}: P_{c p}(X) \rightarrow P_{c p}(X)$, defined by $T_{F}(Y)=\bigcup_{i=1}^{m} F_{i}(Y)$ is called the multi-fractal operator generated by the IMS $F$. A fixed point of $T_{F}$ is, by definition, a multivalued fractal.

Example 37. If $F_{i}: X \rightarrow P_{c p}(X)$ are multivalued $a_{i}$-contractions (for $i \in\{1, \ldots, m\})$ then the multi-fractal operator $T_{F}$ is a single-valued $a$ contraction (where $a=\max _{1 \leq i \leq m} a_{i}$ ) and hence it is a Picard operator. The unique fixed point $A_{F}^{*} \in \bar{P}_{c p}(X)$ of $T_{F}$ is a multivalued fractal. $A_{F}^{*}$ is also called an attractor of the IMS F. (Nadler [29])

Example 38. Let $(X, d)$ be a complete metric space and $\varphi:[0, \infty[\rightarrow[0, \infty[$ such that:
(i) $\varphi$ is continuous and nondecreasing
(ii) $\varphi(0)=0$ and $0<\varphi(t)<t$, for each $t>0$
(iii) $\lim _{t \rightarrow \infty}(t-\varphi(t))=\infty$.

An operator $f: X \rightarrow X$ is said to be strict $\varphi$-contraction or weakly contractive if for each $x, y \in X$ we have $d(f(x), f(y)) \leq \varphi(d(x, y))$.

A multivalued operator $T: X \rightarrow P_{c p}(X)$ is said to be multivalued strict $\varphi$-contraction or multivalued weakly contractive if for each $x, y \in X$ we have $H(T(x), T(y)) \leq \varphi(d(x, y))$.

If $F$ is an IMS of strict $\varphi$-contractions, then the multi-fractal operator $T_{F}$ is a strict $\varphi$-contraction having an unique multivalued fractal $A_{F}^{*}$. Moreover $T_{F}$ is a Picard operator. (Browder [10], Boyd-Wong [9], Andres-Fišer [4], Rhoades [36], Rus [37].)

Example 39. (Petruşel [30]) $F: X \rightarrow P(X)$ is a multivalued Meir-Keeler type operator if:
for each $\eta>0$ there exists $\delta>0$ such that $x, y \in X, \eta \leq d(x, y)<$ $\eta+\delta$ we have $H(F(x), F(y))<\eta$.

It is known that any multivalued Meir-Keeler type operator is contractive (i.e. $H(F(x), F(y))<d(x, y)$, for each $x, y \in X$, with $x \neq y$ ) and hence $u$. s. $c$.

If $F=\left(F_{1}, \ldots, F_{m}\right)$ is an IMS with $F_{i}: X \rightarrow P_{c p}(X)$ satisfying to a Meir-Keeler type condition, for $i \in\{1, \ldots, m\}$ then the multi-fractal operator $T_{F}$ is a (singlevalued) Meir-Keeler operator and has an unique multivalued fractal $A_{F}^{*}$. Moreover, $T_{F}$ is a Picard operator.

For the case of iterated multifunction systems, there are important contributions of Andres-Fišer [4], Andres-Górniwewicz [3], Jachymski [19], La-sota-Myjak [21], [22], etc.

An important abstract result is:
Theorem 40. (Petruşel [31]) Let $F=\left(F_{1}, \ldots, F_{m}\right)$ be an IMS such that $F_{i}: X \rightarrow P_{c p}(X)$ is u. s. c. and MWP operator, for $i \in\{1, \ldots, m\}$.

Consider the Barnsley-Hutchinson multifunction $\tilde{F}: X \rightarrow P_{c p}(X) \tilde{F}(x)=$ $m$ $\bigcup^{m} F_{i}(x), x \in X$, generated by the IMS $F$ and the multi-fractal operator $T_{F}$ ${ }_{i=1}$
generated by the same IMS $F=\left(F_{1}, \ldots, F_{m}\right)$, i.e. $T_{F}: P_{c p}(X) \rightarrow P_{c p}(X)$, $T_{F}(Y)=\bigcup_{i=1}^{m} F_{i}(Y)$, for each $Y \in P_{c p}(X)$.

Then we have:
i) $\tilde{F}$ is u. s. c.
ii) $\tilde{F}$ is MWP operator.
iii) $T_{F}=T_{\tilde{F}}$.

Remark 41. From iii) of the above theorem we can conclude that the dynamics of the IMS $F=\left(F_{1}, \ldots, F_{m}\right)$ coincide with those generated by an unique operator $\tilde{F}$. So we will consider for the rest of the paper the case of an unique multivalued operator.

The following result was established in [34].
Theorem 42. (Petruşel-Rus [34]) Let $(X, d)$ be a complete metric space and $U \subseteq P_{c l}(X)$ be such that $x \in X$ implies $\{x\} \in U$. Let $T: X \rightarrow U$ be an u. s. c. multivalued operator such that $A \in U$ implies $T(A) \in U$. We suppose that $\hat{T}: U \rightarrow U$ is a Picard operator and denote by $A_{T}^{*}$ the unique fixed point of $\hat{T}$. Then:
(i) $F_{T} \subset A_{T}^{*}$ and $A_{T}^{*}=\bigcup_{n \in \mathbb{N}^{*}} T^{n}(x)$, for each $x \in F_{T}$.
(ii) If $T(x) \in P_{c p}(X)$, for each $x \in X$ then $F_{T}$ is a compact set.
(iii) If $T\left(F_{T}\right)=F_{T}$ and $F_{T} \in U$ then $\lim _{n \rightarrow \infty} T^{n}(x)=F_{T}$, for each $x \in X$.
(iv) If $(S F)_{T} \neq \emptyset$ then $F_{T}=(S F)_{T}=\left\{\begin{array}{l}n \rightarrow \infty \\ *\end{array}\right\}$.

Proof. (i) Let $x \in F_{T}$. Then we have $x \in T(x) \subset T^{2}(x) \subset \cdots \subset T^{n}(x) \subset \cdots$. So, $x \in T^{n}(x)$, for each $n \in \mathbb{N}$ and $\lim _{n \rightarrow \infty} T^{n}(x)=\bigcup_{n \in \mathbb{N}^{*}} T^{n}(x)$.
(ii) From the completeness of the space $\left(P_{c p}(X), H\right)$ and the upper semicontinuity of $T$ it follows that $A_{T}^{*} \in P_{c p}(X)$. From (i) $F_{T} \subset A_{T}^{*}$. On the other hand $F_{T}$ is a closed set and hence it is compact too.
(iii) The conclusion follows using the fact that $F_{T}=A_{T}^{*}$.
(iv) Let $x^{*} \in(S F)_{T}$. Then $\left\{x^{*}\right\} \in U$ and $T\left(x^{*}\right)=\left\{x^{*}\right\}$. Hence we have that $A_{T}^{*}=\left\{x^{*}\right\}$ and from (i) we get the conclusion $F_{T}=\left\{x^{*}\right\}$.

Let us discuss now some particular cases:

Corollary 43. If $U=P_{c p}(X)$ and $T$ is a multivalued contraction, then Theorem 42 contain some already known results, but obtained by other techniques, as follows:
a) assertion (ii) was proved by Saint-Raymond [46].
b) assertion (iv) is from I. A. Rus [37].

Corollary 44. If $U=P_{c p}(X)$ and $T$ is a multivalued strict $\varphi$-contraction, then the fractal operator $\hat{T}$ is a (singlevalued) strict $\varphi$-contraction. Hence $\hat{T}$ is a Picard operator and we get all the conclusions of Theorem 42

Corollary 45. If $U=P_{c p}(X)$ and $T$ is a multivalued Meir-Keeler operator, then $\hat{T}$ is a (singlevalued) Meir-Keeler operator and hence a Picard operator. So the conclusions of Theorem 42 hold.

For example, an new interesting result follows, namely that the fixed point set of the Meir-Keeler multifunction with compact values is compact.

In order to consider the cases $U=P_{b, c l}(X)$ and $U=P_{c l}(X)$, let us define another fractal type operator. Again $(X, d)$ is a metric space and $T: X \rightarrow P(X)$ a multivalued operator. Then $\check{T}: P_{c l}(X) \rightarrow P_{c l}(X)$, given by

$$
\check{T}(Y):=\overline{\bigcup_{x \in Y} T(x)}, \text { for } Y \in P_{c l}(X)
$$

is called the extended fractal operator generated by $T$.
We have:
Theorem 46. Let $(X, d)$ be a complete metric space and $U \subseteq P_{c l}(X)$ be such that $x \in X$ implies $\{x\} \in U$. Let $T: X \rightarrow U$ be an u. s. c. multivalued operator such that $A \in U$ implies $\overline{T(A)} \in U$. We suppose that $\check{T}: U \rightarrow U$ is a Picard operator and denote by $A_{T}^{*}$ its unique fixed point. Then:
(i) $F_{T} \subset A_{T}^{*}$ and $A_{T}^{*}=\bigcup_{n \in \mathbb{N}^{*}} T^{n}(x)$, for each $x \in F_{T}$.
(ii) If $\overline{T\left(F_{T}\right)}=F_{T}$ and $F_{T} \in U$ then $\lim _{n \rightarrow \infty} T^{n}(x)=F_{T}$, for each $x \in X$.
(iii) If $(S F)_{T} \neq \emptyset$ then $F_{T}=(S F)_{T}=\left\{x^{*}\right\}$.

Proof. (i) Let $x \in F_{T}$. Then we have $x \in T(x) \subset T^{2}(x) \subset \cdots \subset T^{n}(x) \subset \cdots$. Hence $x \in T^{n}(x)$, for each $n \in \mathbb{N}$. Then we get $x \in \bigcup_{n \in \mathbb{N}^{*}} T^{n}(x)=\lim _{n \rightarrow \infty} T^{n}(x)$. But $A_{T}^{*}$ is the unique fixed point of $\check{T}$ and $\check{T}$ is Picard. So, for each $A_{0} \in U$ the sequence of successive approximations $\check{T}^{n}\left(A_{0}\right)$ converges to $A_{T}^{*}$. Taking $A_{0}=\{x\}$ we obtain that $\check{T}^{n}\left(A_{0}\right)=\check{T}^{n}(\{x\})=T^{n}(x)$. Hence $\lim _{n \rightarrow \infty} T^{n}(x)=$ $A_{T}^{*}$ and we got the desired conclusions: $x \in A_{T}^{*}$ and $\bigcup_{n \in \mathbb{N}^{*}} T^{n}(x)=A_{T}^{*}$.
(ii) $\overline{T\left(F_{T}\right)}=F_{T}$ and $F_{T} \in U$ implies that $\check{T}\left(F_{T}\right)=F_{T}$ and hence $F_{T}=A_{T}^{*}=\lim _{n \rightarrow \infty} T^{n}(x)$, for each $x \in X$.
(iii) If $(S F)_{T} \neq \emptyset$ let $x^{*} \in X$ with $\left\{x^{*}\right\}=T\left(\left\{x^{*}\right\}\right)$. Then $\left\{x^{*}\right\}=$ $\check{T}\left(\left\{x^{*}\right\}\right)$ and so $A_{T}^{*}=\left\{x^{*}\right\}$. Then conclusion follows now by (i).

Remark 47. The above theorem works at least for the following particular cases:

1) If $U=P_{b, c l}(X)$ and $T$ is a multivalued a-contraction. In particular, we obtain that $F_{T}$ is bounded.
2) If $U=P_{b, c l}(X)$ and $T$ is a multivalued strict $\varphi$-contraction. Again, we can observe that $F_{T}$ is bounded.
3) If $U=P_{c l}(X)$ and $T$ is a multivalued a-contraction.
4) If $U=P_{c l}(X)$ and $T$ is a multivalued strict $\varphi$-contraction.

Remark 48. If $U=P_{b, c l}(X)$ or $U=P_{c l}(X)$ and $T$ a multivalued MeirKeeler operator, it is an open question if $\check{T}$ is a Picard operator.

## Acknowledgement

This paper is a concise version of some lessons presented in October 2003 in the Seminario de Análisis Functional of the Departamento de Análisis Matemático, Universidad de Sevilla. I wish to express my deep gratitude to the whole department and in particular to Genaro López and Rafa Espínola for the invitation to Sevilla University and their warm hospitality during the visit.

## References

[1] R. P. Agarwal, D. O'Regan, Fixed-point theorems for multivalued maps with closed values on complete gauge spaces, Appl. Math. Lett., 14 (2001), 831-836.
[2] A. Alesina, S. Massa, D. Roux, Punti uniti di multifunzioni con condizioni di tipo Boyd-Wong, Boll. U. M. I., 8 (1973), 29-34.
[3] J. Andres, L. Górniewicz, On the Banach contraction principle for multivalued mappings, Approximation, optimization and mathematical economics (Pointe-à-Pitre, 1999), 1-23, Physica, Heidelberg, 2001.
[4] J. Andres, J. Fišer, Metric and topological multivalued fractals, Int. J. Bifurc. Chaos, 14 (2004), (to appear).
[5] J. -P. Aubin, H. Frankowska, Set-valued analysis, Birkhauser, Basel, 1990.
[6] Y. M. Ayerbe Toledano, T. Dominguez Benavides, G. López Acedo, Measures of noncompactness in metric fixed point theory, Birkhäuser Verlag, Basel, 1997.
[7] M. F. Barnsley, Fractals everywhere., Academic Press, Boston, 1988.
[8] G. Beer, Topologies on closed and closed convex sets, Kluwer Acad. Publ., Dordrecht, 1993.
[9] D. W. Boyd, J. S. W. Wong, On nonlinear contractions, Proc. A.M.S., 20 (1969), 458-464.
[10] F. E. Browder, On the convergence of successive approximations for nonlinear functional equations, Indag. Math. 30 (1968), 27-35.
[11] H. Covitz, S.B. Nadler jr., Multivalued contraction mappings in generalized metric spaces, Israel J. Math., 8 (1970), 5-11.
[12] M. Fréchet, Les espaces abstraits, Gauthier-Villars, Paris, 1928.
[13] G. Garegnani, C. Zanco, Fixed points of somehow contractive multivalued mappings, Istit. Lombardo Accad. Sci. Lett. Rend. A, 114 (1980), 138-148
[14] V. Glăvan, V. Guţu, On the dynamics of contracting relations, in Analysis and Optimization of Differential systems (Eds. V. Barbu, I. Lasiecka, D. Tiba, C. Varsan), Kluwer Acad. Publ., 2003, 179-188.
[15] K. Goebel, W. A. Kirk, Topics in metric fixed point theory, Cambridge Univ. Press, Cambridge, 1990.
[16] L. Górniewicz, Topological fixed point theory of multivalued mappings, Kluwer Acad. Publ., Dordrecht, 1999.
[17] S. Hu, N. S. Papageorgiou, Handbook of multivalued analysis, Vol. I and II, Kluwer Acad. Publ., Dordrecht, 1997 and 1999.
[18] J. E. Hutchinson, Fractals and self-similarity, Indiana Univ. Math. J., 30 (1981), 713-747.
[19] J. Jachymski, Continuous dependence of attractors of iterated function systems, J. Math. Anal. Appl., 198 (1996), 221-226.
[20] W. A. Kirk, B. Sims (editors), Handbook of metric fixed point theory, Kluwer Acad. Publ., Dordrecht, 2001.
[21] A. Lasota, J. Myjak, Attractors of multifunctions, Bull. Polish Acad. Sci. Math., 48 (2000), 319-334.
[22] A. Lasota, J. Myjak, Semifractals on Polish spaces, Bull. Polish Acad. Sci. Math., 46 (1998), 179-196.
[23] A. Latif, I. Beg, Geometric fixed points for single and multivalued mappings, Demonstratio Math., 30 (1997), 791-800.
[24] T. -C. Lim, On fixed point stability for set-valued contractive mappings with applications to generalized differential equations, J. Math. Anal. Appl., 110 (1985), 436-441.
[25] B. Mandelbrot, Les objets fractals. Forme, hasard et dimension, Flammarion, Paris, 1975, (in French).
[26] S. Massa, Generalized multicontractive mappings, Riv. Mat. Univ. Parma, 6 (1980), 103-110.
[27] A. Meir, E. Keeler, A theorem on contraction mappings, J. Math. Anal. Appl., 28 (1969), 326-329.
[28] A. Muntean, Fixed point principles and applications to mathematical economics, Cluj University Press, 2002.
[29] S. B. Nadler jr., Multivalued contraction mappings, Pacific J. Math., 30 (1969), 475-488.
[30] A. Petruşel, Singlevalued and multivalued Meir-Keeler type operators, Revue D'Analse Num. et de Th. de l'Approx., Tome 30 (No. 1), 2001, 75-80.
[31] A. Petruşel, Multivalued weakly Picard operators and applications, Scienticae Mathematicae Japonicae, 38 (2004), 169-202.
[32] A. Petruşel, I. A. Rus, Dynamics on $\left(P_{c p}(X), H_{d}\right)$ generated by a finite family of multi-valued operators on ( $X, d$ ), Math. Moravica, 5 (2001), 103-110.
[33] A. Petruşel, Dynamic systems, fixed points and fractals, Pure Math. Appl., 13 (2002), 275-281.
[34] A. Petruşel, I. A. Rus, Multivalued weakly Picard and multivalued Picard operators, Proceedings of the International Conference on Fixed Point Theory, Valencia, 13-19 July 2003, to appear.
[35] S. Reich, Fixed point of contractive functions, Boll. U. M. I., 5 (1972), 26-42.
[36] B. E. Rhoades, Some theorems on weakly contractive maps, Nonlinear Analysis, 47 (2001), 2683-2693.
[37] I. A. Rus, Generalized contractions and applications, Cluj Univ. Press, 2001.
[38] I. A. Rus, Fixed point theorems for multivalued mappings in complete metric spaces, Math. Japonica, 20 (1975), 21-24.
[39] I. A. Rus, Weakly Picard mappings, Comment. Math. Univ. Carolinae, 34 (1993), 769-773.
[40] I. A. Rus, Stability of attractor of a $\varphi$-contractions system, Seminar on Fixed Point Theory, Babes-Bolyai Univ., 1998, 31-34.
[41] I. A. Rus, Picard operators and applications, Scientia Mathematicae Japonicae, 58 (2003), 191-219.
[42] I. A. Rus, S. Mureşan, Data dependence of the fixed points set of some weakly Picard operators, "Tiberiu Popoviciu" Itinerant Seminar of Functional Equations, Approximation and Convexity, Cluj-Napoca, 2000, 201-207.
[43] I. A. Rus, A. Petruşel, G. Petruşel, Fixed point theory 1950-2000 : Romanian contributions, House of the Book of Science, Cluj-Napoca, 2002.
[44] I. A. Rus, A. Petruşel and A. Sîntămărian, Data dependence of the fixed point set of multivalued weakly Picard operators, Studia Univ. Babes-Bolyai, Mathematica, 46 (2001), 111-121.
[45] I. A. Rus, A. Petruşel and A. Sîntămărian, Data dependence of the fixed point set of some multivalued weakly Picard operators, Nonlinear Analysis, 52 (2003), no. 8, 1947-1959.
[46] J. Saint-Raymond, Multivalued contractions, Set-Valued Analysis, 2 (1994), 559-571.
[47] A. Sîntămărian, Weakly Picard pairs of multivalued operators, Mathematica, (to appear).
[48] A. Sîntămărian, Metrical strict fixed point theorems for multivalued mappings, Sem. on Fixed Point Theory, Preprint no. 3 (1997), BabeşBolyai Univ. Cluj-Napoca, 27-30.
[49] W. Strother, Fixed points, fixed sets, and M-retracts, Duke Math. J., 22 (1955), 551-556.
[50] E. Tarafdar, G. X. -Z. Yuan, Set-valued contraction mapping principle, Applied Math. Letter, 8 (1995), 79-81.
[51] T. Wang, Fixed point theorems and fixed point stability for multivalued mappings on metric spaces, J. Nanjing Univ., Math. Biq., 6 (1989), 16-23.
[52] H.- K. Xu, $\varepsilon$-chainability and fixed points of set-valued mappings in metric spaces, Math. Japonica, 39 (1994), 353-356.
[53] M. Yamaguti, M. Hata, J. Kigani, Mathematics of fractals, Translations Math.Monograph, Vol. 167, AMS Providence, RI, 1997.
[54] G. X. Z. Yuan, KKM theory and applications in nonlinear analysis, Marcel Dekker, New York, 1999.

## Author's address:

Adrian Petruşel<br>Department of Applied Mathematics<br>Babeş-Bolyai University<br>3400, Cluj-Napoca, Romania<br>E-mail: petrusel@math.ubbcluj.ro

# Equilateral sets in finite-dimensional normed spaces 

Konrad J. Swanepoel *


#### Abstract

This is an expository paper on the largest size of equilateral sets in finite-dimensional normed spaces.


## 1 Introduction

In this paper I discuss some of the more important known results on equilateral sets in finite-dimensional normed spaces with an analytical flavour. I have omitted a discussion of some of the low-dimensional results of a discrete geometric nature. The omission of a topic has no bearing on its importance, only on my prejudices and ignorance. The following topics are outside the scope of this paper: almost-equilateral sets, $k$-distance sets, and equilateral sets in infinite dimensional spaces. However, see the last section for references to these topics.

I have included the proofs that I find fascinating. The arguments are mostly very geometrical, using results from convex geometry and the local theory of Banach spaces, and also from algebraic topology. In the case of the $\ell_{p}$-norm the known results use tools from linear algebra, probability theory, combinatorics and approximation theory.

Equilateral sets have been applied in differential geometry to find minimal surfaces in finite-dimensional normed spaces [24, 28].

Throughout the paper I state problems. Perhaps some of them are easy, but I do not know the answer to any of them.

### 1.1 Definitions

A subset $S$ of a metric space $(X, d)$ is $\lambda$-equilateral $(\lambda>0)$ if $d(x, y)=\lambda$ for all $x, y \in S, x \neq y$. Let $e_{\lambda}(X)$ denote the largest size of a $\lambda$-equilateral

[^8]set in $X$, if it exists, otherwise set $e_{\lambda}(X)=\infty$. If $X$ is compact then it is easily seen that $e_{\lambda}(X)<\infty$ for each $\lambda>0$.

In a normed space $X$, if $S$ is $\lambda$-equilateral, then $\frac{\mu}{\lambda} S$ is $\mu$-equilateral for any $\mu>0$; thus the specific value of $\lambda$ does not matter, and we only write $e(X)$. If $X$ is finite-dimensional, then $e(X)<\infty$. In the remainder of the paper we abbreviate finite-dimensional normed space to Minkowski space. The Banach-Mazur distance between two Banach spaces $X$ and $Y$ is the infimum of all $c \geq 1$ such that there exists a linear isomorphism $T: X \rightarrow Y$ satisfying $\|T\|\left\|T^{-1}\right\| \leq c$. In the finite-dimensional case, this infimum is of course attained. We always use $n$ to denote the dimension of $X$, and $m$ the size of a set. We use $\mathbb{R}$ for the field of real numbers, $\mathbb{R}^{n}$ for the $n$-dimensional real vector space of $n$-dimensional column vectors, and $\mathbf{e}_{i}$ for the $i$ th standard basis vector which has 1 in position $i$ and 0 in all other positions. The transpose of a matrix $A$ is denoted by $A^{\text {tr }}$. For a set $S \subseteq \mathbb{R}^{n}, \operatorname{conv}(S), \operatorname{int}(S), \operatorname{vol}(S)$ denotes its convex hull, interior, and Lebesgue measure (if $S$ is measurable), respectively. By $\log x$ we denote the natural logarithm, and by $\lg x$ the logarithm with base 2 .

### 1.2 Kusner's questions for the $\ell_{p}$ norms

On $\mathbb{R}^{n}$ the $\ell_{p}$ norm is defined by

$$
\left\|\left(x_{1}, x_{2}, \ldots, x_{n}\right)\right\|_{p}:=\left(\sum_{i=1}^{n}\left|x_{i}\right|^{p}\right)^{1 / p}
$$

if $1<p<\infty$, and for $p=\infty$ by

$$
\left\|\left(x_{1}, x_{2}, \ldots, x_{n}\right)\right\|_{\infty}:=\max \left\{\left|x_{i}\right|: i=1, \ldots, n\right\} .
$$

We denote the Minkowski space $\left(\mathbb{R}^{n},\|\cdot\|_{p}\right)$ by $\ell_{p}^{n}$. The $\ell_{p}$ distance between two points $\mathbf{a}, \mathbf{b} \in \mathbb{R}^{n}$ is $\|\mathbf{a}-\mathbf{b}\|_{p}$. Thus Euclidean space is $\ell_{2}^{n}$.

The following lower bounds for $e\left(\ell_{p}^{n}\right)$ are simple:

- Since $\left\{ \pm \mathbf{e}_{i}: i=1, \ldots, n\right\}$ is 2-equilateral in $\ell_{1}^{n}$, we have $e\left(\ell_{1}^{n}\right) \geq 2 n$.
- For any $p \in(1, \infty)$ and an appropriate choice of $\lambda=\lambda(p) \in \mathbb{R}$, the set

$$
\left\{\mathbf{e}_{1}, \mathbf{e}_{2}, \ldots, \mathbf{e}_{n}, \lambda \sum_{i=1}^{n} \mathbf{e}_{i}\right\}
$$

is $2^{1 / p}$-equilateral in $\ell_{p}^{n}$; hence $e\left(\ell_{p}^{n}\right) \geq n+1$.

- $\left\{\left(\varepsilon_{1}, \ldots, \varepsilon_{n}\right): \varepsilon_{i}= \pm 1\right\}$ is 2-equilateral in $\ell_{\infty}^{n}$; thus $e\left(\ell_{\infty}^{n}\right) \geq 2^{n}$.

A theorem of Petty [29] (see Section 2.1) states that $e(X) \leq 2^{n}$ if the dimension of $X$ is $n$; this then gives $e\left(\ell_{\infty}^{n}\right)=2^{n}$. Kusner [20] asked whether $e\left(\ell_{1}^{n}\right)=2 n$ and $e\left(\ell_{p}^{n}\right)=n+1$ holds for all $1<p<\infty$. For $p$ an even integer it is easy to show an upper bound of about $p n$ using linear algebra (observation of Galvin; see Smyth [33]). It is possible to improve this upper bound to about $p n / 2$, and in the particular case of $p=4$ to show $e\left(\ell_{4}^{n}\right)=n+1[37]$ (see Section 6.1). For $1<p<2$ there are examples known showing $e\left(\ell_{p}^{n}\right)>n+1$ if $n$ is sufficiently large (depending on $p$ ) [37] (see Section 3.2). On the other hand if $p$ is sufficiently near 2 (depending on $n$ ), then $e\left(\ell_{p}^{n}\right)=n+1$. Smyth [33] gives a quantitative estimate (see Section 4.2).

The above results are not difficult; to improve the $2^{n}$ upper bound for $e\left(\ell_{p}^{n}\right)$ turned out to need deeper results from analysis and linear algebra. The first breakthrough came when Smyth [33] combined a linear algebra method with the Jackson theorems from approximation theory to prove $e\left(\ell_{p}^{n}\right)<$ $c_{p} n^{(p+1) /(p-1)}$ for $1<p<\infty$. Then Alon and Pudlák [1] combined Smyth's method with a result on the rank of approximations of the identity matrix (the rank lemma; see Section 6.2) to prove $e\left(\ell_{p}^{n}\right)<c_{p} n^{(2 p+2) /(2 p-1)}$ for $1 \leq$ $p<\infty$. For the $\ell_{1}$ norm Alon and Pudlák combined the rank lemma with a probabilistic argument (randomized rounding) to prove $e\left(\ell_{1}^{n}\right)<c n \log n$, and more generally $e\left(\ell_{p}^{n}\right)<c_{p} n \log n$ if $p$ is an odd integer.

Finally we mention that it is known that $e\left(\ell_{1}^{3}\right)=6$ (Bandelt et al. [3]) and $e\left(\ell_{1}^{4}\right)=8$ (Koolen et al. [23]). We do not discuss the proofs of these two results.

Problem 1 (Kusner). Prove (or disprove) that $e\left(\ell_{1}^{n}\right)=2 n$ for $n \geq 5$.
Problem 2 (Kusner). Prove (or disprove) that $e\left(\ell_{p}^{n}\right)=n+1$ for $p>2$ ( $p \neq 4$ ).

Problem 3. Prove (or disprove) that $e\left(\ell_{p}^{n}\right)<c n$ for $1<p<2$.

## 2 General upper bounds

We now need the following notions from convex geometry. A polytope is a convex body that is the convex hull of finitely many points. A face of a convex body $C$ is a set $F \subseteq C$ satisfying the following property:

If a segment $\mathbf{a b} \subseteq C$ intersects $F$, then $\mathbf{a b} \subseteq F$.

The following are well-known (and easily proved) facts from convex geometry:

1. Every point of a convex body $C$ is contained in the relative interior of a unique face of $C$.
2. The faces of a polytope forms a ranked finite lattice, with dimension as rank function.
3. The 1-dimensional faces are called vertices, and the ( $n-1$ )-dimensional faces facets.
4. Any $(n-2)$-dimensional face is contained in exactly two facets.

A parallelotope is a translate of the polytope $\left\{\sum_{i=1}^{n} \lambda_{i} \mathbf{a}_{i}:-1 \leq \lambda \leq\right.$ $1\}$ where $\mathbf{a}_{1}, \ldots, \mathbf{a}_{n}$ are linearly independent. It is easily seen that if a Minkowski space has a parallelotope as unit ball, then it is isometric to $\ell_{\infty}^{n}$.

### 2.1 The theorem of Petty and Soltan

The following theorem was proved by Petty [29], and later also by P. Soltan [34]. A different proof, using the isodiametric inequality, is given in [15].

Theorem 4 (Petty [29] \& Soltan [34]). For any $n$-dimensional normed space $X, e(X) \leq 2^{n}$ with equality iff $X$ is isometric to $\ell_{\infty}^{n}$, in which case any equilateral set of size $2^{n}$ is the vertex set of some ball (which is then a parallelotope).

Proof. Let $\mathbf{p}_{1}, \ldots, \mathbf{p}_{m} \in X$ be 1-equilateral, where $m=e(X)$. Let $P=$ $\operatorname{conv}\left\{\mathbf{p}_{1}, \ldots, \mathbf{p}_{m}\right\}$ be the convex hull of the points $\mathbf{p}_{i}$. If $P$ is not fulldimensional we may use induction on $n$ to obtain $m \leq 2^{n-1}<2^{n}$ (and the cases $n=0$ and $n=1$ are trivial).

Thus without loss of generality $P$ is full-dimensional. For each $i=$ $1, \ldots, m$, let $P_{i}=\frac{1}{2}\left(P+\mathbf{p}_{i}\right)$. Note the following easily proved facts:

1. Each $P_{i} \subseteq P$.
2. Each $P_{i} \subseteq B\left(\mathbf{p}_{i}, \frac{1}{2}\right)$.
3. For any $i \neq j, P_{i} \cap P_{j}$ has empty interior.

Thus the $P_{i}$ pack $P$, and we now calculate volumes:

$$
\sum_{i=1}^{m} \operatorname{vol}\left(P_{i}\right)=\operatorname{vol}\left(\bigcup_{i=1}^{m} P_{i}\right) \leq \operatorname{vol}(P)
$$

Since $\operatorname{vol}\left(P_{i}\right)=\left(\frac{1}{2}\right)^{n} \operatorname{vol}(P)$, we obtain $m \leq 2^{n}$.
If equality holds, $P$ is a polytope in $\mathbb{R}^{n}$ that is perfectly packed or tiled by $2^{n}$ translates of $\frac{1}{2} P$. By the following lemma $P$ must be a parallelotope $\mathbf{v}+$ $\left\{\sum_{i=1}^{n} \lambda_{i} \mathbf{a}_{i}:-1 \leq \lambda \leq 1\right\}$ with vertex set $\left\{\mathbf{p}_{1}, \ldots, \mathbf{p}_{m}\right\}=\mathbf{v}+\left\{\sum_{i=1}^{n} \varepsilon_{i} \mathbf{a}_{i}\right.$ : $\left.\varepsilon_{i}= \pm 1\right\}$. Since $\mathbf{p}_{i}-\mathbf{p}_{j}$ is a unit vector for $i \neq j$, it follows that $\sum_{i=1}^{n} 2 \varepsilon_{i} \mathbf{a}_{i}$ is a unit vector for all choices of $\varepsilon_{i} \in\{-1,0,1\}$ with not all $\varepsilon_{i}=0$. It is then easy to see that the unit ball is the parallelotope $\left\{\sum_{i=1}^{n} \lambda_{i} \mathbf{a}_{i}:-1 \leq \lambda \leq 1\right\}$ (exercise), and the linear map sending $\mathbf{a}_{i} \mapsto \mathbf{e}_{i}$ is an isometry.

Lemma 5 (Groemer [17]). Let $P$ be the convex hull of $\mathbf{v}_{1}, \ldots, \mathbf{v}_{2^{n}} \in \mathbb{R}^{n}$. Suppose that $P=\bigcup_{i=1}^{2^{n}} \frac{1}{2}\left(P+\mathbf{v}_{i}\right)$. Then $P$ is a parallelotope with vertices $\mathbf{v}_{1}, \ldots, \mathbf{v}_{2^{n}}$.

Proof. Exercise.

### 2.2 Strictly convex norms

A normed space is strictly convex if $\|\mathbf{x}+\mathbf{y}\|<\|\mathbf{x}\|+\|\mathbf{y}\|$ for all linearly independent $\mathbf{x}, \mathbf{y}$, or equivalently, if the boundary of the uniy ball does not contain a line segment. If $X$ is a strictly convex Minkowski space, then $e(X) \leq 2^{n}-1$ by Theorem 4, but as far as I know there is no better upper bound for $n \geq 4$.

Problem 6. If $X$ is an n-dimensional strictly convex Minkowski space, $n \geq 4$, prove that $e(X) \leq 2^{n}-2$.

In Section 3.1 we'll see that $e(X)=3$ when $\operatorname{dim} X=2$, and for $\operatorname{dim} X=$ 3, Petty [29] deduced $e(X) \leq 5$ from a result of Grünbaum [19].

Conjecture 1 (Füredi, Lagarias, Morgan [15]). There exists some constant $\varepsilon>0$ such that if $X$ is a strictly convex $n$-dimensional Minkowski space then $e(X) \leq(2-\varepsilon)^{n}$.

There exist strictly convex spaces with equilateral sets of size at least exponential in the dimension.

Theorem 7 (Füredi, Lagarias, Morgan [15]). For each $n \geq 271$ there exists an n-dimensional strictly convex $X$ with an equilateral set of size at least $1.058^{n}$.

In [15] the lower bound $1.02^{n}$ was obtained; however, their argument easily gives the bound as stated here. They constructed a norm $\|\cdot\|$ on $\mathbb{R}^{n}$ with the property that $\|\cdot\|-\varepsilon\|\cdot\|_{2}$ is still a norm, and asked whether the
norm can in addition be $C^{\infty}$ on $\mathbb{R}^{n} \backslash\{0\}$. In our construction the norm will lack the above property, but it will be $C^{\infty}$. However, most likely the norm in the above theorem can be made to have the above property and be $C^{\infty}$.

We need the following very nice high-dimensional phenomenon. It is a special case of the Johnson-Lindenstrauss flattening lemma [22], although this special case is essentially the Gilbert-Varshamov lower bound for binary codes (see [25].)

Lemma 8. For each $\delta>0$ there exist $\varepsilon=\varepsilon(\delta)>0$ and $n_{0}=n_{0}(\delta) \geq 1$ such that for all $n \geq n_{0}$ there exist $m>(1+\varepsilon)^{n}$ vectors $\mathbf{w}_{1}, \ldots, \mathbf{w}_{m} \in \mathbb{R}^{n}$ satisfying

$$
\begin{cases}\left\langle\mathbf{w}_{i}, \mathbf{w}_{i}\right\rangle=1 & \text { for all } i  \tag{1}\\ \left|\left\langle\mathbf{w}_{i}, \mathbf{w}_{j}\right\rangle\right|<\delta & \text { for all distinct } i, j\end{cases}
$$

We may take $\varepsilon=\delta^{2} / 2$ and $n_{0} \geq(120 \log 2) /\left(25 \delta^{4}-\delta^{6}\right)$.
We need the Chernoff inequality (proved in e.g. [25, Theorem 1.4.5]). The binary entropy function is defined by $H(0)=H(1):=0$ and

$$
H(x):=-x \lg x-(1-x) \lg (1-x), \text { for } 0<x<1 .
$$

Lemma 9. For any $\varepsilon>0$ and $n \in \mathbb{N}$,

$$
\begin{equation*}
\sum_{0 \leq k \leq \varepsilon n}\binom{n}{k} \leq 2^{n H(\varepsilon)} \tag{2}
\end{equation*}
$$

Proof of Lemma 8. For $n \geq 1$, define a graph on $\{-1,1\}^{n}$ as vertex set by joining two points if the number of coordinates in which they differ is $\leq \frac{1}{2}(1-\delta) n$ or $\geq \frac{1}{2}(1+\delta) n$. This graph is regular of degree

$$
m:=\sum_{0 \leq k \leq \frac{1}{2}(1-\delta) n}\binom{n}{k}+\sum_{\frac{1}{2}(1+\delta) n \leq k \leq n}\binom{n}{k} \leq-1+2^{n H((1-\delta) / 2)+1},
$$

by (2). If $\mathbf{x}$ and $\mathbf{y}$ are not connected in this graph, then $n^{-1 / 2} \mathbf{x}$ and $n^{-1 / 2} \mathbf{y}$ are unit vectors satisfying $|\langle\mathbf{x}, \mathbf{y}\rangle|<\delta$. We therefore have to show that the graph contains an independent set (a set without an edge between any two vertices) of at least $(1+\epsilon)^{n}$ points. We arbitrarily choose a vertex, delete it and its $m$ neighbours, choose a remaining vertex, delete it and its at most $m$ neighbours, and continue until nothing remains. Thus we obtain an independent set of size at least

$$
2^{n} /(m+1) \leq 2^{n(1-H((1-\delta) / 2))-1}
$$

vertices, which is greater than $\left(1+\delta^{2} / 2\right)^{n}$ if

$$
n f(\delta)>1,
$$

where $f(\delta)=1-H((1-\delta) / 2)-\lg \left(1+\delta^{2} / 2\right)$. But this is true for sufficiently large $n$, since it is easily seen that $f(\delta)>0$ for $0<\delta<1$. That $n \geq$ $(120 \log 2) /\left(25 \delta^{4}-\delta^{6}\right)$ is sufficient can be seen by calculating the Taylor expansion of $f(\delta)$.

## Remarks

1. The proof shows that if $M(n, \delta)$ is the maximum number of unit vectors in $\mathbb{R}^{n}$ satisfying (1), then

$$
\lim _{n \rightarrow \infty} \frac{1}{n} \lg M(n, \delta) \geq 1-H\left(\frac{1-\delta}{2}\right)=\frac{1}{\log 2}\left(\frac{\delta^{2}}{2}+\frac{\delta^{4}}{12}+\ldots\right)
$$

If one repeatedly choose unit vectors on the Euclidean unit sphere $\mathbb{S}^{n-1}$ and delete spherical caps around them, then, as shown by Wyner [40, pp. 1089-1092] (slightly improving an earlier estimate of Shannon [32]),

$$
\lim _{n \rightarrow \infty} \frac{1}{n} \lg M(n, \delta) \geq \lg \left(1-\delta^{2}\right)^{-1 / 2}=\frac{1}{\log 2}\left(\frac{\delta^{2}}{2}+\frac{\delta^{4}}{4}+\ldots\right)
$$

which is the same up to first order, but has a better higher order term. Thus the above greedy argument is not really improved by choosing vectors from the Euclidean unit sphere.
2. Recently, the Gilbert-Varshamov lower bound for binary codes has been slightly improved [21], which will improve the lower bound in Lemma 8 to $c(\delta) n\left(1+\delta^{2} / 2\right)^{n}$.
3. If we prove this lemma by choosing vectors in $\{ \pm 1\}$ randomly, and estimating the probabilities in the obvious way, then we obtain a slightly worse dependence of $\varepsilon=\delta^{2} / 4$.

Proof of Theorem 7. By Lemma 8 there exists $\mathbf{w}_{1}, \ldots, \mathbf{w}_{m}, m \geq 1.058^{n}$, satisfying (1) with $\delta=1 / 3$.

We now show that $P=\operatorname{conv}\left\{\mathbf{w}_{i}-\mathbf{w}_{j}: 1 \leq i, j \leq m\right\}$ is a centrally symmetric polytope with $m(m-1)$ vertices $\mathbf{w}_{i}-\mathbf{w}_{j}, i \neq j$. It is sufficient to note that the hyperplane

$$
\left\{\mathbf{x} \in \mathbb{R}^{n}:\left\langle\mathbf{x}, \mathbf{w}_{i}-\mathbf{w}_{j}\right\rangle=\left\langle\mathbf{w}_{i}-\mathbf{w}_{j}, \mathbf{w}_{i}-\mathbf{w}_{j}\right\rangle\right\}
$$

is a supporting hyperplane of $P$ passing only through $\mathbf{w}_{i}-\mathbf{w}_{j}$. This follows from

$$
\begin{aligned}
& \left\langle\mathbf{w}_{j}-\mathbf{w}_{i}, \mathbf{w}_{i}-\mathbf{w}_{j}\right\rangle, \pm\left\langle\mathbf{w}_{i}-\mathbf{w}_{k}, \mathbf{w}_{i}-\mathbf{w}_{j}\right\rangle, \\
& \pm\left\langle\mathbf{w}_{j}-\mathbf{w}_{k}, \mathbf{w}_{i}-\mathbf{w}_{j}\right\rangle, \pm\left\langle\mathbf{w}_{k}-\mathbf{w}_{\ell}, \mathbf{w}_{i}-\mathbf{w}_{j}\right\rangle \\
\leq & \left\langle\mathbf{w}_{i}-\mathbf{w}_{j}, \mathbf{w}_{i}-\mathbf{w}_{j}\right\rangle \text { for distinct } i, j, k, \ell,
\end{aligned}
$$

which follow easily from (1) and $\delta=1 / 3$.
With $P$ as a unit ball we thus already have $\left\{\mathbf{w}_{i}: i=1, \ldots, m\right\}$ as an equilateral set. We now need to find a strictly convex centrally symmetric body with the vertices of $P$ on its boundary. This is provided by the next lemma.

The following lemma is obvious, but we provide a simple proof.
Lemma 10. Let $S$ be the vertex set of a centrally symmetric polytope in $\mathbb{R}^{n}$. Then there is a smooth, strictly convex norm $\|\cdot\|$ on $\mathbb{R}^{n}$ such that $\|\mathbf{x}\|=1$ for all $\mathbf{x} \in S$.

Proof. Let $S=\left\{ \pm \mathbf{x}_{1}, \ldots, \pm \mathbf{x}_{m}\right\}$. For each $\mathbf{x}_{i}$, choose $\mathbf{y}_{i} \in \mathbb{R}^{n}$ such that $\left\langle\mathbf{y}_{i}, \mathbf{x}_{i}\right\rangle=1$ and $\left|\left\langle\mathbf{y}_{i}, \mathbf{x}_{j}\right\rangle\right|<1$ for all $i \neq j$. The required norm will be

$$
\|\mathbf{x}\|=\left(\sum_{j=1}^{m} \lambda_{j}\left|\left\langle\mathbf{y}_{j}, \mathbf{x}\right\rangle\right|^{p}\right)^{1 / p}
$$

for suitably chosen $\lambda_{j}>0$ and $1<p<\infty$, i.e., we want to imbed $S$ into the unit sphere of $\ell_{p}^{m}$ for some $p$.

Choose $p$ large enough such that

$$
\left|\left\langle\mathbf{y}_{j}, \mathbf{x}_{i}\right\rangle\right|^{p}<\frac{1}{2 m} \text { for all } j \neq i .
$$

Consider the matrix $A=\left[a_{i j}\right]_{i, j=1}^{m}$ with

$$
a_{i j}= \begin{cases}\left|\left\langle\mathbf{y}_{j}, \mathbf{x}_{i}\right\rangle\right|^{p} & \text { if } i \neq j \\ 0 & \text { if } i=j\end{cases}
$$

Considering $A$ as a linear transformation on $\ell_{\infty}^{m}$, we have $\|A\|<\frac{1}{2}$. We want to find a vector $\mathbf{v}=\left(\lambda_{1}, \ldots, \lambda_{m}\right)^{\operatorname{tr}}>0$ such that

$$
(I+A) \mathbf{v}=(1, \ldots, 1)^{\mathrm{tr}}
$$

since then we would have $\left\|\mathbf{x}_{i}\right\|=1$ for all $i$. However, since $\|A\|<\frac{1}{2}, I+A$ is invertible, and then necessarily $\mathbf{v}=(I+A)^{-1}(1, \ldots, 1)^{\mathrm{tr}}$. Also,

$$
\begin{aligned}
\left\|\mathbf{v}-(1, \ldots, 1)^{\mathrm{tr}}\right\|_{\infty} & =\left\|(I+A)^{-1}(1, \ldots, 1)^{\mathrm{tr}}-(1, \ldots, 1)^{\mathrm{tr}}\right\|_{\infty} \\
& \leq\left\|(I+A)^{-1}-I\right\|\|(1, \ldots, 1)\|_{\infty} \\
& =\left\|\sum_{i=1}^{\infty} A^{i}\right\| \leq \sum_{i=1}^{\infty}\|A\|^{i}<1 .
\end{aligned}
$$

Thus, $\lambda_{i}>0$ for all $i$.
Note that in the above lemma we may choose $p$ to be an even integer which gives a norm that is $C^{\infty}$ on $\mathbb{R}^{n} \backslash\{\mathbf{o}\}$.

## 3 Simple lower bounds

We now consider the problem of finding equilateral sets in a general Minkowski space, thus providing a lower bound to $e(X)$.

### 3.1 General Minkowski spaces

Proposition 11. If $\operatorname{dim} X \geq 2$ then $e(X) \geq 3$.
Proof. Exercise. (Hint: Euclid Book I Proposition 1.)
The above proposition combined with Theorem 4 gives the following
Corollary 12. If $\operatorname{dim} X=2$ then

$$
e(X)= \begin{cases}4 & \text { if the unit ball of } X \text { is a parallelogram } \\ 3 & \text { if the unit ball of } X \text { is not a parallelogram } .\end{cases}
$$

In general one would hope for the following conjecture, stated in $[18,28$, 29, 39].

