ON RATES IN MEAN ERGODIC THEOREMS

ALEXANDER GOMILKO, MARKUS HAASE, AND YURI TOMILOV

Abstract. We create a general framework for the study of rates of decay in mean ergodic theorems. As a result, we unify and generalize results due to Assani, Cohen, Cuny, Derriennic, and Lin dealing with rates in mean ergodic theorems in a number of cases. In particular, we prove that the Cesàro means of a power-bounded operator applied to elements from the domain of its abstract one-sided ergodic Hilbert transform decay logarithmically, and this decay is best possible under natural spectral assumptions.

1. Introduction

In recent years considerable emphasis has been put on the study of convergence rates in limit theorems arising in probability and ergodic theory, see e.g. [15, 13, 11] and the references therein. The basic result of interest here is of course the ergodic theorem in its two forms regarding almost everywhere convergence and convergence in norm. The latter\(^1\) is the famous “Mean Ergodic Theorem” of von Neumann (1931), subsequently generalized by Riesz, Yosida, Kakutani, Lorch and Eberlein towards a general theory of (norm-)convergence of the Cesàro averages

\[
A_n(T)x := \frac{1}{n} \sum_{k=0}^{n-1} T^k x, \quad x \in X,
\]

for a bounded linear operator \(T\) on a Banach space \(X\) [24, Chapter 2]. A natural assumption on the operator \(T\) in this context is that \(T\) is power-bounded, i.e., that \(\sup_{n \in \mathbb{N}} \|T^n\| < \infty\). For a power-bounded operator \(T\) it is simple operator theory to show that for \(x, y \in X\)

\[
A_n(T)x \to 0 \iff x \in \text{ran}(I - T)
\]

and

\[
A_n(T)x \to y \implies y \in \ker(I - T).
\]

It follows that \(\ker(I - T) \oplus \text{ran}(I - T)\) is precisely the subspace of \(X\) on which the Cesàro averages converge strongly, and the operator \(T\) is called mean ergodic if this is already the whole space \(X\). A mean ergodic theorem gives conditions under which this is the case, and Lorch’s famous generalization of von Neumann’s theorem simply says that every power-bounded operator on a reflexive space is mean ergodic [24, Chapter 2, Theorem 1.2].

\(^1\)We shall not be concerned with pointwise almost everywhere convergence in this paper.
In this paper we shall be concerned with the study of convergence rates for $A_n(T)x$, where $T$ is a power-bounded operator. If $A_n(T)x \to y$ then $A_n(T)x - y = A_n(T)(x - y)$, and hence in the study of rates in the convergence of Cesàro averages one can restrict oneself to the convergence to zero on the space $\text{ran}(I - T)$. Since one can always consider the restriction of $T$ to this (obviously $T$-invariant) subspace, often it is no loss of generality in assuming that $X = \text{ran}(I - T)$ in the first place. However, we shall always make this assumption explicit when it is needed. Let us collect some known facts.

**Proposition 1.1.** Let $T$ be a power-bounded operator on a Banach space $X$. Then the following statements hold:

a) If $\|A_n(T)x\| = o(1/n)$ then $x = 0$.

b) If $x \in \text{ran}(I - T)$ then $\|A_n(T)x\| = O(1/n)$, and the converse is true if $X$ is reflexive.

c) If there exists a sequence $(r_n)_{n \geq 1}$ of positive numbers such that $r_n \downarrow 0$ and $\|A_n(T)x\| = O(r_n)$ for every $x \in X$, then $I - T$ is invertible.

Part a) is due to Butzer and Westphal [5]. It tells us that we cannot expect better convergence rates than $O(1/n)$. The first assertion in part b) is utterly trivial, and the second is due to Browder [4] (but appears also in [5]). The proof of c) rests on the principle of uniform boundedness, by virtue of which one first concludes that $A_n(T) \to 0$ in norm. Hence $T$ is a so-called uniformly ergodic operator, and Dunford [17] has shown that for such operators $\text{ran}(I - T)$ must be closed, cf. [26]. Uniformly ergodic operators are quite special, and Derriennic [15, p.144] remarks that an ergodic measure preserving transformation on a non-atomic space is never uniformly ergodic. Indeed, it is proved in [14, Proposition W.12] that an isometric lattice isomorphism on a Banach lattice is uniformly ergodic if and only if it is periodic. Hence by Proposition 1.1 there are plenty of examples where the Cesàro averages fail to converge with a uniform rate. For intricate results on the failing of rates see also the recent work [3] of Badea and Müller.

The absence of a global uniform convergence rate leads naturally to the problem of identifying subspaces of $X$ for which one has a specified rate and of describing a possible rate if a subspace is given. Results in this direction have importance for almost everywhere convergence theorems, see e.g. [2, 7, 9, 11, 12, 16], and central limit theorems for Markov chains, see e.g. [13] and the references therein.

Polynomial rates of decay were studied thoroughly in [16] in case of arbitrary Banach space contractions while in the special case of unitary and normal contractions on Hilbert spaces more general rates were investigated in [2] and [9]. For instance, it was proved in [16] that for a power-bounded and mean ergodic operator $T$ and $s \in (0, 1)$ one has

\begin{equation}
\|A_n(T)x\| = o(n^{-s}) \quad \text{as } n \to \infty.
\end{equation}

for every $x \in \text{ran}(I - T)^s$. Moreover, it was shown there that under the same assumptions the condition $x \in \text{ran}(I - T)^s$ is equivalent to the strong and also to the weak convergence of the series

$$\sum_{n=0}^{\infty} \alpha_n^{-s} T^n x,$$

where $(1 - z)^{-s} = \sum_{n \geq 0} \alpha_n^{-s} z^n$ is the power series representation of $(1 - z)^{-s}$. 
Towards logarithmic rates it was proved in [9] that if $T$ is a normal contraction on a Hilbert space, then
\begin{equation}
H_T x := \sum_{k=1}^{\infty} \frac{T^k x}{k} \text{ converges } \implies \|A_n(T)x\| = O\left(\frac{1}{\log n}\right),
\end{equation}
but for more general operators $T$ the validity of (1.3) remained an open problem. (Note that it was proved in [1] that the converse implication in (1.3) does not, in general, hold even for unitary $T$.) The operator $H_T$ appearing here (with its natural domain) is called the one-sided ergodic Hilbert transform, and was studied thoroughly in the mentioned papers by Assani, Cohen, Cuny, Derrienic and Lin. It was an open question for some years whether in general one has $-H_T = \log(I - T)$ if $\text{ran}(I - T) = X$, where the latter operator is the generator of the $C_0$-semigroup $((I - T)s)_{s \geq 0}$. It was solved (in the affirmative) recently, independently in [10] and [20].

In the present paper we shall show that the implication (1.3) holds in general, see Section 3. The result is actually a corollary of a more general statement that also includes the known facts about polynomial rates. In order to explain this result, let us first note that we can frame the underlying question in the following abstract way:

**Problem 1.2.** Suppose that $f(z) = \sum_{k \geq 0} \alpha_k z^k$ is holomorphic on the open unit disc $D$ and suppose that $T$ is some power-bounded operator on a Banach space and $x \in X$, such that $\sum_{k=0}^{\infty} \alpha_k T^k x$ converges (in some sense). Does this imply a certain rate of convergence for $(A_n(T)x)_{n \geq 1}$, and if yes, which is the optimal such rate on the space of all such vectors $x$?

To put it concisely, Problem 1.2 asks whether the subspace $\text{con}(f) := \left\{ x \mid \sum_{k=0}^{\infty} \alpha_k T^k x \text{ converges weakly} \right\}$ realizes a certain rate in the sense of the following definition.

**Definition 1.3.** Let $T$ be a power-bounded operator on a Banach space $X$. A linear subspace $Y \subseteq X$ is said to realize the rate $r = (r_n)_{n \geq 1}$ if
\[
\|A_n(T)x\| = O(r_n) \quad \text{as } n \to \infty
\]
for every $x \in Y$, and for any positive sequence $(\epsilon_n)_{n \geq 1}$ such that $\epsilon_n \to 0$ as $n \to \infty$ there exists $x \in Y$ such that
\begin{equation}
\|A_n(T)x\| \neq O(\epsilon_n r_n) \quad \text{as } n \to \infty.
\end{equation}

For instance, it follows from Proposition 1.1 that $Y = \text{ran}(I - T)$ realizes the rate $r_n = 1/n$, and $\text{ran}(I - T)$ realizes the rate $r_n = 1$ whenever 1 is in the spectrum of $T$.