Conjecture 2. If $\operatorname{dim} X=n$ then $e(X) \geq n+1$.
As seen above this is simple for $n=2$, and the next theorem shows its truth for $n=3$. However, for each $n \geq 4$ this is open, and the best that is known is the theorem of Brass and Dekster giving $e(X) \geq c(\log n)^{1 / 3}$, discussed in Section 5.

Theorem 13 (Petty [29]). If $\operatorname{dim} X \geq 3$ then any equilateral set of size 3 may be extended to an equilateral set of size 4 .

The proof needs the following technical but simple lemma.
Lemma 14. Let $S_{2}$ be the unit circle (i.e. boundary of the unit ball) of a 2 -dimensional Minkowski space. Fix any $\mathbf{u} \in S_{2}$. Let $\mathbf{f}:[0,1] \rightarrow S_{2}$ be a parametrization of an arc of $S_{2}$ from $\mathbf{u}$ to $-\mathbf{u}$. Fix $\mathbf{p}=\lambda \mathbf{u}, \lambda>0$. Then $t \mapsto\|\mathbf{p}-\mathbf{f}(t)\|$ is strictly increasing before it reaches the value $\|\mathbf{p}+\mathbf{u}\|$, i.e. on the interval $\left[0, t_{0}\right]$, where $t_{0}=\min \{t:\|\mathbf{p}-\mathbf{f}(t)\|<\|\mathbf{p}+\mathbf{u}\|\}$.
Proof. Exercise.
Proof of Theorem 13. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be 1-equilateral in $X$. Without loss of generality $\mathbf{a}=\mathbf{o}$. Let $X_{2}=\operatorname{span}\{\mathbf{b}, \mathbf{c}\}$, and let $X_{3}$ be any 3 -dimensional subspace of $X$ containing $X_{2}$. Let $S_{2}$ be the unit circle of $X_{2}$, and let $S_{3}$ be the unit sphere of $X_{3}$.

Consider the mapping $\mathbf{f}: S_{3} \rightarrow \mathbb{R}^{2}$ defined by $\mathbf{f}(x)=(\|\mathbf{x}-\mathbf{b}\|,\|\mathbf{x}-\mathbf{c}\|)$. We have to show that $(1,1)$ is in the range of $\mathbf{f}$. If we can show that the restriction of $\mathbf{f}$ to $S_{2}$ encircles $(1,1)$, then the theorem follows from the following topological argument.
$S_{2}$ can be contracted to a point on $S_{3}$, since $S_{3}$ is simply connected; thus $\mathbf{f}\left(S_{2}\right)$ can also be contracted to a point; if $(1,1)$ is in the interior of the Jordan curve $\mathbf{f}\left(S_{2}\right)$, then at some stage of the contraction the curve must pass through $(1,1)$, since $\mathbb{R}^{2} \backslash\{(1,1)\}$ is not simply connected.

Thus we assume without loss of generality that $(1,1) \notin \mathbf{f}\left(S_{2}\right)$, and we now follow the curve $\mathbf{f}\left(S_{2}\right)$. We start at $\mathbf{f}(\mathbf{b})=(0,1)$ and go to $\mathbf{f}(\mathbf{c})=(1,0)$. By Lemma 14 the $x$-coordinate increases strictly while the $y$-coordinate decreases strictly. Therefore the arc of the curve strictly between $(0,1)$ and $(1,0)$ is contained in the square $0<x, y<1$. Then from $(1,0)$ to $\mathbf{f}(\mathbf{c}-\mathbf{b})=(\|\mathbf{c}-2 \mathbf{b}\|, 1)$ the $y$-coordinate increases strictly, and the $x$ coordinate also increases strictly at least in a neighbourhood of $(1,0)$ if $\|\mathbf{c}-2 \mathbf{b}\|>1$. However, by the triangle inequality we have $\|\mathbf{c}-2 \mathbf{b}\| \geq 1$, and if $\|\mathbf{c}-\mathbf{2 b}\|=1$ then we already have found $(1,1)$ to be in the range of f , which we assumed not to happen.

From $\mathbf{f}(\mathbf{c}-\mathbf{b})$ to $\mathbf{f}(-\mathbf{b})=(2,\|\mathbf{b}+\mathbf{c}\|)$ the $x$-coordinate increases, and from $\mathbf{f}(-\mathbf{b})$ to $\mathbf{f}(-\mathbf{c})=(\|\mathbf{b}+\mathbf{c}\|, 2)$ the $y$-coordinate increases. From $\mathbf{f}(-\mathbf{c})$ to $\mathbf{f}(\mathbf{b}-\mathbf{c})=(1,\|\mathbf{b}-2 \mathbf{c}\|)$ the $x$-coordinate decreases strictly, and the $y$ coordinate decreases. Again by the triangle inequality and assumption we have $\|\mathbf{b}-2 \mathbf{c}\|>1$. Finally, from $\mathbf{f}(\mathbf{b}-2 \mathbf{c})$ to $\mathbf{f}(\mathbf{b})$ the $x$-coordinate decreases strictly, and the $y$-coordinate decreases.

It follows that $\mathbf{f}\left(S_{2}\right)$ encircles $(1,1)$, and the theorem is proved.
Note that if the fourth point which extends the equilateral set is in the plane $X_{2}$ (i.e. if $(1,1) \in \mathbf{f}\left(S_{2}\right)$ in the above proof), then $X_{2}$ is isometric to
$\ell_{\infty}^{2}$, i.e. $S_{2}$ is a parallelogram (by Corollary 12). However, it is impossible for all 2-dimensional sections of the unit ball through the origin to be parallelograms. Thus we may always choose an $X_{2}$ so that the fourth point is outside $X_{2}$ to obtain a non-coplanar equilateral set of four points.

One would hope that ideas similar to the above proof would help to prove Conjecture 2. However, the above proof cannot be naïvely generalized, as the following example of Petty [29] shows.

Define the following norm on $\mathbb{R}^{n}$ :

$$
\left\|\left(x_{1}, \ldots, x_{n}\right)\right\|:=\left|x_{1}\right|+\sqrt{x_{2}^{2}+\cdots+x_{n}^{2}}
$$

Thus $\|\cdot\|$ is the $\ell_{1}$-sum of $\mathbb{R}$ and $\ell_{2}^{n-1}:\left(\mathbb{R}^{n},\|\cdot\|\right)=\mathbb{R} \oplus_{1} \ell_{2}^{n-1}$. Its unit ball is a double cone over an $(n-1)$-dimensional Euclidean ball:

$$
B_{\|\cdot\|}=\operatorname{conv}\left(B_{2}^{n-1} \cup\left\{ \pm \mathbf{e}_{1}\right\}\right)
$$

For any $n \geq 2,\left(\mathbb{R}^{n},\|\cdot\|\right)$ contains a maximal equilateral set of four points. In fact the largest size of an equilateral set containing $\pm \mathbf{e}_{1}$ is 4 . First of all note that

$$
\left\|\mathbf{x}-\mathbf{e}_{1}\right\|=\left\|\mathbf{x}+\mathbf{e}_{1}\right\| \Longleftrightarrow x_{1}=0
$$

Secondly, since $\left\|\mathbf{e}_{1}-\left(-\mathbf{e}_{1}\right)\right\|=2$, we want $\left\|\mathbf{x} \pm \mathbf{e}_{1}\right\|=2$. Since then $x_{1}=0$, we obtain that $1+\sqrt{x_{2}^{2}+\cdots+x_{n}^{2}}=2$ must hold, which gives that $\mathbf{x} \in \mathbb{S}^{n-2}$. Clearly there are at most two points on the Euclidean ( $n-2$ )-dimensional unit sphere at distance 2 (and two such points must be antipodal).

On the other hand any $n$ equilateral points in the ( $n-1$ )-dimensional Euclidean subspace $x_{1}=0$ can always be extended to $n+1$ equilateral points. For example, if we choose the $n$ equilateral points on $\mathbb{S}^{n-2}$ then $\lambda \mathbf{e}_{1}$ will be the $(n+1)$ st point, for some appropriate $\lambda \in \mathbb{R}$.

Problem 15. Calculate e $(X)$ for the Minkowski space in the above example.

## $3.2 \ell_{p}^{n}$ with $1<p<2$

We sketch the proof that $e\left(\ell_{p}^{n}\right)>n+1$ for any $1<p<2$ with $n$ sufficiently large (depending on $p$ ).

Note that the four $\pm 1$-vectors in $\mathbb{R}^{3}$ with an even number of -1 's, i.e.,

$$
\left[\begin{array}{l}
1  \tag{3}\\
1 \\
1
\end{array}\right],\left[\begin{array}{c}
1 \\
-1 \\
-1
\end{array}\right],\left[\begin{array}{c}
-1 \\
1 \\
-1
\end{array}\right],\left[\begin{array}{c}
-1 \\
-1 \\
1
\end{array}\right],
$$

is an equilateral set for any $p$-norm. Thus if we consider $\mathbb{R}^{6}$ to be the direct sum $\mathbb{R}^{3} \oplus \mathbb{R}^{3}$ and put the above four vectors in each copy of $\mathbb{R}^{3}$, we obtain the following 8 vectors in $\mathbb{R}^{6}$ :

$$
\left[\begin{array}{l}
1  \tag{4}\\
1 \\
1 \\
0 \\
0 \\
0
\end{array}\right],\left[\begin{array}{c}
1 \\
-1 \\
-1 \\
0 \\
0 \\
0
\end{array}\right],\left[\begin{array}{c}
-1 \\
1 \\
-1 \\
0 \\
0 \\
0
\end{array}\right],\left[\begin{array}{l}
-1 \\
-1 \\
1 \\
0 \\
0 \\
0
\end{array}\right],\left[\begin{array}{l}
0 \\
0 \\
0 \\
1 \\
1 \\
1
\end{array}\right],\left[\begin{array}{c}
0 \\
0 \\
0 \\
1 \\
-1 \\
-1
\end{array}\right],\left[\begin{array}{c}
0 \\
0 \\
0 \\
-1 \\
1 \\
-1
\end{array}\right],\left[\begin{array}{c}
0 \\
0 \\
0 \\
-1 \\
-1 \\
1
\end{array}\right] .
$$

Then it is easily seen that these 8 vectors are equilateral iff $2 \cdot 2^{p}=6 \Longleftrightarrow$ $p=\log 3 / \log 2$.

The combinatorial reason why this worked is that any two of the vectors in (3) differ in exactly two positions. For a generalization we use Hadamard matrices. An Hadamard matrix is an $n \times n$ matrix $H$ with all entries $\pm 1$ and with orthogonal columns. By multiplying some of the columns of an Hadamard matrix one obtains an Hadamard matrix with the first row containing only 1 's. If we remove this first row, we obtain $n$ column vectors in $\mathbb{R}^{n-1}$ with $\pm 1$ entries such that any two differ in exactly $n / 2$ positions. The vectors in (3) was obtained in this way from a $4 \times 4$ Hadamard matrix. We may now construct $2 n$ vectors in $\mathbb{R}^{2 n-2}=\mathbb{R}^{n-1} \oplus \mathbb{R}^{n-1}$ as before that will be equilateral iff

$$
\frac{n}{2} 2^{p}=2 n-2 \Longleftrightarrow p=2+\lg \frac{n-1}{n}
$$

The only problem remaining is to find Hadamard matrices. It is easy to see that for an $n \times n$ Hadamard matrix to exist, $n$ must be divisible by 4 . It is an open problem whether the converse is true. However, the following well-known construction gives a Hadamard matrix of order $2^{k}$ :

$$
H_{0}=[1], \quad H_{k+1}=\left[\begin{array}{cc}
H_{k} & H_{k} \\
H_{k} & -H_{k}
\end{array}\right], k \geq 0
$$

Thus we have $e\left(\ell_{p}^{n}\right)>n+1$ for $p$ arbitrarily close to 2 .
With a more involved construction the following can be shown:
Theorem 16. For any $1<p<2$ and $n \geq 1$, let

$$
k=\left\lceil\frac{\log \left(1-2^{p-2}\right)^{-1}}{\log 2}\right\rceil-1
$$

Then

$$
e\left(\ell_{p}^{n}\right) \geq\left\lfloor\frac{2^{k+1}}{2^{k+1}-1} n\right\rfloor
$$

In particular, if $n \geq 2^{k+2}-2$ then $e\left(\ell_{p}^{n}\right)>n+1$.
Corollary 17. $e\left(\ell_{p}^{n}\right)>n+1$ if $p<2-\frac{1+o(1)}{(\log 2) n}$.
Compare this with Corollary 25 in Section 4.2. The smallest dimension for which the theorem gives an example of $e\left(\ell_{p}^{n}\right)>n+1$ is $n=6$ (and $1<p \leq \log 3 / \log 2$ ). In [37] there is also a 4 -dimensional example (with $\left.1<p \leq \log \frac{5}{2} / \log 2\right)$.

In the next two sections we show that $e(X)$ goes to infinity with the dimension of $X$ (Brass-Dekster), although the best known lower bound for $e(X)$ is much smaller than linear in $n$. There are three ingredients to the proof: The Cayley-Menger theory of the embedding of metric spaces into Euclidean space, Dvoretzky's theorem, and the Brouwer fixed point theorem.

## 4 Cayley-Menger Theory

In this section we discuss a fragment of the Cayley-Menger theory by giving necessary and sufficient conditions for a metric space of size $n+1$ to be embeddable as an affinely independent set in $\ell_{2}^{n}$. The general theory is by Menger [27]; see also Blumenthal [8, §40].

### 4.1 The Cayley-Menger determinant

Let a metric space on $n+1$ points $\mathbf{p}_{0}, \ldots, \mathbf{p}_{n}$ be given, with distances $\rho_{i j}=d\left(\mathbf{p}_{i}, \mathbf{p}_{j}\right)$. We now derive a necessary condition for $\mathbf{p}_{0}, \ldots, \mathbf{p}_{n}$ to be isometric to an affinely independent subset of $\ell_{2}^{n}$, in terms of $\rho_{i j}$.

So we assume that $\left\{\mathbf{p}_{0}, \ldots, \mathbf{p}_{n}\right\}$ is an affinely independent set in $\ell_{2}^{n}$ with $\left\|\mathbf{p}_{i}-\mathbf{p}_{j}\right\|_{2}=\rho_{i j}$. We write the coordinates of $\mathbf{p}_{i}$ as $\left(p_{i}^{(1)}, p_{i}^{(2)}, \ldots, p_{i}^{(n)}\right)$. Let $\Delta=\operatorname{conv}\left\{\mathbf{p}_{0}, \ldots, \mathbf{p}_{n}\right\}$ denote the simplex with vertices $\mathbf{p}_{i}$. Then

$$
\begin{aligned}
\pm \operatorname{vol}(\Delta) & =\frac{1}{n!}\left|\begin{array}{ccc}
p_{0}^{(1)} & p_{1}^{(1)} & p_{n}^{(1)} \\
p_{0}^{(2)} & p_{1}^{(2)} & p_{n}^{(2)} \\
\vdots & \vdots & \vdots \\
p_{0}^{(n)} & p_{1}^{(n)} & p_{n}^{(n)} \\
1 & 1 & 1
\end{array}\right| \\
& =\frac{1}{n!}\left|\begin{array}{ccc}
\mathbf{p}_{0} & \mathbf{p}_{1} & \mathbf{p}_{n} \\
1 & 1 & 1
\end{array}\right| .
\end{aligned}
$$

Squaring (and adding an extra row and column),

$$
\begin{aligned}
& \operatorname{vol}(\Delta)^{2}=\frac{1}{(n!)^{2}} \operatorname{det}\left[\begin{array}{ccc}
\mathbf{p}_{0}^{\operatorname{tr}} & 1 & 0 \\
\mathbf{p}_{1}^{\operatorname{tr}} & 1 & 0 \\
\vdots & \vdots & \vdots \\
\mathbf{p}_{n}^{\operatorname{tr}} & 1 & 0 \\
\mathbf{o}^{\operatorname{tr}} & 0 & 1
\end{array}\right]\left[\begin{array}{ccc}
\mathbf{p}_{0} & \mathbf{p}_{n} & \mathbf{o} \\
1 & 1 & 0 \\
0 & 0 & 1
\end{array}\right] \\
& =-\frac{1}{(n!)^{2}} \operatorname{det}\left[\begin{array}{ccc}
\mathbf{p}_{0}^{\operatorname{tr}} & 0 & 1 \\
\mathbf{p}_{1}^{\operatorname{tr}} & 0 & 1 \\
\vdots & \vdots & \vdots \\
\mathbf{p}_{n}^{\operatorname{tr}} & 0 & 1 \\
\mathbf{o}^{\operatorname{tr}} & 1 & 0
\end{array}\right]\left[\begin{array}{ccc}
\mathbf{p}_{0} & \mathbf{p}_{n} & \mathbf{o} \\
1 & 1 & 0 \\
0 & 0 & 1
\end{array}\right] \\
& =-\frac{1}{(n!)^{2}}\left|\begin{array}{cccc}
\left\langle\mathbf{p}_{0}, \mathbf{p}_{0}\right\rangle & \left\langle\mathbf{p}_{0}, \mathbf{p}_{1}\right\rangle & \left\langle\mathbf{p}_{0}, \mathbf{p}_{n}\right\rangle & 1 \\
\left\langle\mathbf{p}_{1}, \mathbf{p}_{0}\right\rangle & \left\langle\mathbf{p}_{1}, \mathbf{p}_{1}\right\rangle & \left\langle\mathbf{p}_{1}, \mathbf{p}_{n}\right\rangle & 1 \\
\vdots & \vdots & \vdots & \\
\left\langle\mathbf{p}_{n}, \mathbf{p}_{0}\right\rangle & \left\langle\mathbf{p}_{n}, \mathbf{p}_{1}\right\rangle & \left\langle\mathbf{p}_{n}, \mathbf{p}_{n}\right\rangle & 1 \\
1 & 1 & 1 & 0
\end{array}\right| \\
& =-\frac{1}{2^{n+1}(n!)^{2}}\left|\begin{array}{ccccc}
0 & -\rho_{01}^{2} & & -\rho_{0 n}^{2} & 1 \\
-\rho_{10}^{2} & 0 & & \vdots & \vdots \\
\vdots & & \ddots & \vdots & \vdots \\
-\rho_{n 0}^{2} & & -\rho_{n, n-1}^{2} & 0 & 1 \\
1 & & \cdots & & 0
\end{array}\right| \\
& \text { (interchange last two } \\
& \text { columns of first matrix) } \\
& \begin{array}{l}
\begin{array}{l}
\text { use } 2\left\langle\mathbf{p}_{i}, \mathbf{p}_{j}\right\rangle= \\
\left\langle\mathbf{p}_{i}, \mathbf{p}_{i}\right\rangle+\left\langle\mathbf{p}_{j}, \mathbf{p}_{j}\right\rangle-\rho_{i j}^{2} \\
\text { and subtract multiples } \\
\text { of last row and column) }
\end{array}
\end{array}=\frac{(-1)^{n+1}}{2^{n+1}(n!)^{2}} \operatorname{det}\left[\begin{array}{lll} 
& & \\
& P & \\
& & \vdots \\
1 & & 1 \\
& & 1 \\
\hline
\end{array}\right],
\end{aligned}
$$

where $P=\left[\rho_{i j}^{2}\right]$. We call

$$
\left.\operatorname{CMdet}\left(\mathbf{p}_{0}, \ldots, \mathbf{p}_{n}\right):=\left\lvert\, \begin{array}{ccc} 
& & \\
& P & \\
& & \vdots \\
1 & & 1
\end{array}\right.\right)
$$

the Cayley-Menger determinant of the $(n+1)$-point metric space. We have shown the following.

Proposition 18. If a metric space on $n+1$ points can be embedded into $\ell_{2}^{n}$ as an affinely independent set, then its Cayley-Menger determinant has sign $(-1)^{n+1}$. (Also, if the metric space can be embedded as an affinely dependent set, then its Cayley-Menger determinant is 0 .)

Conversely we have
Theorem 19 (Menger [27]). A metric space on $n+1$ points $\mathbf{p}_{0}, \ldots, \mathbf{p}_{n}$ can be embedded into $\ell_{2}^{n}$ as an affinely independent set if for each $k=1, \ldots, n$, the Cayley-Menger determinant of $\mathbf{p}_{0}, \ldots, \mathbf{p}_{k}$ has sign $(-1)^{k+1}$.

Proof. For $n=1$,

$$
\operatorname{CMdet}\left(\mathbf{p}_{0}, \mathbf{p}_{1}\right)=\left|\begin{array}{ccc}
0 & \rho_{01}^{2} & 1 \\
\rho_{10}^{2} & 0 & 1 \\
1 & 1 & 0
\end{array}\right|=\rho_{01}^{2}+\rho_{10}^{2}>0
$$

Thus the theorem is trivial for $n=1$ (and even more so for $n=0$ ).
Thus we may assume as inductive hypothesis that the theorem is true for $n-1$. Thus $\mathbf{p}_{0}, \ldots, \mathbf{p}_{n-1}$ can be embedded into $\ell_{2}^{n-1} \subset \ell_{2}^{n}$ as an affinely independent set, say $\mathbf{p}_{0} \mapsto \mathbf{o}$ and $\mathbf{p}_{i} \mapsto \mathbf{x}_{i}, i=1, \ldots, n-1$. Then $\operatorname{span}\left\{\mathbf{x}_{1}, \ldots, \mathbf{x}_{n-1}\right\}=\ell_{2}^{n-1}$ and $\mathbf{x}_{1}, \ldots, \mathbf{x}_{n-1}$ are linearly independent.

We have $\left\|\mathbf{x}_{i}\right\|_{2}=\rho_{i 0}$, and $\left\|\mathbf{x}_{i}-\mathbf{x}_{j}\right\|_{2}=\rho_{i j}$ for $1 \leq i, j \leq n-1$, and we have to find an $\mathbf{x} \notin \ell_{2}^{n-1}$ satisfying

$$
\begin{equation*}
\|\mathbf{x}\|_{2}=\rho_{0 n} \tag{5}
\end{equation*}
$$

and

$$
\begin{equation*}
\left\|\mathbf{x}-\mathbf{x}_{i}\right\|_{2}=\rho_{\text {in }} \text { for all } i=1, \ldots, n-1 \tag{6}
\end{equation*}
$$

We write $\mathbf{x}=\mathbf{v}+\lambda \mathbf{e}_{n}$ uniquely, where $\mathbf{v} \in \ell_{2}^{n-1}$ and $\lambda \in \mathbb{R}$. If we square (6) and simplify, we find that $\mathbf{v}$ has to satisfy the following $n-1$ linear equations in $\ell_{2}^{n-1}$ :

$$
2\left\langle\mathbf{x}_{i}, \mathbf{v}\right\rangle=\|\mathbf{x}\|_{2}^{2}+\rho_{0 i}^{2}-\rho_{i n}^{2}, i=1, \ldots, n-1 .
$$

Since the $\mathbf{x}_{i}$ are linearly independent, there is a unique solution $\mathbf{v} \in \ell_{2}^{n-1}$ to the following modified equation:

$$
\begin{equation*}
2\left\langle\mathbf{x}_{i}, \mathbf{v}\right\rangle=\rho_{0 n}^{2}+\rho_{0 i}^{2}-\rho_{i n}^{2}, i=1, \ldots, n-1 \tag{7}
\end{equation*}
$$

and we will have found the required $\mathbf{x}$ once we can satisfy (5), which is equivalent to

$$
\lambda^{2}=\rho_{0 n}^{2}-\|\mathbf{v}\|_{2}^{2}
$$

Thus it remains to prove that $\rho_{0 n}^{2}-\|\mathbf{v}\|_{2}^{2}>0$. To this end we calculate the Cayley-Menger determinant of $\mathbf{o}, \mathbf{x}_{1}, \ldots, \mathbf{x}_{n-1}, \mathbf{v}$, which is 0 by Proposition 18:

$$
\left|\begin{array}{ccccc}
0 & \rho_{01}^{2} & \rho_{0, n-1}^{2} & \|\mathbf{v}\|_{2}^{2} & 1 \\
\rho_{10}^{2} & 0 & \rho_{1, n-1}^{2} & \left\|\mathbf{v}-\mathbf{x}_{1}\right\|_{2}^{2} & 1 \\
\vdots & & \vdots & \vdots & \vdots \\
\rho_{n-1,0}^{2} & & 0 & \left\|\mathbf{v}-\mathbf{x}_{n-1}\right\|_{2}^{2} & 1 \\
\|\mathbf{v}\|_{2}^{2} & \left\|\mathbf{v}-\mathbf{x}_{1}\right\|_{2}^{2} & \left\|\mathbf{v}-\mathbf{x}_{n-1}\right\|_{2}^{2} & 0 & 1 \\
1 & 1 & 1 & 1 & 0
\end{array}\right|=0 .
$$

Multiplying out

$$
\begin{aligned}
\left\|\mathbf{v}-\mathbf{x}_{i}\right\|_{2}^{2} & =\|\mathbf{v}\|_{2}^{2}-2\left\langle\mathbf{x}_{i}, \mathbf{v}\right\rangle+\left\|\mathbf{x}_{i}\right\|_{2}^{2} \\
& =\|\mathbf{v}\|_{2}^{2}-\rho_{0 n}^{2}+\rho_{i n}^{2} \quad(\text { by }(7))
\end{aligned}
$$

and using the last row and column to eliminate $\|\mathbf{v}\|_{2}^{2}-\rho_{0 n}^{2}$, we obtain

$$
\left|\begin{array}{ccccc}
0 & \rho_{01}^{2} & \rho_{0, n-1}^{2} & \rho_{0 n}^{2} & 1 \\
\rho_{10}^{2} & 0 & \rho_{1, n-1}^{2} & \rho_{1 n}^{2} & 1 \\
\vdots & \ddots & \vdots & \vdots & \vdots \\
\rho_{n-1,0}^{2} & \cdots & 0 & \rho_{n-1, n}^{2} & 1 \\
\rho_{n, 0}^{2} & \rho_{n, 1}^{2} & \rho_{n, n-1}^{2} & -2\left(\|\mathbf{v}\|_{2}^{2}-\rho_{0 n}^{2}\right) & 1 \\
1 & 1 & 1 & 1 & 0
\end{array}\right|=0 .
$$

This determinant differs from $\operatorname{CMdet}\left(\mathbf{p}_{0}, \ldots, \mathbf{p}_{n}\right)$ only in the second last column. Thus we may subtract this determinant from $\operatorname{CMdet}\left(\mathbf{p}_{0}, \ldots, \mathbf{p}_{n}\right)$, which we are given has sign $(-1)^{n+1}$, to obtain the following determinant of $\operatorname{sign}(-1)^{n+1}$ :
$\left|\begin{array}{ccccc}0 & \rho_{01}^{2} & \rho_{0, n-1}^{2} & 0 & 1 \\ \rho_{10}^{2} & 0 & \rho_{1, n-1}^{2} & 0 & 1 \\ \vdots & & \vdots & \vdots & \vdots \\ \rho_{n-1,0}^{2} & \cdots & 0 & 0 & 1 \\ \rho_{n, 0}^{2} & \rho_{n, 1}^{2} & \rho_{n, n-1}^{2} & 2\left(\|\mathbf{v}\|_{2}^{2}-\rho_{0 n}^{2}\right) & 1 \\ 1 & 1 & 1 & 1 & 0\end{array}\right|$
$=2\left(\|\mathbf{v}\|_{2}^{2}-\rho_{0 n}^{2}\right) \operatorname{CMdet}\left(\mathbf{p}_{0}, \ldots, \mathbf{p}_{n-1}\right)$ (expanding along 2nd last column).

Since it is also given that $\operatorname{CMdet}\left(\mathbf{p}_{0}, \ldots, \mathbf{p}_{n-1}\right)$ has $\operatorname{sign}(-1)^{n}$, it follows that $\|\mathbf{v}\|_{2}^{2}-\rho_{0 n}^{2}<0$. This gives us a $\lambda \in \mathbb{R}$ such that $x=\mathbf{v}+\lambda \mathbf{e}_{n}$ satisfies (5) as well. The theorem is proved.

### 4.2 Embedding an almost equilateral $n$-simplex into $\ell_{2}^{n}$

We now want to use Theorem 19 to show that an $n$-simplex with all side lengths close to 1 can be embedded isometrically into $\ell_{2}^{n}$. Thus we have to find the sign of the determinant of a matrix with 0 on the diagonal and all off-diagonal entries close to 1 , since this is how the Cayley-Menger determinant will look like.

We start with the following simple observations. Let $J$ be the $n \times n$ matrix with all entries equal to 1 .

Lemma 20. Let $n \geq 1$.

1. The characteristic polynomial of $J$ is $(-1)^{n}(x-n) x^{n-1}$.
2. Therefore the characteristic polynomial of $J-I$ is $(-1)^{n}(x-n+1)(x+$ $1)^{n-1}$, and $\operatorname{det}(J-I)=(-1)^{n-1}(n-1)$.
3. Therefore $J-I$ is invertible, and its inverse has eigenvalues -1 with multiplicity $n-1$, and $1 /(n-1)$.
Proposition 21. Let $A=\left[a_{i j}\right]$ be an $(n+2) \times(n+2)$ matrix with $a_{i i}=0$, $\left|1-a_{i j}\right| \leq 1 /\left(n-\frac{3}{2}\right)$ for $1 \leq i \neq j \leq n+1$, and $a_{i, n+2}=a_{n+2, i}=1$ for $i=1, \ldots, n$. Then $\operatorname{det}(A)$ has sign $(-1)^{n+1}$.

Proof. We first prove that $A$ is invertible. Let $E=\left[\varepsilon_{i j}\right]:=J-I-A$. Then $\varepsilon_{i i}=0$ and $\varepsilon_{i, n+2}=\varepsilon_{n+2, i}=0$ for all $i$, and $\left|\varepsilon_{i j}\right| \leq 1 /\left(n-\frac{3}{2}\right)$ for all $1 \leq i \neq j \leq n+1$. Since $A=J-I-E=(J-I)\left(I-(J-I)^{-1} E\right)$, it is sufficient to prove that $\left\|(J-I)^{-1} E\right\|_{2}<1$, where $\|\cdot\|_{2}$ is the operator norm on $L\left(\ell_{2}^{n} \rightarrow \ell_{2}^{n}\right)$. Firstly, since the operator norm is the maximum modulus of the eigenvalues, we have by Lemma 20 that $\left\|(J-I)^{-1}\right\|_{2}=1$. Secondly, by the Cauchy-Schwartz inequality, $\|E\|_{2} \leq \sqrt{\sum_{i, j=1}^{n+2} \varepsilon_{i j}^{2}}<1$. Thus $A$ is invertible.

To show that $\operatorname{det} A$ has $\operatorname{sign}(-1)^{n+1}$, observe that we may join $J-I$ to $A$ with a curve $A_{t}, 0 \leq t \leq 1$, where each $A_{t}$ satisfies the hypothesis of the proposition. Thus we know from what we have just proved that each $A_{t}$ is invertible. By continuity considerations the sign of $\operatorname{det} A_{t}$ does not change. Therefore, $A$ and $J-I$ have the same sign, which is $(-1)^{n+1}$ by Lemma 20.

Theorem 22 (Dekster \& Wilker [14]). Let $\mathbf{p}_{0}, \ldots, \mathbf{p}_{n}$ be a metric space with distances satisfying

$$
1 \leq d\left(\mathbf{p}_{i}, \mathbf{p}_{j}\right) \leq 1+\frac{1}{n+1}
$$

Then $\mathbf{p}_{0}, \ldots, \mathbf{p}_{n}$ can be embedded into $\ell_{2}^{n}$, and any such embedding must be an affinely independent set.
Proof. We first scale the distances by some $\alpha>0$ in such a way that the given inequality is transformed into

$$
1-\varepsilon \leq\left(\alpha d\left(\mathbf{p}_{i}, \mathbf{p}_{j}\right)\right)^{2} \leq 1+\varepsilon
$$

for some $\varepsilon>0$. Thus $\varepsilon$ and $\alpha$ must satisfy

$$
\frac{1-\varepsilon}{\alpha^{2}}=1 \text { and } \frac{1+\varepsilon}{\alpha^{2}}=\left(\frac{n+2}{n+1}\right)^{2} .
$$

Eliminating $\alpha$ we find

$$
\begin{equation*}
\frac{1+\varepsilon}{1-\varepsilon}=\left(\frac{n+2}{n+1}\right)^{2} \tag{8}
\end{equation*}
$$

We want to apply Theorem 19. Since the CM determinant of $\mathbf{p}_{0}, \ldots, \mathbf{p}_{k}$ is a $(k+2) \times(k+2)$ matrix with $k \leq n$, to show that it has sign $(-1)^{k+1}$, it is sufficient to prove that

$$
\varepsilon \leq \frac{1}{n+\frac{3}{2}}
$$

according to Proposition 21. By (8) we have

$$
\begin{aligned}
\varepsilon & =\frac{\left(\frac{n+2}{n+1}\right)^{2}-1}{\left(\frac{n+2}{n+1}\right)^{2}+1} \\
& =\frac{(n+2)^{2}-(n+1)^{2}}{(n+2)^{2}+(n+1)^{2}}=\frac{2 n+3}{2 n^{2}+6 n+5} \\
& <\frac{2 n+3}{2 n^{2}+6 n+\frac{9}{2}}=\frac{2}{2 n+3} .
\end{aligned}
$$

By Proposition 18 no embedding can be affinely dependent.
Dekster and Wilker found the smallest $\gamma_{n}>1$ such that whenever

$$
1 \leq d\left(\mathbf{p}_{i}, \mathbf{p}_{j}\right) \leq \gamma_{n}
$$

then $\mathbf{p}_{0}, \ldots, \mathbf{p}_{n+1}$ can be embedded into $\ell_{2}^{n}$ : it is

$$
\gamma_{n}= \begin{cases}\sqrt{1+\frac{2 n+2}{n^{2}-2}} & \text { if } n \text { is even } \\ \sqrt{1+\frac{2}{n-1}} & \text { if } n \text { is odd }\end{cases}
$$

Their proof is geometrical. The above analytical proof gives values that differ from $\gamma_{n}$ by $O\left(n^{-2}\right)$ as $n \rightarrow \infty$.

Corollary 23 (Schütte [30]). For any $\mathbf{p}_{1}, \ldots, \mathbf{p}_{n+2} \in \ell_{2}^{n}$ we have

$$
\frac{\max \left\|\mathbf{p}_{i}-\mathbf{p}_{j}\right\|_{2}}{\min _{i \neq j}\left\|\mathbf{p}_{i}-\mathbf{p}_{j}\right\|_{2}}>1+\frac{1}{n+2} .
$$

Schütte in fact determined the smallest $\delta_{n}>1$ such that

$$
\frac{\max \left\|\mathbf{p}_{i}-\mathbf{p}_{j}\right\|_{2}}{\min _{i \neq j}\left\|\mathbf{p}_{i}-\mathbf{p}_{j}\right\|_{2}} \geq 1+\delta_{n}
$$

for all possible $\mathbf{p}_{1}, \ldots, \mathbf{p}_{n+2} \in \ell_{2}^{n}$ : it is $\delta_{n}=\gamma_{n+1}$. This result was also discovered by Seidel [31], and follows from the embedding result of Dekster and Wilker [14]; see [4] for a very simple proof.

Corollary 24. If $d\left(X, \ell_{2}^{n}\right) \leq 1+\frac{1}{n+2}$, then $e(X) \leq n+1$.
Corollary 25 (Smyth [33]). If $|p-2|<\frac{4(1+o(1))}{n \log n}$ then $e\left(\ell_{p}^{n}\right)=n+1$.
Compare this with Corollary 17 in Section 3.2.

## 5 The Theorem of Brass and Dekster

Theorem 26 (Brass [10] \& Dekster [13]). An n-dimensional Minkowski space contains an equilateral set of size $c(\log n)^{1 / 3}$ for some constant $c>0$ and $n$ sufficiently large.

Both Brass' and Dekster's proofs use Dvoretzky's theorem combined with a topological result. We follow Brass' proof, which uses only the Brouwer fixed point theorem.

For the proof we need Dvoretzky's theorem.
Dvoretzky's Theorem. There exists a contant $c>0$ such that for any $\varepsilon>0$, any Minkowski space $X$ of sufficiently large dimension contains an $m$-dimensional subspace with Banach-Mazur distance at most $1+\varepsilon$ to $\ell_{2}^{m}$. We may take $\operatorname{dim} X \geq e^{c m \varepsilon^{-2}}$ for some absolute constant $c>0$.

The estimate for $\operatorname{dim} X$, which is best possible (up to the value of $c$ ), was proven by Gordon [16].

Proof of Theorem 26. Denote the $n$-dimensional space by $X$. Let $m=$ $c_{1}(\log n)^{1 / 3}$ and $\varepsilon=1 /(m+1)$. A simple calculation gives that $e^{c m \varepsilon^{-2}}<n$ if we choose $c_{1}=c_{1}(c)$ sufficiently small. Dvoretzky's theorem gives that $X$ contains an $m$-dimensional subspace $Y$ with Banach-Mazur distance at most $1+\frac{1}{m+1}$ to $\ell_{2}^{m}$. The proof is then finished by applying the following theorem.

Theorem 27 (Brass [10] \& Dekster [13]). Let $X$ be an n-dimensional Minkowski space with Banach-Mazur distance $d\left(X, \ell_{2}^{n}\right) \leq 1+\frac{1}{n+1}$. Then an equilateral set in $X$ of at most $n$ points can be extended to one of $n+1$ points. In particular, $e(X) \geq n+1$.

Proof. By induction on nested subspaces of $X$ it is sufficient to prove that a 1 -equilateral set $\mathbf{p}_{1}, \ldots, \mathbf{p}_{n}$ may be extended. We may assume that a coordinate system has been chosen such that

$$
\|\mathbf{x}\| \leq\|\mathbf{x}\|_{2} \leq(1+\varepsilon)\|\mathbf{x}\|
$$

where $\varepsilon=(n+1)^{-1}$. Let $\mathbf{p}_{1}, \ldots, \mathbf{p}_{n}$ be a 1 -equilateral set in $X$. We want to find a $\mathbf{p} \in X$ such that $\left\|\mathbf{p}-\mathbf{p}_{i}\right\|=1$ for all $i=1, \ldots, n$.

We have

$$
1 \leq\left\|\mathbf{p}_{i}-\mathbf{p}_{j}\right\|_{2} \leq 1+\varepsilon \text { for all distinct } i, j \text {. }
$$

By Theorem $22 \mathbf{p}_{1}, \ldots, \mathbf{p}_{n}$ are affinely independent. For any $\rho_{1}, \ldots, \rho_{n} \in$ $[1,1+\varepsilon]$ there exists a point $\mathbf{x} \in X$ with $\left\|\mathbf{x}-\mathbf{p}_{i}\right\|_{2}=\rho_{i}, i=1, \ldots, n$, by Theorem 22. Furthermore, by fixing a half space bounded by the hyperplane through $\mathbf{p}_{1}, \ldots, \mathbf{p}_{n}$, there is a unique such $\mathbf{x}$ in this half space (by simple linear algebra arguments as in the proof of of Theorem 19; exercise). We write $\mathbf{x}=: \mathbf{f}\left(\rho_{1}, \ldots, \rho_{n}\right)$, and define $\varphi:[0, \varepsilon]^{n} \rightarrow[0, \varepsilon]^{n}$ by sending $\left(\delta_{1}, \ldots, \delta_{n}\right) \mapsto\left(y_{1}, \ldots, y_{n}\right)$, where

$$
y_{i}=\delta_{i}+1-\left\|\mathbf{f}\left(1+\delta_{1}, \ldots, 1+\delta_{n}\right)-\mathbf{p}_{i}\right\|, i=1, \ldots, n .
$$

Since $1+\delta_{i} \leq 1+\varepsilon$, each $y_{i}$ is well-defined. Also,

$$
\begin{aligned}
y_{i} & \leq \delta_{i}+1-(1+\varepsilon)^{-1}\left\|\mathbf{f}\left(1+\delta_{1}, \ldots, 1+\delta_{n}\right)-\mathbf{p}_{i}\right\|_{2} \\
& =\delta_{i}+1-(1+\varepsilon)^{-1}\left(1+\delta_{i}\right) \\
& =\left(1-(1+\varepsilon)^{-1}\right)\left(1+\delta_{i}\right) \\
& \leq\left(1-(1+\varepsilon)^{-1}\right)(1+\varepsilon)=\varepsilon,
\end{aligned}
$$

and

$$
\begin{aligned}
y_{i} & \geq \delta_{i}+1-\left\|\mathbf{f}\left(1+\delta_{1}, \ldots, 1+\delta_{n}\right)-\mathbf{p}_{i}\right\|_{2} \\
& =\delta_{i}+1-\left(1+\delta_{i}\right)=0,
\end{aligned}
$$

which shows that $\varphi$ maps into $[0, \varepsilon]^{n}$.
By Brouwer's fixed point theorem, $\varphi$ has a fixed point $\left(\delta_{1}, \ldots, \delta_{n}\right)$, which gives

$$
\left\|\mathbf{f}\left(1+\delta_{1}, \ldots, 1+\delta_{n}\right)-\mathbf{p}_{i}\right\|=1 \text { for all } i=1, \ldots, n
$$

Corollary 28 (Brass [10] \& Dekster [13]). If $d\left(X, \ell_{2}^{n}\right) \leq 1+\frac{1}{n+2}$ then $e(X)=n+1$.

Proof. Combine the above theorem with Corollary 24.

## 6 The Linear Algebra Method

Linear algebra provides a very powerful counting tool. The basic idea is that whenever one has a set $S$ of $m$ elements, and one wants to find an upper bound to $m$, one constructs a vector $\mathbf{v}_{s}$ in some $N$-dimensional vector space for each element of $s \in S$. If one can prove that the vectors constructed are linearly independent, then one immediately has the upper bound $m \leq N$. If the vectors are not linearly independent, then one may try to prove that the square matrix $A^{\text {tr }} A$ has a high rank, where $A$ is the matrix with column vectors $\mathbf{v}_{s}, s \in S$.

### 6.1 Linear independence

As a first example, one can prove that $e\left(\ell_{2}^{n}\right) \leq n+1$ as follows. Let $S \subset \ell_{2}^{n}$ be 1-equilateral, and define for each $\mathbf{s} \in S$ a polynomial in $n$ variables

$$
P_{\mathbf{s}}(\mathbf{x})=P_{\mathbf{s}}\left(x_{1}, \ldots, x_{n}\right):=1-\|\mathbf{x}-\mathbf{s}\|_{2}^{2} .
$$

Then it is easily seen that $\{1\} \cup\left\{P_{\mathbf{s}}(\mathbf{x}): \mathbf{s} \in S\right\}$ is linearly independent in the vector space $\mathbb{R}\left[x_{1}, \ldots, x_{n}\right]$ of polynomials in $x_{1}, \ldots, x_{n}$ with real coefficients, and secondly that all $P_{\mathbf{s}}(\mathbf{x})$ are in the linear span of $\left\{1, \sum_{i=1}^{n} x_{i}^{2}, x_{1}, \ldots, x_{n}\right\}$. It follows that $1+|S| \leq 2+n$, which is what was required.

In the same way one may prove that $e\left(\ell_{p}^{n}\right) \leq 1+(p-1) n$ if $p$ is an even integer (an observation due to Galvin; see [33]). With some more work the following may be shown:

Theorem 29 (Swanepoel [37]). For each $n \geq 1$, $e\left(\ell_{4}^{n}\right)=n+1$.
Proof. The lower bound has already been shown in Section 1.2.
Let $S$ be a 1-equilateral set in $\ell_{4}^{n}$. For each $\mathbf{s} \in S$, let $P_{\mathbf{s}}(x)=p_{\mathbf{s}}\left(x_{1}, \ldots, x_{n}\right)$ be the following polynomial:

$$
\begin{align*}
P_{\mathbf{s}}(\mathbf{x}) & :=1-\|\mathbf{x}-\mathbf{s}\|_{4}^{4} \\
& =1-\|\mathbf{s}\|_{4}^{4}-\sum_{i=1}^{n} x_{i}^{4}+4 \sum_{i=1}^{n} s_{i} x_{i}^{3}-6 \sum_{i=1}^{n} s_{i}^{2} x_{i}^{2}+4 \sum_{i=1}^{n} s_{i}^{3} x_{i} . \tag{9}
\end{align*}
$$

Thus each $P_{\mathbf{s}}$ is in the linear span of

$$
\left\{1, \sum_{i=1}^{n} x_{i}^{4}\right\} \cup\left\{x_{i}^{k}: i=1, \ldots, n ; k=1,2,3\right\},
$$

which is a subspace of $\mathbb{R}\left[x_{1}, \ldots, x_{n}\right]$ of dimension $2+3 n$. Since $P_{\mathbf{s}}(\mathbf{s})=1$ and $P_{\mathbf{s}}\left(\mathbf{s}^{\prime}\right)=0$ for all distinct $\mathbf{s}, \mathbf{s}^{\prime} \in S$ we have that $\left\{P_{\mathbf{s}}: \mathbf{s} \in S\right\}$ is linearly independent. (So already $|S| \leq 2+3 n$.)

We now prove that

$$
\left\{P_{\mathbf{s}}: \mathbf{s} \in S\right\} \cup\{1\} \cup\left\{x_{i}^{k}: i=1, \ldots, n ; k=1,2\right\}
$$

is still linearly independent, which will give $|S|+1+2 n \leq 2+3 n$, proving the theorem.

Let $\lambda_{\mathbf{s}}(\mathbf{s} \in S), \lambda, \lambda_{i}, \mu_{i}(i=1, \ldots, d) \in \mathbb{R}$ satisfy

$$
\begin{equation*}
\sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}} P_{\mathbf{s}}(\mathbf{x})+\lambda 1+\sum_{i=1}^{n} \lambda_{i} x_{i}+\sum_{i=1}^{n} \mu_{i} x_{i}^{2}=0 \quad \forall x_{1}, \ldots, x_{n} \in \mathbb{R} . \tag{10}
\end{equation*}
$$

Substitute (9) into (10):

$$
\begin{aligned}
& \left(-\sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}}\right) \sum_{i=1}^{n} x_{i}^{4}+4 \sum_{i=1}^{n}\left(\sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}} s_{i}\right) x_{i}^{3}+\sum_{i=1}^{n}\left(\mu_{i}-6 \sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}} s_{i}^{2}\right) x_{i}^{2} \\
& +\sum_{i=1}^{n}\left(\lambda_{i}+4 \sum_{\mathbf{s} \in S} s_{i}^{3}\right) x_{i}+\sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}}-\sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}}\|\mathbf{s}\|_{4}^{4}+\lambda=0 .
\end{aligned}
$$

Since the above is a polynomial in $\mathbb{R}\left[x_{1}, \ldots, x_{n}\right]$ which is identically 0 , the coefficients are also 0 :

$$
\begin{align*}
& \sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}}=0,  \tag{11}\\
& \sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}} s_{i}=0 \quad \forall i=1, \ldots, n,  \tag{12}\\
& \mu_{i}-6 \sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}} s_{i}^{2}=0 \quad \forall i=1, \ldots, n,  \tag{13}\\
& \lambda_{i}+4 \sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}} s_{i}^{3}=0 \quad \forall i=1, \ldots, d,  \tag{14}\\
& \lambda+\sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}}-\sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}}\|\mathbf{s}\|_{4}^{4}=0 . \tag{15}
\end{align*}
$$

By substituting $\mathbf{x}=\mathbf{s} \in S$ into (10) we obtain

$$
\begin{equation*}
\lambda_{\mathbf{s}}+\lambda+\sum_{i=1}^{n} \lambda_{i} s_{i}+\sum_{i=1}^{n} \mu_{i} s_{i}^{2}=0 \quad \forall \mathbf{s} \in S . \tag{16}
\end{equation*}
$$

Now multiply (16) by $\lambda_{\mathbf{s}}$ and sum over all $\mathbf{s} \in S$ :

$$
\sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}}^{2}+\lambda \sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}}+\sum_{i=1}^{n} \lambda_{i}\left(\sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}} s_{i}\right)+\sum_{i=1}^{n} \mu_{i}\left(\sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}} s_{i}^{2}\right)=0 .
$$

Then use (11)-(13) to simplify this expression as follows:

$$
\sum_{\mathbf{s} \in S} \lambda_{\mathbf{s}}^{2}+\frac{1}{6} \sum_{i=1}^{n} \mu_{i}^{2}=0
$$

It follows that all $\lambda_{\mathbf{s}}=0$ and all $\mu_{i}=0$. From (14) it follows that all $\lambda_{i}=0$ and from (15) that $\lambda=0$.

Thus we have linear independence.
In the same way the following may be proved:
Theorem 30 (Swanepoel [37]). For $p$ an even integer and $n \geq 1$ we have

$$
e\left(\ell_{p}^{n}\right) \leq \begin{cases}\left(\frac{p}{2}-1\right) n+1 & \text { if } p \equiv 0 \quad(\bmod 4) \\ \frac{p}{2} n+1 & \text { if } p \equiv 2 \quad(\bmod 4)\end{cases}
$$

### 6.2 Rank arguments

The following lemma is very simple, yet extremely powerful.
Rank Lemma. Let $A=\left[a_{i j}\right]$ be a real symmetric $n \times n$ non-zero matrix. Then

$$
\operatorname{rank} A \geq \frac{\left(\sum_{i=1}^{n} a_{i i}\right)^{2}}{\sum_{i, j=1}^{n} a_{i j}^{2}}
$$

Proof. Let $r=\operatorname{rank} A$, and let $\lambda_{1}, \ldots, \lambda_{r} \in \mathbb{R}$ be all non-zero eigenvalues of $A$. Then $\sum_{i} a_{i i}=\operatorname{trace}(A)=\sum_{i=1}^{r} \lambda_{i}$. Also, $\sum_{i, j} a_{i j}^{2}=\operatorname{trace}\left(A^{2}\right)^{\prime}=$ $\sum_{i=1}^{r} \lambda_{i}^{2}$, since $A^{2}$ has non-zero eigenvalues $\lambda_{i}^{2}$. Therefore,

$$
\frac{\left(\sum_{i=1}^{n} a_{i i}\right)^{2}}{\sum_{i, j=1}^{n} a_{i j}^{2}}=\frac{\left(\sum_{i=1}^{r} \lambda_{i}\right)^{2}}{\sum_{i, j=1}^{r} \lambda_{i}^{2}} \leq r
$$

by the Cauchy-Schwartz inequality.

Corollary 31. Let $A=\left[a_{i j}\right]$ be a real symmetric $n \times n$ matrix with $a_{i i}=1$ for all $i$ and $\left|a_{i j}\right| \leq \varepsilon$ for all $i \neq j$. Then

$$
\operatorname{rank} A \geq \frac{n}{1+(n-1) \varepsilon^{2}}
$$

In particular, if $\varepsilon=n^{-1 / 2}$ then $\operatorname{rank} A \geq n / 2$.
Also, if $\varepsilon<1 /(n-1)$ then the corollary gives that $A$ is invertible, which also follows from the fact that $A$ is then diagonally dominated.

The rank lemma together with a simple rounding argument gives an upper bound for $e\left(\ell_{p}^{n}\right), 1 \leq p<\infty$, that is polynomial in $p$ and $n$. This is due to Alon (Smyth, personal communication).

Theorem 32. For some constant $c>0$ we have that for $1 \leq p<\infty$ and $n \geq 1, e\left(\ell_{p}^{n}\right) \leq c p^{2} n^{2+2 / p}$.
Proof. Let $\mathbf{p}_{1}, \ldots, \mathbf{p}_{m}$ be 1-equilateral in $\ell_{p}^{n}$. Since $\left|p_{i}^{(k)}-p_{j}^{(k)}\right| \leq 1$ for all $i \neq j$ and all $k=1, \ldots, m$, we may assume after a translation that all $p_{i}^{(k)} \in[0,1]$. Choose an integer $N$ such that

$$
(e-1) p n^{1 / p} \sqrt{m} \leq N<N+1 \leq e p n^{1 / p} \sqrt{m} .
$$

(The interval $p n^{1 / p} \sqrt{m} \geq 2$ since $m \geq 4$ without loss of generality.) We now round each $\mathbf{p}_{i}$ to a point in the lattice $\left\{0, \frac{1}{N}, \frac{2}{N}, \ldots, \frac{N}{N}\right\}^{n}$ as follows. Let $q_{i}^{(k)}$ be the integer multiple of $1 / N$ nearest to $p_{i}^{(k)}$, say $q_{i}^{(k)}=\frac{d(i, k)-1}{N}$ where $1 \leq d(i, k) \leq N+1$ (if there is a tie, choose arbitrarily). Let $\mathbf{q}_{i}=$ $\left(q_{i}^{(1)}, \ldots, q_{i}^{(n)}\right)$, and let $Q$ be the $m \times m$ matrix $\left[1-\left\|\mathbf{q}_{i}-\mathbf{q}_{j}\right\|_{p}^{p}\right]_{i, j}^{m}$. Thus $Q$ is an approximation of the identity $I_{m}$, and we now estimate how close. The diagonal contains only 1 's. For each $k,\left|p_{i}^{(k)}-q_{i}^{(k)}\right| \leq 1 / 2 N$, and

$$
\begin{aligned}
\left\|\mathbf{q}_{i}-\mathbf{q}_{j}\right\|_{p} & \leq\left\|\mathbf{p}_{i}-\mathbf{p}_{j}\right\|_{p}+\left\|\mathbf{p}_{i}-\mathbf{q}_{i}\right\|_{p}+\left\|\mathbf{p}_{j}-\mathbf{q}_{j}\right\|_{p} \\
& \leq 1+n^{1 / p} / N
\end{aligned}
$$

and similarly, $\left\|\mathbf{q}_{i}-\mathbf{q}_{j}\right\|_{p} \geq 1-n^{1 / p} / N$. Thus

$$
\left\|\mathbf{q}_{i}-\mathbf{q}_{j}\right\|_{p}^{p} \leq\left(1+\frac{n^{1 / p}}{N}\right)^{p} \leq\left(1+\frac{1}{(e-1) p \sqrt{m}}\right)^{p}<e^{\frac{1}{(e-1) \sqrt{m}}}<1+\frac{1}{\sqrt{m}}
$$

since $e^{x}<1+(e-1) x$ for $0<x<1$. It follows that $1-\left\|\mathbf{q}_{i}-\mathbf{q}_{j}\right\|_{p}^{p}>-1 / \sqrt{m}$. Since $(1+x)^{p}+(1-x)^{p}>2$ for $0<x<1$,

$$
1-\left\|\mathbf{q}_{i}-\mathbf{q}_{j}\right\|_{p}^{p} \leq 1-\left(1-\frac{n^{1 / p}}{N}\right)^{p}<\left(1+\frac{n^{1 / p}}{N}\right)^{p}-1<\frac{1}{\sqrt{m}}
$$

Thus by Corollary $31, \operatorname{rank} Q \geq m / 2$. Note that $Q=\sum_{k=1}^{n} Q_{k}$, where

$$
Q_{k}=\left[\frac{1}{n}-\left|q_{i}^{(k)}-q_{j}^{(k)}\right| p\right]_{i, j=1}^{m} .
$$

Define the $(N+1) \times m$ matrix

$$
A_{k}=\left[\begin{array}{lll}
\mathbf{e}_{d(1, k)} & \mathbf{e}_{d(2, k)} & \mathbf{e}_{d(m, k)}
\end{array}\right],
$$

where $\mathbf{e}_{1}, \ldots, \mathbf{e}_{N+1}$ is the standard basis of $\mathbb{R}^{N+1}\left(\right.$ recall $\left.q_{i}^{(k)}=\frac{d(i, k)-1}{N}\right)$. Secondly, define

$$
B=\left[\frac{1}{n}-\left|\frac{i-1}{N}-\frac{j-1}{N}\right|^{p}\right]_{i, j=1}^{N+1} .
$$

Then it is easily seen that $Q_{k}=A_{k}^{\operatorname{tr}} B A_{k}$. It follows that $\operatorname{rank} Q_{k} \leq N+1$, hence

$$
\operatorname{rank} Q \leq \sum_{k=1}^{n} \operatorname{rank} Q_{k} \leq n(N+1)
$$

Putting the upper and lower bounds for rank $Q$ together we obtain

$$
m / 2<n(N+1) \leq e p n^{1+1 / p} \sqrt{m},
$$

and $m<4 e^{2} p^{2} n^{2+2 / p}$.

## 7 Approximation Theory: Smyth's approach

In the proof of Theorem 32 we essentially approximated the function $f_{p}(t)=|t|^{p}, t \in[-1,1]$ uniformly by step functions (look at the definition of the matrix $B$ in the proof). For $p>1$ the function $f_{p}$ is differentiable $\lfloor p\rfloor$ times, so one would expect that there are better uniform approximations to $f_{p}$ using polynomials or splines instead of step functions. This is indeed the case, and can be used to improve the bound of Theorem 32. Smyth [33] used the theorems of Jackson from approximation theory to find the first non-trivial upper bounds for $e\left(\ell_{p}^{n}\right)$. Alon and Pudlák [1] improved his bounds using the rank lemma. We state the theorem of Jackson that is needed (called Jackson V) and then prove their result. The proofs of the Jackson theorems may be found in many texts on approximation theory, e.g. [11].

Jackson V. There is an absolute constant $c>0$ such that for any $f \in$ $C[-1,1]$

1. there exists a polynomial $P$ of degree at most $n$ such that

$$
\|f-P\|_{\infty} \leq c \omega(f, 1 / n)
$$

where $\omega(f, \delta):=\sup \{|f(x)-f(y)|:|x-y| \leq \delta\}$ is the modulus of continuity of $f$,
2. and if $f$ is $k$ times differentiable and $n \geq k$, then there exists a polynomial $P$ of degree at most $n$ such that

$$
\|f-P\|_{\infty} \leq \frac{c^{k}}{(n+1) n(n-1) \ldots(n-k+2)} \omega\left(f^{(k)}, 1 / n\right)
$$

Lemma 33. For each $1 \leq p<\infty$ there exists a constant $c_{p}>0$ such that for any $d \geq 1$ there is a polynomial $P$ of degree at most $d$ such that $P(0)=0$ and

$$
\left||t|^{p}-P(t)\right| \leq \frac{c_{p}}{d^{p}} \quad \text { for all } t \in[-1,1] \text {. }
$$

We have $c_{p}<(c p)^{p}$ for some universal $c>0$.
Proof. By choosing $c_{p}$ sufficiently large we may assume that $d>2 p$. Then $f_{p}(t):=|t|^{p}$ satisfies $f \in C^{(k)}[-1,1]$ with $k=\lceil p\rceil-1$, and

$$
f^{(\lceil p\rceil-1)}(t)=\operatorname{sgn}^{\lceil p\rceil-1}(t) p(p-1) \ldots(p-\lceil p\rceil+2)|t|^{p+1-\lceil p\rceil} .
$$

Thus

$$
\omega\left(f^{(\lceil p\rceil-1)}, \delta\right)=p(p-1) \ldots(p-\lceil p\rceil+2) \delta^{p+1-\lceil p\rceil}
$$

From Jackson V we obtain a polynomial $P$ of degree at most $d$ with

$$
\begin{aligned}
& \left\|f_{p}-P\right\|_{\infty} \\
& \leq \frac{c^{\lceil p\rceil-1}}{(d+1) d(d-1) \ldots(d-\lceil p\rceil+3)} p(p-1) \ldots(p-\lceil p\rceil+2)(1 / d)^{p+1-\lceil p\rceil} \\
& <\frac{c^{\lceil p\rceil-1} p(p-1) \ldots(p-\lceil p\rceil+2)}{\left(\frac{1}{2} d\right)^{\lceil p\rceil-1}}(1 / d)^{p+1-\lceil p\rceil} \\
& =\frac{c}{d^{p}} .
\end{aligned}
$$

If we subtract the constant term from $P$ we obtain a polynomial $Q$ with $Q(0)=0$ and $\left\|f_{p}-Q\right\|_{\infty} \leq 2 c / d^{p}$.

Smyth [33] approached Kusner's problem with the idea of approximating the $p$-norm with polynomials using Jackson's theorems. He obtained an upper bound $c_{p} n^{(p+1) /(p-1)}$ for $1<p<\infty$ using linear independence. The next theorem was proved by Alon and Pudlák by combining Smyth's approach with the rank lemma.
Theorem 34 (Smyth [33], Alon \& Pudlák [1]). For $1 \leq p<\infty$, $e\left(\ell_{p}^{n}\right) \leq c_{p}^{\prime} n^{(2 p+2) /(2 p-1)}$. We may take $c_{p}^{\prime}=c p$, with $c>0$ and absolute constant.
Proof. Let $\mathbf{p}_{1}, \ldots, \mathbf{p}_{m} \in \ell_{p}^{n}$ be 1-equilateral. Let $c=\max \left(c_{p},\left(2^{1 / p}-1\right)^{-p}\right)$, where $c_{p}$ is the constant from Lemma 33. We fix an integer $d$ such that $c n \sqrt{m}<d^{p}<2 c n \sqrt{m}$ (possible since $\left.c \geq\left(2^{1 / p}-1\right)^{-p}\right)$. Let $P$ be the polynomial from Lemma 33. Define the $m \times m$ matrix $A=\left[a_{i j}\right]$ by $a_{i j}=$ $1-\sum_{k=1}^{n} P\left(p_{i}^{(k)}-p_{j}^{(k)}\right)$. Since $\sum_{k=1}^{n} P\left(p_{i}^{(k)}-p_{j}^{(k)}\right)$ is an approximation of $\left\|\mathbf{p}_{i}-\mathbf{p}_{j}\right\|_{p}^{p}$, the matrix $A$ is an approximation of the identity $I_{m}$. Again we estimate how close. Firstly, since $P(0)=0$, all $a_{i i}=1$. Secondly for any $i \neq j$,

$$
\begin{aligned}
\left|a_{i j}\right| & =\left|\left\|\mathbf{p}_{i}-\mathbf{p}_{j}\right\|_{p}^{p}-\sum_{k=1}^{n} P\left(p_{i}^{(k)}-p_{j}^{(k)}\right)\right| \\
& \leq \sum_{k=1}^{n}| | \mathbf{p}_{i}^{(k)}-\left.\mathbf{p}_{j}^{(k)}\right|^{p}-P\left(\mathbf{p}_{i}^{(k)}-\mathbf{p}_{j}^{(k)}\right) \mid \\
& \leq n \frac{c_{p}}{d^{p}}<\frac{1}{\sqrt{m}}
\end{aligned}
$$

by choice of $d$. By Corollary 31 we have $\operatorname{rank} A \geq m / 2$.
We now find an upper bound for the rank of $A$. For each $i=1,2, \ldots, m$, define the polynomial in $n$ variables

$$
P_{i}\left(x_{1}, \ldots, x_{n}\right):=1-\sum_{k=1}^{n} P\left(p_{i}^{(k)}-x_{i}\right) .
$$

Thus $a_{i j}=P_{i}\left(\mathbf{p}_{j}\right)$. Each $P_{i}$ is in the linear span of

$$
\mathcal{P}=\left\{1, \sum_{k=1}^{n} x_{i}^{d}, x_{1}, \ldots, x_{n}, x_{1}^{2}, \ldots, x_{n}^{2}, \ldots, x_{1}^{d-1}, \ldots, x_{n}^{d-1}\right\} .
$$

Thus the $m$ polynomials $P_{i}$ lie in a $(2+(d-1) n)$-dimensional subspace of polynomials. Then the $i$ th row vector of $A,\left[P_{i}\left(\mathbf{p}_{1}\right), \ldots, P_{i}\left(\mathbf{p}_{m}\right)\right]$, is in the linear span of

$$
\left\{\left(f\left(\mathbf{p}_{1}\right), \ldots, f\left(\mathbf{p}_{m}\right)\right): f \in \mathcal{P}\right\}
$$

which is a subspace of $\mathbb{R}^{m}$ of dimension at most $|\mathcal{P}|=2+(d-1) n \leq d n$. It follows that $\operatorname{rank} A \leq d n$.

Putting the upper and lower bound for rank $A$ together, we find $m / 2 \leq$ $d n$. Since $d^{p}<2 c n \sqrt{m}$, we obtain $m<c_{p}^{\prime} n^{(2 p+2) /(2 p-1)}$.

In the next section we find the best known upper bounds for $e\left(\ell_{1}^{n}\right)$, due to Alon and Pudlák.

## 8 The best known upper bound for $e\left(\ell_{1}^{n}\right)$

Alon and Pudlák [1] proved that $e\left(\ell_{p}^{n}\right) \leq c_{p} n \log n$ if $p$ is an odd integer. This matches the lower bound of $2 n$ apart from the $\log n$ factor and the constant that depends on $p$. We here present their proof for the case $p=1$. The proof for other odd $p$ is simple once this case is understood, see [1] for the detail.

Theorem 35 (Alon and Pudlák [1]). $e\left(\ell_{1}^{n}\right) \leq c n \log n$.
Proof. Let $\mathbf{p}_{1}, \ldots, \mathbf{p}_{m+1}$ be 1-equilateral in $\ell_{1}^{n}$. After a translation we may assume $\mathbf{p}_{m+1}=\mathbf{o}$. Thus it is sufficient to prove $m+1 \leq c n \log n$ given that

$$
\begin{align*}
\sum_{k=1}^{n}\left|p_{i}^{(k)}\right| & =1 \quad \forall i=1, \ldots, m  \tag{17}\\
\text { and } \sum_{k=1}^{n}\left|p_{i}^{(k)}\right| & =1 \quad \forall 1 \leq i \neq j \leq m \tag{18}
\end{align*}
$$

We first show that by doubling the dimension we may assume that all $p_{i}^{(k)} \geq$ 0 . We replace each $p_{i}^{(k)}$ by

$$
\begin{cases}\left(p_{i}^{(k)}, 0\right) & \text { if } p_{i}^{(k)} \geq 0 \\ \left(0,-p_{i}^{(k)}\right) & \text { if } p_{i}^{(k)} \leq 0\end{cases}
$$

Now the points $\mathbf{p}_{1}, \ldots, \mathbf{p}_{m}$ are in $\ell_{1}^{2 n}$, they are still 1-equilateral, $\left\|\mathbf{p}_{i}\right\|_{1}=1$, and now all $p_{i}^{(k)} \geq 0$. Thus we may assume together with (17) and (18) that $p_{i}^{(k)} \geq 0$ holds. (Once we have proven $m \leq c n \log n$ with this assumption, then $m \leq 1+c(2 n) \log (2 n)<3 c n \log n$ in the general case.)

We now express $\left|p_{i}^{(k)}-p_{j}^{(k)}\right|$ in terms of $\min \left(p_{i}^{(k)}, p_{j}^{(k)}\right)$. Since

$$
|a-b|=a+b-2 \min (a, b) \quad \text { for } a, b \in \mathbb{R},
$$

and $\sum_{k=1}^{n} p_{i}^{(k)}=1$, we obtain

$$
I_{m}=\left[1-\left\|\mathbf{p}_{i}-\mathbf{p}_{j}\right\|\right]_{i, j=1}^{m}=\left[-1+2 \sum_{k=1}^{n} \min \left(p_{i}^{(k)}, p_{j}^{(k)}\right)\right]_{i, j=1}^{m}
$$

We now want to approximate $\min (a, b)$ by an inner product $\langle\overline{\mathbf{a}}, \overline{\mathbf{b}}\rangle$ where $\overline{\mathbf{a}}, \overline{\mathbf{b}} \in \mathbb{R}^{N}$. Then $I_{m}$ will be approximated by

$$
A=\left[-1+2\left\langle\left(\overline{\mathbf{p}}_{i}^{(1)}, \ldots, \overline{\mathbf{p}}_{i}^{(n)}\right),\left(\overline{\mathbf{p}}_{j}^{(1)}, \ldots, \overline{\mathbf{p}}_{j}^{(n)}\right)\right\rangle\right]_{i, j=1}^{m}
$$

Then $\operatorname{rank} A \leq 1+N n$, since it is a linear combination of $J_{m}$ of rank 1 and a Gram matrix of vectors in $\mathbb{R}^{N n}$. The approximation will be good enough so that the rank lemma will give rank $A>c_{1} m$. Thus $c_{1} m<N n+1$, and we will have to make sure that $N<c_{2} \log n$. What will in fact happen is that $N$ will be different for each coordinate $k=1, \ldots, n$, say $N_{k}$, and we'll have to make sure $\sum_{k=1}^{n} N_{k}<c_{2} n \log n$.

The rough idea for approximating $\min (a, b)$ by an inner product is the following. For each $x \in[0,1]$ define

$$
\overline{\mathbf{x}}=(\underbrace{\frac{1}{\sqrt{N}}, \ldots, \frac{1}{\sqrt{N}}}_{\lfloor x N\rfloor \text { times }}, 0, \ldots, 0) \in \mathbb{R}^{N} .
$$

Then for $0 \leq a \leq b \leq 1$,

$$
\langle\overline{\mathbf{a}}, \overline{\mathbf{b}}\rangle=\frac{1}{N}\lfloor a N\rfloor=a-\varepsilon=\min (a, b)-\varepsilon,
$$

where $0 \leq \varepsilon \leq 1 / N$. Let us see how far this brings us already. Since

$$
\left|\sum_{k=1}^{n} \min \left(p_{i}^{(k)}, p_{j}^{(k)}\right)-\left\langle\left(\overline{\mathbf{p}}_{i}^{(1)}, \ldots, \overline{\mathbf{p}}_{i}^{(n)}\right),\left(\overline{\mathbf{p}}_{j}^{(1)}, \ldots, \overline{\mathbf{p}}_{j}^{(n)}\right)\right\rangle\right| \leq \frac{n}{N},
$$

the error by which $A$ approximates $I_{m}$ is

$$
a_{i j}-\delta_{i j} \leq \frac{2 n}{N}=\frac{1}{\sqrt{m}}
$$

if we choose $N=2 n \sqrt{m}$. By the rank lemma, $\operatorname{rank} A \geq(1+o(1)) m / 2$. Also, $\operatorname{rank} A \leq N n+1$. Putting the bounds together, we obtain $m<c n^{4}$, already obtained in Theorem 32.

We now fine-tune this idea. For any partition $\mathcal{P}=\left\{0=u_{0}<u_{1}<\ldots<\right.$ $\left.u_{N-1}<u_{N}=1\right\}$ of $[0,1]$ into $N$ intervals, we define for any $x \in[0,1]$ the vector $\mathbf{x}_{\mathcal{P}} \in \mathbb{R}^{N}$, where

$$
x_{\mathcal{P}}^{(j)}= \begin{cases}\sqrt{u_{j}-u_{j-1}}, & j<t, \\ \frac{x-u_{t-1}}{\sqrt{u_{t}-u_{t-1}}}, & j=t, \\ 0, & j>t,\end{cases}
$$

where $t$ is such that $x \in\left(u_{t-1}, u_{t}\right)$.

## Remarks.

1. We'll choose a different partition for each coordinate $k=1, \ldots, n$. Since both $\sum_{k} \min \left(p_{i}^{(k)}, p_{j}^{(k)}\right)$ and $\left\langle\left(\overline{\mathbf{p}}_{i}^{(1)}, \ldots, \overline{\mathbf{p}}_{i}^{(n)}\right),\left(\overline{\mathbf{p}}_{j}^{(1)}, \ldots, \overline{\mathbf{p}}_{j}^{(n)}\right)\right\rangle$ are sums over $k=1, \ldots, n$, the error in each coordinate will also sum up, so we first do the analysis in a single coordinate.
2. We'll choose the partitions so that a coordinate will never hit an endpoint of any $\left(u_{j-1}, u_{j}\right)$.

With the above definition we find for $0 \leq a \leq b \leq 1$ that

$$
\begin{aligned}
& \left\langle\mathbf{a}_{\mathcal{P}}, \mathbf{b}_{\mathcal{P}}\right\rangle=\left(\sqrt{u_{1}-0}\right)^{2}+\left(\sqrt{u_{2}-u_{1}}\right)^{2}+\left(\sqrt{u_{t-1}-u_{t-2}}\right)^{2}+a_{\mathcal{P}}^{(t)} b_{\mathcal{P}}^{(t)}, \\
& \text { where } a \in\left(u_{t-1}, u_{t}\right), \\
& = \begin{cases}u_{t-1}+\frac{a-u_{t-1}}{\sqrt{u_{t}-u_{t-1}}} \sqrt{u_{t}-u_{t-1}} & \text { if } b \notin\left(u_{t-1}, u_{t}\right), \\
u_{t-1}+\frac{a-u_{t-1}}{\sqrt{u_{t}-u_{t-1}}} \frac{b-u_{t-1}}{\sqrt{u_{t}-u_{t-1}}} & \text { if } b \in\left(u_{t-1}, u_{t}\right),\end{cases} \\
& = \begin{cases}\min (a, b), & \text { if } a, b \text { are in different intervals of } \mathcal{P}, \\
u_{t-1}+\frac{\left(a-u_{t-1}\right)\left(b-u_{t-1}\right)}{u_{t}-u_{t-1}} & \text { if } a, b \in\left(u_{t-1}, u_{t}\right) .\end{cases}
\end{aligned}
$$

Thus there is no error if $a$ and $b$ are in different intervals. However, when $a$ and $b$ are in the same interval of $\mathcal{P}$, the error is bad, and we want to get rid of it. We do this by adding another $N$ coordinates: Define

$$
\widehat{\mathbf{x}}_{\mathcal{P}}=\frac{x-u_{t-1}}{\sqrt{u_{t}-u_{t-1}}} \mathbf{e}_{t} \in \mathbb{R}^{N}, \quad \text { with } t \text { such that } x \in\left(u_{t-1}, u_{t}\right) .
$$

Then

$$
\left\langle\widehat{\mathbf{a}}_{\mathcal{P}}, \widehat{\mathbf{b}}_{\mathcal{P}}\right\rangle= \begin{cases}0 & \text { if } a \text { and } b \text { are in different intervals of } \mathcal{P}, \\ \frac{\left(a-u_{t-1}\right)\left(b-u_{t-1}\right)}{u_{t}-u_{t-1}} & \text { if } a, b \in\left(u_{t-1}, u_{t}\right)\end{cases}
$$

Since we want to subtract this, we let the inner product on $\mathbb{R}^{N} \oplus \mathbb{R}^{N}$ be

$$
\left\langle\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right),\left(\mathbf{y}_{1}, \mathbf{y}_{2}\right)\right\rangle=\left\langle\mathbf{x}_{1}, \mathbf{y}_{1}\right\rangle-\left\langle\mathbf{x}_{2}, \mathbf{y}_{2}\right\rangle .
$$

We then obtain for $0 \leq a \leq b \leq 1$ that

$$
\left\langle\left(\mathbf{a}_{\mathcal{P}}, \widehat{\mathbf{a}}_{\mathcal{P}}\right),\left(\mathbf{b}_{\mathcal{P}}, \widehat{\mathbf{b}}_{\mathcal{P}}\right)\right\rangle= \begin{cases}\min (a, b) & \text { if } a, b \text { are in different intervals of } \mathcal{P}, \\ u_{t-1} & \text { if } a, b \in\left(u_{t-1}, u_{t}\right)\end{cases}
$$

Thus if $a$ and $b$ are in the same interval, the approximation is a systematic rounding down. We now add a third vector in $\mathbb{R}^{N}$ to improve the approximation in this case. We do this by randomized rounding: For each interval $\left(u_{t-1}, u_{t}\right)$ we choose randomly and uniformly $\tau \in\left(u_{t-1}, u_{t}\right)$ (a threshold), and for any $x \in\left(u_{t-1}, u_{t}\right)$ we define

$$
\widetilde{\mathbf{x}}_{\mathcal{P}}= \begin{cases}0 & \text { if } x \leq \tau, \\ \sqrt{u_{t}-u_{t-1}} \mathbf{e}_{t} & \text { if } x>\tau\end{cases}
$$

Note that $\widetilde{\mathbf{x}}_{\mathcal{P}}$ is a random variable. Thus if $a$ and $b$ are in different intervals then $\left\langle\widetilde{\mathbf{a}}_{\mathcal{P}}, \widetilde{\mathbf{b}}_{\mathcal{P}}\right\rangle=0$, and if $a$ and $b$ are in the same interval $\left(u_{t-1}, u_{t}\right)$, then (as with $\widehat{\mathbf{x}}_{\mathcal{P}}$ )

$$
\left\langle\tilde{\mathbf{a}}_{\mathcal{P}}, \widetilde{\mathbf{b}}_{\mathcal{P}}\right\rangle= \begin{cases}0 & \text { if } \min (a, b) \leq \tau \\ u_{t}-u_{t-1} & \text { if } \min (a, b)>\tau\end{cases}
$$

We now let the inner product on $\mathbb{R}^{N} \oplus \mathbb{R}^{N} \oplus \mathbb{R}^{N}$ be

$$
\left\langle\left(\mathbf{x}_{1}, \mathbf{x}_{2}, \mathbf{x}_{3}\right),\left(\mathbf{y}_{1}, \mathbf{y}_{2}, \mathbf{y}_{3}\right)\right\rangle=\left\langle\mathbf{x}_{1}, \mathbf{y}_{1}\right\rangle-\left\langle\mathbf{x}_{2}, \mathbf{y}_{2}\right\rangle+\left\langle\mathbf{x}_{3}, \mathbf{y}_{3}\right\rangle,
$$

and for each $x \in[0,1]$, let

$$
\overline{\mathbf{x}}_{\mathcal{P}}=\left(\mathbf{x}_{\mathcal{P}}, \widehat{\mathbf{x}}_{\mathcal{P}}, \widetilde{\mathbf{x}}_{\mathcal{P}}\right) \in \mathbb{R}^{3 N} .
$$

Then for any $a, b \in[0,1]$,
$\left\langle\overline{\mathbf{a}}_{\mathcal{P}}, \overline{\mathbf{b}}_{\mathcal{P}}\right\rangle=\left\{\begin{array}{l}\min (a, b) \quad \text { if } a, b \text { are in different intervals of } \mathcal{P}, \\ \left\{\begin{array}{ll}u_{t-1}, & \text { if } \min (a, b) \leq \tau \\ u_{t}, & \text { if } \min (a, b)>\tau\end{array}, \quad \text { if } a, b \in\left(u_{t-1}, u_{t}\right),\right.\end{array}\right.$ and the error $X=\min (a, b)-\left\langle\overline{\mathbf{a}}_{\mathcal{P}}, \overline{\mathbf{b}}_{\mathcal{P}}\right\rangle$ (depending on $\left.\mathcal{P}, a, b\right)$ satisfies

$$
X= \begin{cases}0 & \text { if } a, b \text { in different intervals } \\ a-u_{t-1} & \text { if } \min (a, b) \leq \tau, a, b \in\left(u_{t-1}, u_{t}\right) \\ a-u_{t} & \text { if } \min (a, b)>\tau, a, b \in\left(u_{t-1}, u_{t}\right)\end{cases}
$$

Thus

$$
|X| \leq u_{t}-u_{t-1}=\text { length }\left(u_{t-1}, u_{t}\right),
$$

the length of the interval. We now calculate the expected value and variance of $X$. Without loss of generality $a \leq b$. Then easily

$$
\begin{aligned}
E(X) & =\operatorname{Prob}(a \leq \tau)\left(a-u_{t-1}\right)+\operatorname{Prob}(a>\tau)\left(a-u_{t-1}\right) \\
& =\frac{u_{t}-a}{u_{t}-u_{t-1}}\left(a-u_{t-1}\right)+\frac{a-u_{t-1}}{u_{t}-u_{t-1}}\left(a-u_{t}\right) \\
& =0
\end{aligned}
$$

and

$$
\begin{aligned}
E\left(X^{2}\right) & =\operatorname{Prob}(a \leq \tau)\left(a-u_{t-1}\right)^{2}+\operatorname{Prob}(a>\tau)\left(a-u_{t-1}\right)^{2} \\
& =\frac{u_{t}-a}{u_{t}-u_{t-1}}\left(a-u_{t-1}\right)^{2}+\frac{a-u_{t-1}}{u_{t}-u_{t-1}}\left(a-u_{t}\right)^{2} \\
& =\frac{\left(a-u_{t-1}\right)\left(u_{t}-a\right)}{u_{t}-u_{t-1}}\left(a-u_{t-1}-a+u_{t}\right) \\
& =\left(a-u_{t-1}\right)\left(u_{t}-a\right) \\
& \leq \frac{1}{4}\left(u_{t}-u_{t-1}\right)^{2}, \text { since } u_{t-1}<a<u_{t} .
\end{aligned}
$$

For each coordinate $k=1, \ldots, n$, we'll (soon) choose a different partition $\mathcal{P}_{k}$ of $[0,1]$ into $N_{k}$ intervals, with an independent random threshold for each interval in each $\mathcal{P}_{k}$. For each $\mathbf{x}=(x(1), \ldots, x(n))$ we let

$$
\overline{\mathbf{x}}=\left(\overline{\mathbf{x}(\mathbf{1})}_{\mathcal{P}_{1}}, \ldots, \overline{\mathbf{x}(\mathbf{n})_{\mathcal{P}_{n}}}\right) \in \mathbb{R}^{3 N_{1}} \oplus \cdots \oplus \mathbb{R}^{3 N_{n}}
$$

Thus for each $\mathbf{p}_{i}$ we now have a $\overline{\mathbf{p}}_{i}=\left(\overline{\mathbf{p}}_{i}^{(1)}, \ldots, \overline{\mathbf{p}}_{i}^{(n)}\right) \in \mathbb{R}^{3 \sum_{k=1}^{n} N_{k}}$ (where in each coordinate a different partition is used, but not denoted anymore). Then we approximate the identity

$$
I_{m}=\left[-1+2 \sum_{k=1}^{n} \min \left(p_{i}^{(k)}, p_{j}^{(k)}\right)\right]_{i, j=1}^{m}
$$

by

$$
A=\left[-1+2\left\langle\overline{\mathbf{p}}_{i}, \overline{\mathbf{p}}_{j}\right\rangle\right]_{i, j=1}^{m}=\left[-1+2 \sum_{k=1}^{n}\left\langle\overline{\mathbf{p}}_{i}^{(k)}, \overline{\mathbf{p}}_{j}^{(k)}\right\rangle\right]_{i, j=1}^{m}
$$

and we let $X_{i, j, k}=\min \left(p_{i}^{(k)}, p_{j}^{(k)}\right)-\left\langle\overline{\mathbf{p}}_{i}^{(k)}, \overline{\mathbf{p}}_{j}^{(k)}\right\rangle$. Note that $\left\{X_{i, j, k}: k=\right.$ $1, \ldots, n\}$ are independent random variables for fixed $i, j$, since $X_{i, j, k} \equiv 0$ if
$p_{i}^{(k)}$ and $p_{j}^{(k)}$ are in different intervals of $\mathcal{P}_{k}$, or depends on a single threshold, with these thresholds independent as $k=1, \ldots, n$.

We want to apply the rank lemma to $A$, so we start estimating $\left(\sum_{i} a_{i i}\right)^{2}$ and $\sum_{i, j} a_{i j}^{2}$. First of all,

$$
\begin{aligned}
a_{i i} & =-1+2 \sum_{k=1}^{n}\left\langle\overline{\mathbf{p}}_{i}^{(k)}, \overline{\mathbf{p}}_{i}^{(k)}\right\rangle \\
& =-1+2 \sum_{k=1}^{n} \min \left(p_{i}^{(k)}, p_{i}^{(k)}\right)-2 \sum_{k=1}^{n} X_{i, i, k} \\
& =1-2 \sum_{k=1}^{n} X_{i, i, k}
\end{aligned}
$$

Similarly, for $i \neq j$,

$$
a_{i j}=-2 \sum_{k=1}^{n} X_{i, j, k} .
$$

Since $\left|X_{i, i, k}\right| \leq$ length $\left(I_{k}\right)$, where $I_{k}$ is the interval of $\mathcal{P}_{k}$ containing $p_{i}^{(k)}$, we obtain

$$
\sum_{k=1}^{n} X_{i, i, k} \leq \sum_{k=1}^{n} \operatorname{length}\left(I_{k}\right),
$$

and

$$
\begin{align*}
\sum_{i=1}^{m} a_{i i} & =m-2 \sum_{i=1}^{m} \sum_{k=1}^{n} X_{k, i, i} \\
& \geq m-2 \sum_{i=1}^{m} \sum_{k=1}^{n} \operatorname{length}\left(I_{k}\right) \\
& =m-2 \sum_{I} \operatorname{hit}(I) \text { length }(I) \tag{19}
\end{align*}
$$

wehre the last sum is over all intervals of all $\mathcal{P}_{k}$, and $\operatorname{hit}(I)$ is the number of $p_{i}^{(k)}$ "hitting" $I$, i.e., the number of ordered pairs $(i, k)$ with $p_{i}^{(k)} \in I$.

Secondly, for $i \neq j, a_{i j}^{2}=4\left(\sum_{k=1}^{n} X_{i, j, k}\right)^{2}$, hence $E\left(a_{i j}^{2}\right)=4 \sum_{k=1}^{n} E\left(X_{i, j, k}^{2}\right)$ due to independence and $E\left(X_{i, j, k}\right)=0$. Similarly,

$$
\begin{aligned}
E\left(a_{i i}^{2}\right) & =1-4 \sum_{k=1}^{n} E\left(X_{i, i, k}\right)+4 \sum_{k=1}^{n} E\left(X_{i, i, k}^{2}\right) \\
& =1+4 \sum_{k=1}^{n} E\left(X_{i, i, k}^{2}\right) .
\end{aligned}
$$

Thus

$$
\begin{aligned}
E\left(\sum_{i, j=1}^{m} a_{i j}^{2}\right) & =\sum_{i=1}^{m} E\left(a_{i i}^{2}\right)+\sum_{\substack{i, j=1 \\
i \neq j}}^{m} E\left(a_{i j}^{2}\right) \\
& =m+4 \sum_{i=1}^{m} \sum_{k=1}^{n} E\left(X_{i, i, k}^{2}\right)+4 \sum_{\substack{i, j=1 \\
i \neq j}}^{m} \sum_{k=1}^{n} E\left(X_{i, j, k}^{2}\right) \\
& =m+4 \sum_{i, j=1}^{m} \sum_{k=1}^{n} E\left(X_{i, j, k}^{2}\right) \\
& \leq m+\sum_{i, j=1}^{m} \sum_{k=1}^{n} \operatorname{length}\left(I_{i, j, k}\right)^{2}, \begin{array}{l}
\text { with } I_{i, j, k} \in \mathcal{P}_{k} \text { such that } \\
p_{i}^{(k)}, p_{j}^{(k)} \in I_{i, j, k} \text { if it exists, } \\
\text { otherwise take length }\left(I_{i, j, k}\right)= \\
0,
\end{array} \\
& =m+\sum_{I} \operatorname{hit}(I)^{2} \operatorname{length}(I)^{2},
\end{aligned}
$$

where the last sum is again over all intervals. Because we have found an upper bound on the expected value of $\sum_{i, j=1}^{m} a_{i j}^{2}$, it follows that there exists a choice of thresholds such that

$$
\begin{equation*}
\sum_{i, j=1}^{m} a_{i j}^{2} \leq m+\sum_{I} \operatorname{hit}(I)^{2} \text { length }(I)^{2} . \tag{20}
\end{equation*}
$$

Fix such a choice of thresholds. (So at this stage we leave randomness behind.) The rank lemma together with (19) and (20) gives

$$
\operatorname{rank} A \geq \frac{\left(m-2 \sum_{I} \operatorname{hit}(I) \text { length }(I)\right)^{2}}{m+\sum_{I} \operatorname{hit}(I)^{2} \text { length }(I)^{2}}
$$

if

$$
\begin{equation*}
m-2 \sum_{I} \operatorname{hit}(I) \text { length }(I) \geq 0 . \tag{21}
\end{equation*}
$$

On the other hand, $\operatorname{rank} A \leq 1+\sum_{k=1}^{n} 3 N_{k}=1+3 \sum_{I} 1$, thus

$$
\begin{equation*}
\frac{\left(m-2 \sum_{I} \operatorname{hit}(I) \text { length }(I)\right)^{2}}{m+\sum_{I} \operatorname{hit}(I)^{2} \text { length }(I)^{2}} \leq 1+3 \sum_{I} 1 . \tag{22}
\end{equation*}
$$

Finally, we fix the partition $\mathcal{P}_{k}$ for each coordinate $k$. We assume that for each $k=1, \ldots, n$, all $p_{i}^{(k)}$ are distinct. (This is not essential: the $p_{i}^{(k)}$
may all be perturbed by a sufficiently small amount so as to weaken the left-hand side of (22) by at most $\varepsilon>0$, for any $\varepsilon>0$.)

We now assume that both $m$ and $n$ are powers of 4 . If we can prove $m \leq c n \log n$ in this case, then $m \leq 16 c n \log n$ in general, since we may round $m$ down to a power of 4 (thus dividing $m$ by at most 4 ), and $n$ up to a power of 4 (thus multiplying $n$ by at most 4). All the logarithms until the end of the proof will be base 2 (and indicated as such). We first divide $\left[0, \frac{1}{n}\right]$ into $\sqrt{m / n}$ equal parts (thus each of length $1 / \sqrt{m n}$ ). These are called base intervals. Without loss of generality no $p_{i}^{(k)}$ is an endpoint of a base interval, again by making an infinitesimal perturbation.

Then start at $t=1$ and let $t$ go down until

$$
\operatorname{hit}([t, 1]) \operatorname{length}([t, 1]) \geq \frac{c n \lg n}{m}
$$

or until $t=2^{-1 / 3}$, whichever comes first. (The constant $c$ will be fixed later; $c=10000$ is sufficient.) If the stopping point $t_{1}$ equals some $p_{i}^{(k)}$, we go down slightly more, without increasing hit $([t, 1])$. Then start with a new interval at $t_{1}$ and go down until

$$
\operatorname{hit}\left(\left[t, t_{1}\right)\right) \operatorname{length}\left(\left[t, t_{1}\right]\right) \geq \frac{c n \lg n}{m}
$$

or until $t=2^{-2 / 3}$, whichever comes first. In general we go down until

$$
\operatorname{hit}\left(\left[t, t_{s-1}\right)\right) \text { length }\left(\left[t, t_{s-1}\right)\right) \geq \frac{c n \lg n}{m}
$$

or until $t=2^{-s / 3}$, for each $s=1, \ldots, 3 \lg n$.
If $\operatorname{hit}(I) \operatorname{length}(I) \geq c n \lg n / m$, we call $I$ a regular interval, otherwise a singular interval.