It is natural — and also suggested by the mentioned results about polynomial and logarithmic rates — to employ the functional calculus for power-bounded operators and its unbounded extension. (For the convenience of the reader, we have put a short summary of this calculus in Section 2 below.) So one asks whether for certain $f$ the domain $\text{dom}(f(T))$ realizes a certain rate. This, however, is not the whole story, because in general it may happen that $\text{dom}(f(T)) \neq \text{con}(f)$. Our work [20] was largely devoted to the latter problem, and we shall make some advance in Theorem 6.1 below. Of course, one cannot expect to solve Problem 1.2 in full generality, that
is, without any restriction on the function $f$. Our investigation of the one-sided ergodic Hilbert transform in [20] lead us to consider so-called admissible functions in the following sense.

**Definition 1.4.** A holomorphic function $f(z) := \sum_{k=0}^{\infty} \alpha_k z^k$ on the open unit disc $\mathbb{D}$ (or, equivalently, the sequence $(\alpha_k)_{k \geq 0}$) is called admissible if $f$ has no zeroes in $\mathbb{D}$ and

$$\frac{1}{f(z)} = \gamma_0 - \sum_{k=1}^{\infty} \gamma_k z^k \quad (z \in \mathbb{D})$$

with $\gamma_k \geq 0$ for all $k \geq 0$.

We postpone a more detailed discussion of this notion to Section 2.3 below. Note that it is proved there that if $f$ is admissible and $g := 1/f$ is as in Definition 1.4, then

$$g(T) := \gamma_0 I - \sum_{k=1}^{\infty} \gamma_k T^k$$

is a well-defined bounded linear operator with

$$\|g(T)\| \leq 2\gamma_0 \cdot \sup_{k \geq 0} \|T^k\|,$$

cf. also Section 2.2. Our main observation in this paper is that if $f$ is admissible with $f(1) = \infty$ and $g := 1/f$, then on ran($g(T)$) the Cesàro averages converge to 0 with a certain rate $r = r[g]$ that can be computed either from the Taylor coefficients of $g$ or from the values of $f$ at certain points. More succinctly, our main results are the following.

**Summary 1.5.** Let $f = \sum_{n=0}^{\infty} \alpha_n z^n$ be admissible such that $f(1) = \infty$, let

$$g(z) = \frac{1}{f(z)} = \gamma_0 - \sum_{k=1}^{\infty} \gamma_k z^k,$$

let $T$ denote a power-bounded operator on a Banach space $X$, and let

$$r_n := \frac{1}{n} \sum_{k=1}^{n} k\gamma_k + \sum_{k>n} \gamma_k = \sum_{k=1}^{\infty} \min(k/n, 1)\gamma_k \quad (n \in \mathbb{N}).$$

1) Then

$$\|A_n(T)x\| = O(r_n) \quad \text{as } n \to \infty$$

for any $x \in \text{ran}(g(T))$. And if $T$ is mean ergodic and $nr_n \to \infty$, then even

$$\|A_n(T)x\| = o(r_n) \quad \text{as } n \to \infty$$

for any $x \in \text{ran}(g(T))$. (Theorem 4.4).

2) One has $r_n \sim g(1 - 1/n)$ (Theorem 5.2).

3) Under certain (mild) spectral conditions the rate $(r_n)_{n \geq 1}$ is realized by ran($g(T)$). In particular, the rate $(r_n)_{n \geq 1}$ on ran($g(T)$) is sharp for a large class of situations (Theorem 5.3).

4) If $y := \sum_{k=0}^{\infty} \alpha_k T^k x$ converges weakly, then $x = g(T)y$ and

$$\|A_n(T)x\| \leq \frac{2M}{f(1 - 1/n)} \|y\| \quad (n \geq 1),$$

with $M := \sup_{n \geq 1} \|T^n\|$. If in addition $n/f(1 - 1/n) \to \infty$ then one even has

$$\|A_n(T)x\| = o\left(\frac{1}{f(1 - 1/n)}\right) \quad \text{as } n \to \infty$$

(Corollary 6.3).
The result on logarithmic rates is obtained from these statements by considering

\[ f(z) = 2 - \log(1 - z) = 2 + \sum_{n \geq 1} \frac{z^n}{n}, \]

see Example 2.4.1 below. Moreover, one can also deduce from Summary 1.5 the mentioned results on polynomial rates obtained originally by a different method in [16, Theorem 2.11 and Corollary 2.15]. Indeed, fix \( s \in (0, 1) \) and consider the function

\[ f_s(z) = (1 - z)^{-s} = \sum_{n=0}^{\infty} \alpha_n^{-s} z^n. \]

Then \( f_s \) is admissible (see Example 2.4.2 below), hence Summary 1.5 can be applied with \( f = f_s \). By 2) one has \( r_n \sim 1/f_s(1 - 1/n) = n^{-s} \); from 1) it follows that

\[ \|A_n(T)x\| = O(n^{-s}) \quad \text{as } n \to \infty \]

for every \( x \in \text{ran}(I - T)^s \) and every power-bounded operator \( T \) on a Banach space. And since \( s \in (0, 1) \), one has \( nr_n \sim n^{1-s} \to \infty \), whence if \( T \) is moreover mean ergodic, then

\[ \|A_n(T)x\| = o(n^{-s}) \quad \text{as } n \to \infty \]

for every \( x \in \text{ran}(I - T)^s \). Part 4) finally shows that if \( y := \sum_{n \geq 0} \alpha_n^{-s} T^n x \) is weakly convergent, then \( x = (I - T)^s y \).

Apart from these known facts, part 3) adds a (previously unknown) condition on \( T \) under which the rate \( n^{-s} \) is optimal for \( A_n(T) \) on \( \text{ran}(I - T)^s \). This is true, e.g., if \( T \) is a non-invertible isometry, like the one-sided shift on \( \ell^p \), \( 1 \leq p \leq \infty \).

\[ \text{2. Preliminaries} \]

In this section we make some terminological and notational conventions (Section 2.1) and devise some background information on functional calculus (Section 2.2) and admissible functions (Section 2.3). There is no harm in skipping this section on first reading.

2.1. Some Notations and Definitions. For a closed linear operator \( A \) on a complex Banach space \( X \) we denote by \( \text{dom}(A) \), \( \text{ran}(A) \), \( \ker(A) \), and \( \sigma(A) \) the \textit{domain}, the \textit{range}, the \textit{kernel}, and the \textit{spectrum} of \( A \), respectively. The norm-closure of the range is written as \( \text{ran}(A) \). The space of bounded linear operators on \( X \) is denoted by \( \mathcal{L}(X) \). The \textit{open unit disc} is denoted by \( \mathbb{D} = \{ z \in \mathbb{C} \mid |z| < 1 \} \), the \textit{torus} by \( T = \{ z \in \mathbb{C} \mid |z| = 1 \} \), and \( \mathbb{N} = \{ 1, 2, 3, \ldots \} \) is the set of \textit{natural numbers}. If \( \alpha = (\alpha_k)_{k=0}^{\infty} \) is a scalar sequence we write \( \hat{\alpha}(z) := \sum_{k=0}^{\infty} \alpha_k z^k \) for the associated power series. For positive sequences \( (r_n)_{n \geq 0} \) and \( (s_n)_{n \geq 0} \) we write \( r_n \sim s_n \) if there is \( c > 0 \) such that \( r_n/c \leq s_n \leq cr_n \) for eventually all \( n \in \mathbb{N} \). Finally, we write

\[ A_n(z) := \frac{1}{n} \sum_{k=0}^{n-1} z^k \quad (n \in \mathbb{N}, z \in \mathbb{D}). \]