Thus for a singular interval $I$ we have

$$
\operatorname{hit}(I) \text { length }(I)<\frac{c n \lg n}{m}, \quad I \text { singular. }
$$

For a regular interval, if the stopping point did not hit a $p_{i}^{(k)}$, then hit $(I)$ length $(I)=c n \lg n / m$. If the stopping point did hit a $p_{i}^{(k)}$, then $\operatorname{hit}(I)$ length $(I)>c n \lg n / m$, but

$$
\text { (hit }(I)-1) \text { length }(I) \leq \frac{c n \lg n}{m}
$$

Thus

$$
\operatorname{hit}(I) \text { length }(I) \geq \frac{c n \lg n}{m}, \quad I \text { regular },
$$

and

$$
\operatorname{hit}(I) \text { length }(I) \leq \frac{2 c n \lg n}{m} \quad \text { if } \operatorname{hit}(I) \geq 2 \text { for regular } I
$$

Later we'll have to take care of regular $I$ with $\operatorname{hit}(I)=1$ separately: for them

$$
\sum_{\substack{I \text { regular } \\ \operatorname{hit}(I)=1}} \operatorname{length}(I)^{2} \leq \sum_{\substack{I \text { regular } \\ \operatorname{hit}(I)=1}} \text { length }(I)<\sum_{I} \text { length }(I)=n .
$$

There are at most $3 n \lg n$ singular intervals. We now bound the number of regular intervals. For each regular interval $I$, let $s_{I}$ be the $s$ for which $I \subseteq\left[2^{-s / 3}, 2^{-(s-1) / 3}\right]$. Then

$$
\begin{align*}
\sum_{\text {regular } I} 2^{-s_{I} / 3} \operatorname{hit}(I) & \leq \sum_{\text {regular } I} \min (I) \operatorname{hit}(I) \\
& <\sum_{I} \sum_{\substack{i=1 \\
p_{i}^{(k)} \in I}}^{m} \min (I), \quad \text { where } k \text { is such that } I \in \mathcal{P}_{k} \\
& <\sum_{I} \sum_{\substack{i=1 \\
p_{i}^{(k)} \in I}}^{m} p_{i}^{(k)}=\sum_{i=1}^{m} \sum_{k=1}^{n} p_{i}^{(k)}=m \tag{23}
\end{align*}
$$

We now bound $\sum_{\text {regular } I} 1 /\left(2^{-s_{I} / 3} \operatorname{hit}(I)\right)$. For each $s=1, \ldots, 3 \lg n$,

$$
\sum_{\substack{\text { regular } I \\
s_{I}=s}} \operatorname{length}(I) \leq\left(2^{-(s-1) / 3}-2^{-s / 3}\right) n \quad \begin{aligned}
& \text { (since for each coordinate the } \\
& \\
& I \text { 's don't overlap) }
\end{aligned}
$$

$$
=\left(2^{1 / 3}-1\right) 2^{2 / 3} n<\frac{1}{3} 2^{-s / 3} n .
$$

Since length $(I) \operatorname{hit}(I) \geq c n \lg n / m$ for regular $I$,

$$
\frac{1}{\operatorname{hit}(I)} \leq \frac{m \text { length }(I)}{c n \lg n}
$$

Thus

$$
\begin{aligned}
\sum_{\substack{\text { regular } \\
s_{I}=s}} \frac{1}{2^{-s / 3} \operatorname{hit}(I)} & \leq \frac{m}{2^{-s / 3} c n \lg n} \sum_{\substack{\text { regular } I \\
s_{I}=s}} \text { length }(I) \\
& <\frac{m}{3 c \lg n}
\end{aligned}
$$

Thus

$$
\begin{equation*}
\sum_{\text {regular } I} \frac{1}{2^{-s_{I} / 3} \operatorname{hit}(I)}<\sum_{s=1}^{3 \lg n} \frac{m}{3 c \lg n}=\frac{m}{c} \tag{24}
\end{equation*}
$$

By Cauchy-Schwartz,

$$
\begin{aligned}
\left(\sum_{\text {regular } I} 1\right)^{2} & =\left(\sum_{\text {regular } I} \sqrt{2^{-s_{I} / 3} \operatorname{hit}(I)} \frac{1}{\sqrt{2^{-s_{I} / 3} \operatorname{hit}(I)}}\right)^{2} \\
& \leq\left(\sum_{\text {regular } I} 2^{-s_{I} / 3} \operatorname{hit}(I)\right)\left(\sum_{\text {regular } I} \frac{1}{2^{-s_{I} / 3} \operatorname{hit}(I)}\right) \\
& <m \frac{m}{c} \quad \text { by }(23) \text { and }(24),
\end{aligned}
$$

and the number of regular intervals is less than $m / \sqrt{c}$. Thus the total number of non-base intervals is at most $3 n \lg n+m / \sqrt{c}$ (and recall that the number of base intervals is $\sqrt{m n}$ ). We are now in a position to estimate the various quantities in (22). We want to show that these estimates imply $m \leq c n \lg n$, so we assume that $m>c n \lg n$ and aim for a contradiction if $c$ is sufficiently large ( 10000 will do). First of all,

$$
\begin{align*}
1+3 \sum_{I} 1 & <1+3\left(\sqrt{m n}+3 n \lg n+\frac{m}{\sqrt{c}}\right) \\
& <1+\frac{3 m}{\sqrt{c \lg n}}+\frac{9 m}{c}+\frac{3 m}{\sqrt{c}} \\
& <\frac{16 m}{\sqrt{c}} \tag{25}
\end{align*}
$$

to be sure. Secondly,

$$
\begin{align*}
m+\sum_{I} \operatorname{hit}(I)^{2} \operatorname{length}(I)^{2}= & m+\sum_{\text {base } I} \operatorname{hit}(I)^{2} \frac{1}{m n}+\sum_{\substack{\text { regular } I \\
\text { hit }(I)=1}} \operatorname{length}(I)^{2} \\
& +\sum_{\substack{\text { non-base } I \\
h_{I} \geq 2 \text { if regular }}}(\operatorname{hit}(I) \operatorname{length}(I))^{2} \\
< & m+n m^{2} \frac{1}{m n}+n+\left(3 n \lg n+\frac{m}{\sqrt{c}}\right)\left(\frac{2 c n \lg n}{m}\right)^{2} \\
< & 2 m+\frac{m}{c \lg n}+\frac{12 m}{c}+\frac{4 m}{\sqrt{c}} \\
< & 2 m+\frac{17 m}{\sqrt{c}} . \tag{26}
\end{align*}
$$

Thirdly,

$$
\begin{aligned}
\sum_{I} \operatorname{hit}(I) \operatorname{length}(I)= & \sum_{\text {base } I} \operatorname{hit}(I) \frac{1}{\sqrt{m n}}+\sum_{\substack{\text { regular } I \\
\text { hit }(I)=1}} \operatorname{length}(I) \\
& +\sum_{\substack{\text { non-base } I \\
h_{I} \geq 2 \text { if regular }}} \operatorname{hit}(I) \text { length }(I) \\
& <n m \frac{1}{\sqrt{m n}}+n+\left(3 n \lg n+\frac{m}{\sqrt{c}}\right)\left(\frac{2 c n \lg n}{m}\right) \\
< & \frac{m}{\sqrt{c \lg n}}+\frac{m}{c \lg n}+\frac{6 m}{c}+\frac{2 m}{\sqrt{c}} \\
< & \frac{10 m}{\sqrt{c}} .
\end{aligned}
$$

Thus if $\sqrt{c} \geq 20$, then (21) is satisfied, and in fact

$$
\begin{equation*}
m-2 \sum_{I} \operatorname{hit}(I) \text { length }(I)>m-\frac{10 m}{\sqrt{c}} . \tag{27}
\end{equation*}
$$

Substituting estimates (25), (26), and (27) into (22) we obtain

$$
\frac{16 m}{\sqrt{c}}>\frac{\left(m-\frac{10 m}{\sqrt{c}}\right)^{2}}{2 m+\frac{17 m}{\sqrt{c}}}, \text { i.e., } \frac{16}{\sqrt{c}}>\frac{\left(1-\frac{10}{\sqrt{c}}\right)^{2}}{2+\frac{17}{\sqrt{c}}}
$$

which is false if $c$ is sufficiently large, e.g. for $\sqrt{c}=100$.
Thus $m<c n \lg n$, and the proof is finished.

## 9 Final remarks

### 9.1 Infinite dimensions

In infinite dimensions almost nothing is known. The most important open question here is the following.

Problem 36. Does there exist an infinite equilateral set in any separable infinite-dimensional normed space?

As observed by Terenzi [38], it follows from the partition calculus of set theory that a normed space of cardinality at least $2^{c}$ has an infinite equilateral set (in fact an uncountable one), where $c=2^{\aleph_{0}}$ is the cardinality of the continuum. The following two easy results are proved in [35].

Theorem 37. Let $\varepsilon>0$ and $X$ an infinite dimensional Banach space. Then $X$ has an equivalent norm with Banach-Mazur distance of at most $2(1+\varepsilon)$ to the original norm, admitting an infinte equilateral set.

Theorem 38. Let $\varepsilon>0$ and $X$ an infinite dimensional superreflexive Banach space. Then $X$ has an equivalent norm with Banach-Mazur distance of at most $1+\epsilon$ to the original norm, admitting an infinite equilateral set.

### 9.2 Generalizations

An $\varepsilon$-almost-equilateral set in a normed space is a set $S$ satisfying $1-\varepsilon \leq$ $\|\mathbf{x}-\mathbf{y}\| \leq 1+\varepsilon$ for all $\mathbf{x}, \mathbf{y} \in S, \mathbf{x} \neq \mathbf{y}$. Such sets have been studied in $[22,9,7,6,5,2]$.

Problem 39. Prove that for each $\varepsilon>0$ there exists $\delta>0$ such that each $n$-dimensional normed space has an $\varepsilon$-almost-equilateral set of size $(1+\delta)^{n}$.

Equilateral sets may be generalized to the notion of a $k$-distance set, i.e., a subset $S$ of a metric space such that $\{d(x, y): x, y \in S, x \neq y\}$ consists of at most $k$ numbers. Then the Euclidean case is already non-trivial, and for normed spaces a few results are known: see [36]. The following conjecture would generalize Petty's theorem:

Conjecture 3 ([36]). A $k$-distance set in an $n$-dimensional normed space has size at most $(k+1)^{n}$. If equality holds for some $k \geq 1$, then the space must be isometric to $\ell_{\infty}^{n}$.

It is known to be true for $n=2$ and for $\ell_{\infty}^{n}$, while for general Minkowski spaces there are some weaker estimates; see [36].

## Acknowledgements

These are lecture notes for a 4 -hour course given on April 20-23, 2004. at the Department of Mathematical Analysis, University of Seville. I thank the University of Seville for their kind invitation and generous hospitality. I thank Rafael Villa, Juan Carlos García-Vázquez, and José Antonio Facenda for valuable hints.

## References

[1] N. Alon and P. Pudlák, Equilateral sets in lscalarproductn ${ }_{p}$, Geom. Funct. Anal. 13 (2003), no. 3, 467-482. MR 1995795
[2] Juan Arias-de Reyna, Keith Ball, and Rafael Villa, Concentration of the distance in finite-dimensional normed spaces, Mathematika 45 (1998), no. 2, 245-252. MR 2000b:46013
[3] H.-J. Bandelt, V. Chepoi and M. Laurent, Embedding into rectilinear spaces, Discrete Comput. Geom 19 (1998), 595-604. MR 99d:51017
[4] I. Bárány, The densest $(n+2)$-set in $R^{n}$, Intuitive Geometry (1991), Coll. Math. Soc. János Bolyai 63, Elsevier, Amsterdam, 1994, pp. 7-10. MR 97a:52012
[5] J. Bastero, A. Peña, and G. Schechtman, Embedding lscalarproductn $\infty^{-}$ cubes in low-dimensional Schatten classes, Geometric aspects of functional analysis (Israel, 1992-1994), Oper. Theory Adv. Appl., vol. 77, Birkhäuser, Basel, 1995, pp. 5-11. MR 96j:46006
[6] Jesús Bastero and Julio Bernués, Applications of deviation inequalities on finite metric sets, Math. Nachr. 153 (1991), 33-41. MR 93a:46017
[7] Jesús Bastero, Julio Bernués, and Nigel Kalton, Embedding lscalarproductn $n_{\infty}$-cubes in finite-dimensional 1-subsymmetric spaces, Rev. Mat. Univ. Complut. Madrid 2 (1989), no. suppl., 47-52, Congress on Functional Analysis (Madrid, 1988). MR 91h:46024
[8] Leonard M. Blumenthal, Theory and applications of distance geometry, Second edition, Chelsea Publishing Co., New York, 1970. MR 42 \#3678
[9] J. Bourgain, V. Milman, and H. Wolfson, On type of metric spaces, Trans. Amer. Math. Soc. 294 (1986), no. 1, 295-317. MR 88h:46033
[10] Peter Braß, On equilateral simplices in normed spaces, Beiträge Algebra Geom. 40 (1999), no. 2, 303-307. MR 2000i:52012
[11] E. W. Cheney, Introduction to approximation theory, AMS Chelsea Publishing, Providence, RI, 1998, Reprint of the second (1982) edition. MR 99f:41001
[12] L. Danzer and B. Grünbaum, Über zwei Probleme bezüglich konvexer Körper von P. ErdHos und von V. L. Klee, Math. Z. 79 (1962), 95-99. MR 25 \#1488
[13] B. V. Dekster, Simplexes with prescribed edge lengths in Minkowski and Banach spaces, Acta Math. Hungar. 86 (2000), no. 4, 343-358. MR 2001b:52001
[14] B. V. Dekster and J. B. Wilker, Edge lengths guaranteed to form a simplex, Arch. Math. (Basel) 49 (1987), no. 4, 351-366. MR 88k:51035
[15] Z. Füredi, J. C. Lagarias, and F. Morgan, Singularities of minimal surfaces and networks and related extremal problems in Minkowski space, Discrete and computational geometry (New Brunswick, NJ, 1989/1990), DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 6, Amer. Math. Soc., Providence, RI, 1991, pp. 95-109. MR 93d:52009
[16] Yehoram Gordon, Some inequalities for Gaussian processes and applications, Israel J. Math. 50 (1985), no. 4, 265-289. MR 87f:60058
[17] Helmut Groemer, Abschätzungen für die Anzahl der konvexen Körper, die einen konvexen Körper berühren, Monatsh. Math. 65 (1961), 74-81. MR 23 \#A2129
[18] Branko Grünbaum, On a conjecture of H. Hadwiger, Pacific J. Math. 11 (1961), 215-219. MR 25 \#1492
[19] Branko Grünbaum, Strictly antipodal sets, Israel J. Math. 1 1963 5-10. MR 28 \#2480
[20] R. K. Guy, An olla-podrida of open problems, often oddly posed, Amer. Math. Monthly 90 (1983), 196-199.
[21] Tao Jiang and Alexander Vardy, Asymptotic Improvement of the Gilbert-Varshamov Bound on the Size of Binary Codes, arXiv:math.CO/0404325, to appear in IEEE Transactions on Information Theory, 2004.
[22] William B. Johnson and Joram Lindenstrauss, Extensions of Lipschitz mappings into a Hilbert space, Conference in modern analysis and probability (New Haven, Conn., 1982), Contemp. Math., vol. 26, Amer. Math. Soc., Providence, RI, 1984, pp. 189-206. MR 86a:46018
[23] Jack Koolen, Monique Laurent and Alexander Schrijver, Equilateral dimension of the rectilinear space, Designs, Codes Cryptogr. 21 (2000), 149-164. MR 2001j:52013
[24] Gary Lawlor and Frank Morgan, Paired calibrations applied to soap films, immiscible fluids, and surfaces or networks minimizing other norms, Pacific J. Math. 166 (1994), no. 1, 55-83. MR 95i:58051
[25] J. H. van Lint, Introduction to Coding Theory, third ed., SpringerVerlag, Berlin, 1992.
[26] Horst Martini, Konrad J. Swanepoel, and Gunter Weiß, The geometry of Minkowski spaces - a survey. I, Expo. Math. 19 (2001), no. 2, 97142. MR 2002h:46015a. Erratum Expo. Math. 19 (2001), no. 4, 364. MR 2002h:46015b
[27] K. Menger, Untersuchungen über allgemeine Metrik, Math. Ann. 100 (1928), 75-163.
[28] Frank Morgan, Minimal surfaces, crystals, shortest networks, and undergraduate research, Math. Intelligencer 14 (1992), no. 3, 37-44. MR 93h:53012
[29] C. M. Petty, Equilateral sets in Minkowski spaces, Proc. Amer. Math. Soc. 29 (1971), 369-374. MR 43 \#1051
[30] K. Schütte, Minimale Durchmesser endlicher Punktmengen mit vorgeschriebenem Mindestabstand, Math. Ann. 150 (1963), 91-98. MR 26 \#5479
[31] J. J. Seidel, Quasiregular two-distance sets, Indag. Math 31 (1969), 64-69.
[32] C. E. Shannon, Probability of error for optimal codes in a Gaussian channel, Bell System Tech. J. 38 (1959), 611-656.
[33] Cliff Smyth, Equilateral or 1-distance sets and Kusner's conjecture, manuscript, 2002.
[34] P. S. Soltan, Analogues of regular simplexes in normed spaces, Dokl. Akad. Nauk SSSR 222 (1975), no. 6, 1303-1305, English translation: Soviet Math. Dokl. 16 (1975), no. 3, 787-789. MR 52 \#4127
[35] K. J. Swanepoel, Combinatorial geometry of Minkowski spaces, PhD Thesis, University of Pretoria, 1997.
[36] K. J. Swanepoel, Cardinalities of $k$-distance sets in Minkowski spaces, Discrete Math. 197/198 (1999), 759-767, 16th British Combinatorial Conference (London, 1997). MR 99k:52028
[37] K. J. Swanepoel, A problem of Kusner on equilateral sets, to appear in Archiv der Mathematik, 2004.
[38] Paolo Terenzi, Equilater sets in Banach spaces, Boll. Un. Mat. Ital. A (7) 3 (1989), no. 1, 119-124. MR 90c:46017
[39] A. C. Thompson, Minkowski geometry, Encyclopedia of Mathematics and its Applications, vol. 63, Cambridge University Press, Cambridge. 1996. MR 97f:52001
[40] A. D. Wyner, Capabilities of bounded discrepancy decoding, Bell System Tech. J. 44 (1965), 1061-1122.

## Author's address:

Konrad J. Swanepoel<br>Department of Mathematics, Applied Mathematics and Astronomy University of South Africa, PO Box 392<br>Pretoria 0003, South Africa.<br>E-mail: swanekj@unisa.ac.za

# Analytic capacity and Calderón-Zygmund theory with non doubling measures 

Xavier Tolsa *


#### Abstract

These notes are the lecture notes of a series of talks given at the Universidad de Sevilla in December 2003. We survey some results of Calderón-Zygmund theory with non doubling measures, and we apply them to prove the semiaddivity of the analytic capacity $\gamma_{+}$. We provide a quite elementary proof which does not use the $T(1)$ theorem. We also review other recent results in connection with the comparability between analytic capacity and the capacity $\gamma_{+}$.


## 1 Introduction

The main purpose of this expository paper is to discuss and review several results on Calderón-Zygmund theory with non doubling measures (also known as non homogeneous Calderón-Zygmund theory) and to show how these results can be applied to problems related to analytic capacity.

In recent years it was shown that many results on Calderón-Zygmund theory remain valid if one does not assume that the underlying measure of the space is doubling. Recall that a Borel measure $\mu$ on $\mathbb{R}^{d}$ is said to be doubling if there exists some constant $C>0$ such that

$$
\mu(B(x, 2 r)) \leq C \mu(B(x, r)) \quad \text { for all } x \in \operatorname{supp}(\mu), r>0 .
$$

One of the main motivations for extending the classical theory to the non doubling context was the solution of several questions related to analytic capacity, like Vitushkin's conjecture or Painlevé's problem. In this type of problems, one considers an arbitrary compact set $E$ in the complex plane and one is interested in finding a Radon measure $\mu$ supported on it such that the Cauchy transform $\mathcal{C}_{\mu}$ (see Section 2 for the precise definition) is

[^9]bounded on $L^{2}(\mu)$. It may happen that the only non zero measures with these properties are non doubling.

In order to study $n$-dimensional Calderón-Zygmund operators (CZO's) in $\mathbb{R}^{d}$, with $0<n \leq d$, we will consider measures $\mu$ satisfying the growth condition

$$
\begin{equation*}
\mu(B(x, r)) \leq C_{0} r^{n} \quad \text { for all } x \in \mathbb{R}^{d}, r>0 \tag{1}
\end{equation*}
$$

Let us remark that this is a quite natural condition, because it is necessary for the $L^{2}(\mu)$ boundedness of any CZO whose kernel $k(x, y)$ satisfies $|k(x, y)| \geq C|x-y|^{-n}$ (see [5, Theorem III.1.4]).

One of the main difficulties that arises when one deals with a non doubling measure $\mu$ is due to the fact that the non centered maximal HardyLittlewood operator

$$
M_{\mu}^{n c} f(x):=\sup \left\{\frac{1}{\mu(\bar{B})} \int_{\bar{B}}|f| d \mu: \bar{B} \text { closed ball, } x \in \bar{B}\right\}
$$

may fail to be of weak type $(1,1)$ (the superindex "nc" stands for non centered). Sometimes the centered version of the operator, that is

$$
M_{\mu} f(x)=\sup _{r>0} \frac{1}{\mu(B(x, r))} \int_{B(x, r)}|f| d \mu
$$

is a good substitute of $M_{\mu}^{n c} f$, because using Besicovitch's covering theorem one can show that $M_{\mu}$ is bounded from $L^{1}(\mu)$ into $L^{1, \infty}(\mu)$, and in $L^{p}(\mu)$, for $1<p \leq \infty$. However, one cannot always use the centered maximal Hardy-Littlewood operator instead of the non centered one. In these cases, other arguments (usually more involved) are required.

This paper is not intended to be a complete survey neither on CalderónZygmund theory with non doubling measures nor on analytic capacity. We recommend the interested reader to have a look at the surveys [6], [28], [54], [26], for example.

Regarding non homogeneous Calderón-Zygmund theory, we will focus our attention on some of the results more directly connected to analytic capacity. In Section 3, for example, we will review the proof of the weak $(1,1)$ boundedness of CZO's which are bounded in $L^{2}(\mu)$, using a CalderónZygmund type decomposition adapted to the non doubling context. We will also give the detailed proof of Cotlar's inequality, which we think that is particularly simple and illuminating. We will state and discuss (but not prove) the $T(1)$ theorem. On the other hand, for reasons of brevity and simplicity we will not pay much attention to $T(b)$ type theorems, although they are important results which play a very important role in connection
with analytic capacity. We ask the reader the to forgive us about this question. Similarly, we will only make some brief comments about other results dealing with the space RBMO, Hardy spaces, commutators, weights. etc.

The second part of the paper is dedicated to analytic capacity. In Section 4, we review some properties of analytic capacity and its connection with the Cauchy transform, Menger curvature, and rectifiability. In Section 5 we obtain several characterizations of the analytic capacity $\gamma_{+}$using some of the results proved or described previously about non homogeneous CalderónZygmund theory. In particular, from one of these characterizations the semiadditivity of $\gamma_{+}$follows in a straightforward way. Moreover, we provide a quite elementary proof of the semiadditivity of $\gamma_{+}$which does not use the $T(1)$ theorem (although we also explain the argument which uses the $T(1)$ theorem).

The proof of the semiadditivity of $\gamma$ and its comparability with $\gamma_{+}$requires much more work and it is out of the scope of this paper. Nevertheless, the last section contains some comments about this topic and related results.

## 2 Preliminaries

An open ball centered at $x$ with radius $r$ is denoted by $B(x, r)$, and a closed ball by $\bar{B}(x, r)$. By a cube $Q$ we mean a closed cube with sides parallel to the axes. We denote its side length by $\ell(Q)$ and its center by $x_{Q}$.

A Radon measure on $\mathbb{R}^{d}$ has growth of degree $n$ (or is of degree $n$ ) if there exists some constant $C_{0}$ such that $\mu(B(x, r)) \leq C_{0} r^{n}$ for all $x \in \mathbb{R}^{d}$, $r>0$. When $n=1$, we say that $\mu$ has linear growth. If there exists some constant $C$ such that

$$
C^{-1} r \leq \mu(B(x, r)) \leq C r \text { for all } x \in \operatorname{supp}(\mu), 0<r \leq \operatorname{diam}(\operatorname{supp}(\mu))
$$

then we say that $\mu$ is $n$-dimensional AD-regular
The space of finite complex Radon measures on $\mathbb{R}^{d}$ is denoted by $M\left(\mathbb{R}^{d}\right)$. This is a Banach space with the norm of the total variation: $\|\mu\|=|\mu|\left(\mathbb{R}^{d}\right)$.

We say that $k(\cdot, \cdot): \mathbb{R}^{d} \times \mathbb{R}^{d} \backslash\left\{(x, y) \in \mathbb{R}^{d} \times \mathbb{R}^{d}: x=y\right\} \rightarrow \mathbb{C}$ is an $n$-dimensional Calderón-Zygmund kernel if there exist constants $C>0$ and $\eta$, with $0<\eta \leq 1$, such that the following inequalities hold for all $x, y \in \mathbb{R}^{d}$,
$x \neq y:$

$$
\begin{align*}
& |k(x, y)| \leq \frac{C}{|x-y|^{n}}, \quad \text { and }  \tag{2}\\
& \left|k(x, y)-k\left(x^{\prime}, y\right)\right| \leq \frac{C\left|x-x^{\prime}\right|^{\eta}}{|x-y|^{n+\eta}} \quad \text { if }\left|x-x^{\prime}\right| \leq|x-y| / 2
\end{align*}
$$

Given a positive or complex Radon measure $\mu$ on $\mathbb{R}^{d}$, we define

$$
\begin{equation*}
T \mu(x):=\int k(x, y) d \mu(y), \quad x \in \mathbb{R}^{d} \backslash \operatorname{supp}(\mu) \tag{3}
\end{equation*}
$$

We say that $T$ is an $n$-dimensional Calderón-Zygmund operator (CZO) with kernel $k(\cdot, \cdot)$. The integral in the definition may not be absolutely convergent if $x \in \operatorname{supp}(\mu)$. For this reason, we consider the following $\varepsilon$-truncated operators $T_{\varepsilon}, \varepsilon>0$ :

$$
T_{\varepsilon} \mu(x):=\int_{|x-y|>\varepsilon} k(x, y) d \mu(y), \quad x \in \mathbb{R}^{d}
$$

Observe that now the integral on the right hand side converges absolutely if, for instance, $|\mu|\left(\mathbb{R}^{d}\right)<\infty$.

Given a fixed positive Radon measure $\mu$ on $\mathbb{R}^{d}$ and $f \in L_{l o c}^{1}(\mu)$, we denote

$$
T_{\mu} f(x):=T(f d \mu)(x) \quad x \in \mathbb{C} \backslash \operatorname{supp}(f d \mu)
$$

and

$$
T_{\mu, \varepsilon} f(x):=T_{\varepsilon}(f d \mu)(x) .
$$

The last definition makes sense for all $x \in \mathbb{R}^{d}$ if, for example, $f \in L^{1}(\mu)$. We say that $T_{\mu}$ is bounded on $L^{2}(\mu)$ if the operators $T_{\mu, \varepsilon}$ are bounded on $L^{2}(\mu)$ uniformly on $\varepsilon>0$. Analogously, with respect to the boundedness from $L^{1}(\mu)$ into $L^{1, \infty}(\mu)$. We also say that $T$ is bounded from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$ if there exists some constant $C$ such that for all $\nu \in M\left(\mathbb{R}^{d}\right)$ and all $\lambda>0$,

$$
\mu\left\{x \in \mathbb{R}^{d}:\left|T_{\varepsilon} \nu\right|>\lambda\right\} \leq \frac{C\|\nu\|}{\lambda}
$$

uniformly on $\varepsilon>0$.
The Cauchy transform is the CZO on $\mathbb{C}$ originated by the kernel

$$
k(x, y):=\frac{1}{y-x}, \quad x, y \in \mathbb{C} .
$$

It is denoted by $\mathcal{C}$. That is to say,

$$
\mathcal{C} \mu(x):=\int \frac{1}{y-x} d \mu(y), \quad x \in \mathbb{C} \backslash \operatorname{supp}(\mu)
$$

As usual, in the paper the letter ' $C$ ' stands for an absolute constant which may change its value at different occurrences. On the other hand, constants with subscripts, such as $C_{1}$, retain its value at different occurrences. The notation $A \lesssim B$ means that there is a positive absolute constant $C$ such that $A \leq C B$. Also, $A \approx B$ is equivalent to $A \lesssim B \lesssim A$.

## 3 Calderón-Zygmund theory with non doubling measures

In this section we will review some results about Calderón-Zygmund theory for non doubling measures $\mu$ in $\mathbb{R}^{d}$ satisfying the growth condition (1) that will be useful in connection with analytic capacity. First, we will describe a Calderón-Zygmund decomposition suitable for this type of measures, and then we will show how one can use it to prove that a CZO which is bounded on $L^{2}(\mu)$ is also of weak type $(1,1)$. Further, we will prove Cotlar's inequality, and we will talk about the $T(1)$ theorem, and other results.

Preliminarily, in next subsection, we deal with the existence and properties of the so called doubling cubes, which play a very important role in this theory.

### 3.1 Doubling cubes

Given $\alpha>1$ and $\beta>\alpha^{n}$, we say that $Q$ is $(\alpha, \beta)$-doubling if $\mu(\alpha Q) \leq$ $\beta \mu(Q)$, where $\alpha Q$ is the cube concentric with $Q$ with side length $\alpha \ell(Q)$. For definiteness, if $\alpha$ and $\beta$ are not specified, by a doubling cube we mean a ( $2,2^{d+1}$ )-doubling cube.

Before proving Theorem 3, we state some remarks about the existence of doubling cubes.

Because $\mu$ satisfies the growth condition (1), there are a lot of "big" doubling cubes. To be precise, given any point $x \in \operatorname{supp}(\mu)$ and $c>0$, there exists some ( $\alpha, \beta$ )-doubling cube $Q$ centered at $x$ with $l(Q) \geq c$. This follows easily from (1) and the fact that $\beta>\alpha^{n}$. Indeed, if there are no doubling cubes centered at $x$ with $l(Q) \geq c$, then $\mu\left(\alpha^{m} Q\right)>\beta^{m} \mu(Q)$ for each $m$, and letting $m \rightarrow \infty$ one sees that (1) cannot hold.

There are a lot of "small" doubling cubes too: if $\beta>\alpha^{d}$, then for $\mu$-a:e. $x \in \mathbb{R}^{d}$ there exists a sequence of $(\alpha, \beta)$-doubling cubes $\left\{Q_{k}\right\}_{k}$ centered at
$x$ with $\ell\left(Q_{k}\right) \rightarrow 0$ as $k \rightarrow \infty$. This is a property that any Radon measure on $\mathbb{R}^{d}$ satisfies (the growth condition (1) is not necessary in this argument). The proof is an easy exercise on geometric measure theory that is left for the reader.

Observe that, by the Lebesgue differentiation theorem, for $\mu$-almost all $x \in \mathbb{R}^{d}$ one can find a sequence of $\left(2,2^{d+1}\right)$-doubling cubes $\left\{Q_{k}\right\}_{k}$ centered at $x$ with $\ell\left(Q_{k}\right) \rightarrow 0$ such that

$$
\lim _{k \rightarrow \infty} \frac{1}{\mu\left(Q_{k}\right)} \int_{Q_{k}} f d \mu=f(x)
$$

As a consequence, for any fixed $\lambda>0$, for $\mu$-almost all $x \in \mathbb{R}^{d}$ such that $|f(x)|>\lambda$, there exists a sequence of cubes $\left\{Q_{k}\right\}_{k}$ centered at $x$ with $\ell\left(Q_{k}\right) \rightarrow 0$ such that

$$
\limsup _{k \rightarrow \infty} \frac{1}{\mu\left(2 Q_{k}\right)} \int_{Q_{k}}|f| d \mu>\frac{\lambda}{2^{d+1}}
$$

In next lemma we prove a very useful estimate from [8] involving non doubling squares which relies on the idea that the mass $\mu$ which lives on non doubling squares must be small.

Lemma 1. If $Q \subset R$ are concentric cubes such that there are no $(\alpha, \beta)$ doubling cubes (with $\beta>\alpha^{n}$ ) of the form $\alpha^{k} Q, k \geq 0$, with $Q \subset \alpha^{k} Q \subset R$, then,

$$
\int_{R \backslash Q} \frac{1}{\left|x-x_{Q}\right|^{n}} d \mu(x) \leq C_{1}
$$

where $C_{1}$ depends only on $\alpha, \beta, n, d$ and $C_{0}$.
Proof. Let $N$ be the least integer such that $R \subset \alpha^{N} Q$. For $0 \leq k \leq N$ we have $\mu\left(\alpha^{k} Q\right) \leq \mu\left(\alpha^{N} Q\right) / \beta^{N-k}$. Then,

$$
\begin{aligned}
\int_{R \backslash Q} \frac{1}{\left|x-x_{Q}\right|^{n}} d \mu(x) & \leq \sum_{k=1}^{N} \int_{\alpha^{k} Q \backslash \alpha^{k-1} Q} \frac{1}{\left|x-x_{Q}\right|^{n}} d \mu(x) \\
& \leq C \sum_{k=1}^{N} \frac{\mu\left(\alpha^{k} Q\right)}{\ell\left(\alpha^{k} Q\right)^{n}} \\
& \leq C \sum_{k=1}^{N} \frac{\beta^{k-N} \mu\left(\alpha^{N} Q\right)}{\alpha^{(k-N) n} \ell\left(\alpha^{N} Q\right)^{n}} \\
& \leq C \frac{\mu\left(\alpha^{N} Q\right)}{\ell\left(\alpha^{N} Q\right)^{n}} \sum_{j=0}^{\infty}\left(\frac{\alpha^{n}}{\beta}\right)^{j} \leq C
\end{aligned}
$$

### 3.2 Calderón-Zygmund decomposition

Lemma 2 (Calderón-Zygmund decomposition). Assume that $\mu$ satisfies (1). For any $f \in L^{1}(\mu)$ and any $\lambda>0$ (with $\lambda>2^{d+1}\|f\|_{L^{1}(\mu)} /\|\mu\|$ if $\|\mu\|<\infty)$ we have:
(a) There exists a family of almost disjoint cubes $\left\{Q_{i}\right\}_{i}$ (that is, $\sum_{i} \chi_{Q_{i}} \leq$ C) such that

$$
\begin{gather*}
\frac{1}{\mu\left(2 Q_{i}\right)} \int_{Q_{i}}|f| d \mu>\frac{\lambda}{2^{d+1}},  \tag{4}\\
\frac{1}{\mu\left(2 \eta Q_{i}\right)} \int_{\eta Q_{i}}|f| d \mu \leq \frac{\lambda}{2^{d+1}} \quad \text { for } \eta>2,  \tag{5}\\
|f| \leq \lambda \quad \text { a.e. }(\mu) \text { on } \mathbb{R}^{d} \backslash \bigcup_{i} Q_{i} . \tag{6}
\end{gather*}
$$

(b) For each $i$, let $R_{i}$ be a $\left(6,6^{n+1}\right)$-doubling cube concentric with $Q_{i}$, with $l\left(R_{i}\right)>4 l\left(Q_{i}\right)$ and denote $w_{i}=\frac{\chi Q_{i}}{\sum_{k} \chi Q_{k}}$. Then, there exists a family of functions $\varphi_{i}$ with $\operatorname{supp}\left(\varphi_{i}\right) \subset R_{i}$ and with constant sign satisfying

$$
\begin{equation*}
\int \varphi_{i} d \mu=\int_{Q_{i}} f w_{i} d \mu \tag{7}
\end{equation*}
$$

$$
\begin{equation*}
\sum_{i}\left|\varphi_{i}\right| \leq B \lambda \tag{8}
\end{equation*}
$$

(where $B$ is some constant), and

$$
\begin{equation*}
\left\|\varphi_{i}\right\|_{L^{\infty}(\mu)} \mu\left(R_{i}\right) \leq C \int_{Q_{i}}|f| d \mu . \tag{9}
\end{equation*}
$$

The lemma above was obtained in [44], where it was used to prove that if a linear operator is bounded from a suitable space of type $H^{1}$ into $L^{1}(\mu)$ and from $L^{\infty}(\mu)$ into a space of type $B M O$, then it is bounded in $L^{p}(\mu)$. for $1<p<\infty$. This Calderón-Zygmund decomposition has also shown to be useful in a variety of other situations (see, for example, [45], [17], [18]).

### 3.3 Weak $(1,1)$ boundedness of Calderón-Zygmund operators

The result below was first obtained in [35], although a previous proof valid only for the Cauchy transform appeared in [42]. Below we reproduce the proof of [45], which is different from the one of [35] and it is based on the Calderón-Zygmund decomposition of Lemma 2.

Theorem 3. Let $\mu$ be a Radon measure on $\mathbb{R}^{d}$ satisfying the growth condition (1). If $T$ is an $n$-dimensional Calderón-Zygmund operator which is bounded in $L^{2}(\mu)$, then it is also bounded from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$. In particular, it is of weak type $(1,1)$. (as far as we know)

Proof. We will show that $T_{\mu}$ is of weak type $(1,1)$. By similar arguments, one gets that $T$ is bounded from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$. In this case, one has to use a version of the Calderón-Zygmund decomposition in the lemma above suitable for complex measures (see the end of the proof for more details).

Let $f \in L^{1}(\mu)$ and $\lambda>0$. It is straightforward to check that we may assume $\lambda>2^{d+1}\|f\|_{L^{1}(\mu)} /\|\mu\|$. Let $\left\{Q_{i}\right\}_{i}$ be the almost disjoint family of cubes of Lemma 2. Let $R_{i}$ be the smallest $\left(6,6^{n+1}\right)$-doubling cube of the form $6^{k} Q_{i}, k \geq 1$. Then we can write $f=g+b$, with

$$
g=f \chi_{\mathbb{R}^{d} \backslash \bigcup_{i} Q_{i}}+\sum_{i} \varphi_{i}
$$

and

$$
b=\sum_{i} b_{i}:=\sum_{i}\left(w_{i} f-\varphi_{i}\right),
$$

where the functions $\varphi_{i}$ satisfy (7), (8) (9) and $w_{i}=\frac{\chi Q_{i}}{\sum_{k} \chi Q_{k}}$.
By (4) we have

$$
\mu\left(\bigcup_{i} 2 Q_{i}\right) \leq \frac{C}{\lambda} \sum_{i} \int_{Q_{i}}|f| d \mu \leq \frac{C}{\lambda} \int|f| d \mu .
$$

So we have to show that

$$
\begin{equation*}
\mu\left\{x \in \mathbb{R}^{d} \backslash \bigcup_{i} 2 Q_{i}:\left|T_{\mu, \varepsilon} f(x)\right|>\lambda\right\} \leq \frac{C}{\lambda} \int|f| d \mu \tag{10}
\end{equation*}
$$

Since $\int b_{i} d \mu=0, \operatorname{supp}\left(b_{i}\right) \subset R_{i}$ and $\left\|b_{i}\right\|_{L^{1}(\mu)} \leq C \int_{Q_{i}}|f| d \mu$, using condition 2 in the definition of a Calderón-Zygmund kernel (which implies Hörmander's condition), we get

$$
\int_{\mathbb{R}^{d} \backslash 2 R_{i}}\left|T_{\mu, \varepsilon} b_{i}\right| d \mu \leq C \int\left|b_{i}\right| d \mu \leq C \int_{Q_{i}}|f| d \mu .
$$

Let us see that

$$
\begin{equation*}
\int_{2 R_{i} \backslash 2 Q_{i}}\left|T_{\mu, \varepsilon} b_{i}\right| d \mu \leq C \int_{Q_{i}}|f| d \mu \tag{11}
\end{equation*}
$$

too. On the one hand, by (9) and using the $L^{2}(\mu)$ boundedness of $T$ and that $R_{i}$ is $\left(6,6^{n+1}\right)$-doubling we get

$$
\begin{aligned}
\int_{2 R_{i}}\left|T_{\mu, \varepsilon} \varphi_{i}\right| d \mu & \leq\left(\int_{2 R_{i}}\left|T_{\mu, \varepsilon} \varphi_{i}\right|^{2} d \mu\right)^{1 / 2} \mu\left(2 R_{i}\right)^{1 / 2} \\
& \leq C\left(\int\left|\varphi_{i}\right|^{2} d \mu\right)^{1 / 2} \mu\left(R_{i}\right)^{1 / 2} \\
& \leq C \int_{Q_{i}}|f| d \mu
\end{aligned}
$$

On the other hand, since $\operatorname{supp}\left(w_{i} f\right) \subset Q_{i}$, if $x \in 2 R_{i} \backslash 2 Q_{i}$, then $\left|T_{\mu, \varepsilon} f(x)\right| \leq$ $C \int_{Q_{i}}|f| d \mu /\left|x-x_{Q_{i}}\right|^{n}$, and so

$$
\int_{2 R_{i} \backslash 2 Q_{i}}\left|T_{\mu, \varepsilon}\left(w_{i} f\right)\right| d \mu \leq C \int_{2 R_{i} \backslash 2 Q_{i}} \frac{1}{\mid x-x_{\left.Q_{i}\right|^{n}}} d \mu(x) \times \int_{Q_{i}}|f| d \mu,
$$

By Lemma 1, the first integral on the right hand side is bounded by some constant independent of $Q_{i}$ and $R_{i}$, since there are no $\left(6,6^{n+1}\right)$-doubling cubes of the form $6^{k} Q_{i}$ between $6 Q_{i}$ and $R_{i}$. Therefore, (11) holds.

Then we have

$$
\begin{aligned}
\int_{\mathbb{R}^{d} \backslash \cup_{k} 2 Q_{k}}\left|T_{\mu, \varepsilon} b\right| d \mu & \leq \sum_{i} \int_{\mathbb{R}^{d} \backslash \cup_{k} 2 Q_{k}}\left|T_{\mu, \varepsilon} b_{i}\right| d \mu \\
& \leq C \sum_{i} \int_{Q_{i}}|f| d \mu \leq C \int|f| d \mu .
\end{aligned}
$$

Therefore,

$$
\begin{equation*}
\mu\left\{x \in \mathbb{R}^{d} \backslash \bigcup_{i} 2 Q_{i}:\left|T_{\mu, \varepsilon} b(x)\right|>\lambda\right\} \leq \frac{C}{\lambda} \int|f| d \mu \tag{12}
\end{equation*}
$$

The corresponding integral for the function $g$ is easier to estimate. Taking into account that $|g| \leq C \lambda$, we get

$$
\begin{equation*}
\mu\left\{x \in \mathbb{R}^{d} \backslash \bigcup_{i} 2 Q_{i}:\left|T_{\mu, \varepsilon} g(x)\right|>\lambda\right\} \leq \frac{C}{\lambda^{2}} \int|g|^{2} d \mu \leq \frac{C}{\lambda} \int|g| d \mu \tag{13}
\end{equation*}
$$

Also, we have

$$
\begin{aligned}
\int|g| d \mu & \leq \int_{\mathbb{R}^{d} \backslash \cup_{i} Q_{i}}|f| d \mu+\sum_{i} \int\left|\varphi_{i}\right| d \mu \\
& \leq \int|f| d \mu+\sum_{i} \int_{Q_{i}}|f| d \mu \leq C \int|f| d \mu
\end{aligned}
$$

Now, by (12) and (13) we get (10).
If we want to show that $T$ is bounded from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$, then in Lemma 2 and in the arguments above $f d \mu$ must be substituted by $d \nu$, with $\nu \in M\left(\mathbb{R}^{d}\right)$, and $|f| d \mu$ by $d|\nu|$. Also, condition (6) of Lemma 2 should be stated as "On $\mathbb{R}^{d} \backslash \bigcup_{i} Q_{i}, \nu$ is absolutely continuous with respect to $\mu$, that is $\nu=f d \nu$, and moreover $|f(x)| \leq \lambda$ a.e. $(\mu) x \in \mathbb{R}^{d} \backslash \bigcup_{i} Q_{i}$ ". With other minor changes, the arguments and estimates above work in this situation too.

### 3.4 Cotlar's inequality

This inequality involves some maximal operators which we proceed to define. The centered maximal Hardy-Littlewood operator applied to $\nu \in$ $M\left(\mathbb{R}^{d}\right)$ is, as usual,

$$
M_{\mu} \nu(x)=\sup _{r>0} \frac{1}{\mu(\bar{B}(x, r))} \int_{\bar{B}(x, r)} d|\nu| .
$$

A useful variant of this operator is the following:

$$
\widetilde{M}_{\mu} \nu(x)=\sup \left\{\frac{1}{\mu(\bar{B}(x, r))} \int_{\bar{B}(x, r)} d|\nu|: r>0, \mu(\bar{B}(x, 5 r)) \leq 5^{d+1} \mu(\bar{B}(x, r))\right\} .
$$

The non centered version of $\widetilde{M}_{\mu}$ is

$$
N_{\mu} \nu(x)=\sup \left\{\frac{1}{\mu(\bar{B})} \int_{\bar{B}} d|\nu|: \bar{B} \text { closed ball, } x \in \bar{B}, \mu(5 \bar{B}) \leq 5^{d+1} \mu(\bar{B})\right\}
$$

For $f \in L_{l o c}^{1}(\mu)$ we set $M_{\mu} f \equiv M_{\mu}(f d \mu), \widetilde{M}_{\mu} f \equiv \widetilde{M}_{\mu}(f d \mu)$, and $N_{\mu} f \equiv$ $N_{\mu}(f d \mu)$, The operators $M_{\mu}$ and $\widetilde{M}_{\mu}$ are bounded in $L^{p}(\mu)$, and from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$. This fact can be proved using Besicovitch's covering theorem for $M_{\mu}$ and $\widetilde{M}_{\mu}$, and Vitali's covering theorem with balls $B(x, 5 r)$ in the case of $N_{\mu}$.

If $T$ is a CZO, the maximal operator $T_{*}$ is

$$
T_{*} \nu(x)=\sup _{\varepsilon>0}\left|T_{\varepsilon} \nu(x)\right| \quad \text { for } \nu \in M\left(\mathbb{R}^{d}\right), x \in \mathbb{R}^{d}
$$

and the $\delta$-truncated maximal operator $T_{*, \delta}$ is

$$
T_{*, \delta} \nu(x)=\sup _{\varepsilon>\delta}\left|T_{\varepsilon} \nu(x)\right| \quad \text { for } \nu \in M\left(\mathbb{R}^{d}\right), x \in \mathbb{R}^{d} .
$$

We also set $T_{\mu, *} f \equiv T_{*}(f d \mu)$ and $T_{\mu, *, \delta} f \equiv T_{*, \delta}(f d \mu)$ for $f \in L_{l o c}^{1}(\mu)$.

Theorem 4 (Cotlar's inequality). Let $\mu$ be a positive Radon measure on $\mathbb{R}^{d}$ with growth of degree $n$. If the $T$ is an $n$-dimensional CZO bounded from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$, then for $0<s \leq 1$ we have
$T_{*, \delta} \nu(x) \leq C_{s}\left(\widetilde{M}_{\mu}\left(\left|T_{\delta} \nu\right|^{s}\right)(x)^{1 / s}+M_{\mu} \nu(x)\right), \quad$ for $\nu \in M\left(\mathbb{R}^{d}\right), x \in \mathbb{R}^{d}$,
where $C_{s}$ depends only on the constant $C_{0}$ in (1), $s, n, d$, and the norm of the $T_{\delta}$ from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$.

Cotlar's inequality with non doubling measures is due to Nazarov, Treil and Volberg [35], although not in the form stated above, which is from [43]

To prove Theorem 4 we will need some lemmas. The first one is Kolmogorov's inequality whose proof can be found in [27, p. 299].
Lemma 5. Let $\mu$ be a positive Radon measure on $\mathbb{R}^{d}$ and $f: \mathbb{R}^{d} \longrightarrow \mathbb{C} a$ Borel function in $L^{1, \infty}(\mu)$. Then for $0<s<1$ and for any $\mu$-measurable set $A \subset \mathbb{R}^{d}$ with $\mu(A)<\infty$,

$$
\left(\frac{1}{\mu(A)} \int_{A}|f|^{s} d \mu\right)^{1 / s} \leq(1-s)^{-1 / s} \frac{\|f\|_{L^{1, \infty}(\mu)}}{\mu(A)}
$$

Also, we have the following result.
Lemma 6. Let $0<r<R$, with $R=5^{N} r$. If $5^{d+1} \mu\left(\bar{B}\left(x, 5^{k-1} r\right)\right) \leq$ $\mu\left(\bar{B}\left(x, 5^{k} r\right)\right)$ for $k=2, \ldots, N$, then we have

$$
\left|T_{R} \nu(x)-T_{r} \nu(x)\right| \lesssim \frac{\mu(\bar{B}(x, R))}{R} M_{\mu} \nu(x),
$$

for each $\nu \in M\left(\mathbb{R}^{d}\right)$.
Compare this result with Lemma 1. In both cases one assumes that there exists a sequence of concentric non doubling balls or squares. Moreover, the proofs are similar.

Proof. We set $\bar{B}_{k}=\bar{B}\left(x, 5^{k} r\right)$ and $K_{0}=\mu(\bar{B}(x, R)) / R$. Then we have

$$
\begin{align*}
\left|T_{R} \nu(x)-T_{r} \nu(x)\right| & =\left|\int_{r<|y-x| \leq 5^{N} r} k(x, y) d \nu(y)\right| \\
& \lesssim \sum_{k=1}^{N} \int_{5^{k-1} r<|y-x| \leq 5^{k} r} \frac{1}{|y-x|^{n}} d|\nu|(y) \\
& \lesssim \sum_{k=1}^{N} \frac{|\nu|\left(\bar{B}_{k}\right)}{\left(5^{k} r\right)^{n}}=\sum_{k=1}^{N} \frac{|\nu|\left(\bar{B}_{k}\right)}{\left(5^{k-N} R\right)^{n}} \tag{15}
\end{align*}
$$

Also, notice that

$$
\mu\left(\bar{B}_{k}\right) \leq 5^{(k-N)(d+1)} \mu\left(\bar{B}_{N}\right)=5^{(k-N)(d+1)} R K_{0}
$$

Therefore,

$$
\frac{1}{5^{(k-N) n} R^{n}} \leq K_{0} \frac{5^{(k-N)(d+1-n)}}{\mu\left(\bar{B}_{k}\right)}
$$

and by (15),

$$
\begin{aligned}
\left|T_{R} \nu(x)-T_{r} \nu(x)\right| & \lesssim K_{0} \sum_{k=1}^{N} \frac{5^{(k-N)(d+1-n)}|\nu|\left(\bar{B}_{k}\right)}{\mu\left(\bar{B}_{k}\right)} \\
& \lesssim K_{0} \sum_{k=1}^{N} 5^{(k-N)(d+1-n)} M_{\mu} \nu(x) \lesssim K_{0} M_{\mu} \nu(x)
\end{aligned}
$$

Combining Lemma 6 with the usual arguments we are going to prove Cotlar's inequality (14).

Proof of Theorem 4 Let $\varepsilon>\delta$ and $x \in \mathbb{R}^{d}$. Since $\mu$ has growth of degree $n$, there exists some $n \geq 1$ such that

$$
\begin{equation*}
\mu\left(\bar{B}\left(x, 5^{n} \varepsilon\right)\right) \leq 5^{d+1} \mu\left(\bar{B}\left(x, 5^{n-1} \varepsilon\right)\right) \tag{16}
\end{equation*}
$$

(see Subsection 3.1). We assume that $n$ is the least integer $\geq 1$ such that (16) holds. Set $\varepsilon^{\prime}=5^{n} \varepsilon$. By Lemma 6,

$$
\left|T_{\varepsilon} \nu(x)-T_{\varepsilon^{\prime} / 5} \nu(x)\right| \leq C M_{\mu} \nu(x) .
$$

Also, it is straightforward to check that $\left|T_{\varepsilon^{\prime} / 5} \nu(x)-T_{\varepsilon^{\prime}} \nu(x)\right| \leq C M_{\mu} \nu(x)$. Therefore,

$$
\left|T_{\varepsilon} \nu(x)-T_{\varepsilon^{\prime}} \nu(x)\right| \leq C M_{\mu} \nu(x)
$$

Thus it only remains to show that

$$
\begin{equation*}
\left|T_{\varepsilon^{\prime}} \nu(x)\right| \leq C_{s}\left(\widetilde{M}_{\mu}\left(\left|T_{\delta} \nu\right|^{s}\right)(x)^{1 / s}+M_{\mu} \nu(x)\right) \tag{17}
\end{equation*}
$$

Since

$$
\begin{equation*}
\mu\left(\bar{B}\left(x, \varepsilon^{\prime}\right)\right) \leq 5^{d+1} \mu\left(\bar{B}\left(x, \varepsilon^{\prime} / 5\right)\right) \tag{18}
\end{equation*}
$$

we can apply the usual argument, as in [27], pp. 299-300, to prove (17). We set

$$
d \nu_{1}=\chi_{\bar{B}\left(x, \varepsilon^{\prime}\right)} d \nu, \quad d \nu_{2}=d \nu-d \nu_{1}
$$

For $y \in \bar{B}\left(x, \varepsilon^{\prime} / 5\right)$, since $\varepsilon^{\prime}>5 \delta$ we have $T_{\varepsilon^{\prime}} \nu_{2}(x)=T_{\delta} \nu_{2}(x)=T \nu_{2}(x)$ and $T_{\delta} \nu_{2}(y)=T \nu_{2}(y)$. Using (1) it is easy to check that $\left|T_{\delta} \nu_{2}(y)-T_{\delta} \nu_{2}(x)\right| \leq$ $C M_{\mu} \nu(x)$. Therefore,

$$
\left|T_{\varepsilon^{\prime}} \nu(x)\right|=\left|T_{\delta} \nu_{2}(x)\right| \begin{gather*}
\leq\left|T_{\delta} \nu_{2}(y)\right|+C_{2} M_{\mu} \nu(x)  \tag{19}\\
\leq\left|T_{\delta} \nu_{1}(y)\right|+\left|T_{\delta} \nu(y)\right|+C_{2} M_{\mu} \nu(x) \tag{20}
\end{gather*}
$$

Assume first $s=1$. If $T_{\varepsilon^{\prime}} \nu(x) \neq 0$, let $0<\lambda<\left|T_{\varepsilon^{\prime}} \nu(x)\right|$. For $y \in \bar{B}\left(x, \varepsilon^{\prime} / 5\right)$, by (20) either $C_{2} M_{\mu} \nu(x)>\lambda / 3$ or $\left|T_{\delta} \nu(y)\right|>\lambda / 3$ or $\left|T_{\delta} \nu_{1}(y)\right|>\lambda / 3$. Therefore, either

$$
\lambda<3 C_{2} M_{\mu} \nu(x)
$$

or
$\bar{B}\left(x, \varepsilon^{\prime} / 5\right)=\left\{y \in \bar{B}\left(x, \varepsilon^{\prime} / 5\right):\left|T_{\delta} \nu(y)\right|>\lambda / 3\right\} \cup\left\{y \in \bar{B}\left(x, \varepsilon^{\prime} / 5\right):\left|T_{\delta} \nu_{1}(y)\right|>\lambda / 3\right\}$.
But we have

$$
\begin{aligned}
\mu\left\{y \in \bar{B}\left(x, \varepsilon^{\prime} / 5\right):\left|T_{\delta} \nu(y)\right|>\lambda / 3\right\} & \leq \frac{3}{\lambda} \int_{\bar{B}\left(x, \varepsilon^{\prime} / 5\right)}\left|T_{\delta} \nu\right| d \mu \\
& \leq \frac{3}{\lambda} \mu\left(\bar{B}\left(x, \varepsilon^{\prime} / 5\right)\right) \widetilde{M}_{\mu}\left(T_{\delta} \nu\right)(x)
\end{aligned}
$$

and by the boundedness of $T_{\delta}$ from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$ and (18),

$$
\begin{aligned}
\mu\left\{y \in \bar{B}\left(x, \varepsilon^{\prime} / 5\right):\left|T_{\delta} \nu_{1}(y)\right|>\lambda / 3\right\} & \lesssim \frac{\left\|\nu_{1}\right\|}{\lambda}=\frac{|\nu|\left(\bar{B}\left(x, \varepsilon^{\prime}\right)\right)}{\lambda} \\
& \lesssim \frac{\mu\left(\bar{B}\left(x, \varepsilon^{\prime} / 5\right)\right)}{\lambda} M_{\mu} \nu(x)
\end{aligned}
$$

In any case we obtain $\lambda<3 \widetilde{M}_{\mu}\left(T_{\delta} \nu\right)(x)+C M_{\mu} \nu(x)$. Since this holds for $0<\lambda<\left|T_{\varepsilon^{\prime}} \nu(x)\right|$, (17) follows when $s=1$.

Assume now $0<s<1$. From (20) we get

$$
\left|T_{\varepsilon^{\prime}} \nu(x)\right|^{s} \leq\left|T_{\delta} \nu_{1}(y)\right|^{s}+\left|T_{\delta} \nu(y)\right|^{s}+C M_{\mu} \nu(x)^{s}
$$

Integrating with respect to $\mu$ and $y \in \bar{B}\left(x, \varepsilon^{\prime} / 5\right)$, dividing by $\mu\left(\bar{B}\left(x, \varepsilon^{\prime} / 5\right)\right)$ and raising to the power $1 / s$ we obtain

$$
\begin{align*}
\left|T_{\varepsilon^{\prime}} \nu(x)\right| \leq & C_{s}\left[\left(\frac{1}{\mu\left(\bar{B}\left(x, \varepsilon^{\prime} / 5\right)\right)} \int_{\bar{B}\left(x, \varepsilon^{\prime} / 5\right)}\left|T_{\delta} \nu_{1}\right|^{s} d \mu\right)^{1 / s}\right. \\
& \left.+\left(\frac{1}{\mu\left(\bar{B}\left(x, \varepsilon^{\prime} / 5\right)\right)} \int_{\bar{B}\left(x, \varepsilon^{\prime} / 5\right)}\left|T_{\delta} \nu\right|^{s} d \mu\right)^{1 / s}+M_{\mu} \nu(x)\right] \tag{21}
\end{align*}
$$

By (18), the second term on the right hand side of (21) can be estimated by $\widetilde{M}_{\mu}\left(\left|T_{\delta} \nu\right|^{s}\right)(x)^{1 / s}$. On the other hand, the first term is estimated using Kolmogorov's inequality, the boundedness of $T_{\delta}$ from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$, and (18):

$$
\begin{aligned}
\left(\frac{1}{\mu\left(\bar{B}\left(x, \varepsilon^{\prime} / 5\right)\right)} \int_{\bar{B}\left(x, \varepsilon^{\prime} / 5\right)}\left|T_{\delta} \nu_{1}\right|^{s} d \mu\right)^{1 / s} \lesssim \frac{\left\|T_{\delta} \nu_{1}\right\|_{L^{1, \infty}(\mu)}}{\mu\left(\bar{B}\left(x, \varepsilon^{\prime} / 5\right)\right)} & \lesssim \frac{\left\|\nu_{1}\right\|}{\mu\left(\bar{B}\left(x, \varepsilon^{\prime} / 5\right)\right)} \\
& \lesssim M_{\mu} \nu(x)
\end{aligned}
$$

Now (17) follows.
A direct consequence of Cotlar's inequality and Theorem 3 is the following result.

Theorem 7. Let $\mu$ be a Radon measure on $\mathbb{R}^{d}$ of degree $n$. If $T$ is an $n$-dimensional CZO bounded in $L^{2}(\mu)$, then $T_{\mu, *}$ is bounded in $L^{p}(\mu), p \in$ $(1, \infty)$, and from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$.

Proof. By Theorem 3, interpolation, and duality, $T_{\mu}$ is bounded in $L^{p}(\mu)$, $p \in(1, \infty)$, and from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$. Then, by Cotlar's inequality it is clear that $T_{*, \delta}$ is bounded in $L^{p}(\mu), p \in(1, \infty)$, uniformly on $\delta>$ 0 . Hence, by monotone convergence, $T_{*}$ is also bounded in $L^{p}(\mu), p \in$ $(1, \infty)$. The boundedness of $T_{*}$ from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$ follows as in the classical doubling case, using Kolmogorov's inequality and taking into account that the non centered version of the maximal operator $\widetilde{M}_{\mu}$ (which is $\left.N_{\mu}\right)$ is bounded from $M\left(\mathbb{R}^{d}\right)$ into $L^{1, \infty}(\mu)$. See [43] for the details.

### 3.5 The $T(1)$ theorem and other results

Let us introduce some notation and definitions. Given $\rho>1$, we say that $f \in L_{\text {loc }}^{1}(\mu)$ belongs to the space $B M O_{\rho}(\mu)$ if

$$
\sup _{Q} \frac{1}{\mu(\rho Q)} \int_{Q}\left|f-m_{Q}(f)\right| d \mu<\infty
$$

where the supremum is taken over all the squares in $\mathbb{R}^{d}$ and $m_{Q}(f)$ is the $\mu$-mean of $f$ over $Q$.

Following [34], a Calderón-Zygmund operator $T_{\mu}$ is said to be weakly bounded if

$$
\left|\left\langle T_{\mu, \varepsilon} \chi_{Q}, \chi_{Q}\right\rangle\right| \leq C \mu(Q) \quad \text { for all the cubes } Q \subset \mathbb{R}^{d} \text {, uniformly on } \varepsilon>0 \text {. }
$$

Notice that if $T_{\mu}$ is antisymmetric, then the left hand side above equals zero and so $T_{\mu}$ is weakly bounded.

Now we are ready to state the $T(1)$ theorem:
Theorem 8. Let $\mu$ be a Radon measure on $\mathbb{R}^{d}$ of degree $n$, and let $T$ be an $n$-dimensional Calderón-Zygmund operator. The following conditions are equivalent:
(a) $T_{\mu}$ is bounded on $L^{2}(\mu)$.
(b) $T_{\mu}$ is weakly bounded and, for some $\rho>1$, we have $T_{\mu, \varepsilon}(1), T_{\mu, \varepsilon}^{*}(1) \in$ $B M O_{\rho}(\mu)$ uniformly on $\varepsilon>0$.
(c) There exists some constant $C_{3}$ such that for all $\varepsilon>0$ and all the cubes $Q \subset \mathbb{R}^{d}$,

$$
\left\|T_{\mu, \varepsilon} \chi_{Q}\right\|_{L^{2}(\mu \mid Q)} \leq C_{3} \mu(Q)^{1 / 2} \quad \text { and } \quad\left\|T_{\mu, \varepsilon}^{*} \chi_{Q}\right\|_{L^{2}(\mu \mid Q)} \leq C_{3} \mu(Q)^{1 / 2}
$$

The classical way of stating the $T(1)$ theorem is the equivalence (a) $\Leftrightarrow$ (b). However, for some applications it is sometimes more practical to state the result in terms of the $L^{2}$ boundedness of $T_{\mu}$ and $T_{\mu}^{*}$ over characteristic functions of cubes, i.e (a) $\Leftrightarrow$ (c).

Theorem 8 is the extension of the classical $T(1)$ theorem of David and Journé to measures of degree $n$ which may be non doubling. The result was proved by Nazarov, Treil and Volberg in [34], although not exactly in the form stated above. An independent proof for the particular case of the Cauchy transform was obtained almost simultaneously in [42]. For the equivalence of conditions (b) and (c) above, the reader should see [47, Remark 7.1 and Lemma 7.3]. Other (more recent) proofs of the $T(1)$ theorem for non doubling measures are in [53] (for the particular case of the Cauchy transform) and in [47].

Let us remark that the boundedness of $T_{\mu}$ on $L^{2}(\mu)$ does not imply the boundedness of $T_{\mu}$ from $L^{\infty}(\mu)$ into $B M O(\mu)$ (this is the space $B M O_{\rho}(\mu)$ with parameter $\rho=1$ ), and in general $T_{\mu, \varepsilon}(1), T_{\mu, \varepsilon}^{*}(1) \notin B M O(\mu)$ uniformly on $\varepsilon>0$. See [53] and [22]. On the contrary, one can show that if $T_{\mu}$ is bounded on $L^{2}(\mu)$, then it is also bounded from $L^{\infty}(\mu)$ into $B M O_{\rho}(\mu)$, for $\rho>1$, by arguments similar to the classical ones for homogeneous spaces. However, the space $B M O_{\rho}(\mu)$ has some drawbacks. For example, it depends on the parameter $\rho$ and it does not satisfy the John-Nirenberg inequality. To solve these problems, in [44] a new space called $R B M O(\mu)$ has been introduced. $R B M O(\mu)$ is a subspace of $B M O_{\rho}(\mu)$ for all $\rho>1$, and it coincides
with $B M O(\mu)$ when $\mu$ is an AD-regular measure. Moreover, $R B M O(\mu)$ satisfies a John-Nirenberg type inequality, and all CZO's which are bounded on $L^{2}(\mu)$ are also bounded from $L^{\infty}(\mu)$ into $R B M O(\mu)$. For these reasons $R B M O(\mu)$ seems to be a good substitute of the classical space $B M O$ for non doubling measures of degree $n$. For the precise definition of $R B M O(\mu)$ and its properties, see [44].

Much more results on Calderón-Zygmund theory with non doubling measures have been proved recently. For example, several $T(b)$ type theorems have been obtained in [8], [7], [36], [37], [38]. There are also results concerning Hardy spaces [46]; weights [12], [21], [39]; commutators [2], [19], [44]; multilinear commutators [18]; fractional integrals [13], [14]; Lipschitz spaces [15]; Triebel-Lizorkin spaces [17]; etc.

## 4 Analytic capacity

### 4.1 Definition

The analytic capacity of a compact set $E \subset \mathbb{C}$ is

$$
\begin{equation*}
\gamma(E):=\sup \left|f^{\prime}(\infty)\right| \tag{22}
\end{equation*}
$$

where the supremum is taken over all analytic functions $f: \mathbb{C} \backslash E \longrightarrow \mathbb{C}$ with $|f| \leq 1$ on $\mathbb{C} \backslash E$, and $f^{\prime}(\infty)=\lim _{z \rightarrow \infty} z(f(z)-f(\infty))$.

The notion of analytic capacity was introduced by Ahlfors [1] in the 1940's in order to study the removability of singularities of bounded analytic functions. A compact set $E \subset \mathbb{C}$ is removable for bounded analytic functions if for any open set $\Omega$ containing $E$, every bounded function analytic on $\Omega \backslash E$ has an analytic extension to $\Omega$. Ahlfors showed that $E$ is removable if and only if $\gamma(E)=0$.

Painlevé's problem consists of characterizing removable singularities for bounded analytic functions in a metric/geometric way. By Ahlfors' result this is equivalent to describe compact sets with positive analytic capacity in metric/geometric terms.

Vitushkin in the 1950's and 1960's showed that analytic capacity plays a central role in problems of uniform rational approximation on compact sets of the complex plane. Many results obtained by Vitushkin in connection with uniform rational approximation are stated in terms of $\gamma$. See [56], or [52] for a more modern approach, for example. Further, because its applications to this type of problems he raised the question of the semiadditivity of $\gamma$. Namely, does there exist an absolute constant $C$ such that

$$
\gamma(E \cup F) \leq C(\gamma(E)+\gamma(F)) ?
$$

### 4.2 Basic properties of analytic capacity

One should keep in mind that, in a sense, analytic capacity measures the size of a set as a non removable singularity for bounded analytic functions. A direct consequence of the definition is that

$$
E \subset F \Rightarrow \gamma(E) \leq \gamma(F)
$$

Moreover, it is also easy to check that analytic capacity is translation invariant:

$$
\gamma(z+E)=\gamma(E) \quad \text { for all } z \in \mathbb{C}
$$

Concerning dilations, we have

$$
\gamma(\lambda E)=|\lambda| \gamma(E) \quad \text { for all } \lambda \in \mathbb{C} .
$$

Further, if $E$ is connected, then

$$
\operatorname{diam}(E) / 4 \leq \gamma(E) \leq \operatorname{diam}(E)
$$

The second inequality follows from the fact that the analytic capacity of a closed disk coincides with its radius, and the first one is a consequence of Koebe's $1 / 4$ theorem (see [11, Chapter VIII] for the details, for example).

### 4.3 Relationship with Hausdorff measure

The relationship between Hausdorff measure and analytic capacity is the following:

- If $\operatorname{dim}_{H}(E)>1$ (here $\operatorname{dim}_{H}$ stands for the Hausdorff dimension), then $\gamma(E)>0$. This result follows easily from Frostman's Lemma.
- $\gamma(E) \leq \mathcal{H}^{1}(E)$, where $\mathcal{H}^{1}$ is the one dimensional Hausdorff measure, or length. This follows from Cauchy's integral formula, and it was proved by Painlevé about one hundred years ago. Observe that, in particular we deduce that if $\operatorname{dim}_{H}(E)<1$, then $\gamma(E)=0$.

By the statements above, it turns out that dimension 1 is the critical dimension in connection with analytic capacity. Moreover, a natural question arises: is it true that $\gamma(E)>0$ if and only if $\mathcal{H}^{1}(E)>0$ ?

Vitushkin showed that the answer is $n o$. Indeed, he constructed a compact set in $\mathbb{C}$ with positive length and vanishing analytic capacity. This set was purely unrectifiable. That is, it intersects any rectifiable curve at most in a set of zero length. Motivated by this example (and others, I guess)

Vitushkin conjectured that pure unrectifiability is a necessary and sufficient condition for vanishing analytic capacity, for sets with finite length.

Guy David [7] showed in 1998 that Vitushkin's conjecture is true:
Theorem 9. Let $E \subset \mathbb{C}$ be compact with $\mathcal{H}^{1}(E)<\infty$. Then, $\gamma(E)=0$ if and only if $E$ is purely unrectifiable.

Let us remark that the "if" part of the theorem is not due to David (it follows from Calderón's theorem on the $L^{2}$ boundedness of the Cauchy transform on Lipschitz graphs). The "only if" part of the theorem, which is more difficult, is the one proved by David. See also [30], [8] and [20] for some preliminary contributions to the proof.

Theorem 9 is the solution of Painleve's problem for sets with finite length. The analogous result is false for sets with infinite length. For this type of sets there is no such a nice geometric solution of Painlevé's problem, and we have to content ourselves with a characterization such as the one in Corollary 18 below (at least, for the moment).

### 4.4 The capacity $\gamma_{+}$and the Cauchy transform

The capacity $\gamma_{+}$of a compact set $E \subset \mathbb{C}$ is

$$
\begin{equation*}
\gamma_{+}(E):=\sup \left\{\mu(E): \operatorname{supp}(\mu) \subset E,\|\mathcal{C} \mu\|_{L^{\infty}(\mathbb{C})} \leq 1\right\} \tag{23}
\end{equation*}
$$

That is, $\gamma_{+}$is defined as $\gamma$ in (22) with the additional constraint that $f$ should coincide with $\mathcal{C} \mu$, where $\mu$ is some positive Radon measure supported on $E$ (observe that $(\mathcal{C} \mu)^{\prime}(\infty)=-\mu(\mathbb{C})$ for any Radon measure $\mu$ ). To be precise, there is another little difference: in (22) we asked $\|f\|_{L^{\infty}(\mathbb{C} \backslash E)} \leq 1$, while in (23) $\|f\|_{L^{\infty}(\mathbb{C})} \leq 1$ (for $f=\mathcal{C} \mu$ ). Trivially, we have $\gamma_{+}(E) \leq \gamma(E)$.

The following lemma relates weak $(1,1)$ estimates for the Cauchy integral operator with $L^{\infty}$ estimates (which in its turn are connected with $\gamma_{+}$and $\gamma)$.

Lemma 10. let $\mu$ be a Radon measure with linear growth on $\mathbb{C}$. The following statements are equivalent:
(a) The Cauchy transform is bounded from $M(\mathbb{C})$ into $L^{1, \infty}(\mu)$.
(b) For any set $A \subset \mathbb{C}$ there exists some function $h$ supported on $A$, with $0 \leq h \leq 1$, such that $\int h d \mu \geq C^{-1} \mu(A)$ and $\left\|\mathcal{C}_{\varepsilon}(h d \mu)\right\|_{L^{\infty}(\mathbb{C})} \leq C$ for all $\varepsilon>0$.

The constant $C$ in (b) depends only on the norm of the Cauchy transform is bounded from $M(\mathbb{C})$ into $L^{1, \infty}(\mu)$, and conversely.

This lemma is a particular case of a result which applies to more general linear operators. The statement (b) should be understood as a weak substitute of the $L^{\infty}(\mu)$ boundedness of the Cauchy integral operator, which does not hold in general.

We will prove the easy implication of the lemma, that is, $(\mathrm{b}) \Rightarrow$ (a). For the other implication, which is due to Davie and Øksendal [10] the reader is referred to [3, Chapter VII].

Proof of $(b) \Rightarrow(a)$. It is enough to show that for any complex measure $\nu \in$ $M(\mathbb{C})$ and any $\lambda>0$,

$$
\mu\left\{x \in \mathbb{C}: \operatorname{Re}\left(\mathcal{C}_{\varepsilon} \nu(x)\right)>\lambda\right\} \leq \frac{C\|\nu\|}{\lambda} .
$$

To this end, let us denote by $A$ the set on the left side above, and let $h$ be a function supported on $A$ fulfilling the properties in the statement (b) of the lemma. Then we have
$\mu(A) \leq C \int h d \mu \leq \frac{C}{\lambda} \operatorname{Re}\left(\int\left(\mathcal{C}_{\varepsilon} \nu\right) h d \mu\right)=\frac{-C}{\lambda} \operatorname{Re}\left(\int \mathcal{C}_{\varepsilon}(h d \mu) d \nu\right) \leq \frac{C\|\nu\|}{\lambda}$.

Remark 11. Notice that if E supports a non zero Radon measure $\mu$ with linear growth such that the Cauchy integral operator $\mathcal{C}_{\mu}$ is bounded on $L^{2}(\mu)$, then there exists some nonzero function $h$ with $0 \leq h \leq \chi_{E}$ such that $\left\|\mathcal{C}_{\varepsilon}(h d \mu)\right\|_{L^{\infty}(\mathbb{C})} \leq C$ uniformly on $\varepsilon$, by Theorem 3 and the preceding lemma. Letting $\varepsilon \rightarrow 0$, we infer that $|\mathcal{C}(h d \mu)(z)| \leq C$ for all $z \notin E$, and so $\gamma(E)>0$.

A more precise result will be proved in Theorem 14 below.

### 4.5 The curvature of a measure

Given three pairwise different points $x, y, z \in \mathbb{C}$, their Menger curvature is

$$
c(x, y, z)=\frac{1}{R(x, y, z)},
$$

where $R(x, y, z)$ is the radius of the circumference passing through $x, y, z$ (with $R(x, y, z)=\infty, c(x, y, z)=0$ if $x, y, z$ lie on a same line). If two among these points coincide, we let $c(x, y, z)=0$. For a positive Radon measure $\mu$, we set

$$
c_{\mu}^{2}(x)=\iint c(x, y, z)^{2} d \mu(y) d \mu(z)
$$

and we define the curvature of $\mu$ as

$$
\begin{equation*}
c^{2}(\mu)=\int c_{\mu}^{2}(x) d \mu(x)=\iiint c(x, y, z)^{2} d \mu(x) d \mu(y) d \mu(z) . \tag{24}
\end{equation*}
$$

The notion of curvature of measures was introduced by Melnikov [32] when he was studying a discrete version of analytic capacity, and it is one of the ideas which is responsible of the big recent advances in connection with analytic capacity. The notion of curvature is connected to the Cauchy transform by the following result, proved by Melnikov and Verdera.

Proposition 12. Let $\mu$ be a Radon measure on $\mathbb{C}$ with linear growth. We have

$$
\begin{equation*}
\left\|\mathcal{C}_{\varepsilon} \mu\right\|_{L^{2}(\mu)}^{2}=\frac{1}{6} c_{\varepsilon}^{2}(\mu)+O(\mu(\mathbb{C})) \tag{25}
\end{equation*}
$$

where $c_{\varepsilon}^{2}(\mu)$ is the $\varepsilon$-truncated version of $c^{2}(\mu)$ (defined as in the right hand side of (24), but with the triple integral over $\{x, y, z \in \mathbb{C}:|x-y|,|y-z|, \mid x-$ $z \mid>\varepsilon\}$ ), and $|O(\mu(\mathbb{C}))| \leq C \mu(\mathbb{C})$.

The identity (25) is remarkable because it relates an analytic notion (the Cauchy transform of a measure) with a metric-geometric one (curvature). We give a sketch of the proof.

Sketch of the proof of Proposition 12. If we don't worry about truncations and the absolute convergence of the integrals, we can write

$$
\begin{aligned}
\|\mathcal{C} \mu\|_{L^{2}(\mu)}^{2} & =\int\left|\int \frac{1}{y-x} d \mu(y)\right|^{2} d \mu(x) \\
& =\iiint \frac{1}{(y-x)(\overline{z-x})} d \mu(y) d \mu(z) d \mu(x)
\end{aligned}
$$

By Fubini (assuming that it can be applied correctly), permuting $x, y, z$, we get,

$$
\|\mathcal{C} \mu\|_{L^{2}(\mu)}^{2}=\frac{1}{6} \iiint \sum_{s \in S_{3}} \frac{1}{\left(z_{s_{2}}-z_{s_{1}}\right)\left(\overline{z_{s_{3}}-z_{s_{1}}}\right)} d \mu\left(z_{1}\right) d \mu\left(z_{2}\right) d \mu\left(z_{3}\right),
$$

where $S_{3}$ is the group of permutations of three elements. An elementary calculation shows that

$$
\sum_{s \in S_{3}} \frac{1}{\left(z_{s_{2}}-z_{s_{1}}\right)\left(\overline{z_{s_{3}}-z_{s_{1}}}\right)}=c\left(z_{1}, z_{2}, z_{3}\right)^{2} .
$$

So we get

$$
\|\mathcal{C} \mu\|_{L^{2}(\mu)}^{2}=\frac{1}{6} c^{2}(\mu) .
$$

To argue rigorously, above one should use the truncated Cauchy transform $\mathcal{C}_{\varepsilon} \mu$ instead of $\mathcal{C} \mu$. Then we obtain

$$
\begin{aligned}
\left\|\mathcal{C}_{\varepsilon} \mu\right\|_{L^{2}(\mu)}^{2} & =\iiint_{\substack{|x-y||\varepsilon \varepsilon\\
| x-z \mid>\varepsilon}} \frac{1}{(y-x)(\overline{z-x})} d \mu(y) d \mu(z) d \mu(x) \\
& =\iiint \int_{\substack{|x-y||\varepsilon \varepsilon\\
| y-z| | \mid>\varepsilon}} \frac{1}{(y-x)(\overline{z-x})} d \mu(y) d \mu(z) d \mu(x)+O(\mu(\mathbb{C}))(26)
\end{aligned}
$$

By the linear growth of $\mu$, it is easy to check that $|O(\mu(\mathbb{C}))| \leq \mu(\mathbb{C})$. As above, using Fubini and permuting $x, y, z$, one shows that the triple integral in (26) equals $c_{\varepsilon}^{2}(\mu) / 6$.

Due to Proposition 12, the $T(1)$ theorem for the Cauchy transform can be rewritten in the following way:

Theorem 13. Let $\mu$ be a Radon measure on $\mathbb{C}$ with linear growth. The Cauchy transform is bounded on $L^{2}(\mu)$ if and only if

$$
c^{2}\left(\mu_{\mid Q}\right) \leq C \mu(Q) \quad \text { for all the squares } Q \subset \mathbb{C} \text {. }
$$

Observe that this result is a restatement of the equivalence (a) $\Leftrightarrow$ (c) in Theorem 8, by an application of (25) to the measure $\mu_{\mid Q}$, for all the squares $Q \subset \mathbb{C}$.

## 5 Semiadditivity of $\gamma_{+}$and its characterization in terms of curvature

We denote by $\Sigma(E)$ the set of Radon measures supported on $E$ such that $\mu(B(x, r)) \leq r$ for all $x \in \mathbb{C}, r>0$.

Theorem 14. For any compact set $E \subset \mathbb{C}$ we have

$$
\begin{aligned}
\gamma_{+}(E) & \approx \sup \left\{\mu(E): \mu \in \Sigma(E),\left\|\mathcal{C}_{\varepsilon} \mu\right\|_{L^{\infty}(\mu)} \leq 1 \forall \varepsilon>0\right\} \\
& \approx \sup \left\{\mu(E): \mu \in \Sigma(E),\left\|\mathcal{C}_{\varepsilon} \mu\right\|_{L^{2}(\mu)}^{2} \leq \mu(E) \forall \varepsilon>0\right\} \\
& \approx \sup \left\{\mu(E): \mu \in \Sigma(E), c^{2}(\mu) \leq \mu(E)\right\} \\
& \approx \sup \left\{\mu(E): \mu \in \Sigma(E),\left\|\mathcal{C}_{\mu}\right\|_{L^{2}(\mu), L^{2}(\mu)} \leq 1 \forall \varepsilon>0\right\} .
\end{aligned}
$$

In the statement above, $\left\|\mathcal{C}_{\mu}\right\|_{L^{2}(\mu), L^{2}(\mu)}$ stands for the operator norm of $\mathcal{C}_{\mu}$ on $L^{2}(\mu)$. That is, $\left\|\mathcal{C}_{\mu}\right\|_{L^{2}(\mu), L^{2}(\mu)}=\sup _{\varepsilon>0}\left\|\mathcal{C}_{\mu, \varepsilon}\right\|_{L^{2}(\mu), L^{2}(\mu)}$.

Proof. We denote

$$
\begin{aligned}
& S_{1}:=\sup \left\{\mu(E): \mu \in \Sigma(E),\left\|\mathcal{C}_{\varepsilon} \mu\right\|_{L^{\infty}(\mu)} \leq 1 \forall \varepsilon>0\right\}, \\
& S_{2}:=\sup \left\{\mu(E): \mu \in \Sigma(E),\left\|\mathcal{C}_{\varepsilon} \mu\right\|_{L^{2}(\mu)}^{2} \leq \mu(E) \forall \varepsilon>0\right\}, \\
& S_{3}:=\sup \left\{\mu(E): \mu \in \Sigma(E), c^{2}(\mu) \leq \mu(E)\right\}, \\
& S_{4}:=\sup \left\{\mu(E): \mu \in \Sigma(E),\left\|\mathcal{C}_{\mu}\right\|_{L^{2}(\mu), L^{2}(\mu)} \leq 1 \forall \varepsilon>0\right\} .
\end{aligned}
$$

We will show that $\gamma_{+}(E) \lesssim S_{1} \lesssim S_{2} \lesssim S_{3} \lesssim S_{4} \lesssim \gamma_{+}(E)$. The inequality $S_{3} \lesssim S_{4}$ requires more work than the others. We will give two proofs of it. One uses the $T(1)$ theorem and the other not (and so it is more elementary). Proof of $\gamma_{+}(E) \lesssim S_{1}$. Let $\mu$ be supported on $E$ such that $\|\mathcal{C} \mu\|_{L^{\infty}(\mathbb{C})} \leq 1$ with $\gamma_{+}(E) \leq 2 \mu(E)$. It is enough to show that $\mu$ has linear growth and $\left\|\mathcal{C}_{\varepsilon} \mu\right\|_{L^{\infty}(\mu)} \leq C$ uniformly on $\varepsilon>0$.

First we will prove the linear growth of $\mu$. For any fixed $x \in \mathbb{C}$, by Fubini it turns out that for almost all $r>0$,

$$
\int_{|z-x|=r} \frac{1}{|z-x|} d \mu(z)<\infty .
$$

For this $r$ we have

$$
\mu(B(x, r))=-\int_{|z-x|=r} \mathcal{C} \mu(z) \frac{d z}{2 \pi i} \leq r .
$$

Now the linear growth of $\mu$ follows easily.
To deal with the $L^{\infty}(\mu)$ norm of $\mathcal{C}_{\varepsilon}$ we use a standard technique: we replace $\mathcal{C}_{\varepsilon}$ by the regularized operator $\tilde{\mathcal{C}}_{\varepsilon}$, defined as

$$
\widetilde{\mathcal{C}}_{\varepsilon} \mu(x)=\int r_{\varepsilon}(y-x) d \mu(y)
$$

where $r_{\varepsilon}$ is the kernel

$$
r_{\varepsilon}(z)=\left\{\begin{array}{cc}
\frac{1}{z} & \text { if }|z|>\varepsilon \\
\frac{\bar{z}}{\varepsilon^{2}} & \text { if }|z| \leq \varepsilon
\end{array}\right.
$$

Then, $\widetilde{\mathcal{C}}_{\varepsilon} \mu$ is the convolution of the complex measure $\mu$ with the uniformly continuous kernel $r_{\varepsilon}$ and so $\widetilde{\mathcal{C}}_{\varepsilon} \mu$ is a continuous function. Also, we have

$$
r_{\varepsilon}(z)=\frac{1}{z} * \frac{\chi_{\varepsilon}}{\pi \varepsilon^{2}},
$$

where $\chi_{\varepsilon}$ is the characteristic function of $B(0, \varepsilon)$. Since $\mu$ is compactly supported, we have the following identity:

$$
\widetilde{\mathcal{C}}_{\varepsilon} \mu=\frac{1}{z} * \frac{\chi_{\varepsilon}}{\pi \varepsilon^{2}} * \mu=\frac{\chi_{\varepsilon}}{\pi \varepsilon^{2}} * \mathcal{C} \mu .
$$

This equality must be understood in the sense of distributions, with $\mathcal{C} \mu$ being a function of $L_{l o c}^{1}(\mathbb{C})$ with respect to Lebesgue planar measure. As a consequence, if $\|\mathcal{C} \mu\|_{L^{\infty}(\mathbb{C})} \leq 1$, we infer that $\left\|\widetilde{\mathcal{C}}_{\varepsilon} \mu\right\|_{L^{\infty}(\mu)} \leq 1$ for all $\varepsilon>0$.

Since $\mu$ has linear growth, we have

$$
\begin{equation*}
\left|\widetilde{\mathcal{C}}_{\varepsilon} \mu(x)-\mathcal{C}_{\varepsilon} \mu(x)\right|=\frac{1}{\varepsilon^{2}}\left|\int_{|y-x|<\varepsilon}(\overline{y-x}) d \mu(y)\right| \leq C, \tag{27}
\end{equation*}
$$

and so $\left\|\mathcal{C}_{\varepsilon} \mu\right\|_{L^{\infty}(\mu)} \leq C$ uniformly on $\varepsilon>0$.
Proof of $S_{1} \lesssim S_{2}$. Trivial.
Proof of $S_{2} \lesssim S_{3}$. This is a direct consequence of Proposition 12.
Proof of $S_{3} \lesssim S_{4}$ using the $T(1)$ theorem. Let $\mu$ supported on $E$ with linear growth such that $c^{2}(\mu) \leq \mu(E)$ and $S_{3} \leq 2 \mu(E)$. We set

$$
A:=\left\{x \in E: c_{\mu}^{2}(x) \leq 2\right\} .
$$

By Tchebychev $\mu(A) \geq \mu(E) / 2$. Moreover, for any set $B \subset \mathbb{C}$,

$$
\begin{aligned}
c^{2}\left(\mu_{\mid B \cap A}\right) & \leq \iiint_{x \in B \cap A} c(x, y, z)^{2} d \mu(x) d \mu(y) d \mu(z) \\
& =\int_{x \in B \cap A} c_{\mu}^{2}(x) d \mu(x) \leq 2 \mu(B) .
\end{aligned}
$$

In particular, this estimate holds when $B$ is any square in $\mathbb{C}$, and so $\mathcal{C}_{\mu_{\mid A}}$ is bounded on $L^{2}\left(\mu_{\mid A}\right)$, by Theorem 13. Thus $S_{4} \gtrsim \mu(A) \approx S_{3}$.
Proof of $S_{3} \lesssim S_{4}$ without using the $T(1)$ theorem. Take $\mu$ supported on $E$ with linear growth such that $c^{2}(\mu) \leq \mu(E)$ and $S_{3} \leq 2 \mu(E)$. To prove $S_{3} \lesssim S_{4}$ we will show that there exists a measure $\nu$ supported on $E$ with linear growth such that $\nu(E) \geq \mu(E) / 4$ and $\left\|\mathcal{C}_{\nu}\right\|_{L^{2}(\nu), L^{2}(\nu)} \leq C$.

Given $C_{4}>0$, let

$$
A_{\varepsilon}:=\left\{x \in E:\left|\mathcal{C}_{\varepsilon} \mu(x)\right| \leq C_{4} \text { and } c_{\mu}^{2}(x) \leq C_{4}^{2}\right\}
$$

Since $\int c_{\mu}^{2}(x) d \mu(x)=c^{2}(\mu) \leq \mu(E)$ and, by Proposition $12, \int\left|\mathcal{C}_{\varepsilon} \mu\right|^{2} d \mu \leq$ $C \mu(E)$, we infer that $\mu\left(A_{\varepsilon}\right) \geq \mu(E) / 2$ if $C_{4}$ is chosen big enough, by Tchebychev.

We want to show that the Cauchy integral operator $\mathcal{C}_{\mu_{\mid A_{\varepsilon}}, \varepsilon}$ is bounded on $L^{2}\left(\mu_{\mid A_{\varepsilon}}\right)$. To this end we introduce an auxiliary "curvature operator": for $x, y \in A_{\varepsilon}$, consider the kernel $k(x, y):=\int c(x, y, z)^{2} d \mu(z)$, and let $T$ be the operator

$$
T f(x)=\int k(x, y) f(y) d \mu(y)
$$

By Schur's lemma, $T$ is bounded on $L^{p}\left(\mu_{\mid A_{\varepsilon}}\right)$ for all $p \in[1, \infty]$, because for all $x \in A_{\varepsilon}$,

$$
\begin{aligned}
\int k(x, y) d \mu_{\mid A_{\varepsilon}}(y) & =\int k(y, x) d \mu_{\mid A_{\varepsilon}}(y) \\
& =\int_{y \in A_{\varepsilon}} c(x, y, z)^{2} d \mu(y) d \mu(z) \leq c_{\mu}^{2}(x) \leq C_{4}^{2}
\end{aligned}
$$

Given a non negative (real) function $f$ supported on $A_{\varepsilon}$, by arguments similar to the ones in the proof of Proposition 12, we have

$$
\begin{aligned}
4 \int\left|\mathcal{C}_{\varepsilon}(f d \mu)\right|^{2} d \mu= & \iiint_{\substack{|x-y|>\varepsilon \\
|x-z|>\varepsilon \\
|y-z|>\varepsilon}} c(x, y, z)^{2} f(x) f(y) d \mu(x) d \mu(y) d \mu(z) \\
& -2 \operatorname{Re} \int\left(\mathcal{C}_{\varepsilon} \mu\right) \overline{\mathcal{C}_{\varepsilon}(f d \mu)} f d \mu+O\left(\|f\|_{L^{2}(\mu)}^{2}\right) .
\end{aligned}
$$

See [53, Lemma 1] for the details, for example. Thus,

$$
\begin{equation*}
\int\left|\mathcal{C}_{\varepsilon}(f d \mu)\right|^{2} d \mu \leq \frac{1}{4}|\langle T f, f\rangle|+\frac{1}{2} \int\left|\left(\mathcal{C}_{\varepsilon} \mu\right) \mathcal{C}_{\varepsilon}(f d \mu) f\right| d \mu+C\|f\|_{L^{2}(\mu)}^{2} \tag{28}
\end{equation*}
$$

To estimate the first term on the right side we use the $L^{2}\left(\mu_{\mid A_{\varepsilon}}\right)$ boundedness of $T$ (recall that $\left.\operatorname{supp}(f) \subset A_{\varepsilon}\right)$ :

$$
|\langle T f, f\rangle| \leq\|T f\|_{L^{2}(\mu)}\|f\|_{L^{2}(\mu)} \leq C\|f\|_{L^{2}(\mu)}^{2}
$$

To deal with the second integral on the right side of (28), notice that $\left|\mathcal{C}_{\varepsilon} \mu\right| \leq$ $C_{4}$ on the support of $f$, and so

$$
\int\left|\left(\mathcal{C}_{\varepsilon} \mu\right) \mathcal{C}_{\varepsilon}(f d \mu) f\right| d \mu \leq C_{4} \int\left|\mathcal{C}_{\varepsilon}(f d \mu) f\right| d \mu \leq C_{4}\left\|\mathcal{C}_{\varepsilon}(f d \mu)\right\|_{L^{2}(\mu)}\|f\|_{L^{2}(\mu)}
$$

By (28) we get

$$
\left\|\mathcal{C}_{\varepsilon}(f d \mu)\right\|_{L^{2}(\mu)}^{2} \leq C\|f\|_{L^{2}(\mu)}^{2}+\frac{C_{4}}{2}\left\|\mathcal{C}_{\varepsilon}(f d \mu)\right\|_{L^{2}(\mu)}\|f\|_{L^{2}(\mu)},
$$

which implies that $\left\|\mathcal{C}_{\varepsilon}(f d \mu)\right\|_{L^{2}(\mu)} \leq C\|f\|_{L^{2}(\mu)}$.
So far we have proved the $L^{2}\left(\mu_{\mid A_{\varepsilon}}\right)$ boundedness of $\mathcal{C}_{\mu_{\mid A_{\varepsilon}}, \varepsilon}$. If $A_{\varepsilon}$ were independent of $\varepsilon$, we would set $\nu:=\mu_{\mid A_{\varepsilon}}$ and we would be done. Unfortunately this is not the case and we have to work a little more. We set

$$
B_{\varepsilon}:=\left\{x \in E:\left|\mathcal{C}_{\varepsilon, *} \mu(x)\right| \leq C_{5} \text { and } c_{\mu}^{2}(x) \leq C_{5}^{2}\right\},
$$

where $C_{5}$ is some constant big enough (with $C_{5}>C_{4}$ ) to be chosen below. By Theorem 7 and the discussion above, we know that $\mathcal{C}_{\varepsilon, *}$ is bounded from $M(\mathbb{C})$ into $L^{1, \infty}\left(\mu_{\mid A_{\varepsilon}}\right)$ (with constants independent of $\varepsilon$ ). Thus,

$$
\mu\left\{x \in A_{\varepsilon}:\left|\mathcal{C}_{\varepsilon, *} \mu(x)\right|>C_{5}\right\} \leq \frac{C \mu(E)}{C_{5}} .
$$

If $C_{5}$ is big enough, the right hand side of the preceding inequality is $\leq$ $\mu(E) / 4 \leq \mu\left(A_{\varepsilon}\right) / 2$. Thus, $\mu\left(B_{\varepsilon}\right) \geq \mu(E) / 4$.