2.2. Functional Calculus. We denote by \( \Lambda^1_{+}(\mathbb{D}) \) the algebra of holomorphic functions on \( \mathbb{D} \) that have absolutely summable Taylor coefficients

\[ \Lambda^1_{+}(\mathbb{D}) = \left\{ g \mid g(z) = \sum_{k \geq 0} \alpha_k z^k, \sum_{k \geq 0} |\alpha_k| < \infty \right\} \]

with the norm

\[ \|g\|_{\Lambda^1_{+}} := \|\alpha\|_1 = \sum_{k \geq 0} |\alpha_k| \quad \text{for } g(z) = \sum_{k \geq 0} \alpha_k z^k \in \Lambda^1_{+}(\mathbb{D}). \]
It is well-known (and easy to see) that for each power-bounded operator $T$ on a Banach space $X$ the assignemnt
g = \sum_{k \geq 0} \alpha_k z^k \mapsto g(T) := \sum_{k \geq 0} \alpha_k T^k
is a continuous algebra homomorphism (a functional calculus) of $A^1_\lambda(\mathbb{D})$ into $\mathcal{L}(X)$, satisfying
\begin{equation}
\|g(T)\| \leq \left( \sup_{n \geq 0} \|T^n\| \right) \|g\|_{A^1_\lambda(\mathbb{D})} \quad (g \in A^1_\lambda(\mathbb{D})).
\end{equation}
For this functional calculus one has a spectral mapping theorem; however, we need only the following weaker statement, which we prove for the convenience of the reader.

**Lemma 2.1 (Spectral inclusion theorem).** Let $g \in A^1_\lambda(\mathbb{D})$ and let $T \in \mathcal{L}(X)$ be a power-bounded operator on a Banach space $X$. Then
\[ g(\sigma(T)) = \{ g(\lambda) \mid \lambda \in \sigma(T) \} \subseteq \sigma(g(T)). \]

**Proof.** Let $\lambda \in \sigma(T)$. Since $T$ is power-bounded, $|\lambda| \leq 1$. Then
\[ g(\lambda) - g(T) = \lim_{n \to \infty} \sum_{k=1}^{n} \alpha_k (\lambda^k - T^k). \]
The convergence here is in operator norm, hence if $g(\lambda) - g(T)$ is invertible then for large $n \in \mathbb{N}$ also
\[ \sum_{k=1}^{n} \alpha_k (\lambda^k - T^k) = (\lambda - T) \sum_{k=1}^{n} \alpha_k \sum_{j=0}^{k} \lambda^{k-j} T^j \]
is invertible. But this implies that $\lambda - T$ is invertible, contradicting our assumption. \hfill \Box

The $A^1_\lambda(\mathbb{D})$-functional calculus has a natural extension to a larger function class, as is described abstractly in [18, Chapter 2]. However, we need only the following fairly intuitive facts:

**Lemma 2.2.**
\begin{itemize}
  \item[a)] If $g \in A^1_\lambda(\mathbb{D})$ has no zeroes in $\mathbb{D}$, $f := 1/g$, and $g(T)$ is injective, then $f(T)$ is defined and one has $f(T) = g(T)^{-1}$.
  \item[b)] If $f_1 = f_2 + g$ for some $g \in A^1_\lambda(\mathbb{D})$, then $f_1(T)$ is defined if and only if $f_2(T)$ is defined, and one has $\text{dom}(f_1(T)) = \text{dom}(f_2(T))$ and $f_1(T) = f_2(T) + g(T)$.
\end{itemize}

**2.3. Admissible functions.** Recall Definition 1.4 of an admissible function/sequence. Bounded admissible sequences $(\alpha_k)_{k \geq 0}$ with $\alpha_0 = 1$ are also called renewal sequences, and there is an extensive literature about them and their role in stochastic processes, see e.g. [22, 23]. We shall review briefly the basic analytic properties.

Using induction and the identities linking the Taylor coefficients of $f$ and $g := 1/f$ one finds that if $f$ is admissible, then $\alpha_0 > 0$ and hence $\gamma_0 = \alpha_0^{-1} > 0$; moreover $\alpha_k \geq 0$ for all $k \geq 1$. In particular, $f(t)$ is positive and increasing for $0 \leq t < 1$, and this yields that $\sum_{k=1}^{\infty} \gamma_k \leq \gamma_0$. In particular, $g$ has absolutely summable Taylor coefficients, i.e., $g \in A^1_\lambda(\mathbb{D})$. So two cases can occur, described by the equivalences
\begin{align}
(2.2) \quad \sum_{k \geq 1} \gamma_k < \gamma_0 & \iff g(1) > 0 \iff f(1) < \infty \iff f \in A^1_\lambda(\mathbb{D}) \\
n(2.3) \quad \sum_{k \geq 1} \gamma_k = \gamma_0 & \iff g(1) = 0 \iff f(1) = \infty \iff f \notin A^1_\lambda(\mathbb{D}).
\end{align}
We remark that every function $g$ that satisfies the natural conditions
\[
g(z) = \gamma_0 - \sum_{k=1}^{\infty} \gamma_k z^k, \quad \gamma_k \geq 0 \quad (k \geq 1), \quad \sum_{k=1}^{\infty} \gamma_k \leq \gamma_0 > 0.
\]
satisfies $g = 1/f$ for some admissible function $f$. (This follows from the fact that $1$ is an extremal point of $\mathbb{D}$.)

This allows to construct admissible functions by virtue of their inverses. A much harder question is whether a given function $f$ given by its Taylor expansion is admissible. The best known criterion is the following result of Kaluza [21].

**Proposition 2.3** (Kaluza). Let $\alpha = (\alpha_k)_{k \geq 0}$ be a positive and decreasing sequence. Suppose that $\alpha$ is log-convex, i.e. $\alpha_0 > 0$ and $\alpha_k^2 \leq \alpha_{k-1} \alpha_{k+1}, k \geq 1$, and let $f(z) := \sum_{k=0}^{\infty} \alpha_k z^k, z \in \mathbb{D}$. Then $f$ is admissible.

**Proof.** Since $f(0) = \alpha_0 > 0$, $-1/f(z)$ can be expanded into the Taylor series
\[
-1/f(z) = \sum_{k=0}^{\infty} \beta_k z^k
\]
around zero. Observe that $\beta_0 = -1/\alpha_0, \beta_1 = \alpha_1/\alpha_0^2$ and
\[
0 = \alpha_n \beta_0 + \sum_{j=1}^{n-1} \alpha_{n-j} \beta_j, \quad 0 = \alpha_{n+1} \beta_0 + \sum_{j=1}^{n} \alpha_{n+1-j} \beta_j
\]
for all $n \geq 1$. Multiplying the first identity by $\alpha_{n+1}$ and the second by $\alpha_n$ and subtracting we obtain
\[
0 = -\alpha_0 \alpha_n \beta_{n+1} + \sum_{j=1}^{n} \left( \alpha_{n-j} \alpha_{n+1} - \alpha_{n+1-j} \alpha_n \right) \beta_j
\]
so that
\[
\beta_{n+1} = \frac{1}{\alpha_0} \sum_{j=1}^{n} \left( \frac{\alpha_{n+1}}{\alpha_n} - \frac{\alpha_{n-j+1}}{\alpha_{n-j}} \right) \alpha_n \beta_j \geq 0
\]
by the log-convexity of $\alpha$. Hence by induction $\beta_k \geq 0$ for all $k \geq 1$. Further, using positivity of $\beta_k$ we obtain
\[
\beta_n = -\frac{1}{\alpha_0} \left( \alpha_n \beta_0 + \sum_{j=1}^{n-1} \alpha_j \beta_{n-j} \right) = \frac{\alpha_n}{\alpha_0^2} - \frac{1}{\alpha_0} \sum_{j=1}^{n-1} \alpha_j \beta_{n-j} \leq \frac{\alpha_n}{\alpha_0^2}
\]
for all $n \geq 1$. Therefore, the radius of convergence of the Taylor series for $1/f$ is at least 1, thus $f$ has no zeroes in $\mathbb{D}$. \hfill \Box

**Examples 2.4.** 1) From Kaluza’s theorem it is immediate that the functions
\[
-\log(1-z) = 1 + \frac{z}{2} + \frac{z^2}{3} + \ldots \quad \text{and} \quad 2 - \log(1-z) = 2 + z + \frac{z^2}{2} + \frac{z^3}{3} + \ldots
\]
are admissible.

2) Consider the functions
\[
f_s(z) := (1-z)^{-s} \quad (s \in (0, 1)).
\]
They are admissible, since the binomial series yields
\[
g_s(z) = 1/f_s(z) = (1-z)^s = \sum_{n=0}^{\infty} \alpha_n^{(s)} z^n
\]
with $\alpha_0^{(s)} = 1$ and
\[
\alpha_n^{(s)} = \binom{s}{n} (-1)^n = \frac{(-s)(1-s) \cdots (n-1-s)}{n!} \leq 0.
\]
(One can also use Kaluza’s theorem here.)
3. The one-sided ergodic Hilbert transform

Because of its central importance we discuss the special case of logarithmic rates first, and postpone our abstract investigations to Sections 4–6.

Let $T$ be a power-bounded operator on a Banach space. Recall from the Introduction that the (unbounded) operator $H_T$, defined by

$$\text{dom}(H_T) := \left\{ x \in X \mid \sum_{k=1}^{\infty} \frac{T^k x}{k} \text{ converges} \right\}, \quad H_T x := \sum_{k=1}^{\infty} \frac{T^k x}{k}$$

is called the one-sided ergodic Hilbert transform associated with $T$. This operator has recently obtained greater attention [7, 9, 10, 12]. In particular, it was conjectured that for $x \in \text{dom} H_T$ one would have a logarithmic decay of the Cesàro means $(A_n(T)x)_{n \geq 1}$. In [20] we suggested to study the operator $H_T$ by writing

$$\sum_{k=1}^{\infty} \frac{z^k}{k} = -\log(1-z) = \frac{-\log(1-z)}{z} = \left[ 1 - \sum_{k=1}^{\infty} \frac{z^k}{k(k+1)} \right] = f(z) - h(z),$$

where $f(z) = -\log(1-z)/z$ is admissible and

$$h(z) = 1 - \sum_{k=1}^{\infty} \frac{1}{k(k+1)} z^k \in \mathcal{A}_1^1(\mathbb{D}),$$

so that $h(T)$ is bounded. By Lemma 2.2.b, $f(T)$ is defined within the extended functional calculus if and only if $-\log(I-T)$ is, with the same domains. By Lemma 2.2.a this is the case if and only if $g(T)$ is injective, and this is the case if and only if $\ker(I-T) = \{0\}$ (cf. Lemma 6.4 below). In this case the operator $\log(I-T)$ can also be characterized as the generator of the holomorphic semigroup of fractional powers $(I-T)^s)_{\Re s > 0}$, cf. [20, p.273]. It has been shown independently in [20, Theorem 6.2] and [10] that if moreover $\text{ran}(I-T) = X$, then $H_T = -\log(I-T)$ with equal domains.

We note (cf. [10]) that one also can base one’s considerations on the representation

$$-\log(1-z) = (2 - \log(1-z)) - 2 = f(z) - 2$$

with $f(z) = 2 - \log(1-z)$ being admissible. Since this appears to us as the simpler alternative, we shall use it here. The following theorem sharpens and generalizes (1.3), proved in [9, Proposition 3.4] for normal contractions on Hilbert spaces.

**Theorem 3.1.** Let $T$ be a power-bounded operator on a Banach space $X$.

a) If $x \in X$ is such that $z := \sum_{k=1}^{\infty} (1/k)T^k x$ converges weakly in $X$, then

$$\|A_n(T)x\| \leq \frac{2eM}{2 + \log n} \|z + 2x\| \quad (n \geq 1),$$

where $M := \sup_{n \geq 0} \|T^n\|$. Moreover,

$$\|A_n(T)x\| = O(1) \quad \text{as } n \to \infty.$$

b) Suppose that $T$ is mean ergodic and $z = 1$ is an accumulation point of

$$\sigma(T) \cap \{ z \in \mathbb{D} \mid \arg(1-z) \leq \theta \}$$

for some angle $\theta \in [0, \pi/2)$. Then for any sequence $(\epsilon_n)_{n \geq 1}$ with $\epsilon_n \to 0$ there is some $x \in \text{dom} H_T$ such that

$$\|A_n(T)x\| \neq O(\epsilon_n / \log n) \quad \text{as } n \to \infty.$$
We note that for fixed \( x \in X \) the weak convergence of the series

\[
\sum_{k=1}^{\infty} \frac{1}{k} T^k x
\]
is equivalent to its strong convergence, cf. [10, Theorem 3.2] or [20, Theorem 6.2]; and the strong convergence clearly implies \( \lim_{n \to \infty} A_n(T)x = 0 \) by Kronecker’s lemma, cf. [16, p. 203]. The novelty of Theorem 3.1 is that a specific rate for this latter convergence is identified and that this rate is recognized as optimal under mild spectral conditions.

We are now going to sketch the proof of Theorem 3.1 for those readers who are mainly interested in this theorem and not so much in our more abstract considerations. However, those are nevertheless essential, because the Taylor coefficients of the relevant functions

\[
-\frac{z}{\log(1-z)} \quad \text{or} \quad \frac{1}{2 - \log(1-z)}
\]
are not known explicitly. Accordingly, we shall have to refer to results from later sections at some stages of the proof.

The proof proceeds in several steps. In the first step we replace the operator \( H_T \) by \( 2I + H_T \) and consider the admissible function \( f(z) := 2 - \log(1-z) \). Let us write

\[
g(z) = 1/f(z) = \frac{1}{2 - \log(1-z)} = 1/2 - \sum_{n=1}^{\infty} \gamma_n z^n.
\]

By admissibility of \( f \) one has \( \gamma_n \geq 0 \) for all \( n \geq 1 \). And since \( f(1) = \infty \) one also has \( \sum_{n=1}^{\infty} \gamma_n = 1/2 \).

Now suppose that \( z = \sum_{n=1}^{\infty} (1/n)T^n x \) weakly and let \( y := 2x + z \). In the second step we observe that then \( x = g(T)y \). Indeed, this holds by letting \( c = g \) in Theorem 6.1 below. It follows that we can estimate

\[
\|A_n(T)x\| = \|A_n(T)g(T)y\| = \|(A_n g)(T)y\| \leq M \|y\| \|A_n g\|_{A^*_1(D)}
\]

by the functional calculus, where \( M := \sup_{n \geq 0} \|T^n\| \).

In the third step we estimate \( \|A_n g\|_{A^*_1(D)} \); this is done by computing the Taylor coefficients of the function \( A_n g \), and then use the positivity of the \( \gamma_n \) to compute the \( A^*_1(D) \)-norm. It turns out that one has

\[
\|A_n g\|_{A^*_1(D)} = 2r_n \quad \text{with} \quad r_n = \sum_{k=1}^{\infty} \min(1, k/n) \gamma_k \quad (n \in \mathbb{N}),
\]
see Lemma 4.2 below.

This yields a rate, but this rate is not very explicit. Thus in the fourth step we prove that

\[
(3.2) \quad \frac{r_n}{e} \leq \frac{1}{2 + \log n} \leq 2r_n.
\]

To this aim, write \( z_n := 1 - 1/n \) and note that

\[
g(z_n) = \frac{1}{2 - \log(1-z_n)} = \frac{1}{2 + \log n} \quad (n \in \mathbb{N}).
\]
To establish (3.2) we write
\[ g(z) = \frac{1}{2} - \sum_{k=1}^{\infty} \gamma_k z^k = \sum_{k=1}^{\infty} \gamma_k - \sum_{k=1}^{\infty} \gamma_k z^k = \sum_{k=1}^{\infty} \gamma_k(1 - z^k) = \sum_{k=1}^{n} \gamma_k(1 - z^k) + \sum_{k=n+1}^{\infty} \gamma_k(1 - z^k); \]
leading to \(|g(z)| \leq |1 - z| \sum_{k=1}^{n} \gamma_k k + 2 \sum_{k=n+1}^{\infty} \gamma_k \) for \(|z| \leq 1\).