We set

$$
B:=\bigcap_{\varepsilon>0} B_{\varepsilon} .
$$

Notice that, by definition, $B_{\varepsilon} \subset B_{\delta}$ if $\varepsilon>\delta$ and so we have

$$
\mu(B)=\lim _{\varepsilon \rightarrow 0} \mu\left(B_{\varepsilon}\right) \geq \frac{1}{4} \mu(E) .
$$

By the same argument used for $A_{\varepsilon}$, it follows that $\mathcal{C}_{\mu_{\mid B_{\varepsilon}}, \varepsilon}$ is bounded on $L^{2}\left(\mu_{B_{\varepsilon}}\right)$ (with constant independent of $\varepsilon$ ), and thus $\mathcal{C}_{\mu_{\mid B}}$ is bounded on $L^{2}\left(\mu_{\mid B}\right)$. If we take $\nu:=\mu_{\mid B}$, we are dome.
Proof of $S_{4} \lesssim \gamma_{+}(E)$. This is a direct consequence of Lemma 10 and the fact that the $L^{2}(\mu)$ boundedness of $\mathcal{C}_{\mu}$ implies its boundedness from $M(\mathbb{C})$ into $L^{1, \infty}(\mu)$, as shown in Theorem 3 .

From the preceding theorem, since the term

$$
\sup \left\{\mu(E): \mu \in \Sigma(E),\left\|\mathcal{C}_{\mu}\right\|_{L^{2}(\mu), L^{2}(\mu)} \leq 1 \forall \varepsilon>0\right\}
$$

is countably semiadditive, we deduce that $\gamma_{+}$is also countably semiadditive.
Corollary 15. The capacity $\gamma_{+}$is countably semiadditive. That is, if $E_{i}$, $i=1,2, \ldots$, is a countable (or finite) family of compact sets, we have

$$
\gamma_{+}\left(\bigcup_{i=1}^{\infty} E_{i}\right) \leq C \sum_{i=1}^{\infty} \gamma_{+}\left(E_{i}\right) .
$$

Another consequence of Theorem 14 is that the capacity $\gamma_{+}$can be characterized in terms of the following potential, introduced by Verdera [53]:

$$
\begin{equation*}
U_{\mu}(x)=\sup _{r>0} \frac{\mu(B(x, r))}{r}+c_{\mu}^{2}(x)^{1 / 2} \tag{29}
\end{equation*}
$$

The precise result is the following.
Corollary 16. For any compact set $E \subset \mathbb{C}$ we have

$$
\gamma_{+}(E) \approx \sup \left\{\mu(E): \mu \in \Sigma(E), U_{\mu}(x) \leq 1 \forall x \in \mathbb{C}\right\} .
$$

The proof of this corollary follows easily from the fact that

$$
\gamma_{+}(E) \approx \sup \left\{\mu(E): \mu \in \Sigma(E), c^{2}(\mu) \leq \mu(E)\right\},
$$

using Tchebychev. The details are left for the reader.
Let us remark that the preceding characterization of $\gamma_{+}$in terms of $U_{\mu}$ is interesting because it suggests that some techniques of potential theory could be useful to study $\gamma_{+}$. See [48] and [53].

## 6 The comparability between $\gamma$ and $\gamma_{+}$, and related results

### 6.1 Comparability between $\gamma$ and $\gamma_{+}$

In [49] the following result has been proved.
Theorem 17. There exists an absolute constant $C$ such that for any compact set $E \subset \mathbb{C}$ we have

$$
\gamma(E) \leq C \gamma_{+}(E)
$$

As a consequence, $\gamma(E) \approx \gamma_{+}(E)$.
An obvious corollary of the preceding result and the characterization of $\gamma_{+}$in terms of curvature obtained in Theorem 14 is the following.

Corollary 18. Let $E \subset \mathbb{C}$ be compact. Then, $\gamma(E)>0$ if and only if $E$ supports a non zero Radon measure with linear growth and finite curvature.

Since we know that $\gamma_{+}$is countably semiadditive, the same happens with $\gamma$ :

Corollary 19. Analytic capacity is countably semiadditive. That is, if $E_{i}$, $i=1,2, \ldots$, is a countable (or finite) family of compact sets, we have

$$
\gamma\left(\bigcup_{i=1}^{\infty} E_{i}\right) \leq C \sum_{i=1}^{\infty} \gamma\left(E_{i}\right)
$$

Notice that, by Theorem 14, to prove Theorem 17 it is enough to show that there exists some measure $\mu$ supported on $E$ with linear growth, satisfying $\mu(E) \approx \gamma(E)$, and such that the Cauchy transform $\mathcal{C}_{\mu}$ is bounded on $L^{2}(\mu)$ with absolute constants. To implement this argument, the main tool used in [49] is the $T(b)$ theorem of Nazarov, Treil and Volberg [36]. To apply this theorem, one has to construct a suitable measure $\mu$ and a function $b \in L^{\infty}(\mu)$ fulfilling some suitable para-accretivity conditions. The construction of $\mu$ and $b$ is the main difficulty which is overcome in [49], by means of a bootstrapping argument which involves the potential $U_{\mu}$ of (29).

Let us remark that the comparability between $\gamma$ and $\gamma_{+}$had been previously proved by P. Jones for compact connected sets by geometric arguments, very different from the ones in [49] (see [40, Chapter 3]). On the other hand, the case of Cantor sets was studied in [25]. The proof of [49] is inspired in part by the ideas in [25].

Corollary 18 yields a characterization of removable sets for bounded analytic functions in terms of curvature of measures. Although this result has a definite geometric flavour, it is not clear if this is a really good geometric characterization. Nevertheless, in [51] it has been shown that the characterization is invariant under bilipschitz mappings, using a corona type decomposition for non doubling measures. See also [16] for an analogous result for some Cantor sets.

### 6.2 Other capacities

In [50], some results analogous to Theorems 14 and 17 have been obtained for the continuous analytic capacity $\alpha$. This capacity, introduced by Vitushkin, is defined like $\gamma$ in (22), with the additional requirement that the functions $f$ considered in the sup should extend continuously to the whole complex plane. In particular, in [50] it is shown that $\alpha$ is semiadditive. This result has some nice consequences for the theory of uniform rational approximation on the complex plane. For example, it implies the so called inner boundary conjecture.

Volberg [57] has proved the natural generalization of Theorem 17 to higher dimensions. In this case, one should consider Lipschitz harmonic
capacity instead of analytic capacity (see [31] for the definition and properties of Lipschitz harmonic capacity). The main difficulty arises from the fact that in this case one does not have any good substitute of the notion of curvature of measures, and then one has to argue with a potential very different from the one defined in (29). See also [24] for related results which avoid the use of any notion similar to curvature.

The techniques in Theorem 17 have also been used by Prat [41] and Mateu, Prat and Verdera [23] to study the capacities $\gamma_{\alpha}$ associated to $\alpha$ dimensional signed Riesz kernels with $\alpha$ non integer:

$$
k(x, y)=\frac{y-x}{|y-x|^{\alpha+1}} .
$$

In [41] it is shown that sets with finite $\alpha$-dimensional Hausdorff measure have vanishing capacity $\gamma_{\alpha}$ when $0<\alpha<1$. Moreover, for these $\alpha$ 's it is proved in [23] that $\gamma_{\alpha}$ is comparable to one of the non linear Wolff's capacities. The case of non integer $\alpha$ with $\alpha>1$ seems much more difficult to study, although in the AD regular situation some results have been obtained [41]. The results in [41] and [23] show that the behavior of $\gamma_{\alpha}$ with $\alpha$ non integer is very different from the one with $\alpha$ integer.

## Acknowledgements

I would like to thank Genaro López for his kind invitation to give the lectures from which these notes arose, and the entire Department of Mathematical Analysis of the Universidad de Sevilla for his hospitality during my visit in December 2003.

## References

[1] L. V. Ahlfors, Bounded analytic functions, Duke Math. J. 14, (1947). 1-11.
[2] W. Chen and E. Sawyer, A note on commutators of fractional integrals with RBMO functions, Illinois J. Math. 46 (2002), 1287-1298.
[3] M. Christ, Lectures on singular integral operators, CBMS Regional Conf. Ser. in Math. 77 Amer. Math. Soc., Providence, 1990.
[4] G. David, Opérateurs intégraux singuliers sur certaines courbes du plan complexe, Ann. Sci. École Norm. Sup. (4)17 (1984), no. 1, 157-189.
[5] G. David, Wavelets and singular integrals on curves and surfaces, Lecture Notes in Math. 1465 (1991), Springer.
[6] G. David, Analytic capacity, Calderón-Zygmund operators, and rectifiability, Publ. Mat. 43 (1999), no. 1, 3-25.
[7] G. David, Unrectifiable 1-sets have vanishing analytic capacity. Revista Mat. Iberoamericana 14 (2) (1998), 369-479.
[8] G. David and P. Mattila, Removable sets for Lipschitz harmonic functions in the plane, Rev. Mat. Iberoamericana 16 (2000), no. 1, 137-215.
[9] A. M. Davie, Analytic capacity and approximation problems, Trans. Amer. Math. Soc. 171 (1972), 409-444.
[10] A. M. Davie and B. Øksendal, Analytic capacity and differentiability properties of finely harmonic functions, Acta Math. 149 (1982), no. 1-2, 127-152.
[11] T. Gamelin, Uniform Algebras, Prentice Hall, Englewood Cliffs N.J., 1969.
[12] J. García-Cuerva and J. M: Martell, Weighted inequalities and vectorvalued Calderón-Zygmund operators on non-homogeneous spaces, Publ. Mat. 44 (2000), 613-640.
[13] J. García-Cuerva and J. M: Martell, Two-weight norm inequalities for maximal operators and fractional integrals on non-homogeneous spaces, Indiana Univ. Math. J. 50 (2001), no. 3, 1241-1280.
[14] J. García-Cuerva and E. Gatto, Boundedness properties of fractional integral operators associated to non-doubling measures, Studia Math. 162 (2004), 245-261
[15] J. García-Cuerva and E. Gatto, Lipschitz spaces and CalderónZygmund operators associated to non-doubling measures, Preprint (2002).
[16] J. Garnett and J. Verdera, Analytic capacity, bilipschitz maps and Cantor sets, Math. Res. Lett. 10 (2003), 515-522.
[17] Y. Han and D. Yang, Triebel-Lizorkin spaces with non doubling measures, to appear in Studia Math.
[18] G. Hu, Y. Meng and D. Yang, Multilinear commutators of singular integrals with non doubling measures, to appear in Integral Equations Operator Theory.
[19] G. Hu, Y. Meng and D. Yang, New atomic characterization of $H^{1}$ space with non-doubling measures and its applications, to appear in Math. Proc. Cambridge Philos. Soc.
[20] J.C. Léger, Menger curvature and rectifiability, Ann. of Math. 149 (1999), 831-869.
[21] J. Martín and M. Milman, Gehring's lemma for nondoubling measures, Michigan Math. J. 47 (2000), 559-573.
[22] J. Mateu, P. Mattila, A. Nicolau, and J. Orobitg, BMO for nondoubling measures, Duke Math. J. 102 (2000), 533-565.
[23] J. Mateu, L. Prat and J. Verdera, The capacities associated to signed Riesz kernels, and Wolff potentials, Preprint (2003). To appear in J. Reine Angew. Math.
[24] J. Mateu and X. Tolsa, Riesz transforms and harmonic Lip ${ }_{1}$-capacity in Cantor sets, Preprint (2003), to appear in Proc. London Math. Soc.
[25] J. Mateu, X. Tolsa and J. Verdera, The planar Cantor sets of zero analytic capacity and the local $T(b)$-Theorem, J. Amer. Math. Soc. 16 (2003), 19-28.
[26] J. Mateu, X. Tolsa and J. Verdera, On the semiadditivity of analytic capacity and planar Cantor sets, Contemp. Math. 320 (2003), 259-278. The planar Cantor sets of zero analytic capacity and the local $T(b)$ Theorem, J. Amer. Math. Soc. 16 (2003), 19-28.
[27] P. Mattila. Geometry of sets and measures in Euclidean spaces, Cambridge Stud. Adv. Math. 44 Cambridge Univ. Press, Cambridge, 1995.
[28] P. Mattila, Singular integrals, analytic capacity and rectifiability, Proceedings of the conference dedicated to Professor Miguel de Guzmßn (El Escorial, 1996). J. Fourier Anal. Appl. 3 (1997), Special Issue, 797-812.
[29] P. Mattila, Haudorff dimension, projections, and Fourier transform, Publ. Mat. 48 (2004), 3-48.
[30] P. Mattila, M.S. Melnikov and J. Verdera, The Cauchy integral, analytic capacity, and uniform rectifiability, Ann. Math. (2) 144 (1996), 127-136.
[31] P. Mattila and P.V. Paramonov. On geometric properties of harmonic Lip -capacity, Pacific J. Math. 171:2 (1995), 469-490.
[32] M. S. Melnikov, Analytic capacity: discrete approach and curvature of a measure, Sbornik: Mathematics 186(6) (1995), 827-846.
[33] M. S. Melnikov and J. Verdera, A geometric proof of the $L^{2}$ boundedness of the Cauchy integral on Lipschitz graphs, Internat. Math. Res. Notices (1995), 325-331.
[34] F. Nazarov, S. Treil and A. Volberg, Cauchy integral and CalderónZygmund operators in non-homogeneous spaces, Internat. Res. Math. Notices 15 (1997), 703-726.
[35] F. Nazarov, S. Treil and A. Volberg, Weak type estimates and Cotlar inequalities for Calderón-Zygmund operators in nonhomogeneous spaces, Int. Math. Res. Notices 9 (1998), 463-487.
[36] F. Nazarov, S. Treil and A. Volberg. How to prove Vitushkin's conjecture by pulling ourselves up by the hair. Preprint (2000).
[37] F. Nazarov, S. Treil and A. Volberg, The $T b$-theorem on nonhomogeneous spaces, Acta Math. 190 (2003), 151-239.
[38] F. Nazarov, S. Treil and A. Volberg, Accretive system $T b$-theorems on nonhomogeneous spaces, Duke Math. J. 113 (2002), 259-312.
[39] J. Orobitg and C. Pérez, $A_{p}$ weights for nondoubling measures in Rscalarproductn and applications, Trans. Amer. Math. Soc. 354 (2002), 2013-2033
[40] H. Pajot, Analytic capacity, rectifiability, Menger curvature and the Cauchy integral, Lecture Notes in Math. 1799 (2002) Springer.
[41] L. Prat, Potential theory of signed Riesz kernels: capacity and Hausdorff measures, Int. Math. Res. Notices 19 (2004) 937-981.
[42] X. Tolsa, $L^{2}$-boundedness of the Cauchy integral operator for continuous measures, Duke Math. J. 98 (2) (1999) 269-304.
[43] X. Tolsa, Cotlar's inequality and existence of principal values for the Cauchy integral without the doubling condition, J. Reine Angew. Math. 502 (1998) 199-235.
[44] X. Tolsa, $B M O, H^{1}$ and Calderón-Zygmund operators for non doubling measures, Math. Ann. 319 (2001) 89-149.
[45] X. Tolsa, A proof of the weak $(1,1)$ inequality for singular integrals with non doubling measures based on a Calderón-Zygmund decomposition, Publ. Mat.45:1 (2001) 163-174.
[46] X. Tolsa, The space Hscalarproduct1 for nondoubling measures in terms of a grand maximal operator, Trans. Amer. Math. Soc. 355 (2003), 315-348
[47] X. Tolsa, Littlewood-Paley theory and the $T(1)$ theorem with non doubling measures, Adv. Math. 164 (2001), 57-116.
[48] X. Tolsa, On the analytic capacity $\gamma_{+}$, Indiana Univ. Math. J. 51 (2) (2002), 317-344.
[49] X. Tolsa, Painlevé's problem and the semiadditivity of analytic capacity, Acta Math.190:1 (2003), 105-149.
[50] X. Tolsa, The semiadditivity of continuous analytic capacity and the inner boundary conjecture. Preprint (2002). To appear in Amer. J. Math.
[51] X. Tolsa, Bilipschitz maps, analytic capacity, and the Cauchy integral, Preprint (2003).
[52] J. Verdera. Removability, capacity and approximation. In Complex Potential Theory, (Montreal, PQ, 1993), NATO Adv. Sci. Int. Ser. C Math. Phys. Sci. 439 Kluwer Academic Publ., Dordrecht, 1994, pp. 419-473.
[53] J. Verdera, On the $T(1)$-theorem for the Cauchy integral, Ark. Mat. 38 (2000), 183-199.
[54] J. Verdera, The fall of the doubling condition in Calderón-Zygmund theory, Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations (El Escorial, 2000). Publ. Mat. 2002 Vol. Extra, 275-292.
[55] A. G. Vitushkin, Estimate of the Cauchy integral, (Russian) Mat. Sb. 71 (113) (1966), 515-534.
[56] A. G. Vitushkin, The analytic capacity of sets in problems of approximation theory, Uspeikhi Mat. Nauk. 22 (6) (1967), 141-199 (Russian); in Russian Math. Surveys 22 (1967), 139-200.
[57] A. Volberg, Calderón-Zygmund capacities and operators on nonhomogeneous spaces. CBMS Regional Conf. Ser. in Math. 100 Amer. Math. Soc., Providence, 2003.

## Author's address:

Xavier Tolsa<br>Institució Catalana de Recerca i Estudis Avançats (ICREA) and Departament de Matemàtiques Universitat Autònoma de Barcelona, Spain E-mail: xtolsa@mat.uab.es

# Discrete and continuous-time models of financial derivatives 

Hong-Kun Xu*


#### Abstract

Quantitative finance is one of the fastest growing areas of applied mathematics. Under the assumption that the underlying asset (stock) follow a binomial model (in the discrete case) or the geometric Brownian motion model (in the continuous-time case), we discuss the pricing problem of financial derivatives written on the stock. Both European and American options are discussed. Replications and hedging are included. The Black-Scholes model with a stochastic volatility is also discussed.


## 1 The Discrete Binomial Models

### 1.1 The Binomial Model

In our financial world, we assume that two assets are traded: one is riskless and called bond, denoted $B$; one is risky and called stock, denoted $S$.

In a discrete model, we assume that the trading of the stock $S$ is carried out at discrete times $i=0,1,2, \cdots, N$. In a binomial model, we assume that, if, at time $i \geq 0$, the stock price is $S_{i}$, then at the next time $i+1$, the stock price takes one of possible two values: either $u S_{i}$ or $d S_{i}$, where $u>0$ and $d>0$ are constants satisfying the following assumptions (in order to be free from arbitrage):

$$
d<1+r<u
$$

where the constant $r>0$ is the interest rate for each period between times $i$ and $i+1$.

[^10]The stock prices $\left\{S_{i}\right\}_{i=0}^{N}$ are thus random variables. We may identify them in the following way (see [7] for more details). Let $\Omega$ denote the sample space of outcomes of tossing a coin $N$ times independently; that is,

$$
\Omega=\left\{\omega=\omega_{1} \cdots \omega_{N}: \omega_{i}=H \text { or } T, 1 \leq i \leq N\right\}
$$

where $H$ represents for a head, and $T$ for a tail. Assume $p=\mathbb{P}\{H\}$ and $q=\mathbb{P}\{T\}$ are the probabilities that $H$ and respectively, $T$ appears when the coin is tossed once.

Now we can assume that each $S_{i}$ is random variable defined on $\Omega$, but there is a convention: for $\omega=\omega_{1} \cdots \omega_{N} \in \Omega$,

$$
S_{i}(\omega)=S_{i}\left(\omega_{1} \cdots \omega_{N}\right)=S_{i}\left(\omega_{1} \cdots \omega_{i}\right)
$$

only depending upon $\omega_{1} \cdots \omega_{i}$, the outcome of the first $i$ tosses of the coin.
Then the binomial model of our stock is

$$
\begin{aligned}
S_{i+1}\left(\omega_{1} \cdots \omega_{i} H\right) & =u S_{i}\left(\omega_{1} \cdots \omega_{i}\right) \\
S_{i+1}\left(\omega_{1} \cdots \omega_{i} T\right) & =d S_{i}\left(\omega_{1} \cdots \omega_{i}\right)
\end{aligned}
$$

Therefore,

$$
S_{i}\left(\omega_{1} \cdots \omega_{N}\right)=S_{0} u^{\# H_{i}} d^{\# T_{i}}
$$

where $\# H_{i}\left(\# T_{i}\right)$ is the number of heads (tails) in the outcome of the first $i$ tosses $\omega_{1} \cdots \omega_{i}$ of the coin; that is,

$$
\begin{aligned}
\# H_{i} & =\#\left\{j: \omega_{j}=H, 1 \leq j \leq i\right\} \\
\# T_{i} & =\#\left\{j: \omega_{j}=T, 1 \leq j \leq i\right\}
\end{aligned}
$$

### 1.2 Risk-Neutral Valuation Formula

Now suppose that we have a European style derivative security written on the stock $S$. For instance, a European call or put option on $S$. Recall that a European call (put) option gives the holder the right, but not the obligation, to purchase (sell) one share of the underlying stock for a predetermined price $K$ (called strike price) at a predetermined time $T$ (called the expiration time).

Since the derivative security has a right and no obligation, it has a value. One of the fundamental problems in quantitative finance is to determine the (fair or no-arbitrage) value of the security (options, in particular). Below we concentrate on European options for simplicity.

We use $h(x)$ to denote the payoff function. Hence

$$
h(x)=(x-K)^{+}= \begin{cases}x-K & \text { if } x>K, \\ 0 & \text { if } x \leq K,\end{cases}
$$

for a European call option, and

$$
h(x)=(K-x)^{+}= \begin{cases}K-x & \text { if } x<K \\ 0 & \text { if } x \geq K\end{cases}
$$

for a European put option.
Since at the expiration time $T$, the value of the option is a (known) random variable; it is the payoff $h\left(S_{T}\right)=\left(S_{T}-K\right)^{+}$for a call, and $h\left(S_{T}\right)=$ $\left(K-S_{T}\right)^{+}$for a put. We can work out the values at all times by a backward induction method.

Denote by $C_{0}, C_{1}, \cdots, C_{N}$ the values of the option at times $i=0,1, \cdots, N ;$ thus $C_{N}=h\left(S_{T}\right)$. So the question is how to determine $C_{i}$ for $i=N-$ $1, \cdots, 1,0$.

Let $\mathcal{F}_{i}$ denote the $\sigma$-algebra of the stock price information up to time $i$; that is, $\mathcal{F}_{i}$ is the $\sigma$-algebra generated by $\left\{S_{j}\right\}_{j=0}^{i}$, or by the first $i$ tosses of the coin. So $S_{i}$ is known with respect to $\mathcal{F}_{i}$; we say that $S_{i}$ is $\mathcal{F}_{i}$-measurable. $\left\{\mathcal{F}_{i}\right\}_{i=0}^{N}$ is an increasing sequence of $\sigma$-algebras and is called a filtration on $\Omega$. Moreover, $\left(\Omega, \mathcal{F},\left\{\mathcal{F}_{i}\right\}_{i=0}^{N}, \mathbb{P}\right)$ is called a filtered probability space. (Here $\mathcal{F}$ denotes the $\sigma$-algebra of all price information up to time $N$; usually, it is the power set of $\Omega$, and $\mathbb{P}$ is the market probability measure.)

Discounting plays a fundamental role in option pricing theory. We use $\left\{\tilde{S}_{i}\right\}_{i=0}^{N}$ and $\left\{\tilde{C}_{i}\right\}_{i=0}^{N}$ to denote the discounted stock price and the option value processes, respectively. That is, for each $0 \leq i \leq N$,

$$
\begin{aligned}
& \tilde{S}_{i}=\frac{1}{(1+r)^{i}} S_{i}, \\
& \tilde{C}_{i}=\frac{1}{(1+r)^{i}} C_{i} .
\end{aligned}
$$

The risk-neutral probabilities $\tilde{p}$ and $\tilde{q}$ play key role, which are the noarbitrage probabilities that the stock price goes up to $u S_{i}$ and down to $d S_{i}$ from time $i$ to the next time $i+1$. Since the risk-neutral probabilities give no tendency for the prices go upward or downward on average, we have, in terms of conditional expectation,

$$
\begin{equation*}
\tilde{\mathbb{E}}\left[\tilde{S}_{i+1} \mid \mathcal{F}_{i}\right]=\tilde{S}_{i}, \quad i=0,1, \cdots, N-1 \tag{1}
\end{equation*}
$$

In other words, the discounted stock price process $\left\{\tilde{S}_{i}\right\}_{i=0}^{N}$ is a martingale (with respect to the risk-neutral probability measure).

We can find $\tilde{p}$ and $\tilde{q}$ from (1). Indeed, since the stock price process is Markovian (no memory of the past histories), we have, noting that

$$
\begin{aligned}
& \tilde{S}_{i+1}\left(\omega_{1} \cdots \omega_{i} H\right)=\frac{1}{1+r} u \tilde{S}_{j}\left(\omega_{1} \cdots \omega_{i}\right), \\
& \tilde{S}_{i+1}\left(\omega_{1} \cdots \omega_{i} T\right)=\frac{1}{1+r} d \tilde{S}_{j}\left(\omega_{1} \cdots \omega_{i}\right), \\
& \tilde{\mathbb{E}}\left[\tilde{S}_{i+1} \mid \mathcal{F}_{i}\right]\left(\omega_{1} \cdots \omega_{i}\right)=\tilde{\mathbb{E}}\left[\tilde{S}_{i+1} \mid \tilde{S}_{i}\right]\left(\omega_{1} \cdots \omega_{i}\right) \\
&=\tilde{p} \tilde{S}_{i+1}\left(\omega_{1} \cdots \omega_{i} H\right)+\tilde{q} \tilde{S}_{i+1}\left(\omega_{1} \cdots \omega_{i} T\right) \\
&=\tilde{p} \frac{1}{1+r} u \tilde{S}_{i}\left(\omega_{1} \cdots \omega_{i}\right)+\tilde{q} \frac{1}{1+r} d \tilde{S}_{i}\left(\omega_{1} \cdots \omega_{i}\right) \\
&=\frac{1}{1+r}(\tilde{p} u+\tilde{q} d) \tilde{S}_{i}\left(\omega_{1} \cdots \omega_{i}\right) .
\end{aligned}
$$

So the martingale property (1) implies that

$$
1=\frac{1}{1+r}(\tilde{p} u+\tilde{q} d) .
$$

Hence

$$
\begin{equation*}
\tilde{p}=\frac{1+r-d}{u-d}, \quad \tilde{q}=\frac{u-1-r}{u-d} . \tag{2}
\end{equation*}
$$

Under the risk-neutral probabilities $\tilde{p}$ and $\tilde{q}$, the discounted value process $\left\{\tilde{C}_{i}\right\}_{i=0}^{N}$ is also a martingale:

$$
\begin{equation*}
\tilde{\mathbb{E}}\left[\tilde{C}_{i+1} \mid \mathcal{F}_{i}\right]=\tilde{C}_{i}, \quad i=0,1, \cdots, N-1 . \tag{3}
\end{equation*}
$$

Since a martingale has a constant expectation, we have the risk-neutral pricing formula:

$$
\begin{equation*}
C_{i}=\tilde{\mathbb{E}}\left[\left.\frac{C_{N}}{(1+r)^{N-i}} \right\rvert\, \mathcal{F}_{i}\right], \quad i=0,1, \cdots, N . \tag{4}
\end{equation*}
$$

In particular, we have the fair price at time 0 which is the premium that the buyer must pay upfront to the seller:

$$
\begin{equation*}
C_{0}=\tilde{\mathbb{E}}\left[\frac{C_{N}}{(1+r)^{N}}\right] . \tag{5}
\end{equation*}
$$

### 1.3 Hedging

To hedge means to eliminate risks. Our portfolio consists of stock and bond. A trading strategy specifies the number of shares of the stock and the number of units of the bond in the portfolio.

Let $\phi=\left(\phi_{n}^{0}, \phi_{n}^{1}\right)_{n=0}^{N}$ be our dynamic trading strategy, where $\phi_{n}^{0}$ is the number of units of bond and $\phi_{n}^{1}$ the number of shares of the stock which are determined at time $n-1$ and held in our portfolio until the next time $n$ (At time $n$, due to arrival of new price information, the trading strategy is re-adjusted to $\left(\phi_{n+1}^{0}, \phi_{n+1}^{1}\right)$ which is held until the next time $n+1$.)

The value process $\left\{V_{n}(\phi)\right\}_{n=0}^{N}$ of this trading strategy $\phi$ is defined as

$$
V_{n}(\phi)=\phi_{n}^{1} S_{n}+\phi_{n}^{0} B_{n}, \quad n=0,1, \cdots, N
$$

where $B_{n}$ is the cash amount at time $n$ in the bond; that is, $B_{n}=(1+r)^{n}$. We write $V_{n}=V_{n}(\phi)$ unless specified otherwise.

A trading strategy $\phi=\left(\phi_{n}^{0}, \phi_{n}^{1}\right)_{n=0}^{N}$ is called self-financing if

$$
\begin{equation*}
\phi_{n}^{1} S_{n}+\phi_{n}^{0} B_{n}=\phi_{n+1}^{1} S_{n}+\phi_{n+1}^{0} B_{n}, \quad n=0,1, \cdots, N-1 . \tag{6}
\end{equation*}
$$

The interpretation is that at time $n$, once the new price $S_{n}$ is quoted, the investor re-adjusts his position from $\left(\phi_{n}^{0}, \phi_{n}^{1}\right)$ to $\left(\phi_{n+1}^{0}, \phi_{n+1}^{1}\right)$, without bringing or consuming any wealth (no new cash input or output), so the value of the portfolio must remain unchanged.

Note that in a trading strategy $\phi=\left(\phi_{n}^{0}, \phi_{n}^{1}\right), \phi_{n}^{1}$ is determined at time $n-1\left(B_{n}\right.$ is deterministic for all $\left.n\right)$. In other words, $\phi_{n}^{1}$ is determined by the stock price information up to time $n-1$, or $\phi_{n}^{1}$ is $\mathcal{F}_{n-1}$-measurable. We say that $\left\{\phi_{n}^{1}\right\}$ is predictable (or previsible).

A trading strategy $\phi=\left(\phi_{n}^{0}, \phi_{n}^{1}\right)$ is admissible if it is self-financing and if $V_{n} \geq 0$ at all times $n \geq 0$.

An admissible trading strategy $\phi=\left(\phi_{n}^{0}, \phi_{n}^{1}\right)$ is said to replicate a derivative security with expiration time $N$ if

$$
V_{N}=C_{N}
$$

where $C_{N}$ is the payoff of the security on expiration. In this case, we also say that $\phi=\left(\phi_{n}^{0}, \phi_{n}^{1}\right)$ is a replicating trading strategy for the security.

The discounted value process $\left\{\tilde{V}_{n}\right\}$ is given by

$$
\tilde{V}_{n}=\frac{1}{B_{n}} V_{n}=\frac{1}{(1+r)^{n}} V_{n}, \quad n=0,1, \cdots, N .
$$

The self-financing property (6) can be rewritten as

$$
\begin{equation*}
\left(\phi_{n}^{1}-\phi_{n+1}^{1}\right) \tilde{S}_{n}=\phi_{n+1}^{0}-\phi_{n}^{0} . \tag{7}
\end{equation*}
$$

It follows that

$$
\begin{align*}
\tilde{V}_{n+1}-\tilde{V}_{n} & =\phi_{n+1}^{1} \tilde{S}_{n+1}+\phi_{n+1}^{0}-\phi_{n}^{1} \tilde{S}_{n}-\phi_{n}^{0} \\
& =\phi_{n+1}^{1}\left(\tilde{S}_{n+1}-\tilde{S}_{n}\right) . \tag{8}
\end{align*}
$$

Thus

$$
\begin{equation*}
\tilde{V}_{n}=V_{0}+\sum_{j=0}^{n-1} \phi_{j+1}^{1}\left(\tilde{S}_{j+1}-\tilde{S}_{j}\right) \tag{9}
\end{equation*}
$$

Eq. (9) says that the discounted value process $\left\{\tilde{V}_{n}\right\}$ is the martingale transform of the discounted stock price of $\left\{\tilde{S}_{n}\right\}$ by the previsible process $\left\{\phi_{n}^{1}\right\}$; hence, the discounted value process $\left\{\tilde{V}_{n}\right\}$ is martingale under the risk-neutral probability measure $\tilde{\mathbb{P}}$ :

$$
\begin{equation*}
\tilde{\mathbb{E}}\left[\tilde{V}_{n+1} \mid \mathcal{F}_{n}\right]=\tilde{V}_{n}, \quad n=0,1, \cdots, N-1 \tag{10}
\end{equation*}
$$

In particular,

$$
\begin{equation*}
\tilde{V}_{n}=\tilde{\mathbb{E}}\left[\tilde{V}_{N} \mid \mathcal{F}_{n}\right], \quad n=0,1, \cdots, N \tag{11}
\end{equation*}
$$

Undoing the discounting gives

$$
\begin{equation*}
V_{n}=\tilde{\mathbb{E}}\left[\left.\frac{1}{(1+r)^{N-n}} V_{N} \right\rvert\, \mathcal{F}_{n}\right], \quad n=0,1, \cdots, N \tag{12}
\end{equation*}
$$

So, if $\phi=\left(\phi_{n}^{0}, \phi_{n}^{1}\right)$ is a replicating trading strategy (i.e., $V_{N}=C_{N}$ ), then from the risk-neutral pricing formula (4) and (12), we find that

$$
C_{n}=V_{n}, \quad n=0,1, \cdots, N .
$$

In another word, at every time, the value of the derivative security is the same as the value of the replicating strategy. Let us now find what is $\phi_{n}^{1}$ in this portfolio. From Eq. (8), we have

$$
\begin{equation*}
\phi_{n}^{1}=\frac{\Delta \tilde{V}_{n}}{\Delta \tilde{S}_{n}}=\frac{\Delta \tilde{C}_{n}}{\Delta \tilde{S}_{n}}, \quad n=1,2, \cdots, N \tag{13}
\end{equation*}
$$

where

$$
\begin{aligned}
\Delta \tilde{V}_{n}=\tilde{V}_{n}-\tilde{V}_{n-1}, & n \geq 1 \\
\Delta \tilde{C}_{n}=\tilde{C}_{n}-\tilde{C}_{n-1}, & n \geq 1 \\
\Delta \tilde{S}_{n}=\tilde{S}_{n}-\tilde{S}_{n-1}, & n \geq 1
\end{aligned}
$$

To summarize we have:

To price and hedge a European style derivative security against the payoff $C_{T}$ at the expiration time $T$, one follows the three steps below:

- Use (2) to find the risk-neutral probability (martingale) measure $\tilde{\mathbb{P}}$ under which the discounted stock price process $\left\{\tilde{S}_{n}\right\}$ is a martingale.
- Form the discounted value process

$$
\tilde{V}_{n}=\frac{1}{(1+r)^{n}} V_{n}=\tilde{\mathbb{E}}\left[\left.\frac{C_{N}}{(1+r)^{N}} \right\rvert\, \mathcal{F}_{n}\right]
$$

- Find the previsible process $\left\{\phi_{n}^{1}\right\}$ determined by

$$
\phi_{n}^{1}=\frac{\Delta \tilde{V}_{n}}{\Delta \tilde{S}_{n}}
$$

which is the holding in the portfolio of the number of shares of the stock at time $n-1$.

Remark 1. (i) The above summary shows us a fact that every European style derivative security against a payoff $C_{N}$ (often also called a contingent claim) can be replicated. Such a market is called complete; otherwise it is incomplete.
(ii) In the above we showed that $\tilde{\mathbb{P}}$ is a probability measure under which every discounted stock price process is a martingale. Note that $\tilde{\mathbb{P}}$ is indeed equivalent to the market probability measure $\mathbb{P}$; that is,

$$
\tilde{\mathbb{P}}(\omega)=0 \quad \Leftrightarrow \quad \mathbb{P}(\omega)=0, \quad \omega \in \Omega .
$$

It has been shown that a no-arbitrage market is complete if and only if there exists a unique probability measure $\tilde{\mathbb{P}}$ which is equivalent to $\mathbb{P}$ and under which every discounted stock price process is a martingale.

### 1.4 American Options

Let us consider an American style derivative security on a stock within the binomial model discussed above. Recall that an American call (put) gives the holder the right, but not the obligation, to purchase (sell) one share of the stock for a predetermined price $K$ (the strike price) at any time before or at a predetermined time $T$ (the expiration time).

Since an American option has more right over its European counterpart, the American option is worth more than (or at least) its European counterpart. However, it has been shown that an American call option has the same value as its European counterpart, if there is no dividend paid out during the lifetime of the option. But an American put option usually has bigger value than its European counterpart.

An American option can be exercised at any time between 0 and $N$, we can identify the American option with a stochastic process $\left\{Z_{n}\right\}_{n=0}^{N}$ which is adapted to the natural filtration $\left\{\mathcal{F}_{n}\right\}$, where $Z_{n}$ is the immediate profit obtained from exercising the option at time $n$. For an American put, $Z_{n}=\left(K-S_{n}\right)^{+}$.

We shall use the backward method to compute the values of the option. Let $\left\{V_{j}\right\}$ be the value process of the American option. Let $g$ be the payoff function (e.g., $g(x)=(K-x)^{+}$for an American put option).

## American Algorithm.

$$
\begin{aligned}
v_{N}(x) & =g(x) \\
v_{j}(x) & =\max \left\{\frac{1}{1+r}\left(\tilde{p} v_{j+1}(u x)+\tilde{q} v_{j+1}(d x)\right), g(x)\right\}, 0 \leq j \leq N-1, \\
V_{j} & =v_{j}\left(S_{j}\right) \text { is the value of the American option at time } j, 0 \leq j \leq N
\end{aligned}
$$

Alternatively,

$$
\begin{aligned}
V_{N} & =Z_{N}, \\
V_{j} & =\max \left\{\frac{1}{1+r} \tilde{\mathbb{E}}\left[V_{j+1} \mid \mathcal{F}_{j}\right], Z_{j}\right\},
\end{aligned}
$$

We see that $V_{j} \geq Z_{j}$ for all $j$; that is, the value of the American option is worth at all times at least the payoff obtained from exercising.

Again if we consider the discounted processes:

$$
\tilde{Z}_{j}=\frac{1}{(1+r)^{j}} Z_{j},
$$

$$
\tilde{V}_{j}=\frac{1}{(1+r)^{j}} V_{j},
$$

we have

Proposition 2. The discounted value process $\left\{\tilde{V}_{j}\right\}$ is a $\tilde{\mathbb{P}}$-supermartingale:

$$
\tilde{\mathbb{E}}\left[\tilde{V}_{j+1} \mid \mathcal{F}_{j}\right] \leq \tilde{V}_{j}, \quad 0 \leq j \leq N-1
$$

Moreover, $\left\{\tilde{V}_{j}\right\}$ is the smallest $\tilde{\mathbb{P}}$-supermartingale such that $V_{j} \geq Z_{j}$ for all $j$.

When should the owner exercise his American option? To answer this question we need the concept of stopping times:

A random variable $\tau: \Omega \rightarrow\{0,1,2, \cdots, N\}$ is a stopping time if, for each $0 \leq n \leq N$,

$$
\{\tau=n\} \in \mathcal{F}_{n}
$$

Example 3. The random variable at which the value of the American option hits the payoff first time is a stopping time; that is,

$$
\begin{equation*}
\tau(\omega):=\min \left\{j: V_{j}(\omega)=Z_{j}(\omega)\right\}, \quad \omega \in \Omega \tag{14}
\end{equation*}
$$

is a stopping time.
If $\left\{X_{n}\right\}$ is a process and if $\tau$ is a stopping time, then the stopped process $\left\{X_{n \wedge \tau}\right\}$ is defined as follows:

$$
\left(X_{n \wedge \tau}\right)(\omega)=X_{n \wedge \tau(\omega)}(\omega) .
$$

It is known that if $\left\{X_{n}\right\}$ is a martingale and if $\tau$ is a stopping time, then the stopped process $\left\{X_{n \wedge \tau}\right\}$ is also a martingale.

Definition 4. A stopping time $\tau_{0}$ is said to be optimal for the American option $\left\{Z_{j}\right\}$ if

$$
\tilde{\mathbb{E}}\left[Z_{\tau_{0}} \mid \mathcal{F}_{0}\right]=\sup _{\tau \in \tau_{0}, N} \tilde{\mathbb{E}}\left[Z_{\tau} \mid \mathcal{F}_{0}\right],
$$

where $\tau_{0, N}$ is the set of all stopping times $\tau$ taking values in $\{0,1, \cdots, N\}$.
Theorem 5. Let $\left\{V_{j}\right\}$ be defined by the American Algorithm. We have
(i) For $0 \leq j \leq N$,

$$
\begin{equation*}
V_{j}=\max _{\tau}\left\{(1+r)^{j} \tilde{\mathbb{E}}\left[(1+r)^{-\tau} Z_{\tau} \mid \mathcal{F}_{j}\right]\right\} \tag{15}
\end{equation*}
$$

where the max is taken over all stopping times $\tau$ such that $j \leq \tau \leq N$.
(ii) The stopping time $\tau \in \tau_{0, N}$ is optimal if

$$
V_{0}=\tilde{\mathbb{E}}\left[(1+r)^{-\tau} Z_{\tau}\right]
$$

In particular, the stopping time $\tau$ defined in (14) is optimal. Indeed, if the owner does not exercise his American option at time $j$ in the state $\omega$ such that $V_{j}(\omega)=Z_{j}(\omega)$, then the seller of the option can immediately consume the positive amount of the difference

$$
V_{j}(\omega)-\frac{1}{1+r} \tilde{\mathbb{E}}\left[V_{j+1} \mid \mathcal{F}_{j}\right](\omega)
$$

without expose to any risk.
In terms of discounted processes, we can rewrite (15) as

$$
\tilde{V}_{j}=\max _{\tau_{j, N}} \tilde{\mathbb{E}}\left[\tilde{Z}_{\tau} \mid \mathcal{F}_{j}\right]
$$

Since $\left\{\tilde{V}_{j}\right\}$ is a supermartingale, we have the decomposition:

$$
\tilde{V}_{j}=\tilde{M}_{j}-\tilde{A}_{j}
$$

where $\left\{\tilde{M}_{j}\right\}$ is a martingale and $\left\{\tilde{A}_{j}\right\}$ is an increasing predictable process null at 0 .

Since the market is complete, there is a self-financing trading strategy $\phi=\left(\phi_{n}^{0}, \phi_{n}^{1}\right)$ replicating $\tilde{M}$ :

$$
\tilde{V}_{N}(\phi)=\tilde{M}_{N}
$$

Since the discounted value process $\left\{\tilde{V}_{j}(\phi)\right\}$ is a $\tilde{\mathbb{P}}$-martingale, we get for any j,

$$
\begin{aligned}
\tilde{V}_{j}(\phi) & =\tilde{\mathbb{P}}\left[\tilde{V}_{N}(\phi) \mid \mathcal{F}_{j}\right] \\
& =\tilde{\mathbb{P}}\left[\tilde{M}_{N} \mid \mathcal{F}_{j}\right] \\
& =\tilde{M}_{j} .
\end{aligned}
$$

So we have

$$
\tilde{V}_{j}=\tilde{V}_{j}(\phi)-\tilde{A}_{j}
$$

and

$$
V_{j}=V_{j}(\phi)-A_{j},
$$

with $A_{j}=(1+r)^{j} \tilde{A}_{j}$. It follows that the writer of the American option can perfectly hedge himself in the following way: Once he receives the premium $V_{0}=V_{0}(\phi)$ at time 0 , he then follows the trading strategy $\phi$ that creates a wealth $V_{j}(\phi)$ at time $j$ which is bigger than or equal to $V_{j}$. So there is no risk for him to payoff the amount $Z_{j}$ when the owner exercises the option at time $j$.

Remark 6. As mentioned above, the holder of an American option should exercise at the stopping time $\tau$ determined in Eq. (14). The holder obviously does not exercise the option at time $j$ if $V_{j}>Z_{j}$ because if he does, he receives $Z_{j}$, while the market value at time $j$ is $V_{j}$. So he must exercise at a time $\tau$ at which $V_{\tau}=Z_{\tau}$. The first time for this to happen is given by the stopping time in Eq. (14).

On the other hand, the largest optimal stopping time is given by

$$
\tau_{\max }= \begin{cases}N, & \text { if } A_{N}=0, \\ \inf \left\{j: A_{j+1} \neq 0\right\}, & \text { if } A_{N} \neq 0 .\end{cases}
$$

## 2 The Black-Scholes Theory

### 2.1 The Geometric Brownian Motion Model

Assume our market consists of two tradable assets: one riskless bond $B$ and one risky asset (stock) $S$.

The behavior of the bond $B$ is modelled by the ODE

$$
\left\{\begin{array}{l}
d B_{t}=r B_{t}, \\
B_{0}=1,
\end{array}\right.
$$

that is,

$$
B_{t}=e^{r t}
$$

where $r>0$ is a constant which is interpreted as the instantaneous interest rate.

Assume the stock $S$ satisfies the stochastic differential equation (SDE)

$$
\begin{equation*}
d S_{t}=S_{t}\left(\mu d t+\sigma d W_{t}\right) \tag{16}
\end{equation*}
$$

Solving the SDE (16) via Ito's Lemma (see the Appendix) we see that the stock price $S_{t}$ follows the geometric Brownian motion (GBM)

$$
\begin{equation*}
S_{t}=S_{0} e^{\left(\mu-\frac{1}{2} \sigma^{2}\right) t+\sigma W_{t}} \tag{17}
\end{equation*}
$$

Assume the market is within the Black-Scholes environment: arbitrage-free, no transaction costs, same rate $r$ for lending and borrowing, no costs for storage.