Inserting \(z_n = 1 - 1/n\) yields \(|g(z_n)| \leq 2r_n\). For the converse estimate we use that \(z_n^j \geq z_n^k\) for \(j \leq k\) and hence
\[ g(z_n) = \sum_{k=1}^{\infty} \gamma_k(1 - z_n^k) \geq \sum_{k=1}^{n} \gamma_k(1 - z_n^k) + (1 - z_n^n) \sum_{k=1}^{\infty} \gamma_k. \]

Now, for \(1 \leq k \leq n\) one has \(1 - z_n^k = (1 - z_n) \sum_{j=0}^{k-1} z_n^j \geq (k/n)z_n^{n-1}\) and one can see by elementary calculus that \(z_n^{n-1} \geq 1/e\) for \(n \geq 2\).

At this stage, the first part of Theorem 3.1.a is proved. For the second part one uses the fact that
\[ nr_n \sim \frac{n}{2 + \log n} \to \infty \quad \text{as } n \to \infty. \]
Hence \(r_n^{-1} A_n(T)g(T)(I - T) = (nr_n)^{-1}g(T)(I - T^n) \to 0\) in norm. A density argument yields \(r_n^{-1} A_n(T)g(T)w \to 0\) for all \(w \in \text{ran}(I - T)\). But since \(A_n(T)x \to 0\), one has \(x \in \text{ran}(I - T)\), and since this space is \(T\)-invariant, it follows that also \(y \in \text{ran}(I - T)\). This completes the proof of Theorem 3.1.a.

The proof of part b) is the fifth step. The estimate
\[ ||A_n(T)g(T)|| \geq \sup_{\lambda \in \sigma(A_n(T)g(T))} |\lambda| \]
and the spectral inclusion \((A_n g)(\sigma(T)) \subseteq (A_n(T)g(T))\) (Lemma 2.1) together imply that
\[ ||A_n(T)g(T)|| \geq \sup_{\lambda \in \sigma(T)} ||A_n g(\lambda)||. \]
Under suitable assumptions on the spectrum of \(T\) one then obtains lower bounds of the form \(||A_n(T)g(T)|| \geq c r_n\), and the principle of uniform boundedness shows that \(\text{ran}(g(T))\) realizes the rate \((r_n)_n\), cf. Definition 1.3. (See Theorem 5.3 for details of this argument.) To finally obtain Theorem 3.1.b in its actual formulation, one has to note again that \(\text{dom}(H_T) = \text{ran}(g(T))\) for a mean-ergodic operator \(T\).

**Remarks 3.2.**

1) Note that the spectral assumptions in b) are in particular satisfied if \(T\) is a non-invertible isometry, e.g. the one-sided shift on \(\ell^p\). Unfortunately, it remains open whether one has optimality, i.e., the conclusion in b), also under weaker spectral assumptions. In particular, in the case of a unitary operator \(T\) on a Hilbert space such that 1 is an accumulation point of \(\sigma(T) \cap \mathbb{T}\).

2) The recent paper [20] when dealing with \(H_T\) suffers at some places from a confusion of signs. For example, it was written sometimes \(H_T = \log(I - T)\) instead of \(H_T = -\log(I - T)\) (e.g., twice on p.266, in Theorem 6.2 and on page 284) and also the formula (6.2) linking \(-\log(1 - z)\) and \(-\log(1 - z)/z\).
on page 283 contains a wrong sign. However, these mistakes are inessential for the results and can easily be corrected.

4. Establishing and Constructing Rates

In the remaining part of this paper, we develop the theory of rates for the Cesàro means on subspaces of the form \( \text{ran}(g(T)) \), where \( f = 1/g \) is an admissible function. The case \( f(1) < \infty \) is uninteresting here, since then \( f \in A^1_1(\mathbb{D}) \) and \( \text{ran}(g(T)) = X \) is the whole space; and according to the discussion in the Introduction one cannot expect a general convergence rate for \((A_n(T)x)_{n \geq 1}\). So it is reasonable to restrict the considerations to the case that \( f(1) = \infty \), i.e., \( g(1) = 0 \). This means that one has

\[
g(z) = \gamma_0 - \sum_{k=1}^{\infty} \gamma_k z^k, \quad \gamma_k \geq 0 \quad (k \geq 1), \quad \sum_{k=1}^{\infty} \gamma_k = \gamma_0 > 0,
\]

which will be our standard assumptions for \( g \). We note first the following fact.

**Lemma 4.1.** Let \( f \) be admissible with \( f(1) = \infty \), so that \( g := 1/f \) satisfies \((4.1)\). Then \( \text{ran}(g(T)) \subseteq \text{ran}(I - T) \) for every power-bounded operator \( T \) on a Banach space.

**Proof.** One has

\[
g(z) = \lim_{n \to \infty} \sum_{k=1}^{n} \gamma_k (1 - z^k) = \lim_{n \to \infty} (1 - z) \sum_{k=1}^{n} \gamma_k \sum_{j=0}^{k-1} z^j
\]
as a limit in \( A^1_1(\mathbb{D}) \).

By Lemma 4.1 and the mean ergodic theorem, one has \( A_n(T)x \to 0 \) as \( n \to \infty \) whenever \( x \in \text{ran}(g(T)) \). We shall see below that this convergence even happens with a certain rate, which is identified in the following lemma.

**Lemma 4.2.** Let \( f \) be admissible such that \( f(1) = \infty \), and let \( g = 1/f \), i.e., \( g \) satisfies \((4.1)\). Then

\[
\|A_n \cdot g\|_{A^1_1(\mathbb{D})} = 2r_n \quad (n \in \mathbb{N}),
\]

where

\[
(4.2) \quad r_n = r_n[g] := \frac{1}{n} \sum_{k=1}^{n} \gamma_k + \sum_{k>1}^{n} \gamma_k = \sum_{k=1}^{\infty} \min(k/n, 1) \gamma_k.
\]

**Proof.** For \( z \in \mathbb{D} \) we have

\[
h_n(z) := nA_n(z)g(z) = (1 + z + \cdots + z^{n-1}) g(z)
\]

\[
= \left( \sum_{j=0}^{n-1} z^j \right) \cdot \left( \gamma_0 - \sum_{k=1}^{\infty} \gamma_k z^k \right) = \gamma_0 \sum_{j=0}^{n-1} z^j - \sum_{j=0}^{n-1} \sum_{k=1}^{j} \gamma_k z^{k+j}
\]

\[
= \gamma_0 \sum_{j=0}^{n-1} z^j - \sum_{j=0}^{n-1} \sum_{k=1}^{\infty} \gamma_k z^k = \gamma_0 \sum_{k=0}^{n-1} z^k - \sum_{j=0}^{n-1} \sum_{k=j}^{\infty} \gamma_k z^{k+j} - \sum_{j=0}^{n-1} \sum_{k=j+1}^{\infty} \gamma_k z^k
\]

\[
= \sum_{k=0}^{n-1} \left( \gamma_0 - \gamma_k \sum_{j=0}^{k-1} \gamma_{k-j} \right) z^k - \sum_{k=0}^{n-1} \left( \sum_{j=0}^{k} \gamma_{k-j} \gamma_j \right) z^k
\]

\[
= \sum_{k=0}^{n-1} \left( \sum_{j=k+1}^{\infty} \gamma_j \right) z^k - \sum_{k=0}^{n-1} \left( \sum_{j=k+1}^{n} \gamma_j \right) z^k.
\]
Since $h_n(1) = ng(1) = 0$, we must have
\[ \sum_{k=n}^{\infty} \sum_{j=k-n+1}^{k} \gamma_j = \sum_{k=0}^{n-1} \sum_{j=k+1}^{\infty} \gamma_j \]
and hence
\[ \|h_n\|_{A^+_1(\mathbb{D})} = \sum_{k=0}^{n-1} \sum_{j=k+1}^{\infty} \gamma_j + \sum_{k=n}^{\infty} \sum_{j=k-n+1}^{k} \gamma_j = 2 \sum_{k=0}^{n-1} \sum_{j=k+1}^{\infty} \gamma_j = 2 \sum_{j=1}^{\infty} \min(j,n) \gamma_j = 2nr_n. \]

Before proceeding further, it is natural to address the question which sequences $r = (r_n)_{n \geq 1}$ arise as $r = r[g]$ (4.2) for some $g$ satisfying (4.1). This will allow us to construct functions $g$ giving rise to prescribed rates of decay of Cesàro means, cf. Theorem 4.4 below.

**Theorem 4.3.** For a sequence $(r_n)_{n \geq 1} \subseteq \mathbb{R}$ the following assertions are equivalent:

(i) $r_1 > 0$ and $(n+1)r_{n+1} - nr_n \searrow 0$;

(ii) There exists a function $g \in A^+_1(\mathbb{D})$ satisfying (4.1) such that $r = r[g]$ as defined in (4.2).

Moreover, if (i) and (ii) are satisfied, then the following assertions are true:

a) $r_n \searrow 0$ as $n \to \infty$ and $nr_n$ increases with $n$.

b) If $nr_n = O(1)$ as $n \to \infty$, then there is $h \in A^+_1(\mathbb{D})$ such that $g(z) = (1-z)h(z)$.

**Proof.** Suppose that (i) holds. Define $\alpha_n := (n+1)r_{n+1} - nr_n$ for $n \geq 0$, assuming that $\alpha_0 = r_1 > 0$. By hypothesis, $\alpha_n \searrow 0$. Define $\gamma_k := \alpha_{k-1} - \alpha_k \geq 0$ for $k \geq 1$. Then $\sum_{k=1}^{\infty} \gamma_k = \lim_{n \to \infty} (\alpha_0 - \alpha_n) = \alpha_0 = r_1$. Finally

\[ \sum_{k=1}^{\infty} \min(k,n) \gamma_k = \frac{1}{n} \sum_{k=1}^{n} \sum_{j=1}^{k} \gamma_k + \sum_{k=n+1}^{\infty} \gamma_k = \frac{1}{n} \sum_{j=1}^{n} \sum_{k=j}^{\infty} \gamma_k + \alpha_n = \frac{1}{n} \sum_{j=1}^{n} (\alpha_{j-1} - \alpha_n) + \alpha_n = \frac{1}{n} \sum_{j=0}^{n-1} \alpha_j = \frac{1}{n} nr_n = r_n. \]

Hence if we let $\gamma_0 := r_1$ and $g(z) = \gamma_0 - \sum_{k \geq 1} \gamma_k z^k$, the function $g$ satisfies (4.1) and $r = r[g]$.

Conversely, suppose that (ii) holds. Then

\[ r_n = \sum_{n=1}^{\infty} \min(k,n+1) \gamma_k \]

where $\gamma_k \geq 0$ and $0 < \gamma_0 = \sum_{k \geq 1} \gamma_k < \infty$. In particular, $r_1 = \gamma_0 > 0$. Since the sequence of functions $u_n(\cdot) := \min(\cdot/n, 1)$ decreases to zero pointwise on $\mathbb{R}_+$, also $r_n \searrow 0$ by (an elementary version of) the monotone convergence theorem. Analogously, since

\[ \min(k,n+1) - \min(k,n) = 1_{[n+1,\infty)}(k) \quad (k \in \mathbb{N}) \]
and 1_[[\[n+1,\infty\) \setminus \{0\}. It follows that \((n+1)r_{n+1} - nr_n \to 0\). In particular, \((n+1)r_{n+1} - nr_n \geq 0\), which means that \(nr_n\) increases with \(n\). This concludes the proof of (i) and (a).

To prove b), suppose that \(nr_n = O(1)\). Then since \(nr_n \geq \sum_{k=1}^n k\gamma_k\), we conclude that \(\sum_{k \geq 1} k\gamma_k < \infty\). Hence we can write

\[
g(z) = \gamma_0 - \sum_{k=1}^{\infty} \gamma_k z^k = \sum_{k=1}^{\infty} \gamma_k (1 - z^k) = (1 - z) \sum_{k=1}^{\infty} \gamma_k \sum_{j=0}^{k-1} z^j
\]

and \(h(z) \in A_+^1(\mathbb{D})\) since 

\[
\|h\|_{A_+^1(\mathbb{D})} = \sum_{j=0}^{\infty} \sum_{k=j+1}^{\infty} \gamma_k = \sum_{k=1}^{\infty} k\gamma_k < \infty.
\]

Assertion b) from Theorem 4.3 implies that if \(nr_n = O(1)\), then for a power-bounded operator \(T\) we have \(g(T) = (I - T)h(T)\), whence \(\text{ran}(g(T)) \subseteq \text{ran}(I - T)\). No new information concerning rates is obtained by allowing this case, so we can restrict to the case that \(nr_n \to \infty\). Then we have the following theorem.

**Theorem 4.4.** Let \(g(z) = \gamma_0 - \sum_{k=1}^{\infty} \gamma_k z^k\) as in (4.1) and let \(r = r[g]\) as in (4.2). If \(T\) is a power-bounded operator on a Banach space \(X\), then

\[
\|A_n(T)x\| = O(r_n) \quad \text{as} \quad n \to \infty
\]

for every \(x \in \text{ran}(g(T))\). And if \(nr_n \to \infty\) and \(T\) is mean ergodic, then

\[
\|A_n(T)x\| = o(r_n) \quad \text{as} \quad n \to \infty
\]

for every \(x \in \text{ran}(g(T))\).

**Proof.** The first assertion follows from the functional calculus estimate

\[
\|A_n(T)g(T)\| \leq M(T) \|A_n g\|_{A_+^1(\mathbb{D})} \leq 2M(T)r_n \quad (n \in \mathbb{N}),
\]

where \(M(T) := \sup_{n \geq 0} \|T^n\|\). Now suppose that \(T\) is mean ergodic and \(nr_n \to \infty\). If \(y - Ty = (I - T)y = 0\), then \(g(T)y = g(1)y = 0\). And if \(y = (I - T)z \in \text{ran}(I - T)\), then

\[
r_n^{-1}A_n(T)g(T)y = r_n^{-1}A_n(T)g(T)(I - T)z = (nr_n)^{-1}g(T)(I - T^n)z \to 0
\]

since \(nr_n \to \infty\). Since \(\sup_n r_n^{-1} \|A_n(T)g(T)\| < \infty\), it follows by density that \(r_n^{-1}A_n(T)g(T)y \to 0\) for all \(y \in \overline{\text{ran}}(I - T) \oplus \ker(I - T) = X\), by mean ergodicity of \(T\). \(\square\)

The following example shows that for a general power-bounded operator \(T\) the rate \(r[g] = (r_n)_{n \geq 1}\) is optimal and one does not have the \(o(r_n)\)-rate without an extra assumption.

**Example 4.5.** Let \(X = A_+^1(\mathbb{D})\), and let \(T\) be the “forward shift”, i.e.,

\[
(Tf)(z) := zf(z) \quad (z \in \mathbb{D}).
\]

Then \(T\) is clearly a non-invertible isometry, and it is easily seen that \(g(T)1 = g\) for any \(g \in A_+^1(\mathbb{D})\). In particular, if \(f\) is admissible and \(g = 1/f\) then

\[
\|A_n(T)g(T)1\| = \|A_n g\|_{A_+^1(\mathbb{D})} = 2r_n \quad (n \in \mathbb{N}).
\]

In particular, \(\|A_n(T)g(T)1\| \neq o(r_n)\).
5. Realizing Rates

So far, the rate \((r_n)_{n \geq 1}\) is given in terms of the Taylor coefficients of \(g\). However, in situations of interest these are often unknown, so it seems desirable to be able to read off the \(r_n\) (or at least their asymptotic behaviour) from the values of \(f\) or \(g\) at certain points. To achieve this, we begin with an elementary lemma.

Lemma 5.1. Let \(0 \leq \alpha \leq 1\) and \(n \in \mathbb{N}\). If \(z \in \mathbb{D}\) is such that

\[
|1 - z| \leq \frac{1}{n} \quad \text{and} \quad 1 - |z| \geq \frac{\alpha}{n}
\]

then \(1 - |z|^k \geq (\alpha/e) \min(k/n, 1)\) for all \(k \geq 1\).