Suppose now we have a European contingent claim $h_{T}$ based on the stock $S$ with expiration date $T$ and strike price of $K$. (For example, $h_{T}=$ $\left(S_{T}-K\right)^{+}$if the claim is a European call option, and $h_{T}=\left(K-S_{T}\right)^{+}$if it is a European put option.) Such a claim $h_{T}$ is a nonnegative random variable which is $\mathcal{F}_{T}$-measurable, where $\left\{\mathcal{F}_{t}\right\}_{t \geq 0}$ is the natural filtration generated by the standard $\mathbb{P}$-Brownian motion $\left\{W_{t}\right\}_{t \geq 0}$. It is assumed throughout that $h_{T}$ is square-integrable.

The GBS model (17) is the limit of the discrete binomial model discussed in the previous section. To see this, we fix a time $t \in(0, T)$ and equally divide the interval $[0, t]$ into $N$ subintervals: $[(i-1) \delta, i \delta](i=1,2, \cdots, N)$, where $\delta=\delta_{t}=t / N$.

Let us discretize our stock $S_{t}$. Suppose that if the current price of the stock is $s$, then the new price for the next time is given by $u s$ or $d s$, where

$$
u=e^{\mu \delta+\sigma \sqrt{\delta}}, \quad d=e^{\mu \delta-\sigma \sqrt{\delta}},
$$

where $\mu>0$ and $\sigma>0$ are two constants. ( $\mu$ is called the drift and $\sigma$ the volatility of the stock.)

Let $X_{N}$ denote the number of heads in the first $N$ tosses of a coin (then the number of tails is $N-X_{N}$ ) so that the stock price is

$$
S_{t}=s u^{X_{N}} d^{N-X_{N}}=S_{0} e^{(\mu \delta+\sigma \sqrt{\delta}) X_{N}} e^{(\mu \delta-\sigma \sqrt{\delta})\left(N-X_{N}\right)} .
$$

Manipulations give

$$
\begin{equation*}
S_{t}=S_{0} \exp \left(\mu t+\sigma \sqrt{t}\left(\frac{2 X_{N}-N}{\sqrt{N}}\right)\right) \tag{18}
\end{equation*}
$$

The Central Limit Theorem. Let $\left\{\xi_{n}\right\}$ be a sequence of independent identically distributed random variables under the probability measure $\mathbb{P}$ with finite mean $\nu$ and finite variance $\sigma^{2}>0$. Let $Y_{n}=\sum_{i=1}^{n} \xi_{i}$. Then

$$
\frac{Y_{n}-n \nu}{\sqrt{n} \sigma} \longrightarrow Y \sim N(0,1)
$$

where the convergence is in distribution.
Now our $X_{N}$ is the sum of $N$ independent random variables $\sum_{i=1}^{n} \xi_{i}$, where

$$
\begin{gathered}
\xi_{i}= \begin{cases}1, & \text { if the toss is an } H, \\
0, & \text { if the toss is a } T,\end{cases} \\
\mathbb{P}\left\{\xi_{i}=1\right\}=\mathbb{P}\left\{\xi_{i}=0\right\}=\frac{1}{2} .
\end{gathered}
$$

Noticing

$$
\mathbb{E}\left[\xi_{i}\right]=\frac{1}{2}, \quad \operatorname{var}\left[\xi_{i}\right]=\frac{1}{4}
$$

we have by the central Limit Theorem (mean $\nu=\frac{1}{2}$ and variance $\sigma^{2}=\frac{1}{4}$ )

$$
\begin{equation*}
\frac{2 X_{N}-N}{\sqrt{N}} \longrightarrow Z \sim N(0,1) \quad \text { as } N \rightarrow \infty \tag{19}
\end{equation*}
$$

in distribution. This means that, in the limit, $\log \left(S_{t} / S_{0}\right)$ is normally distributed with mean $\mu t$ and variance $\sigma^{2} t$.

The risk-neutral probabilities $\tilde{p}$ and $\tilde{q}$ are given by

$$
\begin{aligned}
\tilde{p} & =\frac{1+r-d}{u-d} \\
& =\frac{1+r-e^{\mu \delta-\sigma \sqrt{\delta}}}{e^{\mu \delta+\sigma \sqrt{\delta}}-e^{\mu \delta-\sigma \sqrt{\delta}}} \\
& \approx \frac{1}{2}\left(1-\sqrt{\delta}\left(\frac{\mu+\frac{1}{2} \sigma^{2}-r}{\sigma}\right)\right)
\end{aligned}
$$

and

$$
\tilde{q} \approx \frac{1}{2}\left(1+\sqrt{\delta}\left(\frac{\mu+\frac{1}{2} \sigma^{2}-r}{\sigma}\right)\right) .
$$

Under the probability measure $\tilde{\mathbb{P}}$ determined by the probabilities $\tilde{p}, \tilde{q}$, we have $X_{N}$ is still binomially distributed, and $\tilde{\mathbb{E}}\left[X_{N}\right]=N \tilde{p}$ and $\operatorname{var}\left(X_{N}\right)=$ $N \tilde{p} \tilde{q}$. Thus,

$$
\begin{aligned}
\tilde{\mathbb{E}}\left[\frac{2 X_{N}-N}{\sqrt{N}}\right] & =\frac{2 N \tilde{p}-N}{\sqrt{N}} \\
& =\sqrt{N}(2 \tilde{p}-1) \\
& \approx-\sqrt{t} \cdot \frac{\mu+\frac{1}{2} \sigma^{2}-r}{\sigma}
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{var}\left[\frac{2 X_{N}-N}{\sqrt{N}}\right] & =\frac{4}{N} \operatorname{var}\left[X_{N}\right] \\
& =4 \tilde{p} \tilde{q} \\
& \approx 1-\delta\left(\frac{\mu+\frac{1}{2} \sigma^{2}-r}{\sigma}\right)^{2} \rightarrow 1
\end{aligned}
$$

So letting $N \rightarrow \infty$ (or $\delta \rightarrow 0$ ) in (19) and using the Central Limit Theorem again, we obtain that $\log \left(S_{t} / S_{0}\right)$ is normally distributed with mean $-\sqrt{t}$. $\frac{\mu+\frac{1}{2} \sigma^{2}-r}{\sigma}$ and variance $\sigma^{2} t$; thus we can write $S_{t}$ as

$$
\begin{equation*}
S_{t}=S_{0} \exp \left(\sigma \sqrt{t} Z+\left(r-\frac{1}{2} \sigma^{2}\right) t\right) \tag{20}
\end{equation*}
$$

where $Z \sim N(0,1)$ under $\tilde{\mathbb{P}}$; i.e., $Z$ is normally distributed with mean 0 and variance 1 .

### 2.2 Self-Financing Trading Strategy

We now introduce trading strategies and their value processes.
Definition 7. A trading strategy is a process $\phi=\left(\phi_{t}^{0}, \phi_{t}^{1}\right)$ which is adapted to the natural filtration $\left\{\mathcal{F}_{t}\right\}$ and where $\phi_{t}^{0}$ and $\phi_{t}^{1}$ are the quantities of the bond and of the stock held in the portfolio at time $t$, respectively. The value process of the portfolio is defined as

$$
V_{t}(\phi)=\phi_{t}^{0} B_{t}+\phi_{t}^{1} S_{t} .
$$

A trading strategy $\phi_{t}:=\left(\phi_{t}^{0}, \phi_{t}^{1}\right)$ is said to be self-financing if

1. $\int_{0}^{T}\left|\phi_{t}^{0}\right| d t+\int_{0}^{T}\left(\phi_{t}^{1}\right)^{2} d t<\infty$;
2. $\phi_{t}^{0} B_{t}+\phi_{t}^{1} S_{t}=\phi_{0}^{0} B_{0}+\phi_{0} S_{0}+\int_{0}^{t} \phi_{u}^{0} d B_{u}+\int_{0}^{t} \phi_{u}^{1} d S_{u}$ a.s., or in differential form,

$$
d V_{t}(\phi)=\phi_{t}^{0} d B_{t}+\phi_{t}^{1} d S_{t} .
$$

Consider the discounted price process

$$
\tilde{S}_{t}=e^{-r t} S_{t}
$$

The discounted value process of a trading strategy $\phi$ is

$$
\tilde{V}_{t}(\phi)=e^{-r t} V_{t}(\phi)=\phi_{t}^{0}+\phi_{t}^{1} \tilde{S}_{t} .
$$

In terms of $\tilde{V}_{t}(\phi)$, a self-financing trading strategy can be characterized as follows.

Proposition 8. A trading strategy $\phi_{t}:=\left(\phi_{t}^{0}, \phi_{t}^{1}\right)$ such that $\int_{0}^{T}\left|\phi_{t}^{0}\right| d t+$ $\int_{0}^{T}\left(\phi_{t}^{1}\right)^{2} d t<\infty$ is self-financing if and only if, for all $t \in[0, T]$,

$$
\tilde{V}_{t}(\phi)=V_{0}(\phi)+\int_{0}^{t} \phi_{u}^{1} d \tilde{S}_{u}, \text { a.s. }
$$

### 2.3 Pricing and Hedging

We now show that the process of discounted prices, $\left\{\tilde{S}_{t}\right\}$, is a martingale under some probability measure $\tilde{\mathbb{P}}$ equivalent to $\mathbb{P}$. Indeed, using the integration by parts formula we get (note that $B_{t}=e^{r t}$ and $d\left\langle e^{-r t}, S_{t}\right\rangle=0$ )

$$
\begin{aligned}
d \tilde{S}_{t} & =-r e^{-r t} S_{t} d t+e^{-r t} d S_{t} \\
& =\tilde{S}_{t}\left((\mu-r) d t+\sigma d W_{t}\right)
\end{aligned}
$$

Put

$$
\begin{gather*}
\tilde{W}_{t}=W_{t}+\frac{(\mu-r) t}{\sigma} \\
d \tilde{S}_{t}=\tilde{S}_{t} \sigma d \tilde{W}_{t} \tag{21}
\end{gather*}
$$

By Girsanov's theorem, there is a probability measure $\tilde{\mathbb{P}}$ equivalent to $\mathbb{P}$ under which $\left\{\tilde{S}_{t}\right\}$ is a martingale. Solving SDE (21) gets

$$
\begin{equation*}
\tilde{S}_{t}=S_{0} \exp \left(\sigma \tilde{W}_{t}-\sigma^{2} t / 2\right) \tag{22}
\end{equation*}
$$

Note that under $\tilde{\mathbb{P}}$, the drift parameter $\mu$ does not appear.
Definition 9. A European claim $h_{T}$ is said to be attainable if there is a self-financing strategy $\phi$ that replicates $h_{T}$; that is,

$$
h_{T}=V_{T}(\phi) .
$$

(We always assume that $h_{T}$ is square-integrable.)
Theorem 10. In the Black-Scholes model, any European claim $h_{T}$ is attainable, under $\tilde{\mathbb{P}}$, and the time $t$ value of the replicating strategy is

$$
V_{t}=\tilde{\mathbb{E}}\left[e^{-r(T-t)} h_{T} \mid \mathcal{F}_{t}\right]
$$

Proof. Let $\phi=\left(\phi_{t}^{0}, \phi_{t}^{1}\right)$ be a self-financing strategy. The value of $\phi$ is

$$
V_{n}(\phi)=\phi_{t}^{0} e^{r t}+\phi_{t}^{1} S_{t}
$$

If $\phi$ replicates $h_{T}$, then $V_{N}(\phi)=h_{T}$. The discounted process

$$
\tilde{V}_{t}(\phi)=\phi_{t}^{0}+\phi_{t}^{1} \tilde{S}_{t}
$$

satisfies

$$
\tilde{V}_{t}(\phi)=V_{0}+\int_{0}^{t} \phi_{u}^{1} \sigma \tilde{S}_{u} d \tilde{W}_{u}
$$

or

$$
d \tilde{V}_{t}(\phi)=\phi_{t}^{1} \sigma \tilde{S}_{t} d \tilde{W}_{t}
$$

Hence, $\left\{\tilde{V}_{t}(\phi)\right\}$ is a martingale under $\tilde{\mathbb{P}}$. It follows that $\left(\tilde{V}_{t}=\tilde{V}_{t}(\phi)\right)$

$$
\tilde{V}_{t}=\tilde{\mathbb{E}}\left[\tilde{V}_{T} \mid \mathcal{F}_{t}\right]
$$

Since $V_{T}=h_{T}$, we obtain

$$
\tilde{V}_{t}=\tilde{\mathbb{E}}\left[h_{T} \mid \mathcal{F}_{t}\right]
$$

Equivalently,

$$
\begin{equation*}
V_{t}=\tilde{\mathbb{E}}\left[e^{-r(T-t)} h_{T} \mid \mathcal{F}_{t}\right] \tag{23}
\end{equation*}
$$

Next we prove the existence of such a strategy $\phi$. To this end, consider the martingale defined by

$$
M_{t}=\tilde{\mathbb{E}}\left[e^{-r T} h_{T} \mid \mathcal{F}_{t}\right]
$$

By the Brownian Martingale Representation Theorem, we have an adapted process $\left(\psi_{t}\right)$ such that $\tilde{\mathbb{E}}\left[\int_{0}^{T} \psi_{t}^{2} d t\right]<\infty$ and

$$
M_{t}=M_{0}+\int_{0}^{t} \psi_{u} d \tilde{W}_{u}
$$

Comparing to (23), we put

$$
\phi_{t}^{1}=\psi_{t} /\left(\sigma \tilde{S}_{t}\right) \text { and } \phi_{t}^{0}=M_{t}-\phi_{t}^{1} \tilde{S}_{t}
$$

Then $\phi=\left(\phi_{t}^{0}, \phi_{t}^{1}\right)$ is self-financing, and its time $t$ value is

$$
\begin{aligned}
V_{t}(\phi) & =\phi_{t}^{0} e^{r t}+\phi_{t}^{1} S_{t} \\
& =\left(M_{t}-\phi_{t}^{1} \tilde{S}_{t}\right) e^{r t}+\phi_{t}^{1} \tilde{S}_{t} e^{r t} \\
& =e^{r t} M_{t} \\
& =\tilde{\mathbb{E}}\left[e^{-r(T-t)} h_{T} \mid \mathcal{F}_{t}\right]
\end{aligned}
$$

as required.

When our contingent claim $h_{T}$ depends only on the stock price at the expiration time $T, h_{T}=h\left(S_{T}\right)$, we can express the time $t$ value of the claim as a function of $t$ and $S_{t}$. Indeed, since (under $\tilde{\mathbb{P}}$ )

$$
S_{T}=S_{t} e^{\left(r-\frac{1}{2} \sigma^{2}\right)(T-t)+\sigma\left(\tilde{W}_{T}-\tilde{W}_{t}\right)}
$$

we have

$$
\begin{aligned}
V_{t} & =\tilde{\mathbb{E}}\left[e^{-r(T-t)} f\left(S_{T}\right) \mid \mathcal{F}_{t}\right] \\
& =\tilde{\mathbb{E}}\left[\left.e^{-r(T-t)} f\left(S_{t} e^{\left(r-\frac{1}{2} \sigma^{2}\right)(T-t)+\sigma\left(\tilde{W}_{T}-\tilde{W}_{t}\right)}\right) \right\rvert\, \mathcal{F}_{t}\right] \\
& =\tilde{\mathbb{E}}\left[e^{-r(T-t)} f\left(S_{t} e^{\left(r-\frac{1}{2} \sigma^{2}\right)(T-t)+\sigma\left(\tilde{W}_{T}-\tilde{W}_{t}\right)}\right)\right]
\end{aligned}
$$

the last equality being due to the fact that $S_{t}$ is $\mathcal{F}_{t}$-measurable and $\tilde{W}_{T}-\tilde{W}_{t}$ is independent of $\mathcal{F}_{t}$. Let

$$
F(t, x)=\tilde{\mathbb{E}}\left[e^{-r(T-t)} f\left(x e^{\left(r-\frac{1}{2} \sigma^{2}\right)(T-t)+\sigma\left(\tilde{W}_{T}-\tilde{W}_{t}\right)}\right)\right]
$$

Then

$$
\begin{equation*}
V_{t}=F\left(t, S_{t}\right) \tag{24}
\end{equation*}
$$

Since $\tilde{W}_{T}-\tilde{W}_{t} \sim N(0, \sqrt{T-t})$, we can write $\tilde{W}_{T}-\tilde{W}_{t}=\sqrt{T-t} Z$, where $Z$ is a standard normal random variable with mean 0 and variance 1 . It follows that

$$
\begin{equation*}
F(t, x)=e^{-r(T-t)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2 \pi}} f\left(x e^{\left(r-\frac{1}{2} \sigma^{2}\right)(T-t)+\sigma \sqrt{T-t} u}\right) e^{-\frac{1}{2} u^{2}} d u \tag{25}
\end{equation*}
$$

In particular, if the claim is a European call option, then $f(x)=(x-K)^{+}$. Hence the function $F(t, x)$ in (25) becomes

$$
\begin{equation*}
F(t, x)=e^{-r(T-t)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2 \pi}}\left(x e^{\left(r-\frac{1}{2} \sigma^{2}\right)(T-t)+\sigma \sqrt{T-t} u}-K\right)^{+} e^{-\frac{1}{2} u^{2}} d u \tag{26}
\end{equation*}
$$

Put

$$
\begin{aligned}
\theta & =T-t \\
d_{1} & =\frac{\log (x / K)+\left(r+\frac{1}{2} \sigma^{2}\right) \theta}{\sigma \sqrt{\theta}} \\
d_{2} & =\frac{\log (x / K)+\left(r-\frac{1}{2} \sigma^{2}\right) \theta}{\sigma \sqrt{T-t}}=d_{1}-\sigma \sqrt{\theta} .
\end{aligned}
$$

Noting that

$$
x e^{\left(r-\frac{1}{2} \sigma^{2}\right) \theta+\sigma \sqrt{\theta} u}-K \geq 0
$$

if and only if

$$
u \geq \frac{\log (K / x)-\left(r-\frac{1}{2} \sigma^{2}\right) \theta}{\sigma \sqrt{\theta}}=-d_{2},
$$

we obtain

$$
\begin{aligned}
F(t, x) & =\int_{-d_{2}}^{\infty}\left(x e^{-\frac{1}{2} \sigma^{2} \theta+\sigma \sqrt{\theta} u}-K e^{-r \theta}\right) \frac{1}{\sqrt{2 \pi}} e^{-\frac{1}{2} u^{2}} d u \\
& =\int_{-\infty}^{d_{2}}\left(x e^{-\frac{1}{2} \sigma^{2} \theta-\sigma \sqrt{\theta} u}-K e^{-r \theta}\right) \frac{1}{\sqrt{2 \pi}} e^{-\frac{1}{2} u^{2}} d u \\
& =x \int_{-\infty}^{d_{2}} e^{-\frac{1}{2} \sigma^{2} \theta-\sigma \sqrt{\theta} u} \frac{1}{\sqrt{2 \pi}} e^{-\frac{1}{2} u^{2}} d u-K e^{-r \theta} \int_{-\infty}^{d_{2}} \frac{1}{\sqrt{2 \pi}} e^{-\frac{1}{2} u^{2}} d u .
\end{aligned}
$$

The first integral equals

$$
\int_{-\infty}^{d_{2}} e^{-\frac{1}{2}(u+\sigma \sqrt{\theta})^{2}} \frac{1}{\sqrt{2 \pi}} d u=\int_{-\infty}^{d_{1}} \frac{1}{\sqrt{2 \pi}} e^{-\frac{1}{2} v^{2}} d v
$$

after using the substitution $v=u+\sigma \sqrt{\theta}$. Therefore, we obtain

$$
\begin{equation*}
F(t, x)=x N\left(d_{1}\right)-K e^{-r(T-t)} N\left(d_{2}\right), \tag{27}
\end{equation*}
$$

where

$$
N(d)=\int_{-\infty}^{d} \frac{1}{\sqrt{2 \pi}} e^{-\frac{1}{2} v^{2}} d v
$$

Similarly, for a European put option, since $f(x)=(K-x)^{+}$, repeating the argument above, we have that the time $t$ value $P\left(t, S_{t}\right)$ of the put is

$$
P\left(t, S_{t}\right)=F\left(t, S_{t}\right)
$$

where

$$
\begin{equation*}
F(t, x)=K e^{-r(T-t)} N\left(-d_{2}\right)-x N\left(-d_{1}\right) . \tag{28}
\end{equation*}
$$

Alternatively, this can be derived from the put-call parity

$$
P_{t}+S_{t}-C_{t}=K e^{-r(T-t)}
$$

where $P_{t}$ and $C_{t}$ are the time $t$ values of the European put and respectively, call options with the same strike $K$ and expiration time $T$.

### 2.4 Explicit Hedging

We can find an explicit hedging in the case where the contingent claim is of the form $h_{T}=h\left(S_{T}\right)$ (path independent). Suppose $\phi$ is a replicating portfolio and denote by $\left\{V_{t}\right\}$ its value process. Thus, $V_{T}=h_{T}=h\left(S_{T}\right)$. As before, denote by $\left\{\tilde{V}_{t}\right\}$ the discounted value process. Since $\left\{\tilde{V}_{t}\right\}$ is a martingale, we must have that for any time $t$,

$$
\tilde{V}_{t}=e^{-r t} F\left(t, S_{t}\right):=\tilde{F}\left(t, \tilde{S}_{t}\right),
$$

where

$$
\tilde{F}(t, x)=e^{-r t} F\left(t, x e^{r t}\right)
$$

Apply Ito's formula to $\tilde{F}(t, x)$ to get

$$
d \tilde{F}\left(t, \tilde{S}_{t}\right)=\tilde{F}_{t}\left(t, \tilde{S}_{t}\right) d t+\tilde{F}_{x}\left(t, \tilde{S}_{t}\right) d \tilde{S}_{t}+\frac{1}{2} \tilde{F}_{x x}\left(t, \tilde{S}_{t}\right) d\langle\tilde{S}, \tilde{S}\rangle_{t}
$$

Under the probability measure $\tilde{\mathbb{P}}, d \tilde{S}_{t}=\sigma \tilde{S}_{t} d \tilde{W}_{t}$, we have

$$
d\langle\tilde{S}, \tilde{S}\rangle_{t}=\sigma^{2} \tilde{S}_{t}^{2} d t
$$

It follows that

$$
d \tilde{V}_{t}=\tilde{F}_{x}\left(t, \tilde{S}_{t}\right) \sigma \tilde{S}_{t} d \tilde{W}_{t}+\left(\tilde{F}_{t}\left(t, \tilde{S}_{t}\right)+\frac{1}{2} \sigma^{2} \tilde{S}_{t}^{2}\right) d t
$$

But, $\left\{\tilde{V}_{t}\right\}$ is a martingale, the drift term must equal 0 ; hence,

$$
d \tilde{V}_{t}=\tilde{F}_{x}\left(t, \tilde{S}_{t}\right) \sigma \tilde{S}_{t} d \tilde{W}_{t}=\tilde{F}_{x}\left(t, \tilde{S}_{t}\right) d \tilde{S}_{t}
$$

So the hedging portfolio $\phi=\left(\phi_{t}^{0}, \phi_{t}^{1}\right)$ is determined by choosing

$$
\phi_{t}^{1}=\tilde{F}_{x}\left(t, \tilde{S}_{t}\right)=F_{x}\left(t, S_{t}\right)
$$

and

$$
\phi_{t}^{0}=\tilde{V}_{t}-\phi_{t}^{1} \tilde{S}_{t} .
$$

It is not hard to check that the portfolio $\phi=\left(\phi_{t}^{0}, \phi_{t}^{1}\right)$ chosen above is selffinancing and its discounted value $\tilde{V}_{t}=\tilde{F}\left(t, \tilde{S}_{t}\right)$.

Remark 11. In the case of a European call, we have

$$
\phi_{t}^{1}=\frac{\partial F}{\partial x}\left(t, S_{t}\right)=N\left(d_{1}\right),
$$

and in the put case,

$$
\phi_{t}^{1}=\frac{\partial F}{\partial x}\left(t, S_{t}\right)=-N\left(-d_{1}\right) .
$$

### 2.5 A PDE approach

Let $V\left(t, S_{t}\right)$ denote the time $t$ value of the claim. In order to find $V\left(t, S_{t}\right)$, we consider a portfolio $\phi=\left(\phi_{t}^{0}, \phi_{t}^{1}\right)$. Let $V_{t}(\phi)$ denote the value of the portfolio at time $t$; that is,

$$
V_{t}(\phi)=\phi_{t}^{0} e^{r t}+\phi_{t}^{1} S_{t} .
$$

We require that out portfolio replicate the claim; i.e., $V_{T}(\phi)=h_{T}=$ $V\left(T, S_{T}\right)$ ). By the arbitrage argument, at each time $t \leq T$ we must have $V_{t}(\phi)=V\left(t, S_{t}\right)$. Hence

$$
V\left(t, S_{t}\right)=\phi_{t}^{0} e^{r t}+\phi_{t}^{1} S_{t} .
$$

Remember that our portfolio is self-financing and out stock $S_{t}$ follows the GBM: $d S_{t}=S_{t}\left(\mu d t+\sigma d W_{t}\right)$. Apply Ito's formula to get

$$
\begin{align*}
& \left(\frac{\partial V}{\partial t}+\mu S_{t} \frac{\partial V}{\partial S}+\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} V}{\partial S^{2}}\right) d t+\sigma S_{t} \frac{\partial V}{\partial S} d W_{t} \\
& =\left(\phi_{t}^{1} \mu S_{t}+\phi_{t}^{0} r e^{r t}\right) d t+a \sigma S_{t} d W_{t} . \tag{29}
\end{align*}
$$

The uncertainty is caused by the randomness of the Brownian motion $W$. Thus to avoid uncertainty we must get rid of the terms in the last equations which are related to $d W_{t}$. Equating the coefficients of $d W_{t}$ of both sides of the last equation gives us

$$
\phi_{t}^{1}=\frac{\partial V}{\partial x}\left(t, S_{t}\right) .
$$

Thus

$$
\phi_{t}^{0}=e^{-r t}\left[V\left(t, S_{t}\right)-\phi_{t}^{1} S_{t}\right] .
$$

Returning to the $\operatorname{SDE}(29)$, we see that $V\left(t, S_{t}\right)$ solves the PDE

$$
\begin{equation*}
\frac{\partial V}{\partial t}+\frac{1}{2} \sigma^{2} S_{t}^{2} \frac{\partial^{2} V}{\partial S^{2}}+r S_{t} \frac{\partial V}{\partial S}-r V=0 \tag{30}
\end{equation*}
$$

This is called the Black-Scholes partial differential equation. $V$ satisfies the boundary (final) condition:

$$
V\left(T, S_{T}\right)=h_{T}
$$

So the boundary-value problem for the Black-Scholes PDE is indeed a backward BVP. The closed form solution is given by (27) (for a call) and respectively, by (28) (for a put).

We can deduce the Black-Scholes equation (30) via the Feynman-Kac theorem. Recall that under $\tilde{\mathbb{P}}$, we have the stock price $S_{t}$ satisfies

$$
d S_{t}=S_{t}\left(r d t+\sigma d \tilde{W}_{t}\right)
$$

and the value function of the European claim is

$$
V(t, x)=\mathbb{E}\left[e^{-r(T-t)} f\left(S_{T}\right) \mid S_{t}=x\right]
$$

Let

$$
G(t, x)=e^{r(T-t)} V(t, x)=\mathbb{E}\left[f\left(S_{T}\right) \mid S_{t}=x\right] .
$$

Then by the Feynman-Kac theorem, $G(t, x)$ satisfies the PDE:

$$
\frac{\partial G}{\partial t}+r x \frac{\partial G}{\partial x}+\frac{1}{2} \sigma^{2} x^{2} \frac{\partial^{2} G}{\partial x^{2}}=0
$$

This is easily seen to be the Black-Scholes equation (30).

### 2.6 American Options

Recall that an American option gives the holder the right, but not the obligation, to exercise his right (buying or selling) at any time up to the expiration time $T$. Thus, an American option can be defined as an adapted nonnegative process $\left\{h_{t}\right\}_{0 \leq t \leq T}$. That is, if the option is exercised at time $t$, the owner receives the amount of $h_{t}$. We consider path-independent American options; that is, we assume that for each $t \in[0, T], h_{t}=h\left(S_{t}\right)$, where $h: \mathbb{R}^{+} \rightarrow \mathbb{R}^{+}$is a continuous function with the property: $h(x) \leq A+B x$ for $x \in \mathbb{R}^{+}$, where $A$ and $B$ are constants. For example, $h(x)=(x-K)^{+}$ for an American call and $h(x)=(K-x)^{+}$for an American put.

Definition 12. A trading strategy with consumption is an adapted process $\phi=\left(\phi_{t}^{0}, \phi_{t}^{1}\right)_{0 \leq t \leq T}$ such that

1. $\int_{0}^{T}\left|\phi_{t}^{0}\right| d t+\int_{0}^{T}\left(\phi_{t}^{1}\right)^{2} d t<\infty$;
2. $\phi_{t}^{0} e^{r t}+\phi_{t}^{1} S_{t}=\phi_{0}^{0}+\phi_{0}^{1} S_{0}+\int_{0}^{t} \phi_{u}^{0} d e^{r u}+\int_{0}^{t} \phi_{u}^{1} d S_{u}-C_{t}$ for all $t \in[0, T]$, where $\left\{C_{t}\right\}_{0 \leq t \leq T}$ is an adapted, continuous, nondecreasing process null at time 0 . ( $C_{t}$ correspondents to the cumulative consumption up to time $t$.)

As before, $V_{t}(\phi)=\phi_{t}^{0} e^{r t}+\phi_{t}^{1} S_{t}$ is the value process of the strategy $\phi$, and $\tilde{V}_{t}(\phi)=\phi_{t}^{0}+\phi_{t}^{1} \tilde{S}_{t}$ is the discounted value process.

Definition 13. A trading strategy with consumption $\phi=\left(\phi_{t}^{0}, \phi_{t}^{1}\right)_{0 \leq t \leq T}$ is said to hedge an American option defined by $\left\{h\left(S_{t}\right)\right\}$ if

$$
V_{t}(\phi) \geq h\left(S_{t}\right), \quad t \in[0, T] .
$$

Denote by $\Phi^{h}$ the collection of trading strategies with consumption which hedge the American option defined by $\left\{h\left(S_{t}\right)\right\}$. Thus if the writer of the option follows a trading strategy $\phi \in \Phi^{h}$, then he does not lose at any time $t$ since the time $t$ value, $V_{t}(\phi)$, is always at least the payoff at time $t, h\left(S_{t}\right)$. The problem is to find the minimal cost required to hedge an American option.
Theorem 14. Define $u:[0, T] \times \mathbb{R}^{+} \rightarrow \mathbb{R}$ by

$$
\begin{equation*}
u(t, x)=\sup _{\tau \in \tau_{t, T}} \tilde{\mathbb{E}}\left[e^{-r(\tau-t)} h\left(x \exp \left(\left(r-\sigma^{2} / 2\right)(\tau-t)+\sigma\left(\tilde{W}_{\tau}-\tilde{W}_{t}\right)\right)\right)\right], \tag{31}
\end{equation*}
$$

where $\tau_{t, T}$ is the set of stopping times $\tau$ taking values in $[t, T]$. Then
(i) $V_{t}(\phi) \geq u\left(t, S_{t}\right)$ for all $\phi \in \Phi^{h}$;
(ii) there exists a trading strategy $\tilde{\phi}$ in $\Phi^{h}$ such that $V_{t}(\tilde{\phi})=u\left(t, S_{t}\right)$ for all $t \in[0, T]$.
The function $u\left(t, S_{t}\right)$ is regarded the value of the American option at time $t$ since it is the minimal value of a trading strategy hedging the option. Note that the time 0 value is given by

$$
\begin{aligned}
u\left(0, S_{0}\right) & =\sup _{\tau \in \tau_{0, T}} \tilde{\mathbb{E}}\left[e^{-r \tau} h\left(S_{0} \exp \left(\left(r-\sigma^{2} / 2\right) \tau+\sigma \tilde{W}_{\tau}\right)\right)\right] \\
& =\sup _{\tau \in \tau_{0}, T} \tilde{\mathbb{E}}\left[e^{-r \tau} h\left(S_{\tau}\right)\right]
\end{aligned}
$$

Theorem 15. Consider an American call option. We have $u(t, x)=F(t, x)$, where $u(t, x)$ is given by (31) with $h(x)=(x-K)^{+}$and $F(t, x)$ is given by (27).

Consider now an American put option which, in general, does not have a closed form solution. Denote by $P(t, x)$ the value function of the put. Since $h(x)=(K-x)^{+}$, we have by (31)

$$
\begin{equation*}
P(t, x)=\sup _{\tau \in \tau, T} \tilde{\mathbb{E}}\left[\left(K e^{-r(\tau-t)}-x \exp \left(-\frac{1}{2} \sigma^{2}(\tau-t)+\sigma\left(\tilde{W}_{\tau}-\tilde{W}_{t}\right)\right)\right)^{+}\right] \tag{32}
\end{equation*}
$$

For American put there is a $S_{f}(t)$-free boundary (not known a priori), with the properties:

- if $S_{t}>S_{t}(f)$, it is optimal to hold on the option;
- if $S_{t} \leq S_{t}(f)$, it is optimal to exercise the option to get the payoff $\left(K-S_{t}\right)^{+}=P\left(t, S_{t}\right)$.

It can be proved that $P(t, x)$ is differentiable and at the boundary $S_{t}(f)$ we have $\frac{\partial P}{\partial t}=-1$.

Theorem 16. There exists a function $S_{f}:[0, T] \rightarrow(0, \infty)$ with the property

1. for $0 \leq x \leq S_{t}(f)$ and $0 \leq t \leq T$,

$$
P(t, x)=K-x \text { and } \frac{\partial P}{\partial t}+\frac{1}{2} \sigma^{2} x^{2} \frac{\partial^{2} P}{\partial x^{2}}+r x \frac{\partial P}{\partial x}-r P<0 ;
$$

2. for $0 \leq t \leq T$ and $S_{f}(t)<x<\infty$,

$$
P(t, x)>(K-x)^{+} \text {and } \frac{\partial P}{\partial t}+\frac{1}{2} \sigma^{2} x^{2} \frac{\partial^{2} P}{\partial x^{2}}+r x \frac{\partial P}{\partial x}-r P=0 .
$$

At the boundary $\left\{(t, x): x=S_{t}(f)\right\}, P(t, x)$ is continuously differentiable in $x$ and continuous in $t$, and

$$
\begin{gathered}
P\left(t, S_{t}(f)\right)=\left(K-S_{t}(f)\right)^{+}, \\
\frac{\partial P}{\partial x}\left(t, S_{t}(f)\right)=-1 .
\end{gathered}
$$

In addition, $P$ satisfies the final condition

$$
P(T, x)=\left(K-S_{T}\right)^{+} .
$$

Let us introduce the Black-Scholes differential operator

$$
\mathcal{L}_{B S}=\frac{\partial}{\partial t}+\frac{1}{2} \sigma^{2} x^{2} \frac{\partial^{2}}{\partial x^{2}}+r x \frac{\partial}{\partial x}-r .
$$

We can restate the above partial differential inequality formulation of American put as a linear complementarity problem:

$$
\mathcal{L}_{B S} P(t, x) \cdot\left(P(t, x)-(K-x)^{+}\right)=0
$$

subject to

$$
\begin{aligned}
& \mathcal{L}_{B S} P(t, x) \leq 0, \quad P(t, x)-(K-x)^{+} \geq 0, \quad P(T, x)=(K-x)^{+}, \\
& P(t, x) \rightarrow \infty \text { as } x \rightarrow \infty, \quad \text { and } P(t, x) \text { and } \frac{\partial P}{\partial x}(t, x) \text { are continuous. }
\end{aligned}
$$

### 2.7 American Perpetual Put Options

An American perpetual put option is an American put option such that the holder can exercise at any time (i.e., the time to expiration is $\infty$ ). Thus, the value function $P(t, x)$ is time-independent and $\frac{\partial P}{\partial t}=0$. Also, the exercise boundary is of the form $S_{t}(f)=\alpha$, where $\alpha$ is a constant to be determined. Therefore, the Black-Scholes equation is reduced to

$$
\begin{equation*}
\frac{1}{2} \sigma^{2} x^{2} \frac{d^{2} P}{d x^{2}}+r x \frac{d P}{d x}-r P=0, \quad x \in(\alpha, \infty) \tag{33}
\end{equation*}
$$

The general solution to Eq. (33) is

$$
P(x)=c_{1} x^{\beta}+c_{2} x^{\gamma}
$$

where $c_{1}, c_{2}, \alpha, \beta$ are constants. Substituting into (33) to see that $\beta$ and $\gamma$ are the roots of the quadratic equation

$$
\frac{1}{2} \sigma^{2} u(u-1)+r u-r=0
$$

These are 1 and $-2 r / \sigma^{2}$. It thus follows that

$$
\begin{equation*}
P(x)=c_{1} x+c_{2} x^{-\frac{2 r}{\sigma^{2}}} \tag{34}
\end{equation*}
$$

The boundary conditions are now

$$
P(\alpha)=K-\alpha, \quad \lim _{x \downarrow \alpha} \frac{d P}{d x}=-1, \quad \lim _{x \rightarrow \infty} P(x)=0 .
$$

The last condition immediately implies that $c_{1}=0$; hence

$$
P(x)=c_{2} x^{-\frac{2 r}{\sigma^{2}}} .
$$

The first two conditions then imply

$$
\begin{aligned}
c_{2} \frac{2 r}{\sigma^{2}} \alpha^{-2 r / \sigma^{2}} & =K-\alpha, \\
c_{2} \alpha^{-2 r / \sigma^{2}} & =\alpha .
\end{aligned}
$$

Solving these equations yields

$$
\alpha=\frac{2 r K}{2 r+\sigma^{2}} \text { and } c_{2}=(K-\alpha) \alpha^{2 r / \sigma^{2}}
$$

Hence

$$
P(x)= \begin{cases}(K-\alpha)\left(\frac{\alpha}{x}\right)^{2 r / \sigma^{2}}, & x \in(\alpha, \infty), \\ K-x, & x \in[0, \alpha] .\end{cases}
$$

The above argument can be used to price perpetual American lookback options. A perpetual American lookback put (call) option gives the holder the right to sell (buy) one share of the underlying asset (stock) at the maximum (minimum) realized price at any time during the lifetime of the option (assuming no dividends paid). Lookback options are path-dependent. Let $M=M_{t}=\max \left\{S_{i}: 0 \leq i \leq t\right\}$ be the maximum stock price realized up to time $t$.

The value function $P$ of the perpetual American lookback put option solves the partial differential equation

$$
\begin{equation*}
\frac{1}{2} \sigma^{2} x^{2} \frac{d^{2} P}{d x^{2}}+r x \frac{d P}{d x}-r P=0, \quad \alpha<x<M \tag{35}
\end{equation*}
$$

subject to the boundary value conditions

$$
\left\{\begin{array}{l}
P(\alpha)=M-\alpha  \tag{36}\\
\lim _{x \downarrow \alpha} \frac{d P}{d x}=-1, \\
\frac{d P}{d M}=0 \text { at } x=M .
\end{array}\right.
$$

The general solution $P$ to Eq. (35) is given by (34). The boundary conditions in (36) imply that $c_{1}>0$ is uniquely determined by the equation

$$
\begin{aligned}
c_{1}\left(1+c_{1}\right)^{\mu} & =\left(\frac{\mu}{1+\mu}\right)^{1+\mu}, \text { and } \\
c_{2} & =\frac{M^{1+\mu}}{\mu} c_{1} \\
\alpha & =M\left(\frac{c_{1}}{1+c_{1}}\right)^{1 /(1+\mu)}
\end{aligned}
$$

where $\mu=2 r / \sigma^{2}$.

## 3 Stochastic Volatility

In the Black-Scholes model, it is assumed that the interest rate $r$ and the volatility $\sigma$ are constant. But in practice, these assumptions are never
satisfied. Empirical studies show that the volatility $\sigma$ changes randomly over time. So the Black-Scholes model can be modified to the following model with stochastic volatility:

$$
\begin{equation*}
d S_{t}=S_{t}\left(\mu d t+\sigma_{t} d W_{t}\right) \tag{37}
\end{equation*}
$$

where $\left\{\sigma_{t}\right\}_{t \geq 0}$ is the volatility process which is positive and which is not necessarily perfectly correlated with the Brownian motion $\left\{W_{t}\right\}_{t \geq 0}$ driving the stock price process $\left\{S_{t}\right\}$ in (17). (For more details of this section, the reader is referred to [4]; see also [1] and [2].)

### 3.1 Mean-Reverting Models for Stochastic Volatility

The stochastic volatility process $\left\{\sigma_{t}\right\}_{t \geq 0}$ is typically taken to be meanreverting which refers to the typical time that a process needs to take to get back to the mean level of its invariant distribution (the long-run distribution of the process). Mathematically, the stochastic volatility process $\left\{\sigma_{t}\right\}_{t \geq 0}$ is given as

$$
\begin{equation*}
\sigma_{t}=f\left(X_{t}\right) \tag{38}
\end{equation*}
$$

where the process $\left\{X_{t}\right\}$ is a process of the form

$$
\begin{equation*}
X_{t}=\alpha\left(m-X_{t}\right) d t+g\left(t, X_{t}\right) d \hat{Z}_{t} \tag{39}
\end{equation*}
$$

where $\left\{\hat{Z}_{t}\right\}_{t \geq 0}$ is a Brownian motion correlated with $\left\{W_{t}\right\}_{t \geq 0}, \alpha$ is the rate of mean reversion, $m$ is the long-run mean level of $X$, and $g$ is a function.

The Brownian motion $\left\{\hat{Z}_{t}\right\}_{t \geq 0}$ is rewritten in the form

$$
\begin{equation*}
\hat{Z}_{t}=\rho W_{t}+\sqrt{1-\rho^{2}} Z_{t} \tag{40}
\end{equation*}
$$

where $\left\{Z_{t}\right\}_{t \geq 0}$ is a standard Brownian motion independent of $\left\{W_{t}\right\}_{t \geq 0}$, and $\rho \in[-1,1]$ is the correlated coefficient and it is often negative.

Recall that the Ornstein-Uhlenbeck (OU) process $\left\{X_{t}\right\}$ is defined by

$$
\begin{equation*}
d X_{t}=\alpha\left(m-X_{t}\right) d t+\beta d \hat{Z}_{t} \tag{41}
\end{equation*}
$$

where $\alpha, \beta$ and $m$ are constant. Solving the SDE gives (assuming $X_{0}=x$ )

$$
\begin{equation*}
X_{t}=m+(x-m) e^{-\alpha t}+\beta \int_{0}^{t} e^{-\alpha(t-s)} d \hat{Z}_{t} \tag{42}
\end{equation*}
$$

Recall also that the Cox-Ingersol-Ross (CIR) process is defined as the process

$$
\begin{equation*}
d X_{t}=\kappa\left(m-X_{t}\right) d t+\gamma \sqrt{X_{t}} d \hat{Z}_{t} \tag{43}
\end{equation*}
$$

where $\kappa, v$ and $m$ are constants. Note that both OU and CIR processes are mean-reverting.

We now list some mean-reverting stochastic volatility models by specifying the function $f$ in (38), the process $\left\{X_{t}\right\}$ in (39), and the correlated coefficient $\rho$ in (40).
(i) The Scott model: $\left\{\begin{aligned} f(x) & =e^{x} \\ d X_{t} & =\alpha\left(m-X_{t}\right) d t+\beta d \hat{Z}_{t} \quad \text { (OU) } \\ \rho & =0 .\end{aligned}\right.$
(ii) The Stein-Stein model: $\left\{\begin{aligned} f(x) & =|x| \\ d X_{t} & =\alpha\left(m-X_{t}\right) d t+\beta d \hat{Z}_{t} \\ \rho & =0 .\end{aligned}\right.$
(iii) The Ball-Roma model: $\left\{\begin{aligned} f(x) & =\sqrt{x} \\ d X_{t} & =\kappa\left(m-X_{t}\right) d t+\gamma \sqrt{X_{t}} d \hat{Z}_{t} \\ \rho & =0 .\end{aligned}\right.$
(iv) The Heston model: $\left\{\begin{aligned} f(x) & =\sqrt{x} \\ d X_{t} & =\kappa\left(m-X_{t}\right) d t+\gamma \sqrt{X_{t}} d \hat{Z}_{t} \\ \rho & \neq 0 .\end{aligned}\right.$

### 3.2 Option pricing with Mean-Reverting Stochastic Volatility

Assume again our financial market consists of two tradable assets. One is the risk-free bond $B$ and the other is a stock $B$. The dynamics of the bond $B$ are deterministic and follow the ODE

$$
d B_{t}=r B_{t}, \quad B_{0}=1,
$$

that is,

$$
B_{t}=e^{r t}
$$

where $r$ is assumed to be constant. The risky stock $S$ now follows a geometric Brownian motion with a mean-reverting stochastic volatility:

$$
\left\{\begin{align*}
d S_{t} & =\mu S_{t} d t+\sigma_{t} S_{t} d W_{t}  \tag{44}\\
\sigma_{t} & =f\left(X_{t}\right) \\
d X_{t} & =\alpha\left(m-X_{t}\right) d t+\beta d \hat{Z}_{t}
\end{align*}\right.
$$

where, as before, $\alpha, \beta$ and $m$ are constant, and $\hat{Z}_{t}$ is given in (40).

Suppose we have a European option (contingent claim) $h$ on $\left\{S_{t}\right\}$ which expires at time $T$. Note that $h$ is $\mathcal{F}_{T}$-measurable and square-integrable. We ask the question how to price and hedge the claim.

Since we now have two sources of randomness, the option's value function $P$ depends on two state variables. That is, $P=P(t, s, x)$ which satisfies a partial differential equation in two space variable $s$ and $x$ (see below). It is therefore insufficient to use the underlying asset only to hedge the option. Indeed, if the $d W_{t}$ term can be balanced, the $d Z_{t}$ cannot. The idea to overcome this difficulty is to use another option with different expiration time to hedge.