Proof. If \(k \geq n\) then \(|z|^k \leq |z|^n\) and hence \(1 - |z|^k \geq 1 - |z|^n\). Therefore it suffices to consider \(1 \leq k \leq n\). In this case

\[
1 - |z|^k = (1 - |z|) \sum_{j=0}^{k-1} |z|^j \geq (1 - |z|)k|z|^{n-1}
\]

\[
\geq (\alpha k/n) (1 - |z|)^n - 1 \geq (\alpha k/n)(1 - (1/n))^{n-1} \geq \frac{\alpha k}{en}
\]

since

\[
\inf_{t > 0} \left( 1 - \frac{1}{t + 1} \right)^t = \lim_{t \to \infty} \left( 1 - \frac{1}{t + 1} \right)^t = \frac{1}{e}
\]

\(\square\)

Let \(g(z) = \gamma_0 - \sum_{k=1}^{\infty} \gamma_k z^k\) satisfy our standard assumptions (4.1), and let \(r = r[g]\) as in (4.2). Then we can write

\[
g(z) = \sum_{k=1}^{\infty} \gamma_k(1 - z^k) = \sum_{k=1}^{n} \gamma_k(1 - z) \sum_{j=0}^{k-1} z^j + \sum_{k=n+1}^{\infty} \gamma_k(1 - z^k)
\]

for all \(z \in \overline{D}\) and all \(n \geq 1\). Hence if we know that \(|1 - w| \leq \beta/n\) for some \(w \in \overline{D}\), \(n \in \mathbb{N}\) and \(\beta > 0\), then we can estimate

\[
|g(w)| \leq \sum_{k=1}^{n} \gamma_k |1 - w| \sum_{j=0}^{k-1} |w|^j + \sum_{k=n+1}^{\infty} \gamma_k |1 - w^k|
\]

\[
\leq \beta \sum_{k=1}^{n} (k/n) \gamma_k + 2 \sum_{k=n+1}^{\infty} \gamma_k \leq \max(\beta, 2) r_n.
\]

And if we know that \(|1 - w| \leq 1/n\) and \(1 - |w| \geq \alpha/n\), then

\[
|g(w)| \geq \text{Re} g(w) = \sum_{k=1}^{\infty} \gamma_k (1 - \text{Re} z^k) \geq \sum_{k=1}^{\infty} \gamma_k (1 - |z|^k) \geq (\alpha/e) r_n
\]

by Lemma 5.1. So by choosing \(w = w_n\) in a suitable way, we are able to read off the rate \(r = r[g]\) from the values \(|g(w_n)|\), as it is stated in the next theorem.

Theorem 5.2. Let \(f\) be admissible such that \(f(1) = \infty\), let \(g = 1/f\) and let \(r = r[g]\) be the associated rate sequence as in (4.2). If \(0 < \alpha \leq 1\) and \((w_n)_{n \in \mathbb{N}} \subseteq D\) are such that \(|1 - w_n| \leq 1/n\) and \(1 - |w_n| \geq \alpha/n\) for all \(n\), then

\[
(1/2) |g(w_n)| \leq r_n \leq (e/\alpha) |g(w_n)| \quad (n \in \mathbb{N}).
\]
So Theorem 4.4 gives an asymptotic estimate for the Cesàro averages $A_n(T)x$, when $x \in \text{ran}(g(T))$, and Theorem 5.2 allows to identify $r_n \sim g(1 - 1/n)$, for instance. The obvious question now is, under which conditions on $T$ the rate $r[g]$ is actually optimal on $\text{ran}(g(T))$. That an extra condition is needed is shown by letting $T = 0$; then $g(T) = I$ and $A_n(T)x = (1/n)x$ for all $n \geq 1$. It seems therefore reasonable to require that $T$ has spectrum on a set where we can reconstruct the rate $r[g]$ from values of $g$. The following is the best we can do at this moment.

**Theorem 5.3.** Let $f$ be an admissible function on $\mathbb{D}$ such that $f(1) = \infty$, let $g = 1/f$ and let $r[g] = (r_n)_{n \geq 1}$ the associated rate sequence. Let $T$ be a power-bounded operator on a Banach space $X$. Suppose in addition that $z = 1$ is an accumulation point of

$$\sigma(T) \cap \{ z \in \mathbb{D} \mid \arg(1 - z) \leq \theta \}$$

for some angle $\theta \in [0, \pi/2)$. Then $g(T)$ realizes the rate $(r_n)_{n \geq 1}$ in the sense of Definition 1.3.

**Proof.** We need some preliminary geometrical considerations. If $z \in \mathbb{D}$ and $\varphi := \arg(1 - z) \leq \theta$, then

$$- \text{Im} z = \text{Im}(1 - z) = |1 - z| \sin \varphi$$

and

$$1 - \text{Re} z = \text{Re}(1 - z) = |1 - z| \cos \varphi.$$

This yields

$$|z|^2 = |1 - z|^2 \sin^2 \varphi + (1 - |1 - z| \cos \varphi)^2$$

$$= 1 + |1 - z| (|1 - z| - 2 \cos \varphi) \leq 1 + |1 - z| (|1 - z| - 2 \cos \theta).$$

Hence

$$1 - |z| \geq 1 - |z|^2 \geq |1 - z| (2 \cos \theta - |1 - z|).$$

By hypothesis there is an infinite set $J \subseteq \mathbb{N}$ such that for each $n \in J$ one has $\cos \theta \geq 1/n$ and one finds $z_n \in \sigma(T)$ such that

$$\arg(1 - z_n) \leq \theta \quad \text{and} \quad \frac{1}{2n} \leq |1 - z_n| \leq \frac{1}{n}.$$

By the geometrical considerations from above we conclude that

$$1 - |z| \geq \frac{1}{2n} (2 \cos \theta - \frac{1}{n}) \geq \frac{\cos \theta}{2n}$$

for each $n \in J$. Hence, by the spectral inclusion Lemma 2.1 and Lemma 5.1 one has

$$\|A_n(T)g(T)\| = \|(A_n \cdot g)(T)\| \geq \sup_{\lambda \in \sigma((A_n \cdot g)(T))} |\mu|$$

$$\geq \sup_{\lambda \in \sigma(T)} |(A_n \cdot g)(\lambda)| \geq |A_n(z_n)g(z_n)| = \frac{|1 - z_n^n|}{n |1 - z_n|} |g(z_n)|$$

$$\geq \frac{1}{n |1 - z_n|} |g(z_n)| \geq \frac{\cos \theta}{2e} \cdot \frac{n}{2e} \cdot \frac{\cos \theta}{2e} r_n = \left(\frac{\cos \theta}{2e}\right)^2 r_n$$

for each $n \in J$. Consequently, if $(\epsilon_n)_{n \geq 1}$ is any positive sequence such that $\epsilon_n \to 0$, then the family of operators

$$\frac{1}{\epsilon_n r_n} A_n(T)g(T), \quad n \in \mathbb{N},$$
is not uniformly norm bounded. By the principle of uniform boundedness there
must exist a vector \( y \in X \) such that
\[
\sup_{n \in \mathbb{N}} \frac{1}{c_n r_n} \| A_n(T)g(T)y \| = \infty.
\]
Hence, with \( x = g(T)y \in \text{ran } g(T) \) we have
\[
\| A_n(T)x \| \neq O(\epsilon_n r_n) \quad \text{as } n \to \infty,
\]
as was to be shown. \( \square \)

6. Convergence of power series and associated rates

Finally, let us return to our starting point, Problem 1.2.

**Theorem 6.1.** Let \( f = \sum_{k=0}^{\infty} \alpha_k z^k \) be a holomorphic function on the open unit disc \( \mathbb{D} \), and let \( T \) be a power-bounded operator on a Banach space \( X \). Suppose that \( x \in X \) is such that \( y := \sum_{k=0}^{\infty} \alpha_k T^k x \) converges weakly in \( X \).

If \( e \in A^1_+(\mathbb{D}) \) is such that \( (ef) \in A^1_+(\mathbb{D}) \), too, then \( e(T)y = (ef)(T)x \).

**Proof.** Let \( C > 0 \) be such that
\[
\left\| \sum_{m=0}^{\infty} \alpha_k T^k x \right\| \leq C
\]
for all \( 0 \leq n \leq m \). Write \( e(z) = \sum_{j=0}^{\infty} \beta_j z^j \). Then
\[
(ef)(z) = \sum_{n=0}^{\infty} \left( \sum_{j=0}^{\infty} \alpha_{n-j} \beta_j \right) z^n \quad (z \in \mathbb{D}).
\]
Since \( \sum_{j=0}^{M} \beta_j T^j \to e(T) \) in operator norm, it follows that
\[
\left( \sum_{j=0}^{M} \beta_j T^j \right) \left( \sum_{k=0}^{M} \alpha_k T^k x \right) \to e(T)y \quad \text{weakly, as } M \to \infty.
\]
On the other hand,
\[
\left( \sum_{j=0}^{M} \beta_j T^j \right) \left( \sum_{k=0}^{M} \alpha_k T^k x \right) = \sum_{k,j=0}^{M} \alpha_j \beta_k T^{j+k} x
\]
\[
= \sum_{n=0}^{M} \left( \sum_{j=0}^{n} \alpha_j \beta_{n-j} \right) T^n x + \sum_{j=0}^{M} \beta_j T^j \sum_{k=M-j+1}^{M} \alpha_k T^k x.
\]
The first summand tends to \( (ef)(T)x \) in norm, so it suffices to show that the second summand tends weakly to zero as \( M \to \infty \). Fix \( \epsilon > 0 \) and find \( M_0 \in \mathbb{N} \) such that \( \sum_{j>M_0} |\beta_j| \leq \epsilon \). Then write
\[
\sum_{j=0}^{M} \beta_j T^j \sum_{k=M-j+1}^{M} \alpha_k T^k x = \sum_{j=0}^{M_0} \beta_j T^j \sum_{k=M-j+1}^{M} \alpha_k T^k x
\]
\[
+ \sum_{j=M_0+1}^{M} \beta_j T^j \sum_{k=M-j+1}^{M} \alpha_k T^k x.
\]
The first summand on the right-hand side tends weakly to zero as \( M \to \infty \), since \( (\sum_{k=0}^{M} \alpha_k T^k x)_{m \geq 0} \) is weakly Cauchy; the second has norm less than \( C\epsilon \). This concludes the proof. \( \square \)
Remark 6.2. One should compare Theorem 6.1 with [20, Theorem 5.1]. Suppose \( f \) is as in Theorem 6.1 and suppose in addition that \( f \) is regularizable, i.e., that there is \( e \in A^1_+(\mathbb{D}) \) such that \( (ef) \in A^1_+(\mathbb{D}) \) and \( e(T) \) is injective. Hence Theorem 6.1 shows that

(6.1) \( \sum_{k=0}^{n} \alpha_k T^k x \to y \) weakly \( \Rightarrow x \in \operatorname{dom}(f(T)), f(T)x = y. \)

In [20, Theorem 5.1] we needed that \( (1 - z)f \in A^1_+(\mathbb{D}) \), but could do under the much weaker hypothesis of weak Cesàro-summability of the series \( \sum_{k=0}^{\infty} \alpha_k T^k x \).

However, within the more general approach to abstract functional calculi developed in [19], Theorem 6.1 yields that the implication (6.1) holds true (with \( f(T) \) denoting the so-called upper-extension \( \Phi(f) \) introduced in [19, Section 2.1]) even without the assumption that \( f \) is regularizable.

For an admissible function \( f \), we can take \( e := 1/f \) in Theorem 6.1 and combining it with Theorems 4.4 and 5.2 we obtain the following corollary.

**Corollary 6.3.** Let \( f(z) = \sum_{k=0}^{\infty} \alpha_k z^k \) be an admissible function with \( f(1) = \infty \). Let \( T \) be a power-bounded operator on a Banach space \( X \), let \( M := \sup_{n \geq 0} \| T^n \| \), and let \( x \in X \) be such that \( y := \sum_{k=0}^{\infty} \alpha_k T^k x \) converges (weakly).

Then \( x = (1/f)(T)y \) and \( \| A_n(T)x \| \leq \frac{2eM}{f(1 - 1/n)} \| y \| \) (\( n \in \mathbb{N} \)). If in addition \( n/f(1 - 1/n) \to \infty \) as \( n \to \infty \), then

\[ \| A_n(T)x \| = o \left( \frac{1}{f(1 - 1/n)} \right) \quad \text{as } n \to \infty. \]

**Proof.** We let \( e := 1/f \) and apply Theorem 6.1 to conclude that \( x = g(T)y \). Then the estimate

\[ \| A_n(T)x \| \leq \frac{2eM}{f(1 - 1/n)} \| y \| \]

follows readily from the proof of Theorem 4.4 and from Theorem 5.2. Now suppose that \( n/f(1 - 1/n) \to \infty \). Then, by Theorem 5.2, \( nr_n \to \infty \). Furthermore, \( x \in Y := \operatorname{ran}(I - T) \) (Lemma 4.1) and \( Y \) is \( T \)-invariant, so \( y \in Y \) as well. This means that we can suppose without loss of generality that \( T \) is mean ergodic. Hence the second part of Theorem 4.4 yields that

\[ \| A_n(T)x \| = o(r_n) = o(1/f(1 - 1/n)) \quad \text{as } n \to \infty, \]

again by Theorem 5.2. \( \square \)

Of course it is of interest when \( f \) as in Corollary 6.3 is regularizable, i.e., when \( (1/f)(T) \) is injective. Here we have the following result, suggested to us by Michael Lin.

**Lemma 6.4.** Let \( g \in A^1_+(\mathbb{D}) \) satisfy (4.1) and let \( T \) be a power-bounded operator on a Banach space \( X \). Then \( \ker g(T) = \ker(I - T^N) \), where

\[ N := \gcd\{k \in \mathbb{N} \mid \gamma_k \neq 0\}, \]

where \( \gcd A \) denotes the greatest common divisor of the elements of a set \( A \subseteq \mathbb{N} \).
Proof. Without loss of generality we may suppose that $\gamma_0 = 1$. If $T^N x = x$ then $\gamma_k T^k x = \gamma_k x$ for all $k \in \mathbb{N}$ and hence $g(T)x = x - \sum_{k=1}^{\infty} \gamma_k T^k x = x - \sum_{k=1}^{\infty} \gamma_k x = 0$. To prove the converse, let $Y := \ker g(T)$ and $S := \left. T \right|_{Y}$. Then $\sum_{k=1}^{\infty} \gamma_k S^k = I$. By a well-known result of Kakutani [24, Chapter 2, Lemma 1.13], $I$ is an extreme point of the unit ball of $\mathcal{L}(Y)$. Consequently, $S^k = I$ whenever $\gamma_k \neq 0$. The Euclidean algorithm yields that if $n, m \in \mathbb{N}$ and $S^n = S^m = I$, then $S^{\gcd(n,m)} = I$. Iterating this observation we finally arrive at $S^N = I$, which means that $\ker g(T) \subseteq \ker (I - T^N)$.

Acknowledgements

The authors are indebted to Michael Lin for his very careful reading of an earlier version of this paper and his useful and inspiring comments. The author would also like to thank the referee for his/her valuable suggestions.

References


Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, ul. Chopina 12/18, 87-100 Toruń, Poland, and Institute of Mathematics, Polish Academy of Sciences, Śniadeckich 8, 00-956 Warszawa, Poland

E-mail address: gomilko@mat.uni.torun.pl

Delft Institute of Applied Mathematics, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands

E-mail address: m.h.a.haase@tudelft.nl

Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, ul. Chopina 12/18, 87-100 Toruń, Poland, and Institute of Mathematics, Polish Academy of Sciences, Śniadeckich 8, 00-956 Warszawa, Poland

E-mail address: tomilov@mat.uni.torun.pl