Consider $P^{(1)}(t, x, y)$ to be the price function of a European option with expiration time $T_{1}$ and with payoff $h\left(S_{T_{1}}\right)$. We are looking for a portfolio $\left(a_{t}, b_{t}, c_{t}\right)_{t \geq 0}$ satisfying

$$
\begin{equation*}
P^{(1)}\left(T_{1}, S_{T_{1}}, X_{T_{1}}\right)=a_{T_{1}} S_{T_{1}}+b_{T_{1}} e^{r T_{1}}+c_{T_{1}} P^{(2)}\left(T_{1}, S_{T_{1}}, X_{T_{1}}\right), \tag{45}
\end{equation*}
$$

where $P^{(2)}(t, x, y)$ is the price function of another European option with the same payoff function $h$ as $P^{(1)}$, but having a longer expiration time $T_{2}$ (i.e., $T_{2}>T_{1}$ ).

If we consider a portfolio consisting of the stock, bond and option (with expiration time $T_{2}$ ), then the right side of (45) shows that the time $T_{1}$ value of the trading strategy $\left(a_{t}, b_{t}, c_{t}\right)$ is the payoff $P^{(1)}$ of the option with expiration time $T_{1}$. As the strategy is self-financing, we have

$$
\begin{equation*}
d P^{(1)}\left(t, S_{t}, X_{t}\right)=a_{t} d S_{t}+b_{t} d e^{r t}+c_{t} d P^{(2)}\left(t, S_{t}, X_{t}\right) \tag{46}
\end{equation*}
$$

If such a portfolio exists, then by the arbitrage argument, we must have

$$
\begin{equation*}
P^{(1)}\left(t, S_{t}, X_{t}\right)=a_{t} S_{t}+b_{t} e^{r t}+c_{t} P^{(2)}\left(t, S_{t}, X_{t}\right), \quad t<T_{1} . \tag{47}
\end{equation*}
$$

We need Ito's formula in the two dimensional space stated below:

$$
\begin{align*}
d g\left(t, S_{t}, X_{t}\right)= & \frac{\partial g}{\partial t} d t+\frac{\partial g}{\partial s} d S_{t}+\frac{\partial g}{\partial x} d X_{t} \\
& +\frac{1}{2}\left(\frac{\partial^{2} g}{\partial s^{2}} d\langle S\rangle_{t}+2 \frac{\partial^{2} g}{\partial s \partial x} d\langle S, X\rangle_{t}+\frac{\partial^{2} g}{\partial x^{2}} d\langle X\rangle_{t}\right) \tag{48}
\end{align*}
$$

Note that we have

$$
d\langle S\rangle_{t}=\sigma_{t}^{2} S_{t}^{2} d t, \quad d\langle S, X\rangle_{t}=\rho \beta \sigma_{t} S_{t} d t, \quad \text { and } \quad d\langle X\rangle_{t}=\beta^{2} d t
$$

Apply Ito's formula (48) to $P^{(1)}$ to get

$$
\begin{aligned}
d P^{(1)}\left(t, S_{t}, X_{t}\right)= & \frac{\partial P^{(1)}}{\partial t} d t+\frac{\partial P^{(1)}}{\partial s} d S_{t}+\frac{\partial P^{(1)}}{\partial x} d X_{t} \\
& +\frac{1}{2} \frac{\partial^{2} P^{(1)}}{\partial s^{2}} d\langle S\rangle_{t}+\frac{\partial^{2} P^{(1)}}{\partial s \partial x} d\langle S, X\rangle_{t}+\frac{1}{2} \frac{\partial^{2} P^{(1)}}{\partial x^{2}} d\langle X\rangle_{t}
\end{aligned}
$$

where all the partial derivatives are evaluated at $\left(t, S_{t}, X_{t}\right)$ (the same for the partial derivatives evaluated below).

If we introduce the differential operator $\mathcal{M}_{1}$ by

$$
\mathcal{M}_{1}=\frac{1}{2} f(x)^{2} s^{2} \frac{\partial^{2}}{\partial s^{2}}+\rho \beta s f(x) \frac{\partial^{2}}{\partial s \partial x}+\frac{1}{2} \beta^{2} \frac{\partial^{2}}{\partial x^{2}},
$$

then we can write

$$
\begin{equation*}
d P^{(1)}\left(t, S_{t}, X_{t}\right)=\left(\frac{\partial P^{(1)}}{\partial t}+\mathcal{M}_{1} P^{(1)}\right) d t+\frac{\partial P^{(1)}}{\partial s} d S_{t}+\frac{\partial P^{(1)}}{\partial x} d X_{t} . \tag{49}
\end{equation*}
$$

Replacing $P^{(1)}$ by $P^{(2)}$ obtains

$$
\begin{equation*}
d P^{(2)}\left(t, S_{t}, X_{t}\right)=\left(\frac{\partial P^{(2)}}{\partial t}+\mathcal{M}_{1} P^{(2)}\right) d t+\frac{\partial P^{(2)}}{\partial s} d S_{t}+\frac{\partial P^{(2)}}{\partial x} d X_{t} \tag{50}
\end{equation*}
$$

Substitute (49) and (50) into (46) and rearrange the terms to get

$$
\begin{align*}
& \left(\frac{\partial P^{(1)}}{\partial t}+\mathcal{M}_{1} P^{(1)}\right) d t+\frac{\partial P^{(1)}}{\partial s} d S_{t}+\frac{\partial P^{(1)}}{\partial x} d X_{t} \\
& =\left(a_{t}+c_{t} \frac{\partial P^{(2)}}{\partial s}\right) d S_{t}+c_{t} \frac{\partial P^{(2)}}{\partial x} d X_{t} \\
& \quad+\left[c_{t}\left(\frac{\partial}{\partial t}+\mathcal{M}_{1}\right) P^{(2)}+b_{t} r e^{r t}\right] d t \tag{51}
\end{align*}
$$

Equating the $d X_{t}$ terms on both sides of (51), we get

$$
\begin{equation*}
c_{t}=\frac{\partial P^{(1)} / \partial x}{\partial P^{(2)} / \partial x} . \tag{52}
\end{equation*}
$$

Matching the coefficients of the $d S_{t}$ terms in (51), we get

$$
\begin{equation*}
a_{t}=\frac{\partial P^{(1)}}{\partial s}-c_{t} \frac{\partial P^{(2)}}{\partial s} \tag{53}
\end{equation*}
$$

Finally the $d t$ terms in (51) yields

$$
\frac{\partial P^{(1)}}{\partial t}+\mathcal{M}_{1} P^{(1)}=c_{t}\left(\frac{\partial}{\partial t}+\mathcal{M}_{1}\right) P^{(2)}+b_{t} r e^{r t}
$$

Substituting (52) and (53) into the last equation we get

$$
\begin{aligned}
& \left(\frac{\partial P^{(1)}}{\partial x}\right)^{-1}\left(\left(\frac{\partial}{\partial t}+\mathcal{M}_{1}\right) P^{(1)}+r\left(S_{t} \frac{\partial P^{(1)}}{\partial s}-P^{(1)}\right)\right) \\
& =\left(\frac{\partial P^{(2)}}{\partial x}\right)^{-1}\left(\left(\frac{\partial}{\partial t}+\mathcal{M}_{1}\right) P^{(2)}+r\left(S_{t} \frac{\partial P^{(2)}}{\partial s}-P^{(2)}\right)\right) .
\end{aligned}
$$

If we introduce the differential operator $\mathcal{M}_{2}$ by

$$
\mathcal{M}_{2}=\frac{\partial}{\partial t}+\mathcal{M}_{1}+r\left(s \frac{\partial}{\partial s}-1\right)
$$

then we can rewrite the last equation as

$$
\begin{equation*}
\left(\frac{\partial P^{(1)}}{\partial x}\right)^{-1} \mathcal{M}_{2} P^{(1)}\left(t, S_{t}, X_{t}\right)=\left(\frac{\partial P^{(2)}}{\partial x}\right)^{-1} \mathcal{M}_{2} P^{(2)}\left(t, S_{t}, X_{t}\right) \tag{54}
\end{equation*}
$$

Now observe that the left side of (54) depends on $T_{1}$ only, while the right side of (54) depends on $T_{2}$ only. So for (54) to hold, we must have that each side of (54) depends upon neither $T_{1}$ nor $T_{2}$. In other words, (54) is independent of the expiration time $T$. Let us denote the function of either side of (54) by

$$
\begin{equation*}
\alpha(x-m)+\beta\left(\frac{\rho(\mu-r)}{f(x)}+\gamma(t, s, x) \sqrt{1-\rho^{2}}\right) \tag{55}
\end{equation*}
$$

where $\gamma(t, s, x)$ is an arbitrary function.
Now the pricing function $P(t, s, x)$ satisfies the partial differential equation

$$
\begin{align*}
& \frac{\partial P}{\partial t}+\frac{1}{2} f(x)^{2} s^{2} \frac{\partial^{2} P}{\partial s^{2}}+\rho \beta s f(x) \frac{\partial^{2} P}{\partial s \partial x}+\frac{1}{2} \beta^{2} \frac{\partial^{2} P}{\partial x^{2}} \\
& +r\left(s \frac{\partial P}{\partial s}-P\right)+(\alpha(m-x)-\beta \Lambda(t, s, x)) \frac{\partial P}{\partial x}=0 \tag{56}
\end{align*}
$$

where

$$
\begin{equation*}
\Lambda(t, s, x)=\frac{\rho(\mu-r)}{f(x)}+\gamma(t, s, x) \sqrt{1-\rho^{2}} . \tag{57}
\end{equation*}
$$

The terminal condition of Eq. (56) is

$$
P(T, s, x)=h(s), x \in \mathbb{R} .
$$

We can regroup the terms in Eq. (56) and get

$$
\begin{aligned}
& \underbrace{\frac{\partial}{\partial t}+\frac{1}{2} f(x)^{2} s^{2} \frac{\partial^{2}}{\partial s^{2}}+r\left(s \frac{\partial}{\partial s}-1\right)}_{\mathcal{L}_{B S}^{f(x)}}+\underbrace{\rho \beta s f(x) \frac{\partial^{2}}{\partial s \partial x}}_{\text {correlation }} \\
& +\underbrace{\frac{1}{2} \beta^{2} \frac{\partial^{2}}{\partial x^{2}}+\alpha(m-x) \frac{\partial}{\partial x}}_{\mathcal{L}_{O U}}-\underbrace{\beta \Lambda \frac{\partial}{\partial x}}_{\text {premium }} .
\end{aligned}
$$

Here $\mathcal{L}_{B S}^{f(x)}$ is $\mathcal{L}_{B S}$ with $\sigma$ replaced with $f(x)$.
The function $\gamma$ in (55) is the risk premium factor from the second source of randomness $\left\{Z_{t}\right\}$ that drives the volatility. In the perfect correlated case (i.e., $|\rho|=1$ ) it does not appear (this is obvious since $\hat{Z}_{t}= \pm W_{t}$ ).

By the two-dimensional Ito formula (48) and the partial differential equation (56) satisfied by $P$, we calculate

$$
\begin{aligned}
d P\left(t, S_{t}, X_{t}\right)= & {\left[\frac{\mu-r}{f(x)}\left(s f(x) \frac{\partial P}{\partial s}+\beta \rho \frac{\partial P}{\partial x}\right)+r P+\gamma \beta \sqrt{1-\rho^{2}} \frac{\partial P}{\partial x}\right] d t } \\
& +\left(s f(x) \frac{\partial P}{\partial s}+\beta \rho \frac{\partial P}{\partial x}\right) d W_{t}+\beta \sqrt{1-\rho^{2}} \frac{\partial P}{\partial x} d Z_{t}
\end{aligned}
$$

This expression says that an infinitesimal fractional increase in the volatility risk $\beta$ increases the infinitesimal rate of return on the option by $\gamma$ times that fraction, in addition to the increase from the excess return-to-risk ratio $(\mu-r) / f(x)$.

### 3.3 Martingale Approach

Recall that if there is an equivalent martingale measure $\tilde{\mathbb{P}}$ under which the discounted stock price $\tilde{S}_{t}=e^{-r t} S_{t}$ is a martingale, then the time $t$ value of the European derivative security with payoff $h_{T}$ at expiration time $T$ is given by

$$
\begin{equation*}
V_{t}=\tilde{\mathbb{E}}\left[e^{-(T-t)} h_{T} \mid \mathcal{F}_{t}\right], \quad 0 \leq t \leq T . \tag{58}
\end{equation*}
$$

In order to find an equivalent martingale measure $\left\{\tilde{W}_{t}\right\}$, we deduce from (44)

$$
\begin{equation*}
d \tilde{S}_{t}=\tilde{S}_{t} f\left(X_{t}\right)\left(\frac{\mu-r}{f\left(X_{t}\right)} d t+d W_{t}\right) \tag{59}
\end{equation*}
$$

So our $\left\{\tilde{W}_{t}\right\}$ satisfies

$$
d \tilde{W}_{t}=\frac{\mu-r}{f\left(X_{t}\right)} d t+d W_{t}
$$

or

$$
\begin{equation*}
\tilde{W}_{t}=\int_{0}^{t} \frac{\mu-r}{f\left(X_{u}\right)} d u+W_{t} \tag{60}
\end{equation*}
$$

By Girsanov's theorem, we know that $\left\{\tilde{W}_{t}\right\}$ will be a standard Brownian motion provided $\left\{f\left(X_{t}\right)\right\}$ satisfies certain regularity condition required in Girsanov's theorem. How about the second Brownian motion $\left\{Z_{t}\right\}$ ? Indeed, we need a two-dimensional version of Girsanov's theorem: If $\left\{\gamma_{t}\right\}$ is an adapted process with sufficient regularity, then we have the process $\left\{\tilde{Z}_{t}\right\}$ defined by

$$
\begin{equation*}
\tilde{Z}_{t}=\int_{0}^{t} \gamma_{u} d u+Z_{t} \tag{61}
\end{equation*}
$$

is also a standard Brownian motion under a martingale measure $\tilde{\mathbb{P}}^{\gamma}$ equivalent to $\mathbb{P}$ and defined by

$$
\frac{d \tilde{\mathbb{P}}^{\gamma}}{d \mathbb{P}}=\exp \left(-\frac{1}{2} \int_{0}^{t}\left[\left(\theta_{u}^{1}\right)^{2}+\left(\theta_{u}^{2}\right)^{2}\right] d u-\int_{0}^{T} \theta_{u}^{1} d W_{u}-\int_{0}^{T} \theta_{u}^{2} d Z_{u}\right),
$$

where

$$
\theta_{t}^{1}=\frac{\mu-r}{f\left(X_{t}\right)} \quad \text { and } \quad \theta_{t}^{2}=\gamma_{t}
$$

Note that the conditions that $f$ be bounded away from 0 and that $\left\{\gamma_{t}\right\}$ be bounded are sufficient for $\tilde{\mathbb{P}}^{\gamma}$ to be an equivalent martingale measure. Under $\tilde{\mathbb{P}}^{\gamma}$, Eq. (44) becomes

$$
\left\{\begin{align*}
d S_{t} & =r S_{t} d t+f\left(X_{t}\right) S_{t} d \tilde{W}_{t}  \tag{62}\\
d X_{t} & =\left[\alpha\left(m-X_{t}\right)-\beta\left(\frac{\rho(\mu-r)}{f\left(X_{t}\right)}+\gamma_{t} \sqrt{1-\rho^{2}}\right)\right] d t+\beta d \hat{\tilde{Z}}_{t} \\
\hat{\tilde{Z}}_{t} & =\rho \tilde{W}_{t}+\sqrt{1-\rho^{2}} \tilde{Z}_{t}
\end{align*}\right.
$$

Thus there exist more than one (infinitely many indeed) equivalent martingale measures $\tilde{\mathbb{P}}^{\gamma}$ which give the no-arbitrage prices of the derivative as follows

$$
\begin{equation*}
V_{t}^{(\gamma)}=\tilde{\mathbb{E}}^{(\gamma)}\left[e^{-(T-t)} h_{T} \mid \mathcal{F}_{t}\right], \quad 0 \leq t \leq T \tag{63}
\end{equation*}
$$

The process $\left\{\gamma_{t}\right\}$ is called the risk premium factor or the market price of volatility risk from the second source of randomness $Z$ that drives the stochastic volatility.

If $\gamma_{t}=\gamma\left(t, S_{t}, X_{t}\right)$ is Markovian, the partial differential equation (56) can be derived from the Feynman-Kac theorem.

## Appendix A: Elements in Stochastic Calculus

## A1. Normal Random Variables

A real random variable $X$ is said to be a standard normal random variable, written $X \sim N(0,1)$, if its distribution function is given by

$$
\mathbb{P}\{X \leq t\}=\int_{-\infty}^{t} \frac{1}{\sqrt{2 \pi}} e^{-\frac{1}{2} v^{2}} d v, \quad t \in \mathbb{R}
$$

If $X$ is a standard normal random variable and if $\mu$ and $\sigma \neq 0$ are two real numbers, then the random variable $Y=\mu+\sigma X$ is a normal random variable with mean $\mu$ and variance $\sigma^{2}$, written $Y \sim N\left(\mu, \sigma^{2}\right)$. That is,

$$
\mathbb{P}\{Y \leq t\}=\int_{-\infty}^{t} \frac{1}{\sqrt{2 \pi} \sigma} \exp \left\{-\frac{(v-\mu)^{2}}{2 \sigma^{2}}\right\} d v, \quad t \in \mathbb{R}
$$

## A2. Conditional Expectations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $X$ be an integrable random variable, and let $\mathcal{G}$ be a sub $\sigma$-algebra of $\mathcal{F}$. Then the conditional expectation of $X$ given $\mathcal{G}$ is defined to be a random variable $\mathbb{E}[X \mid \mathcal{G}]$ such that

- $\mathbb{E}[X \mid \mathcal{G}]$ is $\mathcal{G}$-measurable;
- for any $A \in \mathcal{G}$,

$$
\int_{A} \mathbb{E}[X \mid \mathcal{G}] d \mathbb{P}=\int_{A} X d \mathbb{P}
$$

Proposition A1 (Properties of Conditional Expectations)

1. $\mathbb{E}[a X+b Y \mid \mathcal{G}]=a \mathbb{E}[X \mid \mathcal{G}]+b \mathbb{E}[Y \mid \mathcal{G}]$ (Linearity);
2. $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]]=\mathbb{E}[X]$;
3. $\mathbb{E}[X Y \mid \mathcal{G}]=Y \mathbb{E}[X \mid \mathcal{G}]$ if $Y$ is $\mathcal{G}$-measurable (taking out what is known);
4. $\mathbb{E}[X \mid \mathcal{G}]=\mathbb{E}[X]$ if $X$ is $\mathcal{G}$-measurable (an independent condition drops out);
5. $\mathbb{E}[\mathbb{E}[X|\mathcal{G}| \mathcal{H}]]=\mathbb{E}[X \mid \mathcal{H}]$ if $\mathcal{H} \subset \mathcal{G}$ (tower property);
6. If $X \geq 0$, then $\mathbb{E}[X \mid \mathcal{G}] \geq 0$ (positivity);
7. (Jensen's Inequality) If $\phi: \mathbb{R} \rightarrow \mathbb{R}$ is a convex function and if $X$ is a random variable such that both $X$ and $\phi(X)$ are integrable, then

$$
\phi(\mathbb{E}[X \mid \mathcal{G}]) \leq \mathbb{E}[\phi(X) \mid \mathcal{G}] \quad \text { a.s. }
$$

for any sub $\sigma$-algebra $\mathcal{G}$ of $\mathcal{F}$.

## A3. Stochastic Processes

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

## Definition A2

- A family $\left\{\mathcal{F}_{t}\right\}_{t \geq 0}$ of sub $\sigma$-fields of $\mathcal{F}$ is called a filtration provided

$$
\mathcal{F}_{t} \subset \mathcal{F}_{s} \quad \text { whenever } t \leq s
$$

- A stochastic process $X=\left\{X_{t}\right\}_{t \geq 0}$ is a family of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$.
- A stochastic process $X=\left\{X_{t}\right\}_{t \geq 0}$ is said to be adapted if, for each $t>0, X_{t}$ is $\mathcal{F}_{t}$-measurable.
- $t \mapsto X_{t}(\omega)$, for each fixed $\omega \in \Omega$, is called a (sample) path of $X$.
- A random variable $\tau: \Omega \rightarrow[0, \infty]$ is called a stopping time if, for each $t \geq 0$,

$$
\{\omega: \tau(\omega) \leq t\} \in \mathcal{F}_{t} .
$$

- The $\sigma$-algebra $\mathcal{F}_{\tau}$ associated with a stopping time $\tau$ is defined as

$$
\mathcal{F}_{\tau}=\left\{A \in \mathcal{F}: A \cap\{\tau \leq t\} \in \mathcal{F}_{t} \text { for all } t \geq 0\right\}
$$

Let $\mathbb{E}$ denote the expectation associated with $\mathbb{P}$. Let $\left\{X_{t}\right\}_{t \geq 0}$ be a stochastic process adapted to $\mathcal{F}$ and $\mathbb{E}\left[\left|X_{t}\right|\right]<\infty$ for each $t$. Then $\left\{X_{t}\right\}_{t \geq 0}$ is said to be a

- martingale if $\mathbb{E}\left[X_{t} \mid \mathcal{F}_{s}\right]=X_{s}$ for $t \geq s ;$
- supermartingale if $\mathbb{E}\left[X_{t} \mid \mathcal{F}_{s}\right] \leq X_{s}$ for $t \geq s$;
- submartingale if $\mathbb{E}\left[X_{t} \mid \mathcal{F}_{s}\right] \geq X_{s}$ for $t \geq s$.


## Proposition A3 (Properties of stopping times)

1. If $\tau$ is a stopping time, then $\tau$ is $\mathcal{F}_{\tau}$-measurable.
2. If $\tau$ is a stopping time, finite almost surely, and $\left(X_{t}\right)_{t \geq 0}$ is a continuous adapted process, then $X_{\tau}$ is $\mathcal{F}_{\tau}$-measurable.
3. If $\tau$ and $\eta$ are stopping times such that $\tau \leq \eta$, then $\mathcal{F}_{\tau} \subset \mathcal{F}_{\eta}$.
4. If $\tau$ and $\eta$ are stopping times, then $\tau \wedge \eta=\inf \{\tau, \eta\}$ is a stoping time. In particular, if $\tau$ is a stopping and $s$ is a deterministic time, then $\tau \wedge s$ is a stopping time.

## A4. Brownian Motions

Definition A4 (Brownian motion) A stochastic motion $W=\left(W_{t}\right)_{t \geq 0}$ is said to be a (standard) Brownian motion (or Wiener Process) if

- $W_{0}=0$ a.s.
- Independent normally distributed increments: If $0=t_{0}<t_{1}<t_{2}<$ $\cdots<t_{n}$, then, for $1 \leq j \leq n$, $\left\{W_{t_{j}}-W_{t_{j-1}}\right\}$ is independent; $\mathbb{E}\left[W_{t_{j}}-W_{t_{j-1}}\right]=0 ;$ $\operatorname{var}\left(W_{t_{j}}-W_{t_{j-1}}\right)=t_{j}-t_{j-1}$.

Proposition A5 (Basic Properties of Brownian motion) Let $\left\{\mathcal{F}_{t}\right\}_{t \geq 0}$ be the natural filtration generated by a standard Brownian motion $W=\left(W_{t}\right)_{t \geq 0}$. Then

- $\left\{W_{t}\right\}$ is a martingale.
- $\left\{W_{t}^{2}-t\right\}$ is a martingale. Thus $\mathbb{E}\left[W_{t}^{2}\right]=t$.
- $\exp \left(\sigma W_{t}-\left(\sigma^{2} / 2\right) t\right)$ is a martingale.
- For each $t>0, W$ is of infinite first variation:

$$
\operatorname{Var}(W):=\sup _{\tau}\left\{\sum\left|W_{t_{i}}-W_{t_{i-1}}\right|\right\}=\infty
$$

where the sup is taken over all partitions $\tau$ of $[0, t]$.

- The quadratic variation of $W$ is

$$
\langle W\rangle_{t}=t
$$

where

$$
\langle W\rangle_{t}:=\lim _{n \rightarrow \infty} \sum_{t_{i} \in \tau_{n}, t_{i}<t}\left[W_{t_{i}}-W_{t_{i-1}}\right]^{2}
$$

where $\left\{\tau_{n}\right\}$ is a sequence of partitions of $[0, t]$ such that the mesh $\left|\tau_{n}\right| \rightarrow 0$.

The martingale property can be extended to bounded stopping times.

$$
\mathbb{E}\left[X_{\tau_{2}} \mid \mathcal{F}_{\tau_{1}}\right]=X_{\tau_{1}} \quad \mathbb{P} \text { a.s. }
$$

Similar inequalities hold for super- and sub-martingales.
Definition A7 (Hitting time)

$$
T_{a}=\inf \left\{s \geq 0: W_{s}=a\right\}, \quad a \in \mathbb{R}
$$

(Convention: if there is no such $s$ that $W_{s}=a$, then $T_{a}=\infty$.)

It is known that $T_{a}$ is a stopping time and for $\lambda>0$,

$$
\mathbb{E}\left[\exp \left(-\lambda T_{a}\right)\right]=\exp (-\sqrt{2 \lambda}|a|)
$$

Theorem 18. (Doob's Inequality) If $\left\{X_{t}\right\}_{t \geq 0}$ is a continuous martingale, then

$$
\mathbb{E}\left[\sup _{0 \leq t \leq T}\left|X_{t}\right|^{2}\right] \leq 4 \cdot \mathbb{E}\left[\left|M_{T}\right|^{2}\right]
$$

## A5. Ito Calculus

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\left\{W_{t}\right\}$ be a $\mathbb{P}-$ Brownian motion. The filtration $\left\{\mathcal{F}_{t}\right\}$ is the natural one; that is, for each $t \geq 0, \mathcal{F}_{t}$ is the $\sigma$-algebra generated by $\left\{W_{s}: s \leq t\right\}$ (completed by the $\mathbb{P}$-null sets) so that $\left\{W_{t}\right\}$ is a $\mathbb{P}$-martingale adapted to the filtration $\left\{\mathcal{F}_{t}\right\}$.

Let $\left\{X_{t}\right\}_{0 \leq t \leq T}$ be a stochastic process adapted to $\left\{\mathcal{F}_{t}\right\}$ such that

$$
\mathbb{E}\left[\int_{0}^{T}\left(X_{t}\right)^{2} d t\right]<+\infty
$$

The Ito integral of $X$ w.r.t. $W$ is defined as

$$
\int_{0}^{t} X_{s} d W_{s}:=\lim _{n \rightarrow \infty} \sum_{i=1}^{n} X_{t_{i-1}}\left(W_{t_{i}}-W_{t_{i-1}}\right)
$$

as the mesh of the subdivision goes to zero, where $t_{0}=0<t_{1}<\cdots<t_{n}=t$ is a partition of $[0, t]$.
Proposition A9 (Properties of Ito Integral)

1. $\int_{0}^{t} X_{s} d W_{s}$ is $\mathcal{F}_{t}$-measurable;
2. integral is linear in terms of integrand $X$;
3. $\left\{\int_{0}^{t} X_{s} d W_{s}\right\}_{t \geq 0}$ is a martingale;
4. $\int_{0}^{t} X_{s} d W_{s}$ is continuous in $t$, for almost surely $\omega$. item there holds the Ito isometry:

$$
\mathbb{E}\left[\left(\int_{0}^{t} X_{s} d W_{s}\right)^{2}\right]=\mathbb{E}\left[\int_{0}^{t} X_{s}^{2} d s\right]
$$

Definition A10 (Ito's process) An Ito's process is a stochastic process of the form

$$
X_{t}=X_{0}+\int_{0}^{t} K_{s} d s+\int_{0}^{t} H_{s} d W_{s}
$$

or in differential form

$$
d X_{t}=K_{t} d t+H_{t} d W_{t}
$$

where

- $X_{0}$ is $F_{0}$-measurable;
- $\left(K_{t}\right)$ and $\left(H_{t}\right)$ are adapted;
- $\int_{0}^{T}\left|K_{s}\right| d s<\infty$ (i.e., $\left.K_{t} \in L^{1}[0, T]\right)$ a.s.
- $\int_{0}^{T}\left|H_{s}\right|^{2} d s<\infty$ (i.e. $\left.H_{t} \in L^{2}[0, T]\right)$ a.s.

Remark 19. (i) The processes $\left\{K_{t}\right\}$ and $\left\{H_{t}\right\}$ are unique in the definition of Ito's process.
(ii) Ito's process is not a martingale unless $K=0 d t \times d \mathbb{P}$ a.e.

Theorem 20. (Ito's formula) Let $f \in C^{1,2}$ and let $X$ be a stochastic process given by

$$
d X_{t}=\mu\left(t, X_{t}\right) d t+\sigma\left(t, X_{t}\right) d W_{t}
$$

Then

$$
\begin{aligned}
d f\left(t, X_{t}\right)= & f_{t}^{\prime}\left(t, X_{t}\right) d t+f_{x}^{\prime}\left(t, X_{t}\right) d X_{t} \\
& +\frac{1}{2} f_{x x}^{\prime \prime}\left(t, X_{t}\right) d\langle X\rangle_{t},
\end{aligned}
$$

or in integral form

$$
\begin{aligned}
f\left(t, X_{t}\right)= & f\left(0, X_{0}\right)+\int_{0}^{t} f_{s}^{\prime}\left(s, X_{s}\right) d s \\
& +\int_{0}^{t} f_{x}^{\prime}\left(s, X_{s}\right) d X_{s}+\frac{1}{2} \int_{0}^{t} f_{x x}^{\prime \prime}\left(s, X_{s}\right) d\langle X\rangle_{s}
\end{aligned}
$$

Here $d\langle X\rangle_{t}$ is the quadratic variation of $X_{t}$. We have

$$
d\langle X\rangle_{t}=\sigma^{2}\left(t, X_{t}\right) d t
$$

Hence Ito's formula is rewritten as

$$
\begin{aligned}
d f\left(t, X_{t}\right)= & {\left[f_{t}^{\prime}\left(t, X_{t}\right)+\mu\left(t, X_{t}\right) f_{x}^{\prime}\left(t, X_{t}\right)+\frac{1}{2} \sigma^{2}\left(t, X_{t}\right) f_{x x}^{\prime \prime}\left(t, X_{t}\right)\right] d t } \\
& +\sigma^{2}\left(t, X_{t}\right) f_{x}^{\prime}\left(t, X_{t}\right) d W_{t} .
\end{aligned}
$$

The product rule is below. Let $X$ and $Y$ be stochastic processes given by

$$
d X_{t}=\mu_{X}\left(t, X_{t}\right) d t+\sigma_{X}\left(t, X_{t}\right) d W_{t}
$$

and

$$
d Y_{t}=\mu_{Y}\left(t, Y_{t}\right) d t+\sigma_{Y}\left(t, Y_{t}\right) d W_{t}
$$

then

$$
\begin{aligned}
d\left(X_{t} Y_{t}\right) & =X_{t} d Y_{t}+Y_{t} d X_{t}+d\langle X, Y\rangle_{t} \\
& =X_{t} d Y_{t}+Y_{t} d X_{t}+\sigma_{X}\left(t, X_{t}\right) \sigma_{Y}\left(t, Y_{t}\right) d t .
\end{aligned}
$$

Example 21. Take $f(x)=x^{2}$ and $X_{t}=W_{t}$. Ito's formula

$$
f\left(X_{t}\right)=f\left(X_{0}\right)+\int_{0}^{t} f^{\prime}\left(X_{s}\right) d X_{s}+\frac{1}{2} \int_{0}^{t} f^{\prime \prime}\left(X_{s}\right) d\left\langle X_{s}\right\rangle
$$

gives us (since $d\left\langle X_{s}\right\rangle=d s$ )

$$
W_{t}^{2}=2 \int_{0}^{t} W_{s} d s+t
$$

In differential form, this is written

$$
d W_{t}^{2}=2 W_{t} d W_{t}+t
$$

## A6. Stochastic Differential Equations

Consider the stochastic differential equation (SDE)

$$
X_{t}=Z+\int_{0}^{t} b\left(s, X_{s}\right) d s+\int_{0}^{t} \sigma\left(s, X_{s}\right) d W_{s}
$$

or in differential form

$$
\left\{\begin{array}{l}
d X_{t}=b\left(t, X_{t}\right) d t+\sigma\left(t, X_{t}\right) d W_{t} \\
X_{0}=Z
\end{array}\right.
$$

Theorem 22. Suppose b, $\sigma$ are continuous and for some constant $K>0$,

1. $|b(t, x)-b(t, y)|+|\sigma(t, x)-\sigma(t, y)| \leq K|x-y|$,
2. $|b(t, x)|+|\sigma(t, x)| \leq K(1+|x|)$,
3. $\mathbb{E}\left[Z^{2}\right]<\infty$.

Then for any $T>0$, there exists a unique solution $\left(X_{s}\right)_{0 \leq s \leq T}$ in the interval $[0, T]$ satisfying

$$
\mathbb{E}\left(\sup _{0 \leq s \leq T}\left|X_{s}\right|^{2}\right)<\infty
$$

Example 23. The Ornstein-Ulhenbeck process is the unique solution of the SDE:

$$
\left\{\begin{array}{l}
d X_{t}=-c X_{t} d t+\sigma d W_{t} \\
X_{0}=x
\end{array}\right.
$$

Consider the process $\left\{Y_{t}\right\}$ given by

$$
Y_{t}=e^{c t} X_{t}
$$

By the integration by parts formula, we get

$$
\begin{aligned}
d Y_{t}= & d\left(e^{c t}\right) X_{t}+e^{c t} d X_{t}+\underbrace{d\left\langle c t, X_{t}\right\rangle}_{=0} \\
& =c e^{c t} X_{t} d t+e^{c t}\left(-c X_{t} d t+\sigma d W_{t}\right) \\
& =\sigma e^{c t} d W_{t} .
\end{aligned}
$$

Since $Y_{0}=x$, we obtain

$$
\begin{gathered}
Y_{t}=x+\int_{0}^{t} \sigma e^{c s} d W_{s} \\
X_{t}=x e^{-c t}+\sigma e^{-c t} \int_{0}^{t} e^{c s} d W_{s}
\end{gathered}
$$

It follows that

$$
\mathbb{E}\left[X_{t}\right]=x e^{-c t}
$$

and

$$
\begin{aligned}
\operatorname{var}\left(X_{t}\right) & =\mathbb{E}\left[\left(X_{t}-\mathbb{E}\left[X_{t}\right]\right)^{2}\right] \\
& =\sigma^{2} e^{-2 c t} \mathbb{E}\left[\left(\int_{0}^{t} e^{c s} d W_{s}\right)^{2}\right] \quad \text { by Ito's isometry formula } \\
& =\sigma^{2} e^{-2 c t} \mathbb{E}\left[\int_{0}^{t} e^{2 c s} d s\right] \\
& =\frac{\sigma^{2}}{2 c}\left(1-e^{-2 c t}\right)
\end{aligned}
$$

Example 24. Consider the $S D E$

$$
\left\{\begin{array}{l}
d S_{t}=S_{t}\left(\mu d t+\sigma d W_{t}\right) \\
S_{0}=s_{0}
\end{array}\right.
$$

where $\mu \in \mathbb{R}$ and $\sigma>0$ are constants. A solution is an adapted process $\left\{S_{t}\right\}$ such that the integrals $\int_{0}^{t} S_{s} d s$ and $\int_{0}^{t} S_{s} d W_{s}$ exist and satisfy

$$
S_{t}=x_{0}+\int_{0}^{t} \mu S_{s} d s+\int_{0}^{t} \sigma S_{s} d W_{s}
$$

Under certain measurability conditions, $\left\{S_{t}\right\}$ is an Ito process. Applying Ito's formula to the function $f(x)=\log (x)$ we obtain (noting $d\left\langle S_{t}\right\rangle=$ $\left.\sigma^{2} S_{t}^{2} d t\right)$

$$
\begin{aligned}
\log \left(S_{t}\right) & =\log \left(s_{0}\right)+\int_{0}^{t} \frac{1}{S_{s}} d S_{s}+\frac{1}{2} \int_{0}^{t} \frac{-1}{S_{s}^{2}} \sigma^{2} S_{s}^{2} d s \\
& =\log \left(s_{0}\right)+\int_{0}^{t} \frac{1}{S_{s}}\left(\mu d s+\sigma d W_{s}\right)-\frac{1}{2} \int_{0}^{t} \frac{1}{2} \sigma^{2} t \\
& =\log \left(s_{0}\right)+\left(\mu-\frac{1}{2} \sigma^{2}\right)+\sigma W_{t} .
\end{aligned}
$$

So we have

$$
S_{t}=s_{0} \exp \left(\left(\mu-\frac{1}{2} \sigma^{2}\right)+\sigma W_{t}\right) .
$$

Since $W_{t} \sim N(0, t)$, we can rewrite $W_{t}=\sqrt{t} Z$, where $Z \sim N(0,1)$. Then we have

$$
S_{t}=s_{0} \exp \left(\left(\mu-\frac{1}{2} \sigma^{2}\right)+\sigma \sqrt{t} Z\right) .
$$

This is precisely (replacing $r$ with $\mu$ ) the model that we derived from the discrete binomial model at the beginning of this talk. This model of the stock price is called the lognormal or the geometric Brownian model (GBM) which was suggested by Samuelson and studied by Fisher Black and Myron Scholes (1973).

## A7. The Girsanov Theorem

Theorem 25. (The Girsanov Theorem) Let $\left\{W_{t}\right\}_{t \geq 0}$ be a $\mathbb{P}$-martingale with respect to the natural filtration. Let $\left\{\theta_{t}\right\}_{t \geq 0}$ be an adapted process such that the process $\left(L_{t}\right)_{0 \leq t \leq T}$ defined by

$$
L_{t}=\exp \left(-\int_{0}^{t} \theta_{s} d W_{s}-\frac{1}{2} \int_{0}^{t} \theta_{s}^{2} d s\right)
$$

is a martingale.
Define a probability measure $\tilde{\mathbb{P}}$ by

$$
\tilde{\mathbb{P}}[A]=\int_{A} L_{T} d \mathbb{P} .
$$

Then, under $\tilde{\mathbb{P}}$, the process $\{\tilde{W}\}_{0 \leq t \leq T}$ defined by

$$
\tilde{W}_{t}=W_{t}+\int_{0}^{t} \theta_{s} d s
$$

is a standard Brownian motion.
Remark 26. (i) Note that

$$
\left.\frac{d \tilde{\mathbb{P}}}{d \mathbb{P}}\right|_{\mathcal{F}_{t}}=L_{t}
$$

is the Radon-Nikodym derivative of $\tilde{\mathbb{P}}$ with respect to $\mathbb{P}$.
(ii) A sufficient condition for $\left(L_{t}\right)_{0 \leq t \leq T}$ to be a martingale is that the Novikov condition below holds

$$
\mathbb{E}\left[\exp \left(\frac{1}{2} \int_{0}^{T} \theta_{t}^{2} d t\right)\right]<\infty
$$

## A8. The Brwonian Martingale Representation Theorem

Theorem 27. (The Brwonian Martingale Representation Theorem) Let $\left\{W_{t}\right\}_{t \geq 0}$ be a $\mathbb{P}$-martingale with respect to the natural filtration $\left\{\mathcal{F}_{t}\right\}$. Let $\left\{M_{t}\right\}_{t \geq 0}$ be a square-integral $\left(\mathbb{P},\left\{\mathcal{F}_{t}\right\}\right)$-martingale. Then there exists a $\left\{\mathcal{F}_{t}\right\}$ predictable process $\left\{\theta_{t}\right\}_{t \geq 0}$ such that

$$
M_{t}=M_{0}+\int_{0}^{t} \theta_{s} d W_{s} \quad \text { a.s. }
$$

## A9. The Feynman-Kac Theorem

There are two approaches to the option pricing theory; one using martingale and one using PDEs. The solution given the former approach is expressed in terms of conditional expectations, while that given by the latter approach in terms of PDEs. The Feynman-Kac theorem unifies the two solutions.

Theorem 28. (The Feynman-Kac Theorem) Assume the function F solves the boundary value problem

$$
\left\{\begin{array}{l}
\frac{\partial F}{\partial t}(t, x)+\mu(t, x) \frac{\partial F}{\partial x}+\frac{1}{2} \sigma^{2}(t, x) \frac{\partial^{2} F}{\partial x^{2}}(t, x)=0,0 \leq t \leq T \\
F(T, x)=\phi(x)
\end{array}\right.
$$

Define a stochastic process $\left\{X_{t}\right\}_{0 \leq t \leq T}$ by the SDE

$$
d X_{t}=\mu\left(t, X_{t}\right) d t+\sigma\left(t, X_{t}\right) d W_{t}, 0 \leq t \leq T
$$

where $\left\{W_{t}\right\}_{t \geq 0}$ is a standard Brownian motion. Suppose

$$
\int_{0}^{T} \mathbb{E}\left[\left(\sigma\left(t, X_{t}\right) \frac{\partial F}{\partial x}\left(t, X_{t}\right)\right)^{2}\right] d s<\infty
$$

Then

$$
F(t, x)=\mathbb{E}\left[\phi\left(X_{T}\right) \mid X_{t}=x\right] .
$$

Example 29. Consider the boundary value problem

$$
\left\{\begin{array}{l}
\frac{\partial F}{\partial t}(t, x)+\frac{1}{2} \frac{\partial^{2} F}{\partial x^{2}}(t, x)=0, \\
F(T, x)=\phi(x) .
\end{array}\right.
$$

We have $\mu=0$ and $\sigma=1$. So the corresponding SDE is

$$
d X_{t}=d W_{t} \quad \Longrightarrow \quad X_{t}=W_{t}
$$

and by the Feynman-Kac theorem, we have

$$
F(t, x)=\mathbb{E}\left[\phi\left(W_{T}\right) \mid W_{t}=x\right] .
$$

But, since the transition density for $W_{t}$ is

$$
p(t, x, y)=\frac{1}{\sqrt{2 \pi t}} \exp \left(-\frac{(x-y)^{2}}{2 t}\right)
$$

which is the fundamental solution to the equation

$$
\frac{\partial u}{\partial t}=\frac{1}{2} \frac{\partial^{2} u}{\partial x^{2}}
$$

we have

$$
F(t, x)=\int_{-\infty}^{\infty} \frac{1}{\sqrt{2 \pi(T-t)}} \exp \left(-\frac{(x-y)^{2}}{2(T-t)}\right) \phi(y) d y
$$

More details on stochastic processes and stochastic differential equations can be found in [3], [5], and [6].

## References

[1] Y. Achdou and N. Tchou, The Black and Scholes equation with stochastic volatility. Variational methods, C. R. Acad. Sci. Paris, t. 332, Ser. I, 2001, 1031-1036.
[2] Y. Achdou and N. Tchou, Variational analysis for the Black and Scholes equation with stochastic volatility, Math. Model. Numer. Anal. 36(2002), 373-395.
[3] A. Etheridge, A Course in Financial Calculus, Cambrdge Univerity Press, 2002.
[4] J.-P. Fouque, G. Papanicolau, and R. Sircar, Derivatives in Financial Markets with Stochastic Volatility, Cambrdge Univerity Press, 2000.
[5] D. Lamberton and B. Lapeyre, Introduction to Stochastic Calculus Applied to Finance, Chapman \& Hall, 1996.
[6] B. Oksendale, Stochastic Differential Equations (5th Edition), Springer, 1998.
[7] S. E. Shreve, Stochastic Calculus for Finance I: The Binomial Asset Pricing Model, Springer, 2004.

## Author's address:

Hong-Kun Xu
School of Mathematical Sciences
University of KwaZulu-Natal
Westville Campus, Private Bag X54001
Durban 4000, South Africa
E-mail: xuhk@ukzn.ac.za

Editorial Universidad de Sevilla


[^0]:    *The author is partially supported by Proyecto BMF2002-04013

[^1]:    2000 Mathematics Subject Classification. Primary 47H10, 47H09, 46E30
    Key words and phrases. Convex sets; fixed point property.

[^2]:    *Partially supported by by BFM 2003-03893-C02-02
    2000 Mathematics Subject Classification. 34A60, 34C25, 34C27, 47H10
    Key words and phrases. Integral solutions, periodic solutions, Accretive operators, nonexpansive mappings, fixed points

[^3]:    *Supported by Universidad Nacional del Litoral \& Instituto de Matemática del Litoral, Santa Fe, Argentina

    2000 Mathematics Subject Classification. 42B20
    Key words and phrases. Fractional Integral operator, Fractional Derivative operator, Integral operator of functional order, Derivative operator of functional order CalderónZygmund operators, Besov spaces, Triebel-Lizorkin spaces, Spaces of homogeneous type

[^4]:    2000 Mathematics Subject Classification. 46E30

[^5]:    *Partially supported by MCYT Grant BFM2001-0189
    2000 Mathematics Subject Classification. Primary 42B20, 42B25.
    Key words and phrases. Calderón-Zygmund singular integral operators, Muckenhoupt weights, maximal functions.

[^6]:    *The author was partially supported by MCYT Grant BFM2001-0189.
    2000 Mathematics Subject Classification. 42B20, 42B25, 46E30, 42B35
    Key words and phrases. Extrapolation of weighted norm inequalities, rearrangement invariant function spaces, modular inequalities, maximal functions, singular integrals, commutators.

[^7]:    2000 Mathematics Subject Classification. 47H10, 54H25, 47H04, 28 A 80.
    Key words and phrases. fixed point, strict fixed point, successive approximations, (weakly) Picard operator, multivalued (weakly) Picard operator, fractal operator.

[^8]:    *This material is based upon work supported by the National Research Foundation under Grant number 2053752.

[^9]:    *Partially supported by grants BFM2000-0361 and MTM2004-00519 (Spain), 2001-SGR-00431 (Generalitat de Catalunya), and HPRN-2000-0116 (European Union)

[^10]:    *Supported in part by the National Research Foundation of South Africa.
    2000 Mathematics Subject Classification. Primary 91B28; Secondary 91B70, 60H05.
    Key words and phrases. Binomial model, risk neutral evaluation, European option, American option, martingale, geometric Brownian motion, Black-Scholes, stochastic calculus, stochastic volatility.

